GaThErInG 4 GArDNeR 11

G4G11 Exchange Book

VOLUME 1

ATLANTA, GEORGIA

March 19 - 23, 2014

Published by:
Gathering 4 Gardner

G4G11 Exchange Book

VOLUME 1

The Gift Exchange is an integral part of the Gathering 4 Gardner biennial conferences. Gathering participants exchange gifts, papers, puzzles and other interesting artifacts. This book contains gift exchange papers from the conference held in Atlanta, Georgia from Wednesday, March 19th to Sunday, March 23rd, 2014. It combines all of the papers offered as exchange gifts in two volumes.

Acknowledgments

Gathering 4 Gardner would like to offer thanks to the following individuals:

Freddy Bendekgey – for editing and laying out the pages of this book
Katie LaSeur – for project management and art direction
John Miller – for helping manage the papers and following up with authors
Nancy Blachman – for getting this book started

There are many things you can buy or make yourself, but there are some things that can only be created by a group. This book is such an item.

Preface

by John Miller

G4G11 Themes were John Horton Conway, 11, and parallel lines (||).

Papers mentioning or based on 11 were: ***Elevenses—Being 11 Easy Teasers for G4G11*** by Chris Maslanka; ***Langford's Problem, Remixed*** by John Miller; and ***E1leven*** a dice game, by Cary Staples.

Sand Painting by Gary Greenfield, and ***Still Life with Glider*** by Robert Bosch reference Conway's game of Life. ***Conway and the 3x+1 Problem Continued*** by Gary Greenfield is self-evident. ***John Horton Conway***, by Jeremiah & Karen Farrell, describes a puzzle game using the 10 different letters in Conway's name.

Many papers had accompanying presentations, listed on the G4G website via YouTube. Examples: Adam Atkinson's ***Telephone Calls and the Brontosaurus*** asks an entertaining question about the length and cost of phone calls. Dana Mackenzie's ***Sun Bin's Legacy*** relates a strategy to win a match of three horse races, and then gives a game-theoretic approach to the generalized problem. Kenneth Brecker examines some strange physics in ***The Pseudosphere Uphill Roller***. His roller was displayed in gallery where he engaged visitors.

While some papers were contributed without any presentation, many fine presentations were made without corresponding papers. In particular we note Solomon W. Golomb ***Reptile Sets of Polyominoes***. G4G11 was his last Gathering.

The game of Dots & Boxes was a popular topic. Elwyn Berlekamp gave a wonderfully intuitive analysis, of which we have the video only. Jason Colwell's paper ***A Strategy for Borders***, defines a variant, based on canals and borders.

There were a number of notably serious papers, for example: Lisa Rougetet contributed a well-researched ***Prehistory of Nim***. Ron Taylor's ***Color Addition Across the Spectrum of Mathematics*** proposes and delves into a dominoes-like game based on addition of colors. Bill Liles' ***A New Blackjack Problem*** poses strategies for holding and playing multiple hands. Robert Vallin's ***An Introduction to Gilbreath Numbers*** references Perci Diaconis, Martin Gardner, Colm Mulcahy and others.

As expected, some papers were very colorful and visual: ***Tiling Tetris Boards*** by Steve Butler, ***The Pennyhedron Revisited*** by George Bell, and ***Tilings and Geometric Problems*** by Jaap Scherphuis.

Puzzles were well represented. Among the many papers were: ***Linkage Puzzle Font*** by Erik and Martin Demaine and ***Solving Puzzles Backwards*** by Anany Levitin.

In keeping with the spirit of the Gathering, some papers were just plain fun, such as ***Dr. Matrix Talks to Owen O'Shea***.

Table of Contents | VOLUME 1

Table of Contents (cont.)

*At the author's request, this paper is not included in the online version of this publication.

MAGIC

MATH

G4G11 Schedule

Date: Thursday, 20/Mar/2014

8:30am - 10:00am	**ThuAM1: Thursday AM Early** Location: Ritz Carlton Large Meeting Room **The Conway Immobilizer** Winkler, Peter **Quintetra Blocks** Kostick, John and Jane **The Pennyhedron** Bell, George **Widespreading the Word: promoting math to a wide audience.** Gill, Eoin; Donegan, Sheila **Seven touching infinite cylinders** Bozóki, Sándor; Lee, Tsung-Lin; Rónyai, Lajos **Colliding Masses and the digits of Pi** Hess, Dick **Pythagorize the Flatiron** Lawrence, Cindy **Camoens, Pimenta and the improbable sonnet** Simões, Carlota; Coelho, Nuno
10:30am - 12:00pm	**ThuAM2: Thursday AM Late** Location: Ritz Carlton Large Meeting Room **Sinan's Screens: Networks of Intersecting Polygons in Ottoman Architecture** Bier, Carol **Rep-tile Sets of Polyominoes** Golomb, Solomon W. **Mathenaeum** Whitney, Glen **The Golden Meaning** Bellos, Alex **Wordplay / Slipperiness of Language / N-tendres** Goldklang, Lew **Slice Knots and Conway's Skein Theory** Kauffman, Louis **Crossed Stick Puzzle Design** Muñiz, Alexandre **Grille ciphers** Serra, Michael **Maybe Fair Dice** Kisenwether, Joseph
1:30pm - 3:30pm	**ThuPM1: Thursday PM Early** Location: Ritz Carlton Large Meeting Room **String Theory: (Pythagoras of Samos + Athanasius Kircher)^2 = Jimi Hendrix** Sheppard, Philip **Polyhedral Computing Applied to Spatial Puzzle Design** Alexe, Sorin

Three short introductions.
Uehara, Ryuhei

Can John Conway, Retrolife, the number 11 and a mathematical magic trick be coherently addressed in 5 minutes?
Elran, Yossi

Regular Hexaflexagon Faces
Anderson, Thomas; McLean, Thomas Bruce; Pajoohesh, Homeira; Smith, Chasen

Magic and the Periodic Table of the Elements
Duran, Miquel; Blasco, Fernando

A drawing method on Conway's game of Life
Iwai, Masayoshi

Sun Bin's Legacy
MacKenzie, Dana

What happened few minutes before day 0?
Rougetet, Lisa

Introducing Gilbreath Numbers
Vallin, Robert

The Music of the Icosahedron
Orman, Hilarie

The Level 2 Menger Sponge with Playing Cards
Wilder, Jim

My Clothes Tell Secrets
Lee, Elan

4:00pm - 5:30pm

ThuPM2: Thursday PM Late
Location: Ritz Carlton Large Meeting Room

Toy models
Tokieda, Tadashi

The Algebra of Tetraflexagons
Yackel, Carolyn; Beier, Julie

The Pseudosphere Uphill Roller
Brecher, Kenneth

Beading the Seven-Color Torus
Goldstine, Susan

Telephone Calls and the Brontosaurus
Atkinson, Adam

The State of the Art of Modern Mazes
Fisher, Adrian

11 Unexpected Parallels
Jones, Kate

The Music of the Polygons
Miller, John

Date: Friday, 21/Mar/2014

8:30am - 10:00am	**FriAM1: Friday AM Early** Location: **Ritz Carlton Large Meeting Room**
	The design of a reconfigurable maze **Kaplan, Craig**
	KenKen...The Fastest Growing Logic (and Math!) Puzzle Since Sudoku **Fuhrer, Robert**
	The Unique 4-7-11 Octahedral Monoicosahedron **Banchoff, Thomas**
	The ConSequence of Elevens in Parallel **Marasco, Joe**
	TSP Mazes **Chartier, Tim**
	Peculiar integer triangles containing an edge of 11 **Hosoya, Haruo**
	Character Assassination **Roberts, Siobhan; Conway, John**
	Menger Menger Menger **Taalman, Laura**
	Gerbert's Abacus **Silva, Jorge Nuno; Santos, Carlos**
	The Japanese theorem for cyclic polygons **Richeson, Dave**
10:30am - 11:30am	**FriAM2: Friday AM Late** Location: **Ritz Carlton Large Meeting Room**
	It is better to open a mind with wonder than close it with belief **Menna, Lisa**
	Holographic Visualization for Mathematics and Science **Newswanger, Craig**
	The Game of Light **Kocik, Jerzy {Jurek}**
	Conway and The 3x+1 Problem Continued **Greenfield, Gary**
	On a KenKen from Bit-player **Nacin, David**
	Sculpture Activities **Hart, George**
12:00pm - 5:00pm	**FriPM: Friday PM at Sarah's House**

Date: Saturday, 22/Mar/2014

8:30am - 10:00am	**SatAM1: Saturday AM Early** Location: **Ritz Carlton Large Meeting Room**
	An Irregular Hexaflexagon Schwartz, Ann
	Hyperbolic Fractal Tilings and Surfaces Fathauer, Robert
	What is the G4G - Celebration of Mind? Thompson, Tanya; Morgan, Chris
	The Programmable Galton Board: A Shameless Shill Propp, James
	World in the Balance Crease, Robert
	The World's Favourite Number Bellos, Alex
	JMA Outstanding Paper Award Kaplan, Craig
	Computer aided curved origami design Mitani, Jun
	Quintessence: Puzzling the 120-cell Schleimer, Saul; Segerman, Henry
	Revisiting the Mutilated Chessboard Wright, Colin
10:30am - 12:00pm	**SatAM2: Saturday AM Late** Location: **Ritz Carlton Large Meeting Room**
	TBD Conway, John Horton
	Math Anxiety Camp and My New Beads Fisher, Gwen Laura
	How Magicians fool our brain Hjulstad, Kristine
	The Magic Square Green, Lennart
	Marble Runs and Turing Machines Bickford, Neil
	How To Work My G8 dissection puzzle Gosper, R. William
	What if you find 115 puzzles with no solutions? Knoppers, Peter
	How Puzzles Made Us Human Mutalik, Pradeep
	Fair dice Sherman, Scott
	Dots and Boxes Berlekamp, Elwyn

Time	Session
1:30pm - 3:30pm	**SatPM1: Saturday PM Early** Location: **Ritz Carlton Large Meeting Room**

The Awesome Powers of 11
Crease, Robert; Crease, Alexander

Evileven
Oberg, Bruce

From Twisty-Puzzle Fractals to Penrose Tiles
Hearn, Bob

Optical Illusions of Theodore Deland
Mullins, Bill

Hat Puzzles
Khovanova, Tanya

Eleven = ONCE. Magic for blind people.
Blasco, Fernando; Blasco-Uceda, Fernando

John-Art: The Stochastic Geometry of John Shier
Cipra, Barry

A Deterministic Finite Automaton for determining triangle orientation in a General Order Regular Flexagon
Iacob, Emil

Solving Puzzles Backwards
Levitin, Anany

The Eleven Clocks Problem
Roby, Tom

Recent significant achievements on puzzles in Japan
Takashima, Naoaki

Proportion Systems
Harriss, Edmund

Time	Session
4:00pm - 5:30pm	**SatPM2: Saturday PM Late** Location: **Ritz Carlton Large Meeting Room**

TBD
Holman, Pablos

Geometric Surfaces as Visual Instruments
Matsuura, Akihiro

INTO MYSTERY revealed
Clough, Eric

The Martin Gardner Centennial Initiative
Mulcahy, Colm

Mathematics Awareness Month 2014 Celebrates Martin Gardner
Torrence, Eve; Torrence, Bruce; Mulcahy, Colm

Martin Gardner, yeoman/2c
Richards, Dana

Tales plus Q&A about life with Armand T. Ringer
Gardner, Jim

Date: Sunday, 23/Mar/2014

8:30am - 10:30am

SunAM1: Sunday AM Early
Location: **Ritz Carlton Large Meeting Room**

The Magic of the Superheroes of Sight
Schwab, Ivan R.

Who invented the McIntosh Apple?
McManus, Mickey

Clark Richert, Artist
Hildebrandt, Paul

Conway's Impact on the Theory of Random Tilings
Propp, James

The Neuroscience of Curiosity
Antonick, Gary

Game-of-Life Mosaics
Bosch, Robert

YESGO, an unusual Go game and its problems
Kotani, Yoshiyuki

Making a Binary Computer with 10,000 Dominoes
Parker, Matt

Symmetries in Portugal
Silva, Jorge Nuno; Carvalho, Alda; Santos, Carlos

Making a Real 5×5×5
Hoff, Carl

A Box of Invisibility
Bexfield, Simon

Superfractals
Strickland, Henry

The Quaternion Symmetry Group You've Never Heard Of
Hart, Vi

11:00am - 1:00pm

SunAM2: Sunday AM Late
Location: **Ritz Carlton Large Meeting Room**

A Tribute To Raymond Smullyan
Rosenhouse, Jason

Langford's Problem Remixed
Miller, John

A Simple MatheMagics Trick
Brittain, Skona

A Hair-Tie 120-cell
Hawksley, Andrea Johanna

The Julie Robinson Mathematics Festival
Blachman, Nancy

Magnus Popkous-Bucky Meets Borromeo Through The Amazing Geometry Machine
Esterle, Richard

Square in the Bag and Other Puzzles for Classroom
Iwasawa, Hirokazu

	PHANTASCOPES **Rowett, Tim**
	Solid Three-Dimensional Anamorphic Alice Using A Wave Mirror **Mortillaro, Karen**
	Paper Pentasia **Hayes, Barry**
	11 views of 11 **Kepner, Margaret**
2:00pm - 3:00pm	**GiftExchange: G4G11 Gift Exchange** Papers submitted as part of the G4G11 Gift Exchange
	"Conway Cup" and "John 'Horned' Conway" (2 contributions) **Allen, Amina; Stranahan, Mike**
	"Conway Cup" and "John 'Horned' Conway" (2 contributions) **Allen, Amina; Stranahan, Mike**
	"Exploring the Secrets of The Universe (in Red & White)" **Hinnant, Wm Vandorn**
	/dev/joe Crescents and Vortices **DeVincentis, Joseph**
	4 x 4 = 20 **Rowett, Tim**
	6-card Ball **Hawksley, Andrea Johanna**
	6^3 Board Burr **Knoppers, Peter**
	A cheap way to construct classic puzzles **Blasco, Fernando**
	A Cluster Analysis of Richard's PBEM Server **Hurd, Lyman**
	A Corona with Decagrams on a Late 12th Century Persian Monument **Bier, Carol**
	A Deceptive Tessellation **Gosper, R. William**
	A New Twist on an Old Puzzle **Jones, Robert**
	A Satisfiability-Based Cover Puzzle **Bailey, Duane**
	A Scanner Darkly **Hibbert, Chris**
	A Strategy for Borders -- a variant of Dots & Boxes **Colwell, Jason**
	Alternate Magic Square **Rodgers, Tyler**
	An Irregular Hexaflexagon **Schwartz, Ann**
	An sl(3) Liedoku **Nacin, David**

Basic Strategies to solve Disentanglement Puzzles
Götz, Markus

Beaded Celtic Knots: Rings, Rosettes & Beaded Beads
Fisher, Gwen Laura

Being Paid to Eat Pizza
Atkinson, Adam

Binary Computing with Domino Logic Gates on a Large Scale
Parker, Matt; Steckles, Katie

Bobby Fischer Against the World - soundtrack for G4G friends
Sheppard, Philip

Circles of 11
Nelson, Eric; Rapson, Ole

Clara's Rocket
Orndorff, Robert

Color Addition Across the Spectrum of Mathematics
Taylor, Ron

Constructing Acrostic Puzzles
Henle, Frederick Valentin

Cutting the disk: a geometric dissection puzzle
Morgan, Christopher

DecaStar lasercut notecards
Lawrence, Cindy

Dc30 Ring
Schleimer, Saul; Segerman, Henry

Deland's Optical Illusions
Mullins, Bill

Didactive Magic
Hjulstad, Kristine

e1leven .. a game.
Staples, Carolyn "Cary"

Elevenses
Kepner, Margaret

Elmo's Excellent Excursion
Hess, Dick

From Twisty Puzzle Fractals to Penrose Tiles
Hearn, Bob

G4G Tribute
Greenfield, Gary

G4G11 Cube
Bright, Lambert & Sandy

g4g11 star
Bosch, Derek

Gabriel's paper horn
Richeson, Dave

Generating music from the Thue Morse and the Kolakosi sequences
Calkin, Neil; Bowers, Brian

Geometric Proliferation
Merow, Katharine

Gift Exchange
McLean, Thomas Bruce

Hat Puzzles
Khovanova, Tanya

Hinged Triangle-Square
Taalman, Laura

John Horton Conway: A Series of Puzzles and Games
Farrell, Jeremiah; Farrell, Karen

Langford's Problem Remixed
Miller, John

Linkage Puzzle Font
Demaine, Erik D.; Demaine, Martin L.

Loss Aversion Theory and Stock Price
Hassett, Stephen

Magic and the Periodic Table of the Elements
Duran, Miquel

Manifold
Morin-Drouin, Jerome

Masyu 11x11 – Conway's Game of Life
Baxter, Nick

Mathemagical Deck of Cards
Hirth, Tiago; Silva, Jorge Nuno; Freitas, Pedro; Simões, Carlota

---

Mathematics Awareness Month 2014 Poster
Torrence, Bruce; Torrence, Eve

Maybe Fair Die
Kisenwether, Joseph

Maze of Martin
Chartier, Tim

---

mind reading math trick
Menna, Lisa

Modular Magic: Eleven Cubes
Jones, Kate

Morning Math with Red Queen
Muller, Marnie

New twist on the buttonhole puzzle
Bobroff, Saul

---

non-tetrahedral 4 sided die
Sherman, Scott

Oddly Enough
Aronson, Simon

Paper
Liles, William

Paper on recent significant achivements on puzzles in Japan
Takashima, Naoaki

Pi Day themed contribution
Couzin, Mary

Pirate Math Sampler
Serra, Michael

PLAY WITH 2014
Halici, Emrehan

Polygons Folding to Three Different Boxes
Uehara, Ryuhei

Pot Pourri Bell's Elles
Brokenshire, Laurie; Bell, Donald

Preface and Annotated Table of Contents from Martin Gardner in the 21st Century
Hopkins, Brian

Puzzles: Invented Designed or Discovered
Dalgety, James

Pythagorean Earrings
Vinson, Jade

Quintetra Blocks
Kostick, John and Jane

Recreational Mathematics Magazine
Pereira dos Santos, Carlos

Recreational science for blind people
Blasco-Uceda, Fernando; Blasco, Fernando

Rendering Hendecagrams
Vorthmann, Scott

Retrolife Puzzles and the Apex Magic Trick for G4G11
Elran, Yossi

Scenes from the Mathenaeum
Whitney, Glen

SELF-REFERENTIAL 9 DIGITS REPDIGIT NUMBER
Kurchan, Rodolfo

Seven touching infinite cylinders
Bozóki, Sándor; Lee, Tsung-Lin; Rónyai, Lajos

Seven-Color Bead Crochet Torus Map Pattern
Goldstine, Susan

Slipagons Postcard and Staroid Postcard and Paper on same
Engel, Doug

Solving All 164,604,041,664 Symmetric Positions of the Rubik's Cube in the Quarter Turn Metric
Rokicki, Tomas

Solving Puzzes Backwards
Levitin, Anany

Some neat geometry problems arising from tesselations.
Scherphuis, Jaap

Square in the Bag and Other Puzzles for Classroom
Iwasawa, Hirokazu

Star Polygon Shortbread for the Puzzling Palate
Baker, Ellie

---

still Life with Glider
Bosch, Robert

Stone Piles - A Variation of Nim
Morrill, Ryan

Sun Bin's Legacy
Mackenzie, Dana

Super Cube 11
Manderscheid, Roger

Symmetries in Lisbon, Portugal
Carvalho, Alda

Telephone Calls and the Brontosaurus
Atkinson, Adam

Test
Yung, Veronica

Tetraflexagon Trading Cards
Beier, Julie

Tetraflexagons
Yackel, Carolyn

The Annotated Goodbye Old Girl
Richards, Dana

The Battersea Power Station Puzzle
Singmaster, David

The Checker Shadow Illusion
Rossetti, Dave

The Dragonfly magic of targeting and capture
Schwab, Ivan R.

The Eleven Clocks Problem
Roby, Tom

The game of light
Kocik, Jurek

The Granny Knot and the Square Knot
Kauffman, Louis

The Parking Lot Puzzle
Bickford, Neil

The Planar Tetrarhons Tetrahedron
Hoff, Carl

The Pseudosphere Uphill Roller
Brecher, Kenneth

The Puzzle that Goes to Eleven
Stephens, James

The Stochastic Geometry of John Shier
Cipra, Barry

The UMAP Journal 34 (4) (2013)
Marasco, Joe

The Unique 4-7-11 Octahedral Monoicosahedron
Banchoff, Thomas

THREE PUZZLES
Schoen, Alan

Tiebreaker Dice
Bogart, Rod

Tiling Tetris boards
Butler, Steve; Ekstrand, Jason; Osborne, Steven

Trick-opening boxes from Sri Lanka
de Vreugd, Frans

Tunnel-Cube
Hart, George

Two Parity Puzzles Related to Generalized Space-Filling Peano Curve Constructions and Some Beautiful Silk Scarves
McKenna, Doug

Unique Polyhedra Dice
Fathauer, Robert

Using Calcudoku in Modern Algebra
Nacin, David

When is the next Thanksgivukkah?
Levy, Doron

Who Wrote Martin Gardner's Autobiography?
Kearn, Vickie; Gardner, Jim

Wordplay / Slipperiness of Language / N-tendres
Goldklang, Lew

YESGO, an unusual Go game and its problems
Kotani, Yoshiyuki

Zometool gift
Hildebrandt, Paul

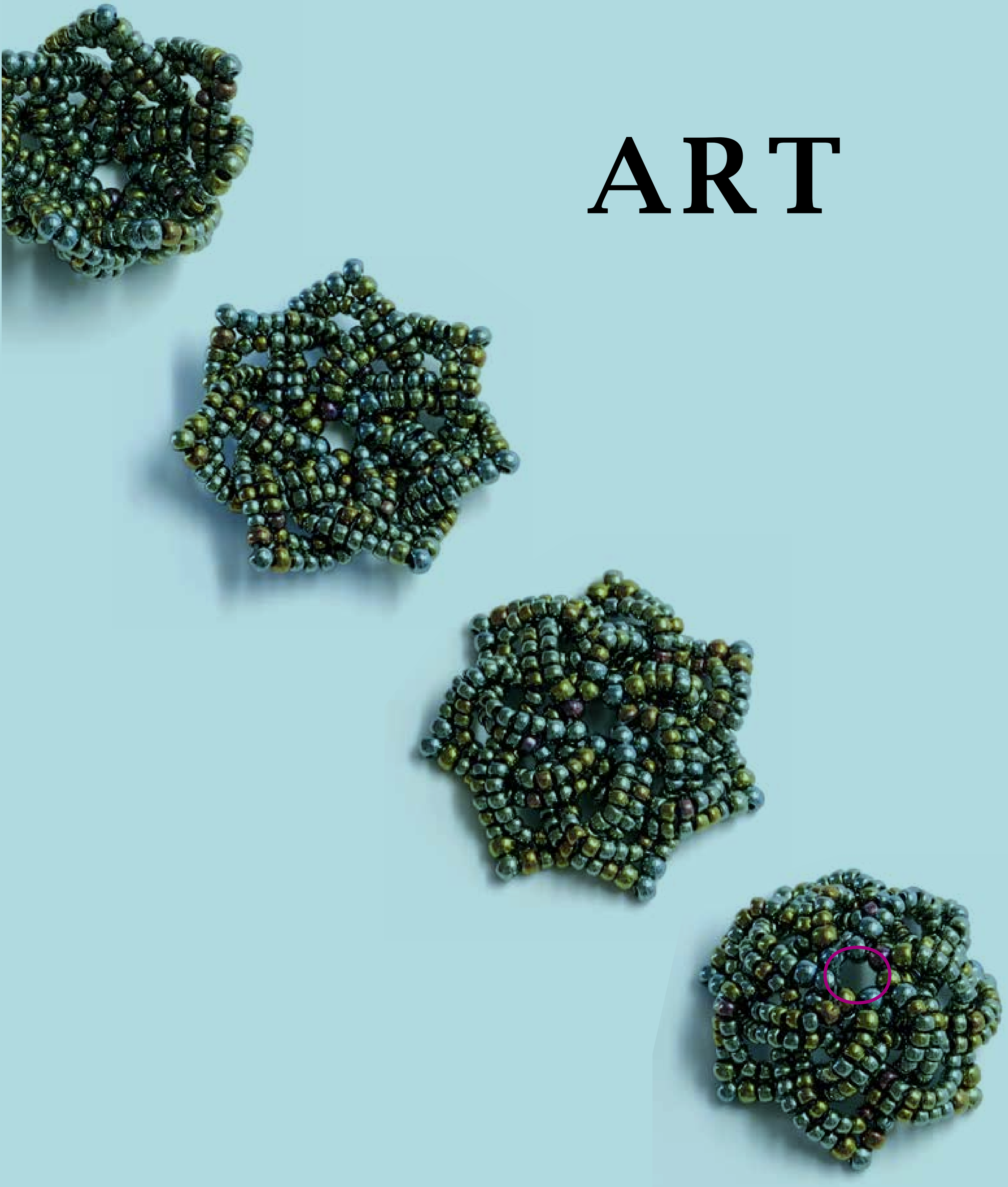
SUDOKU becomes SLIDOKU
Nightingale, Simon

The Prehistory of the Game Nim
Rougetet, Lisa

The Conway Immobilizer
Winkler, Peter

A Simple MatheMagics Trick
Brittain, Skona

ART

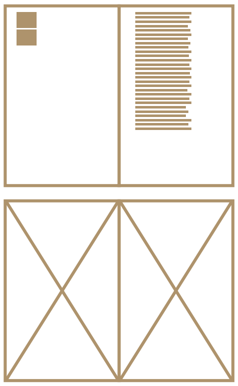


Beaded Celtic Knots | Gwen Laura Fisher | Page 40

Homework

by Alex Bellos

106–109



Homework

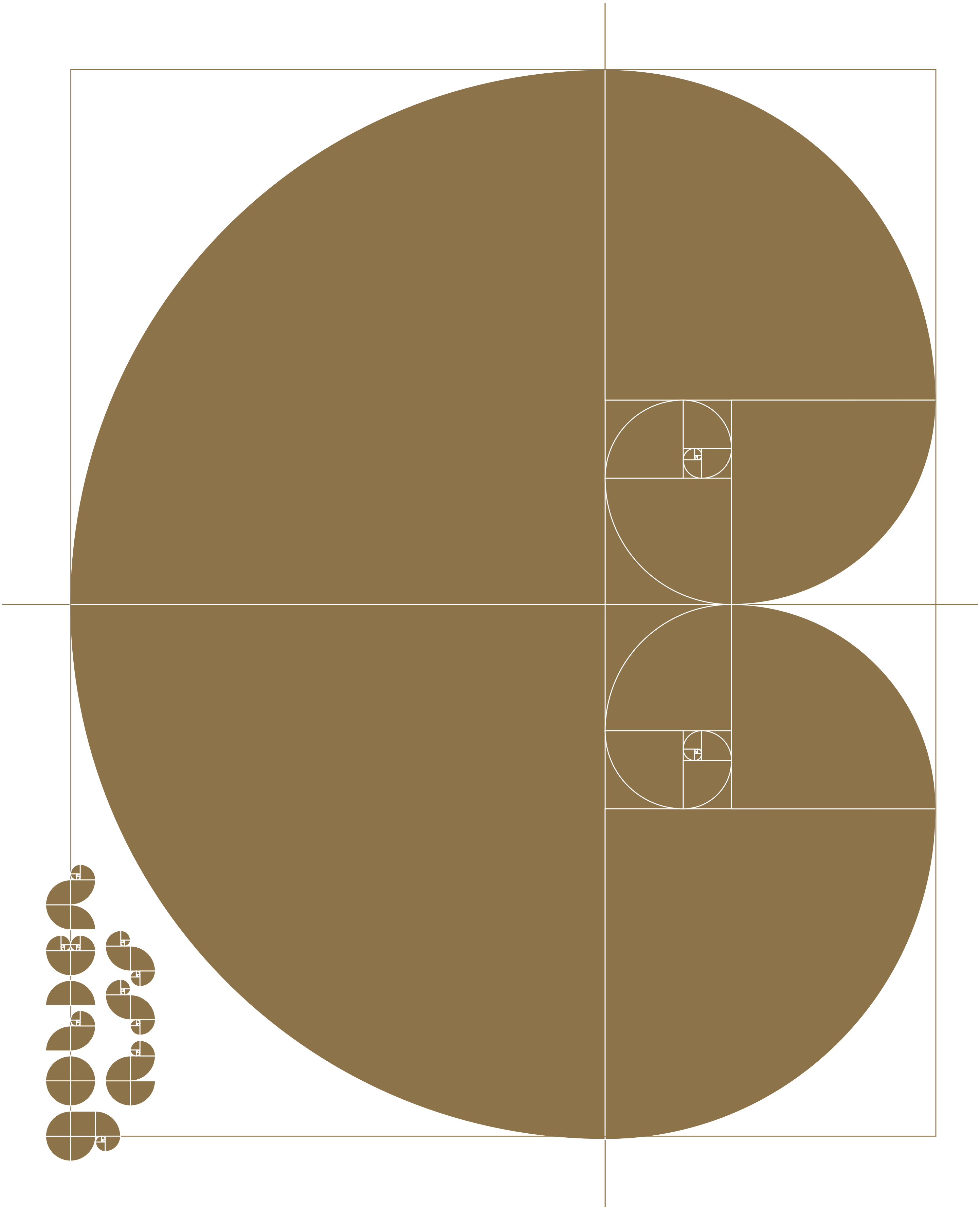
Jerzy Skakun
b1973/Poland

Joanna Górska
b1976/Poland

www.facebook.com/HomeworkDesign
www.homework.com.pl

When we were asked to participate in this project, my first thought was of my grandfather who passed away a few years ago. He was a maths fanatic. I remembered his desk drawer where there were rulers, squares, callipers, protractors, pencils and a rubber. My grandfather and I would often play and draw circles and weird geometric figures. I did not follow my grandfather and become a mathematician. However, I have found there is a lot of maths in graphics, especially in vector design.

Mindful of the Renaissance notion that the golden ratio is 'divine' and the secret of true beauty, I wanted to check this out. I drew a logarithmic spiral and found that this beautiful shape is used to depict many beautiful things: hair curls, shells, etc. I flipped these elements and repeated them, noticing that the mirrored spiral created the shape of an idealised apple, which when rotated, became the ideal ass – I have used maths to depict ideal beauty, albeit a little obscene! Mathematical perfection becomes a perversion. The letters for the title are also based on this logarithmic spiral.



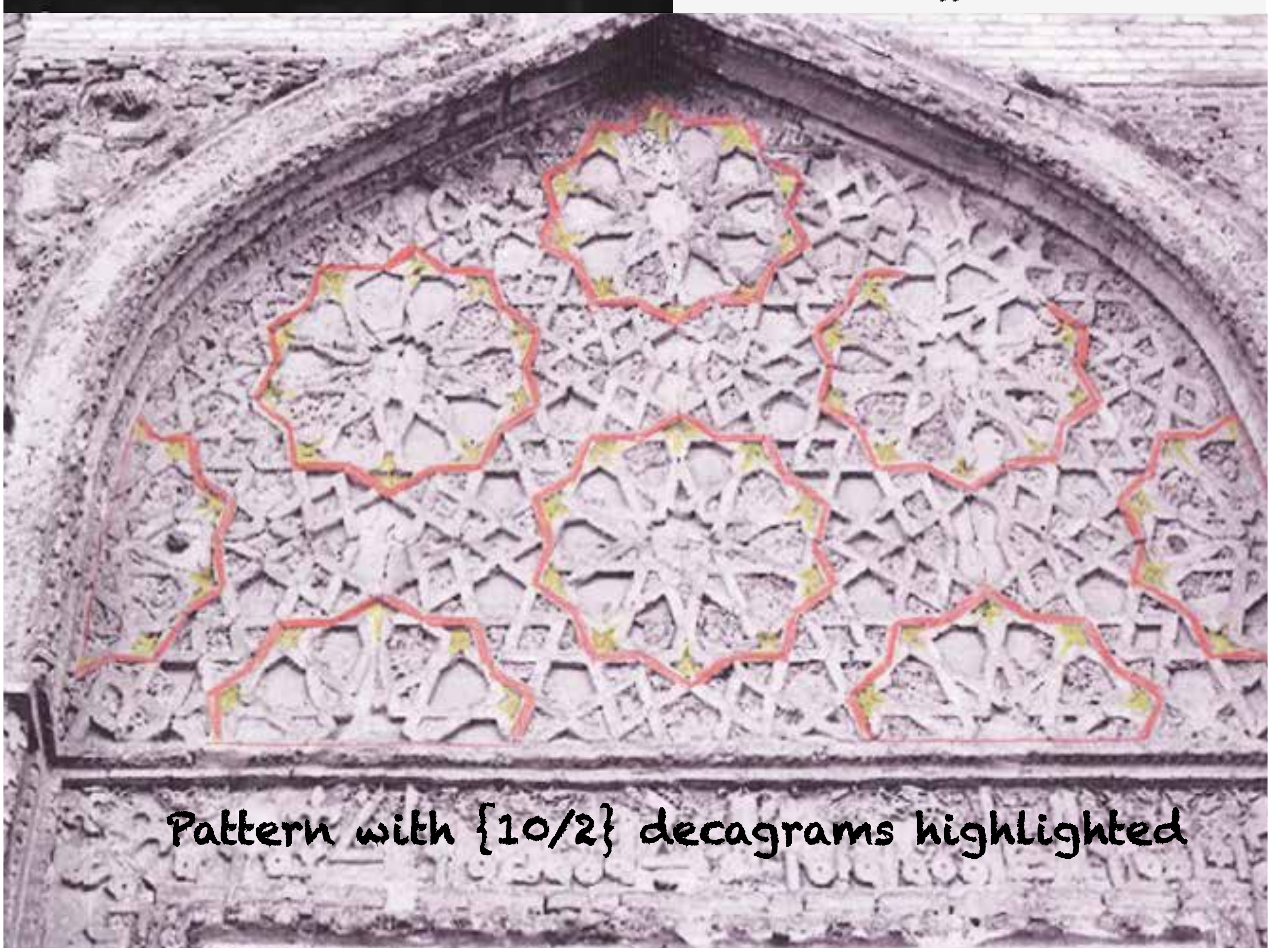
A CORONA WITH DECAGRAMS
On a Late 12th Century Persian Monument

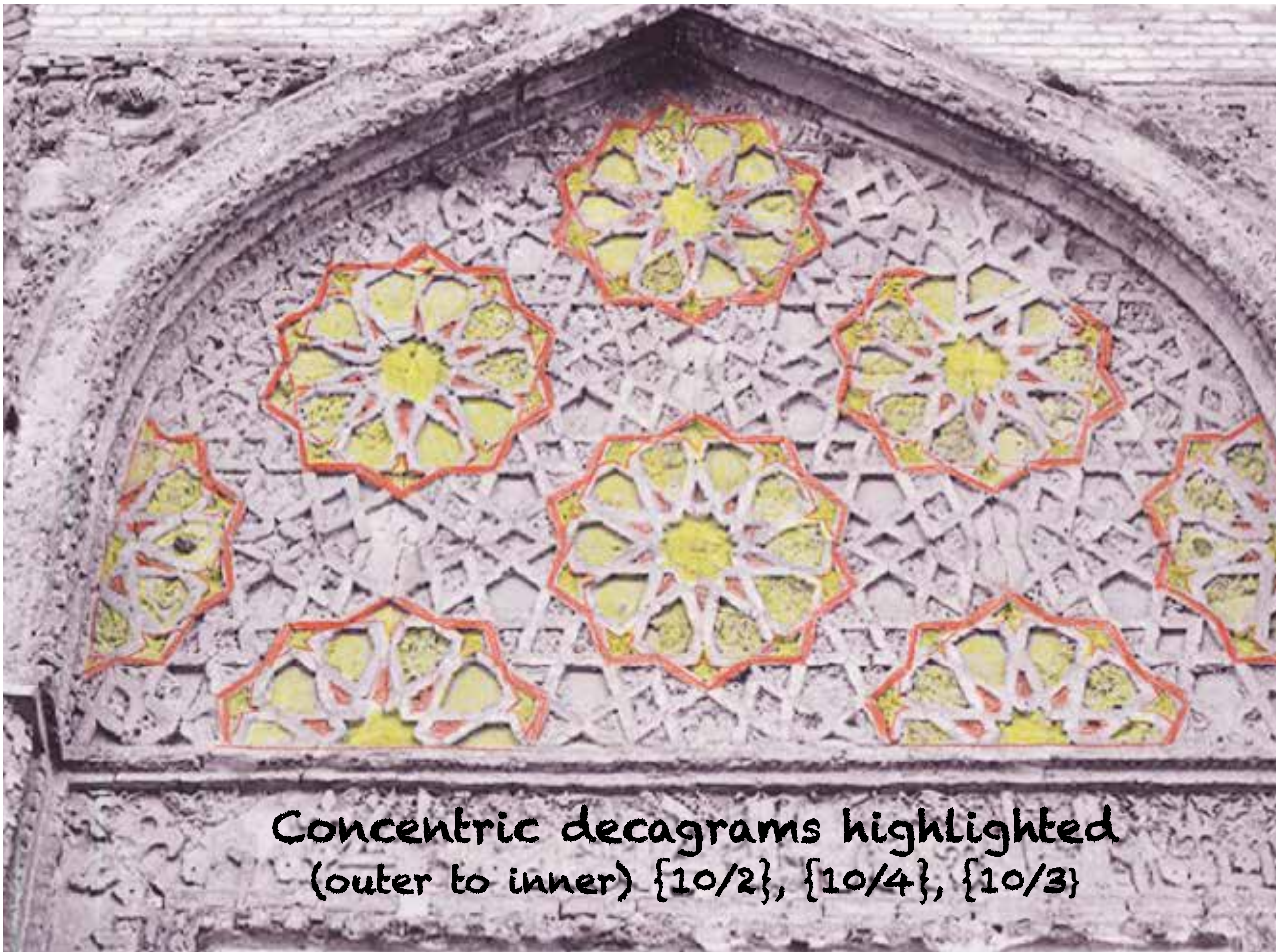
Analytical Pattern Sketches

Carol Bier

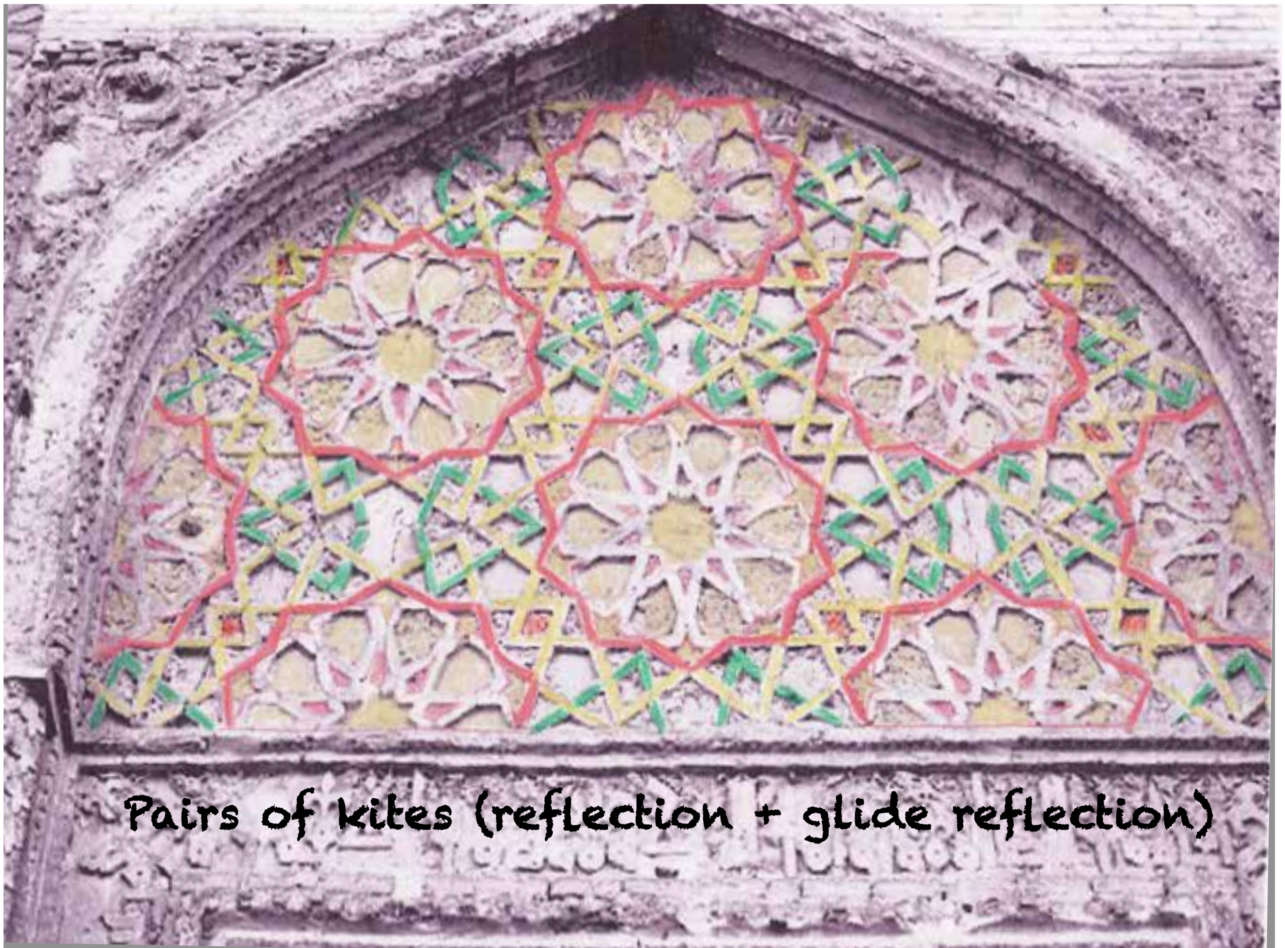
G4G11
19 March 2014

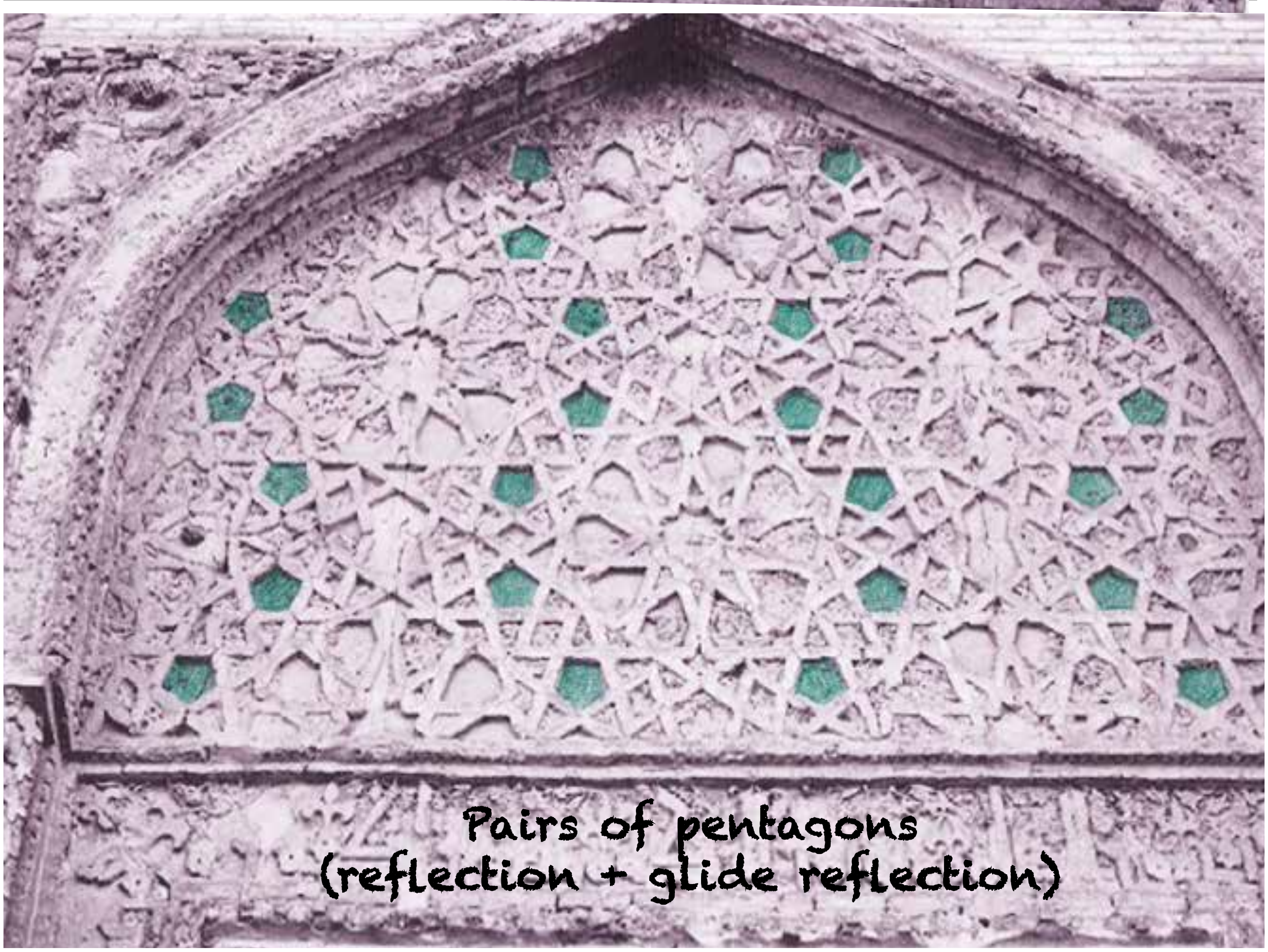
Gonbad-e Alaviyyan, Hamadan, Iran

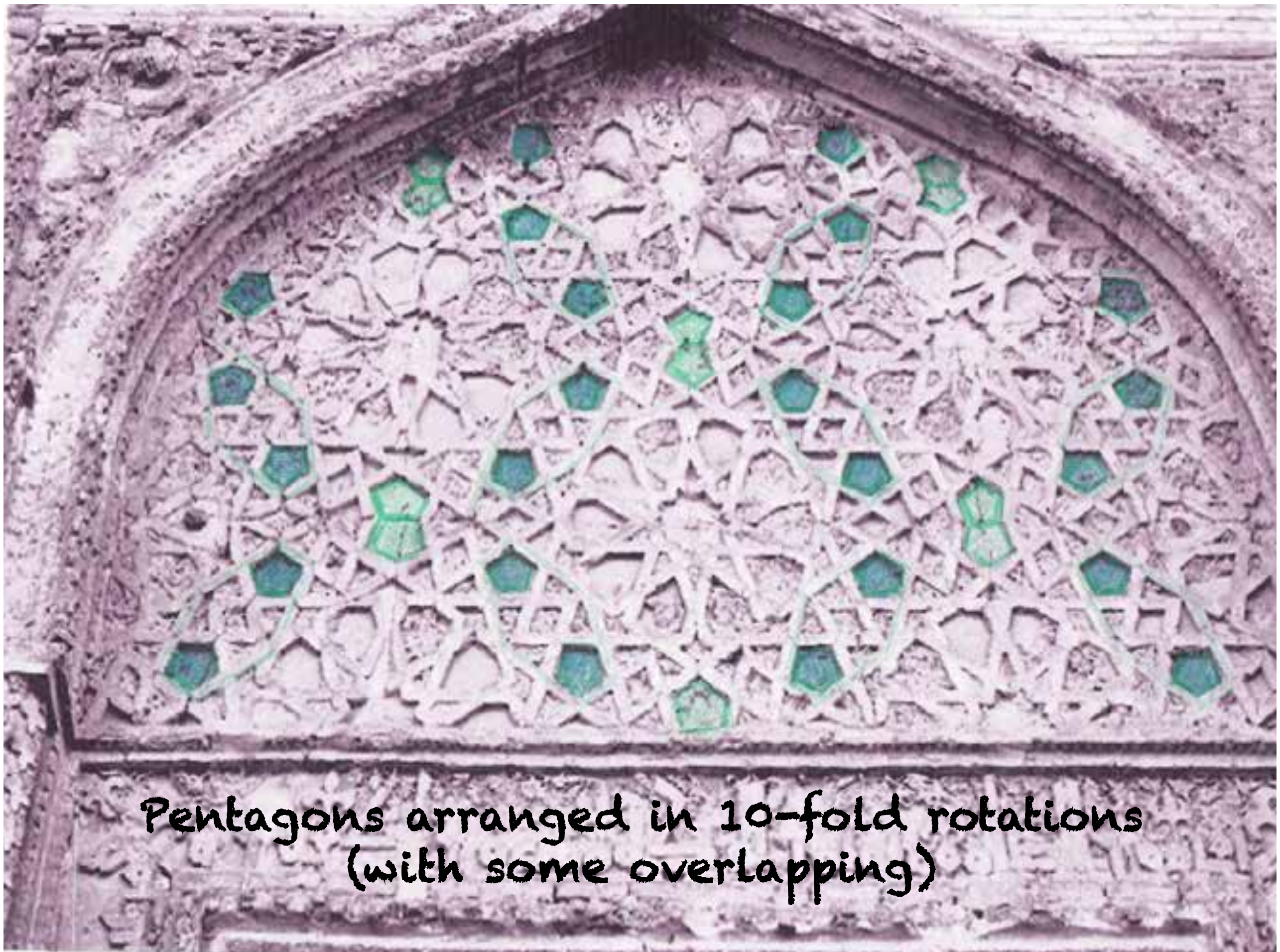



Concentric decagrams highlighted
(outer to inner) {10/2}, {10/4}, {10/3}

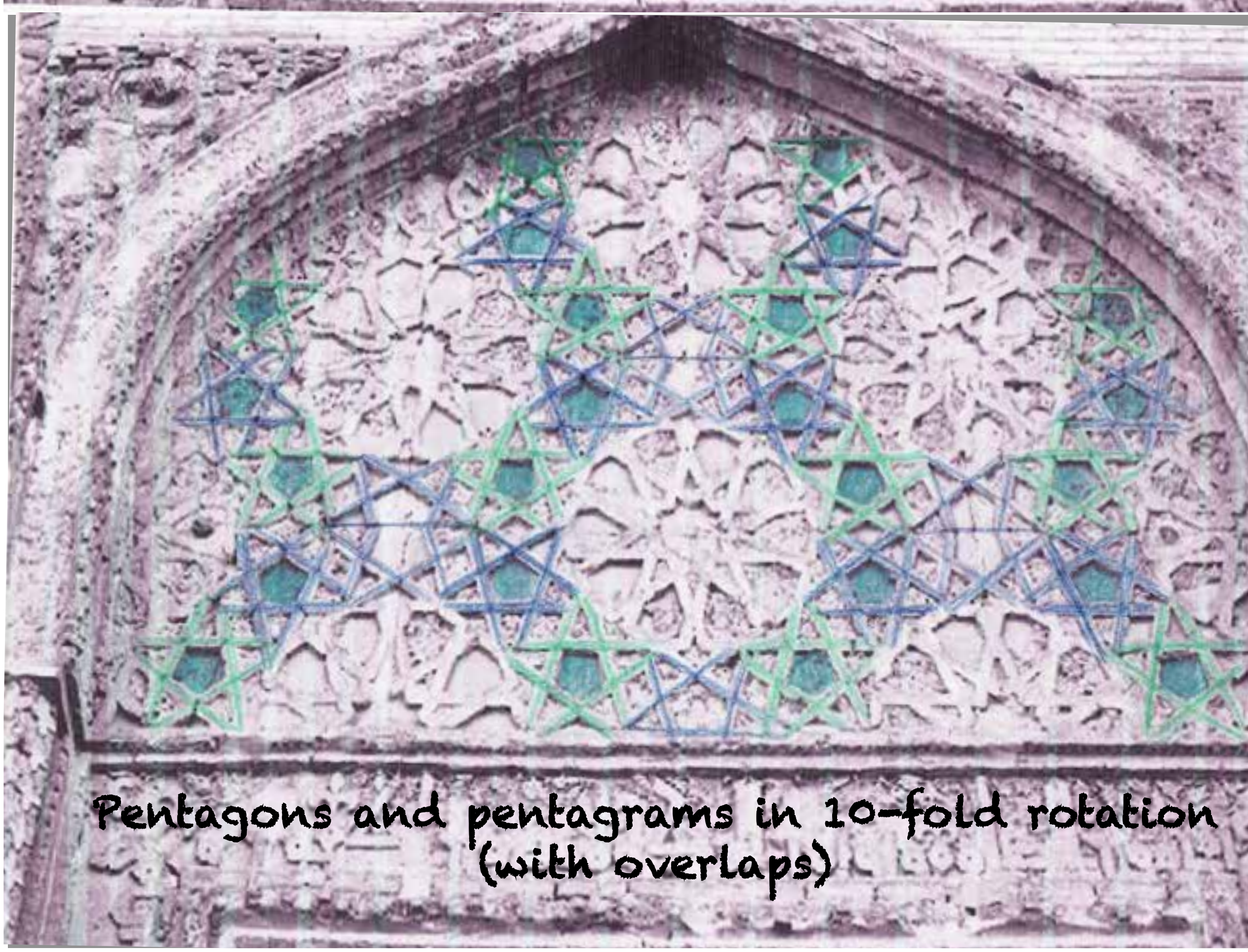

One more concentric decagram {10/3}
But note overlapping!


Pairs of kites (reflection + glide reflection)

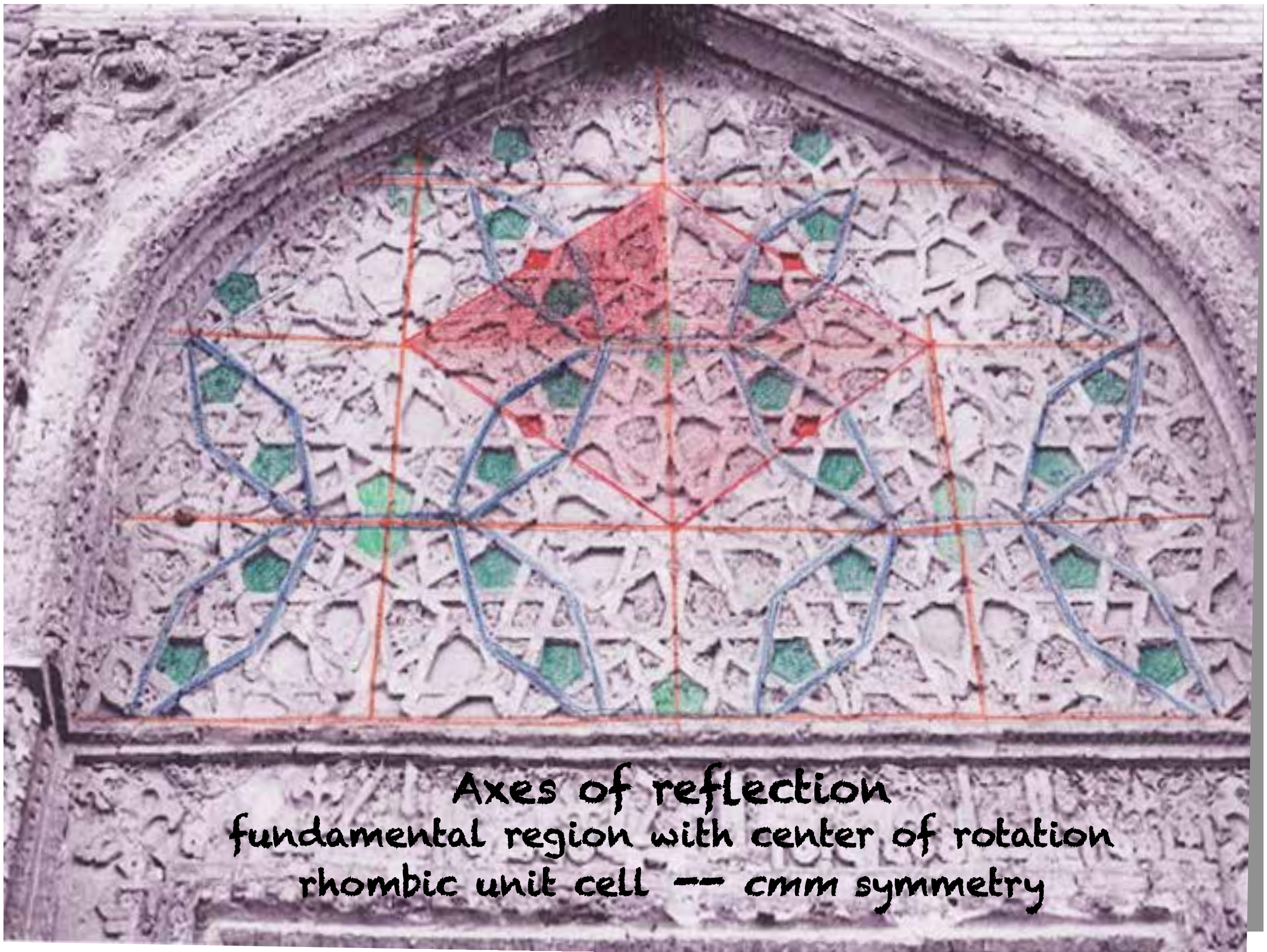

Pairs of pentagons
(reflection + glide reflection)



Pentagons arranged in 10-fold rotations
(with some overlapping)



Pentagons and pentagrams in 10-fold rotation
(with overlaps)



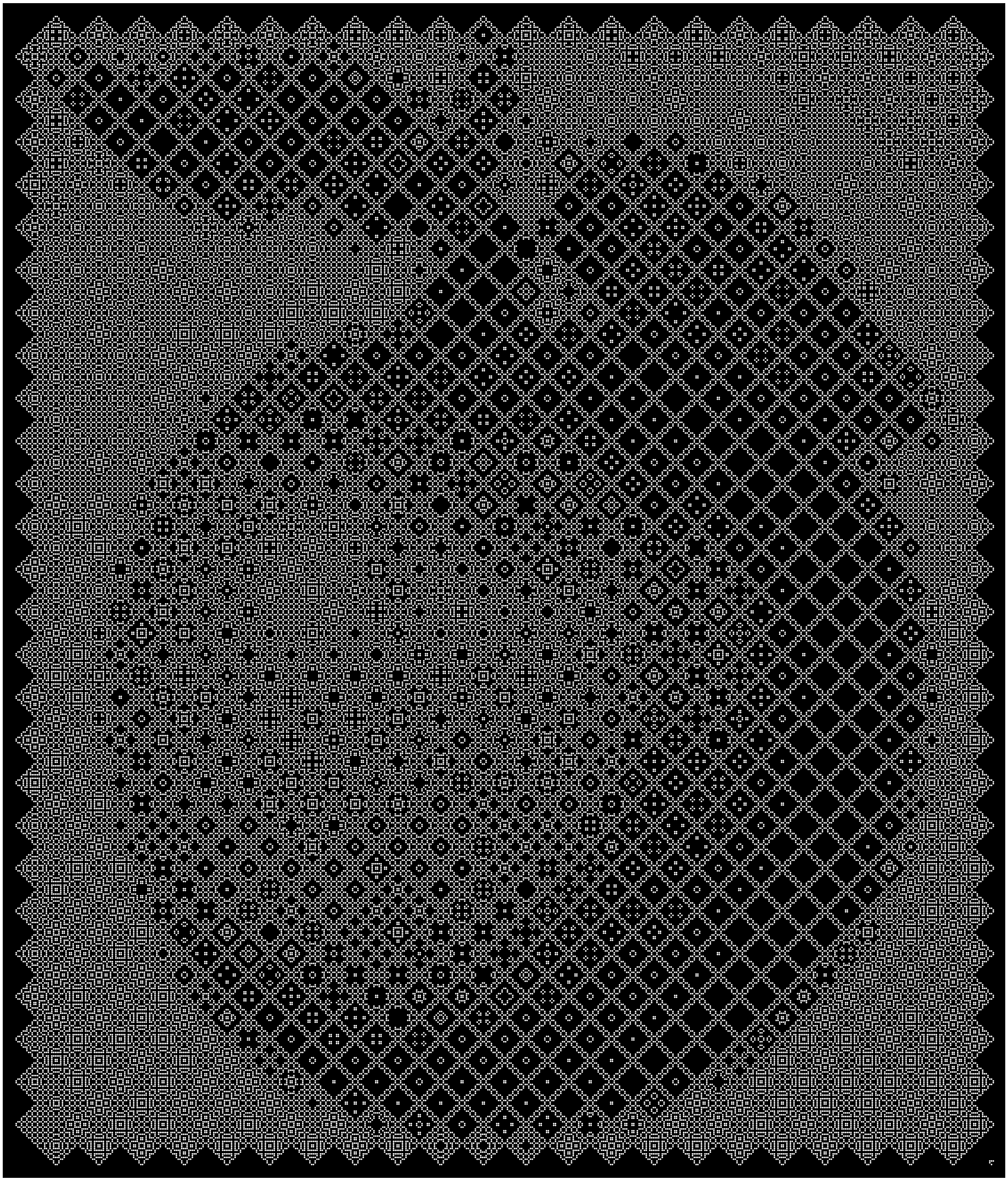
Axes of reflection
fundamental region with center of rotation
rhombic unit cell -- *cmm* symmetry

Implied rhombic grid reveals a CORONA
With centered DECAGRAMS!

Still Life with Glider

by Robert Bosch | Oberlin College

A Still Life (Magritte's "Ceci n'est pas une pomme") as a Still Life, a stable pattern in Conway's Game of Life. But then a glider crashes into it and demolishes it.



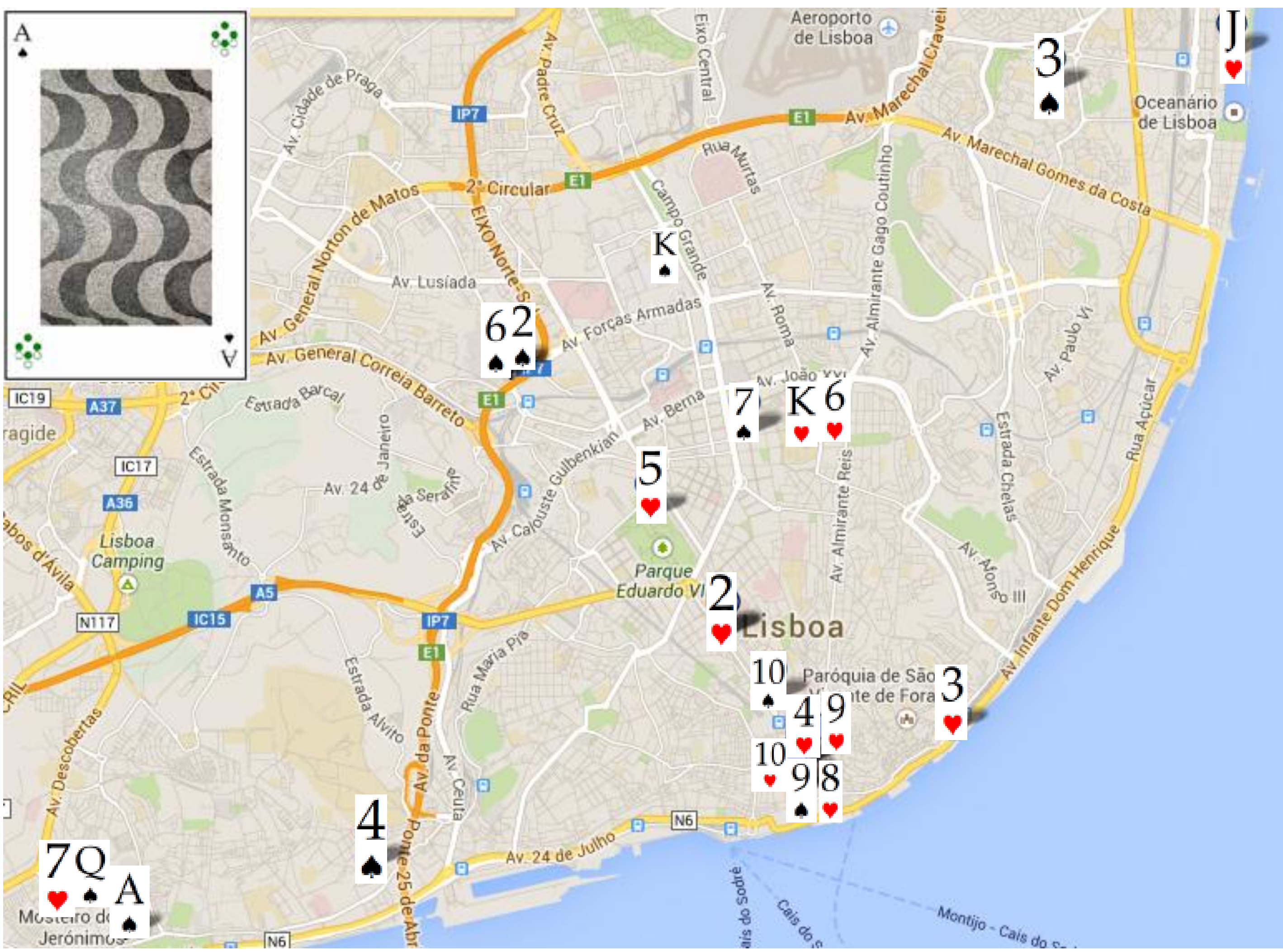
still Life with Glider

Robert Bosch 2014

Alda Carvalho
pwp.net.ipl.pt/dem.isel/acarvalho

Symmetries in Lisbon, Portugal

Come to Lisbon and use Ludus deck to find some symmetries in traditional Portuguese pavements.



★ JOKER

Calçadas de Lisboa

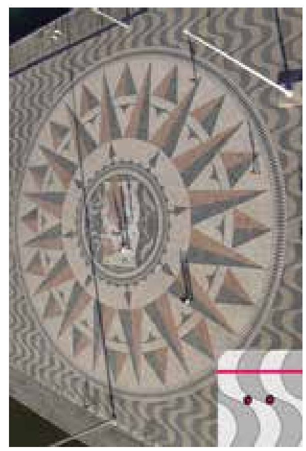
Alda Carvalho
Carlos Pereira dos Santos
Jorge Nuno Silva

★ JOKER

A ♠

A ♠

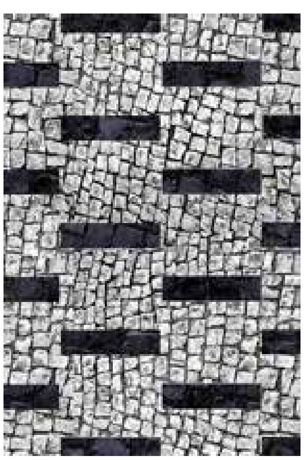
A ♣ Padrão dos Descobrimentos



22*

A ♣

K ♠



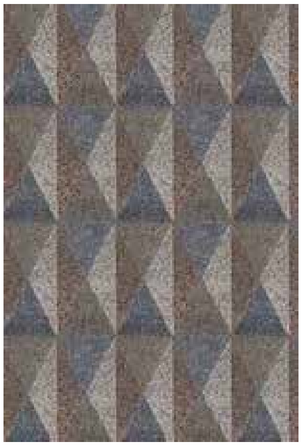
K ♠

K ♣ Biblioteca Nacional Campo Grande

2 * 22

K ♣

Q ♠



Q ♠

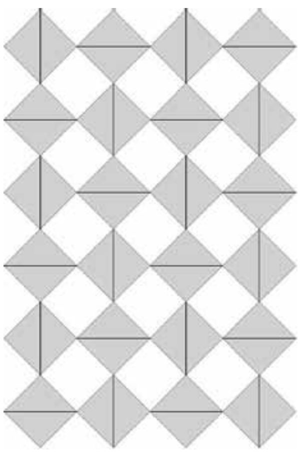
Q ♣

Planetário, Belém

XX

Q ♣

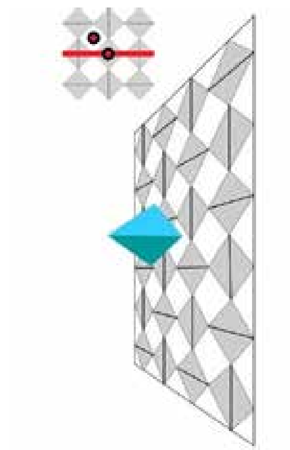
J ♠



J ♠

J ♣

Praça do Tetraedro



4 * 2

J ♣

10 ♠

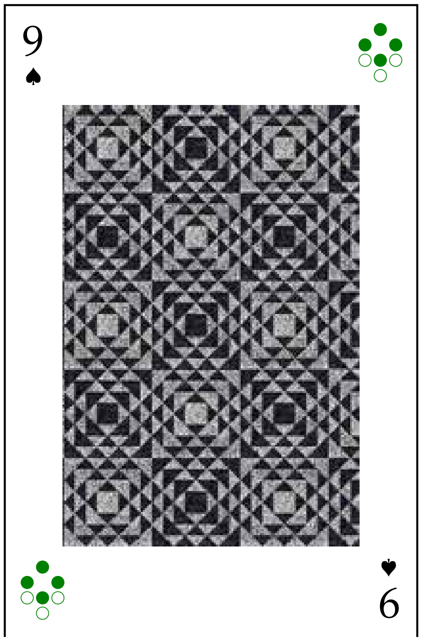
10 ♠

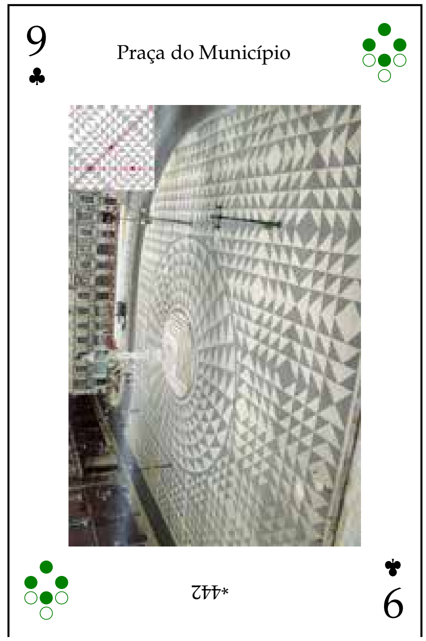
10 ♣

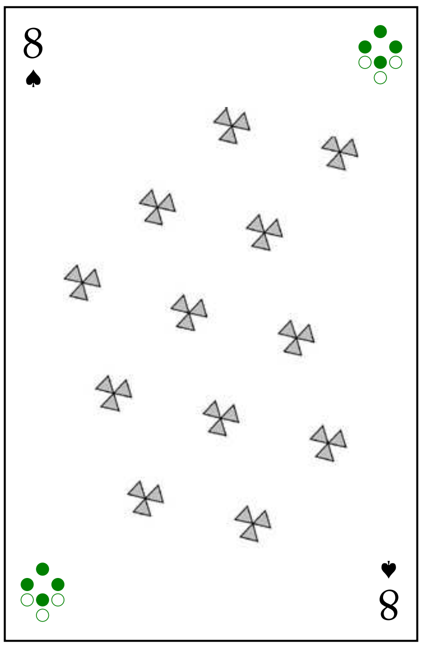
Praça dos Restauradores

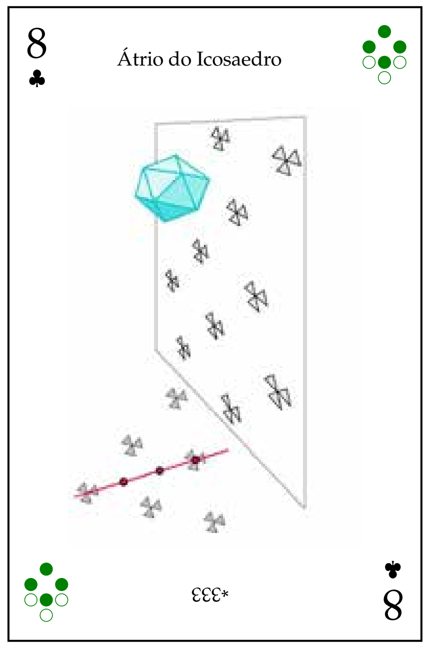
442

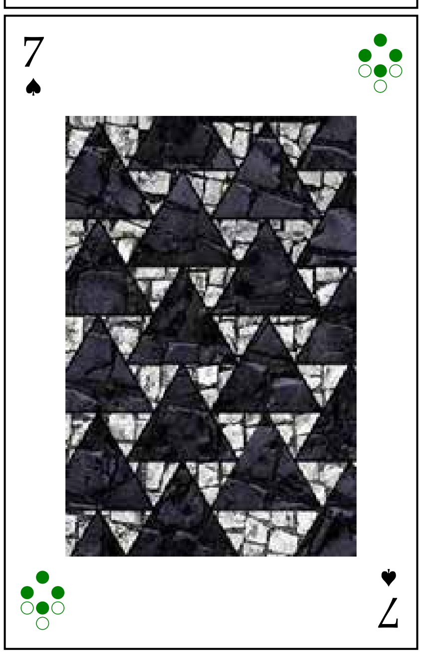
10 ♣


9
♠

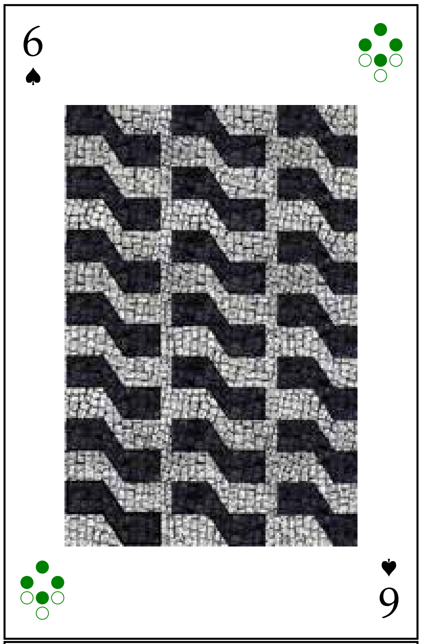

9
♣
Praça do Município
*442


8
♠

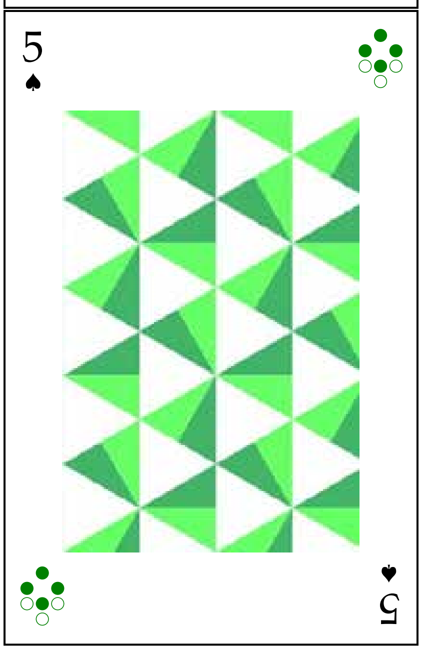

8
♣
Átrio do Icosaedro
*333

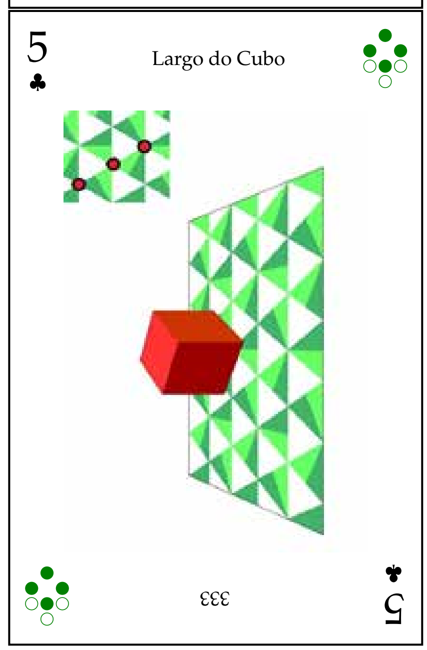

7
♠

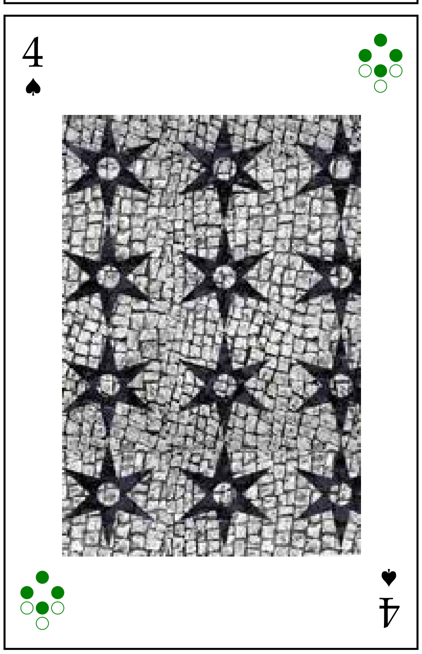

7
♣
Avenida Miguel Bombarda
3 * 3

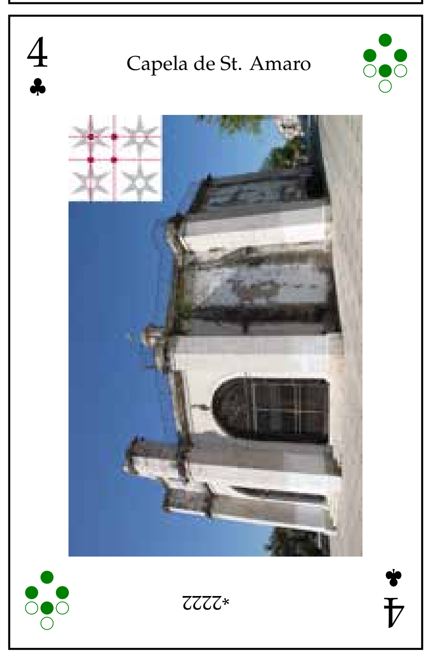

6
♠

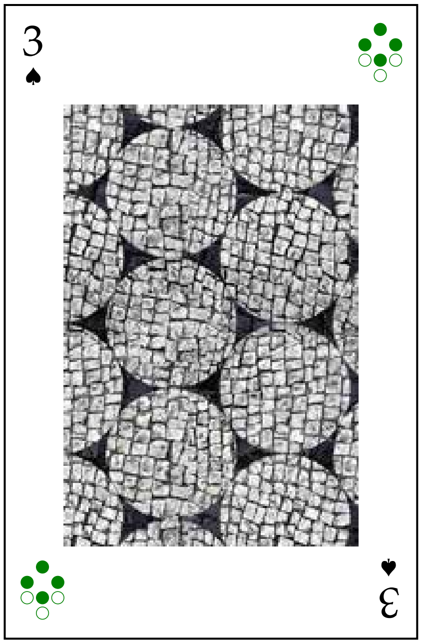

6
♣
Sete Rios
2222

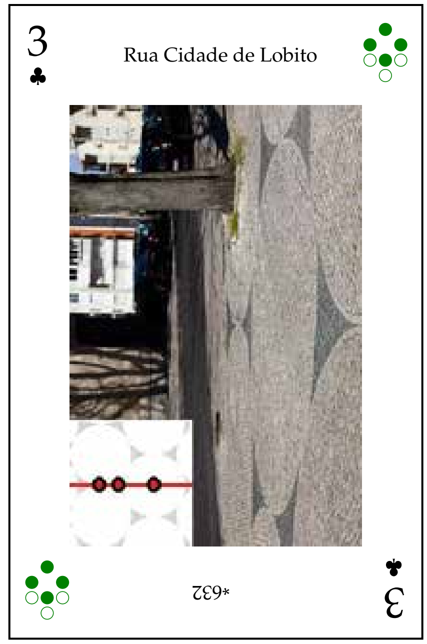

5
♠


5
♣
Largo do Cubo
333


4
♠

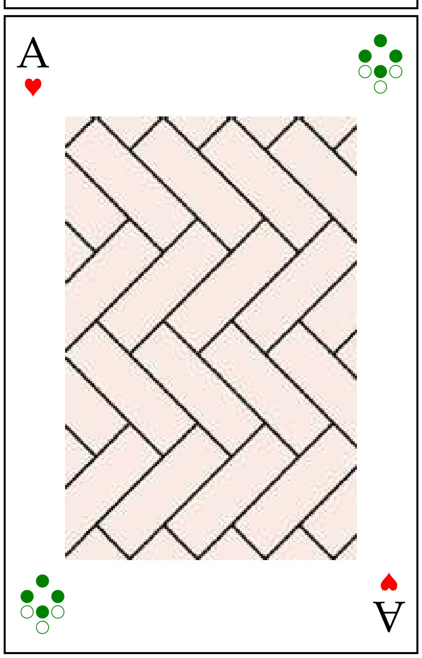

4
♣
Capela de St. Amaro
*2222

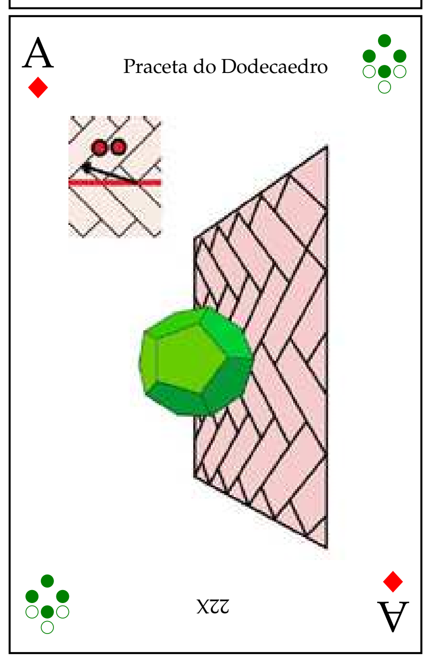

3
3

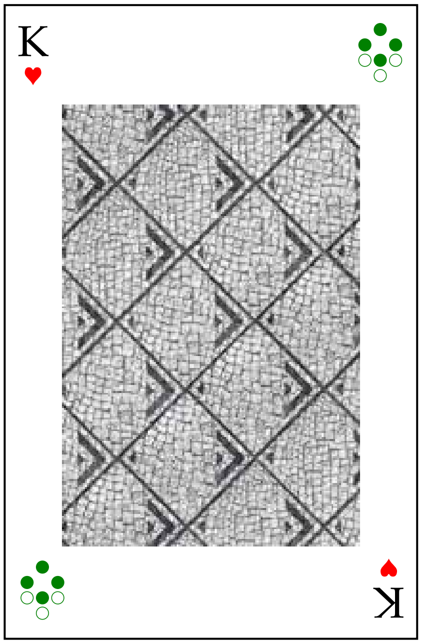

3
Rua Cidade de Lobito
*632
3

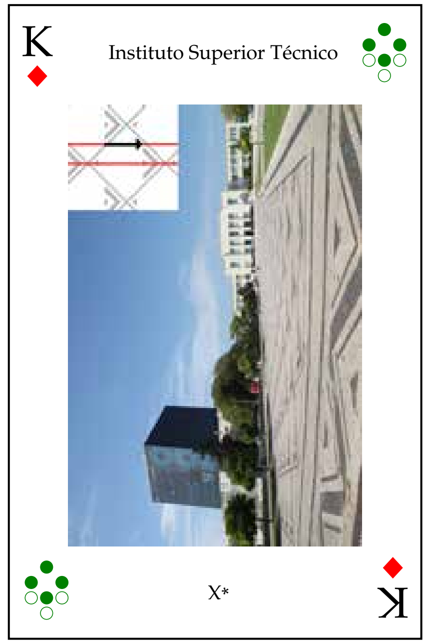

2
2

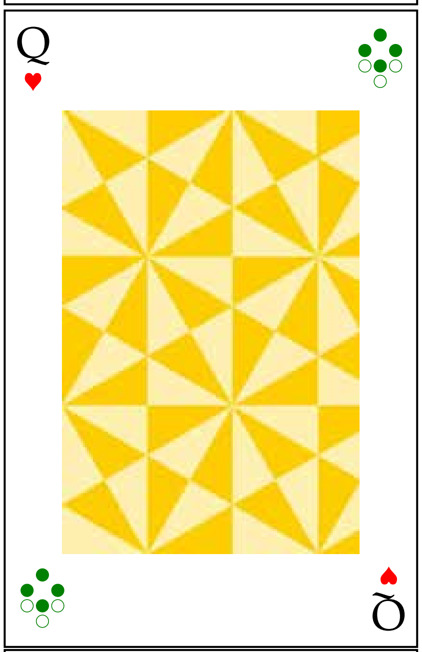

2
Jardim Zoológico
**
2

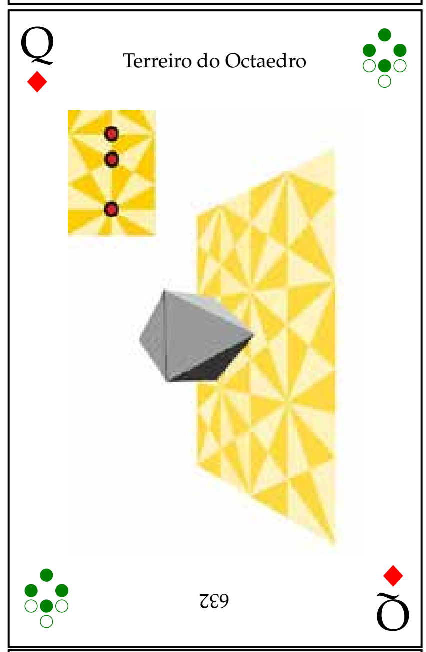

A
A


A
Praceta do Dodecaedro
X22
A


K
K


K
Instituto Superior Técnico
K

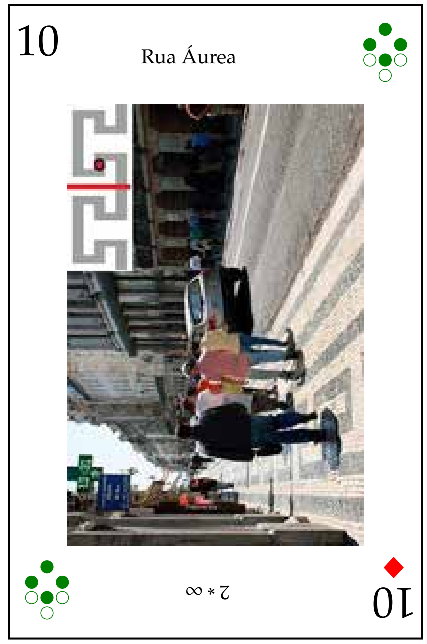

Q
Q

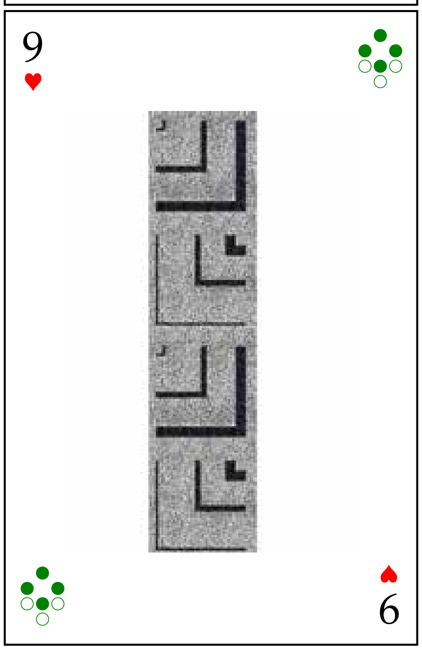

Q
Terreiro do Octaedro
Q


J
J


J
Parque das Nações
J


10


10
Rua Áurea


9


9
Rua dos Correeiros


8


8
Rua Nova do Almada

7 ♥

7 ♥

7 ♦

Praça do Império
Faixa

∞×

7 ♦

6 ♥

6 ♥

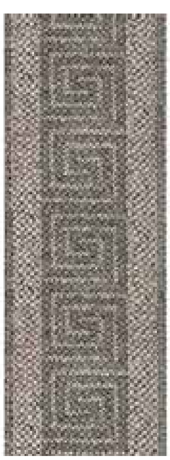
6 ♦

Instituto Superior Técnico
Faixa Lateral

∞*

6 ♦

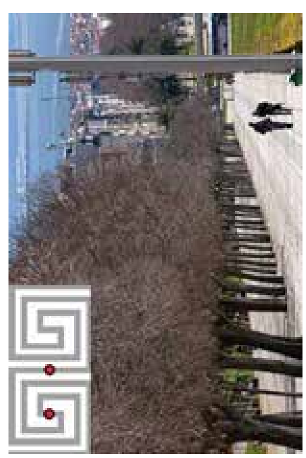
5 ♥



5 ♥

5 ♦

Parque Eduardo VII



22∞

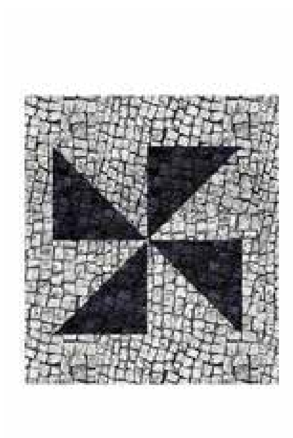
5 ♦

4 ♥

4 ♦ Igreja de São Julião

*∞∞

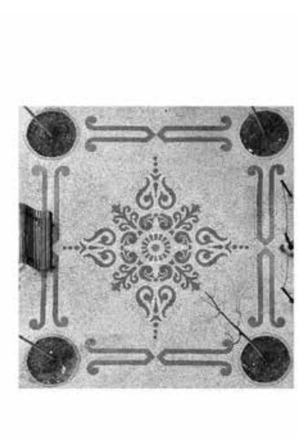
3 ♥



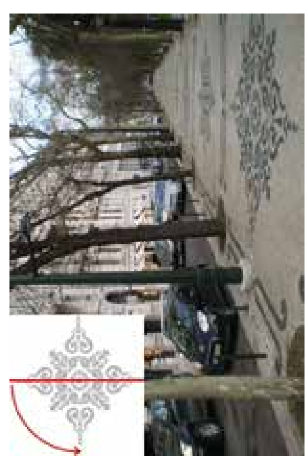
3 ♦ Estação de Santa Apolónia

C_4

2 ♥



2 ♦ Avenida da Liberdade



D_4

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author’s request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

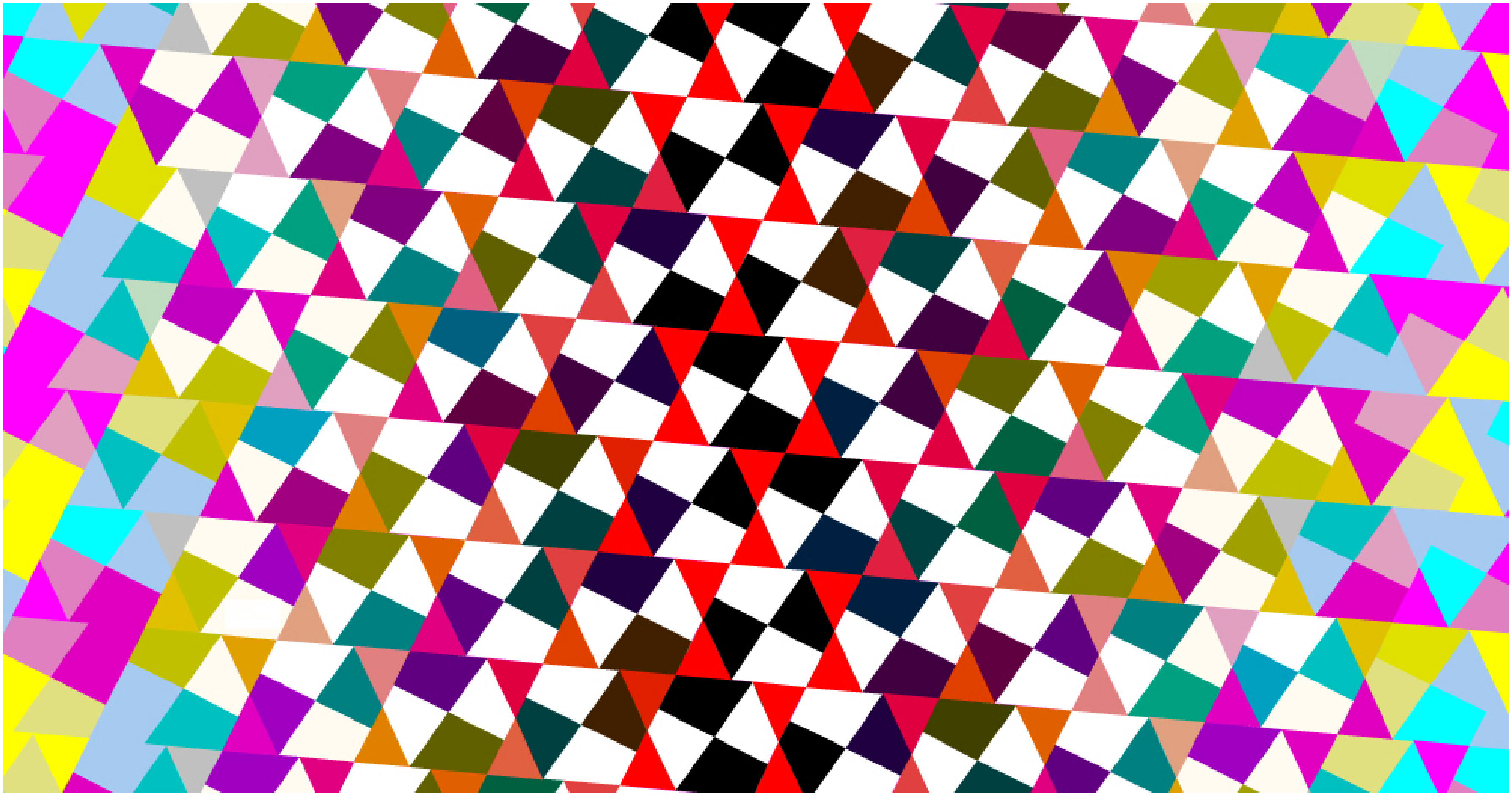
At the author's request, this paper is not included in the online version of this publication. The blank pages have been included to maintain the same pagination as the original printed version.

Self-Referential Breakfast

by Lew Goldklang

A Tiling Trompe l'Oeil

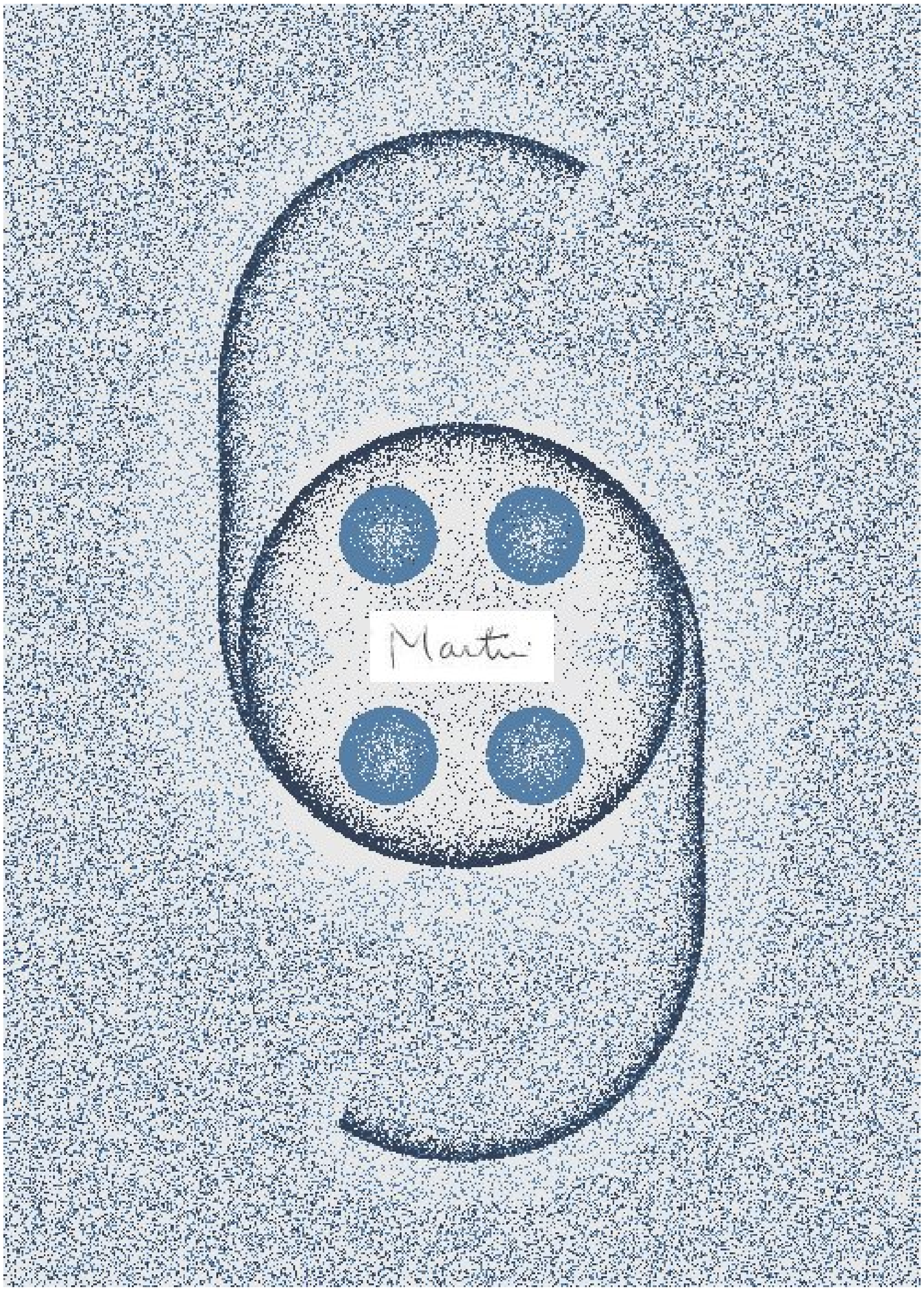
Here is a periodic tiling by two quadrilaterals and a triangle. Is it the refinement of a tiling by regular polygons (hexagons, squares, or equilateral triangles)?



Judging by the high contrast area near the center, it seems impossible to decompose this figure into regular polygons. But the pastel region on the right reveals a tiling of equilateral triangles. And on the left, a tiling of squares! Apparently neither, yet in fact both.

Martin Gardner was big on both tilings and illusions, so this would have been great grist for his column.

--Bill Gosper



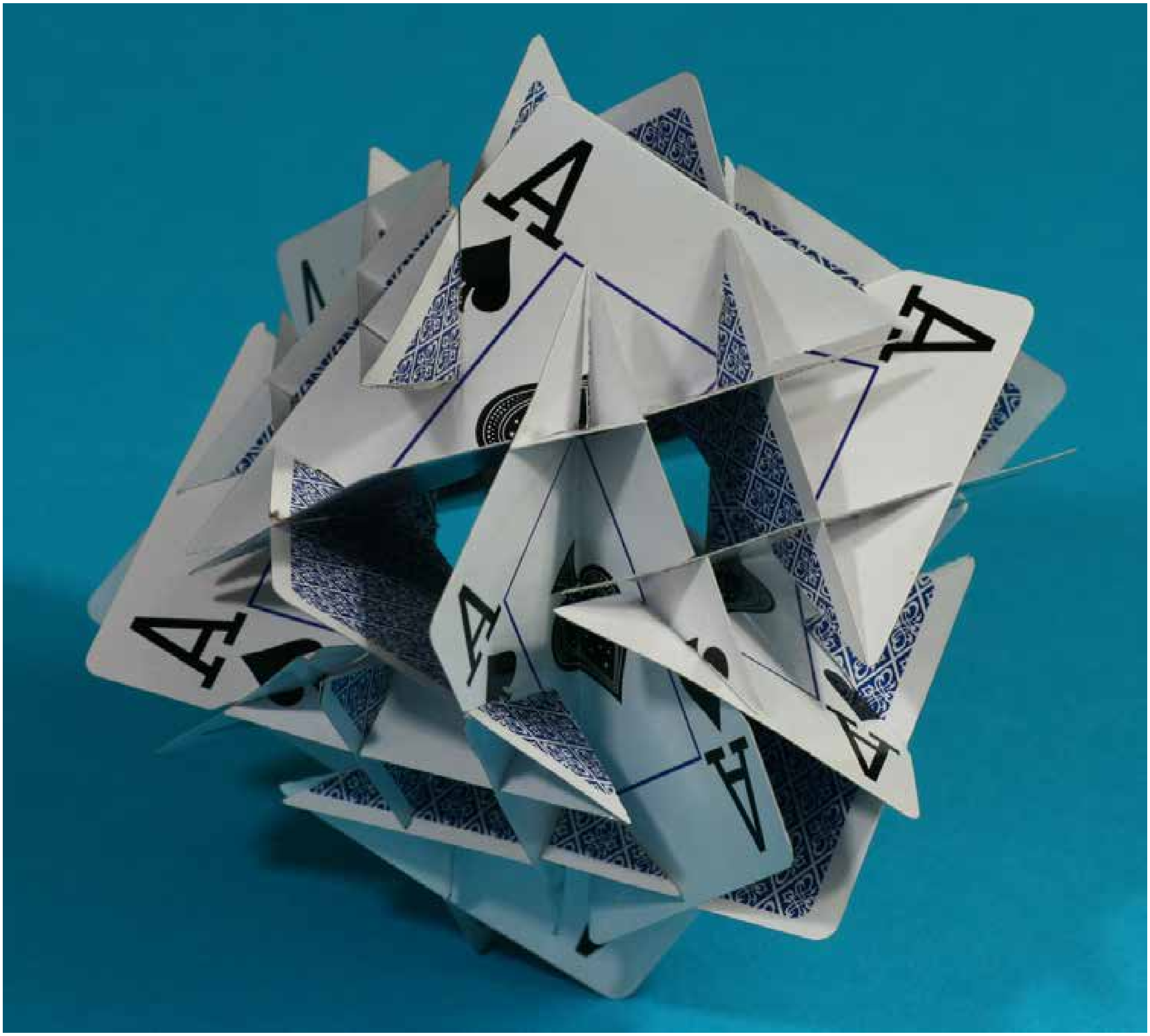
Gary Greenfield G4G Tribute ©2014

Conway's Game of Life (M. Gardner, Scientific American, Oct. 1970 and Feb. 1971) can be realized as stationary automata that sense and change state. Turk's tur-mites (A.K. Dewdney, Scientific American, Sep. 1989) are mobile automata that sense, paint, and move. Sand Painting Artists (P. Urbano, LNCS 625, 2011) model the nest building behavior of T. albipennis ants via mobile automata that pick-up, carry, and drop. G4G Tribute is realized by having 3,000 sand painting artists gather colored sand grains into five circles and two arcs over the course of 200,000 time steps. Martin Gardner's signature from a letter dated 11/1/06 appears in the center.

Tunnel-Cube

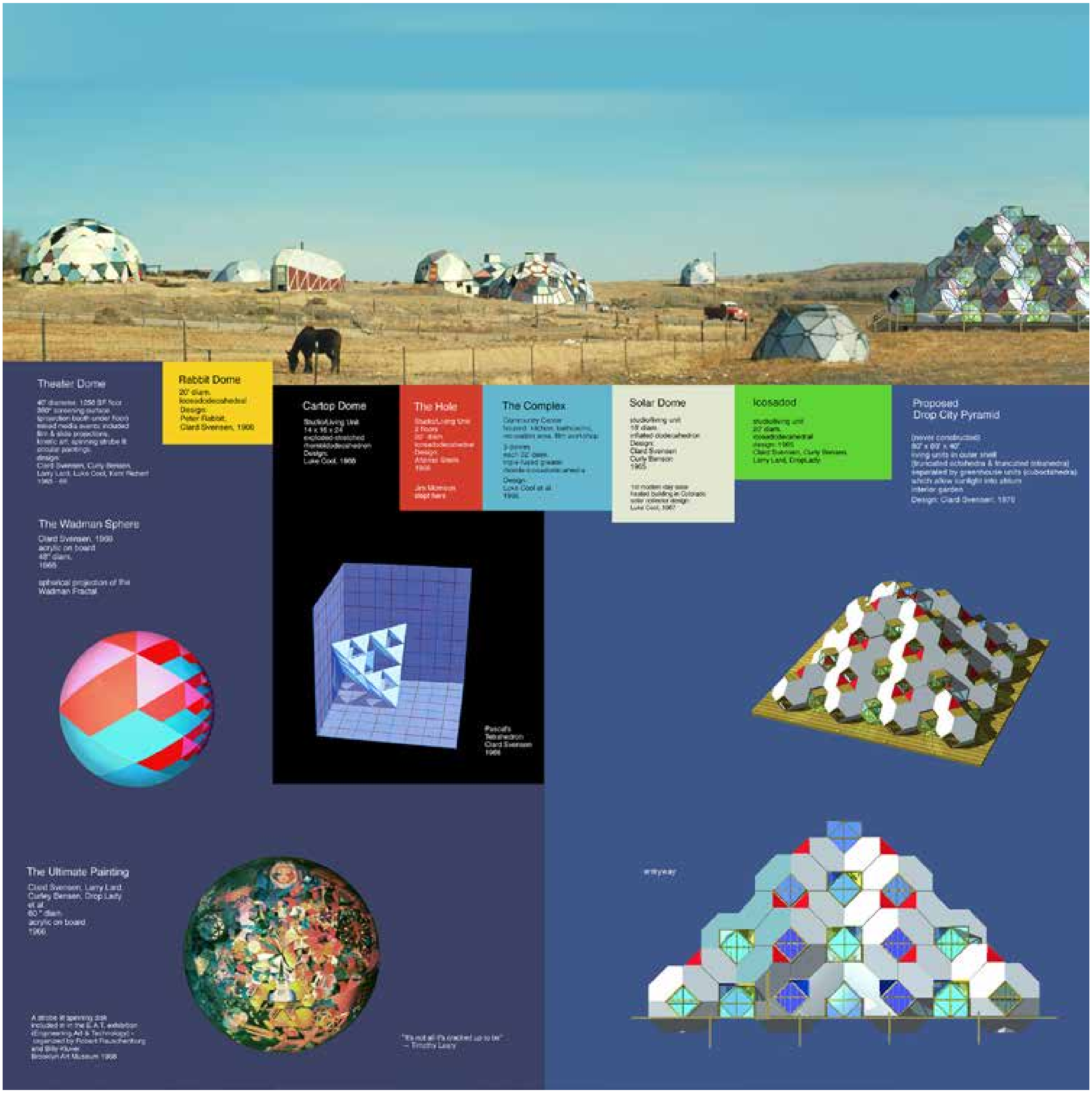
by George Hart

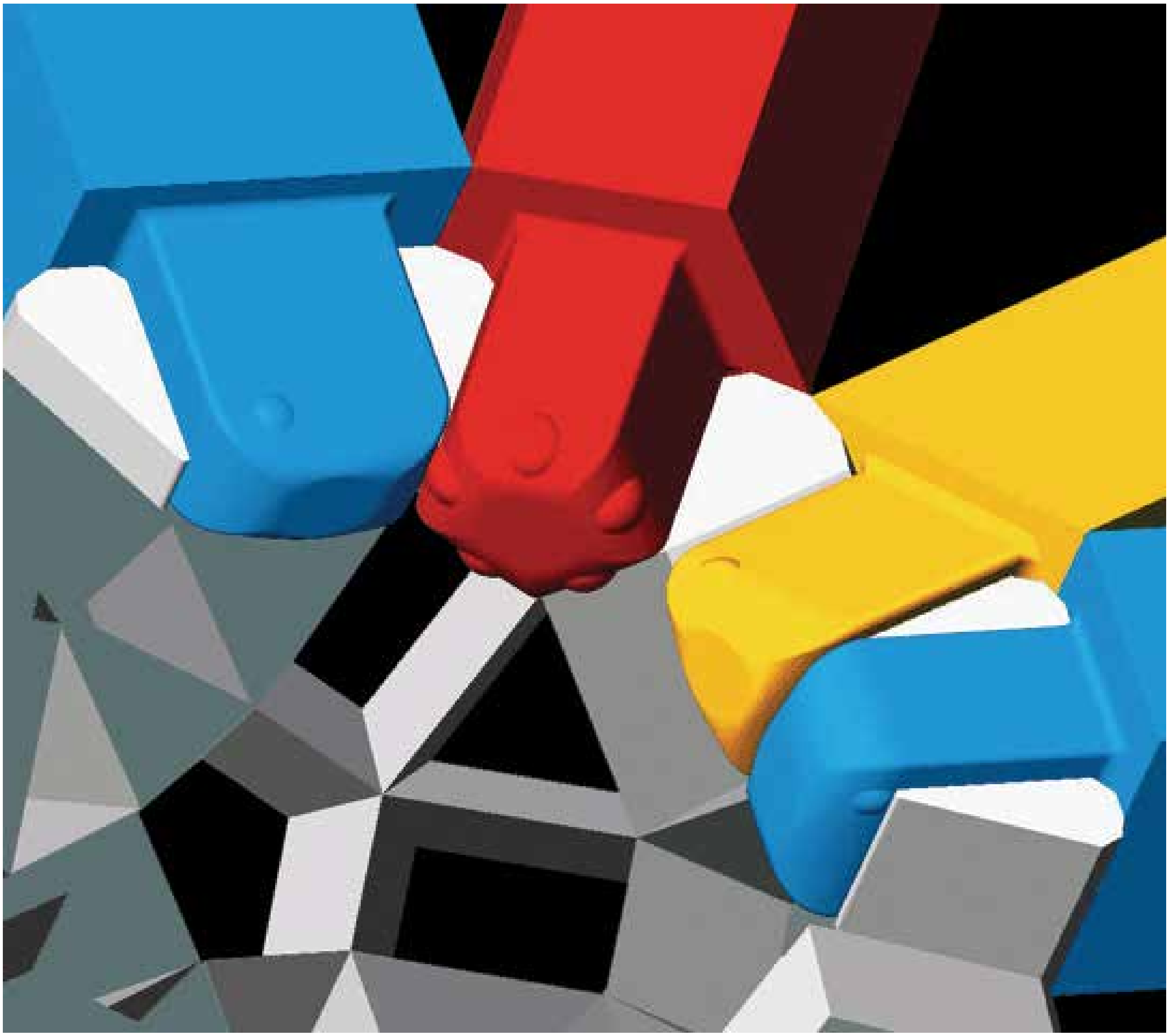
The Tunnel-Cube is a physical sculpture/puzzle consisting of twelve playing cards carefully pre-cut with slots in a way that allows them to be assembled into a mathematical construction. The result is a cube with fourteen tunnels into the center (six square tunnels in the centers of the faces and eight triangular tunnels at the corners).
Assembly instructions are available at http://georgehart.com/g4g11/



Drop City for Rule SM

by Paul Hildebrandt

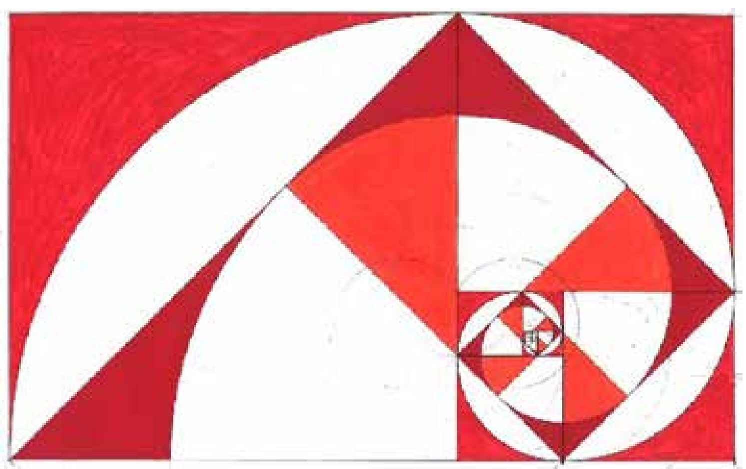




Exploring the Secrets of the Universe (in Red & White)

by Wm Vandorn Hinnant | Winston-Salem State University

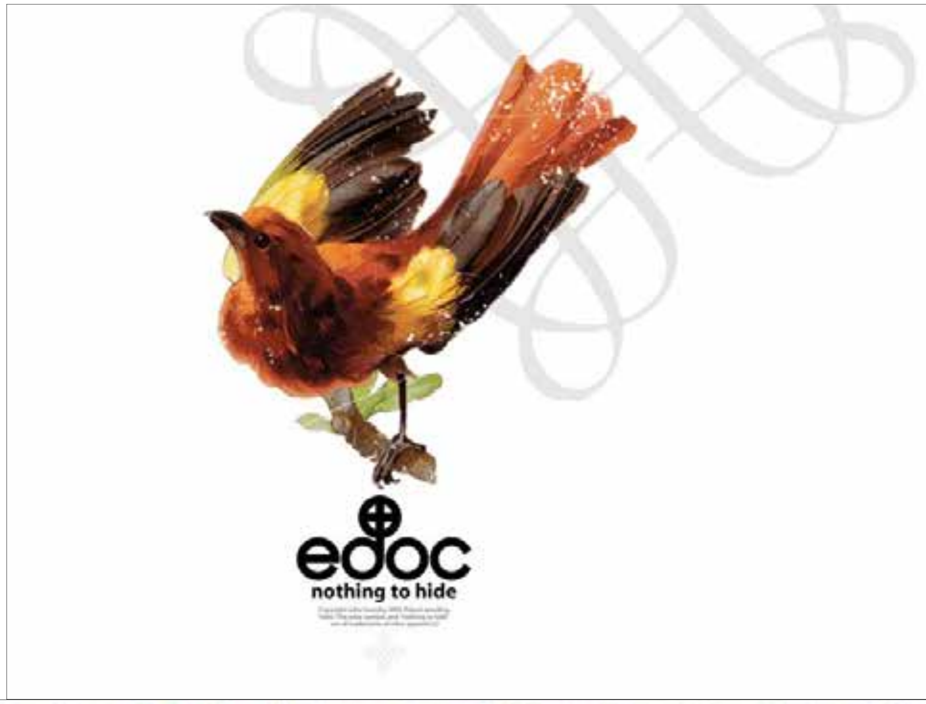
This is a business card sized color reproduction of the original graphic that was created from a black & white drawing that was then hand-colored.



My Clothes Tell Secrets

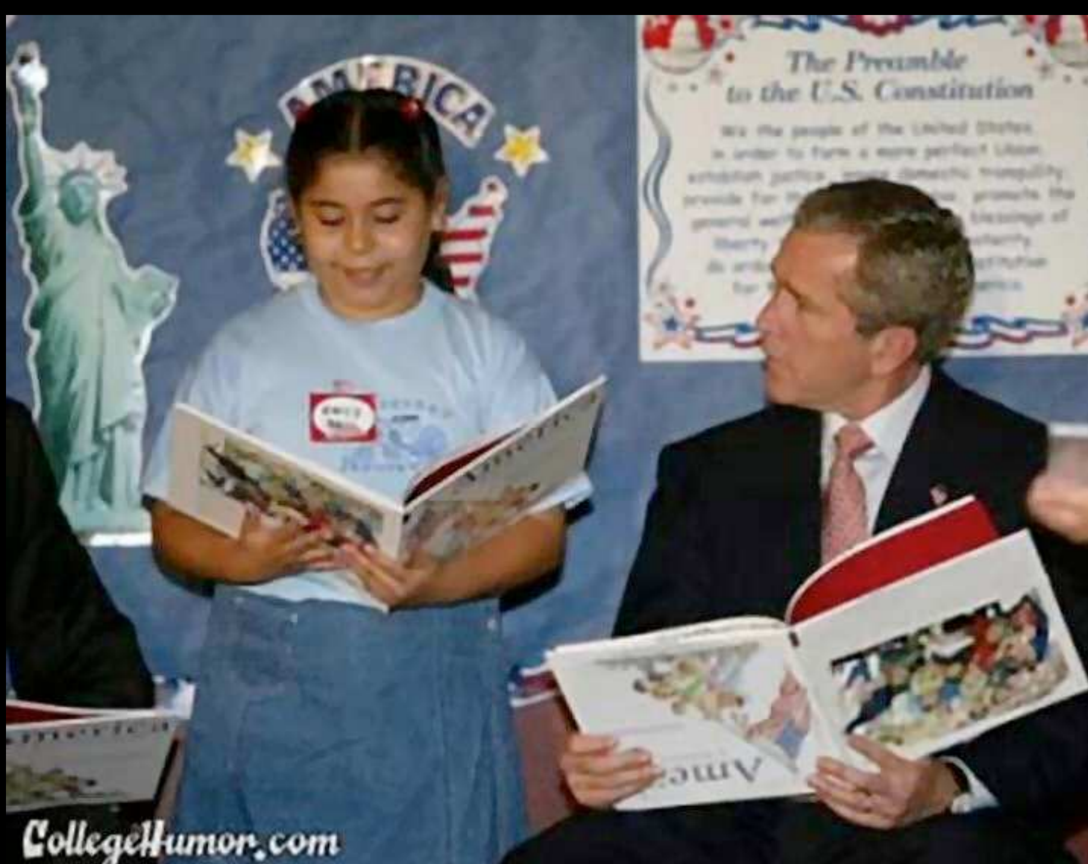
by Elan Lee

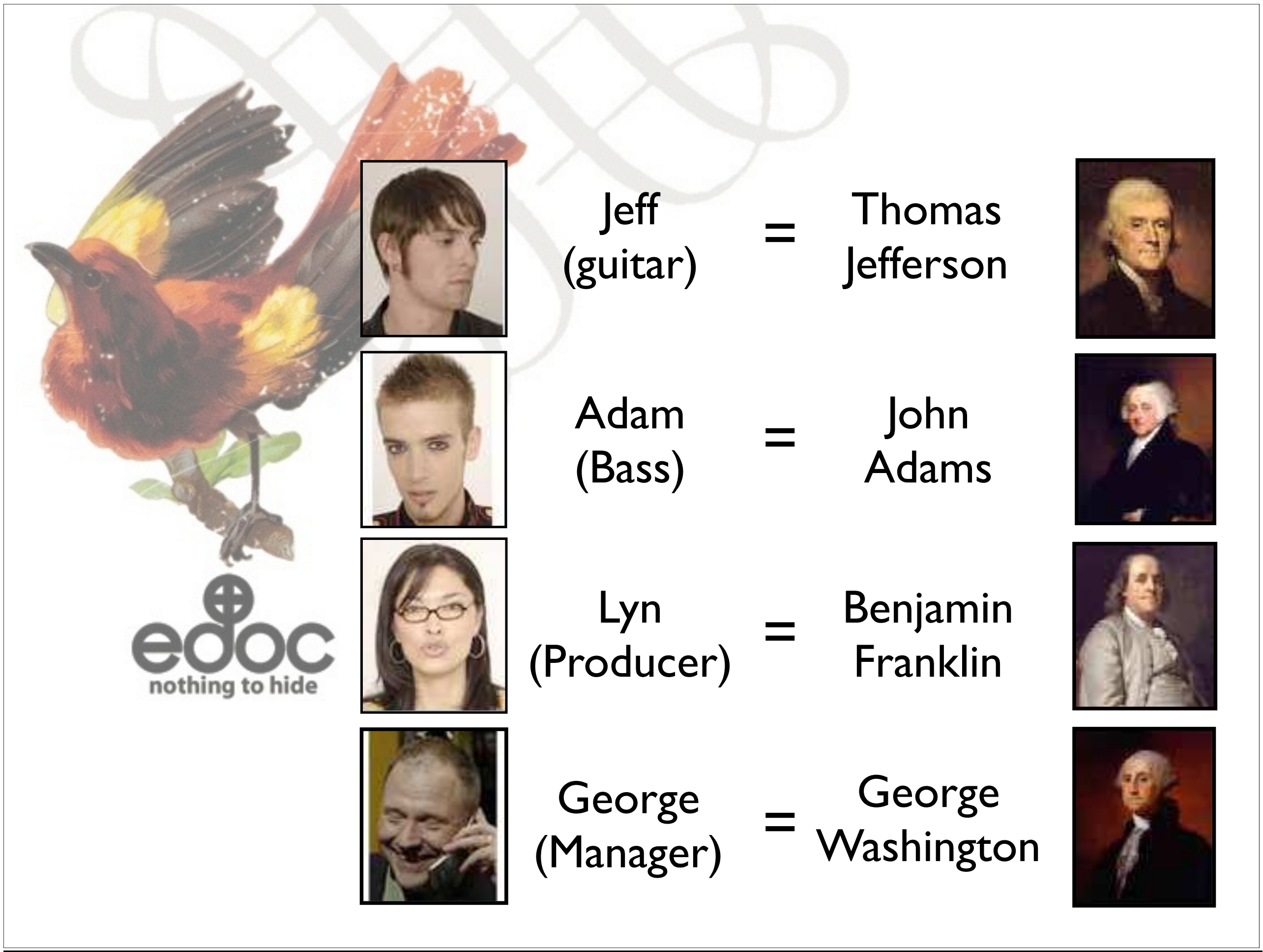
Elan Lee discusses his company EDOC Laundry. A clothing startup that embedded secret codes and hidden messages in a line of t-shirts to tell an episodic interactive story.



5


edoc
nothing to hide
About • Laundry • Contact • Stores
My clothes tell secrets.

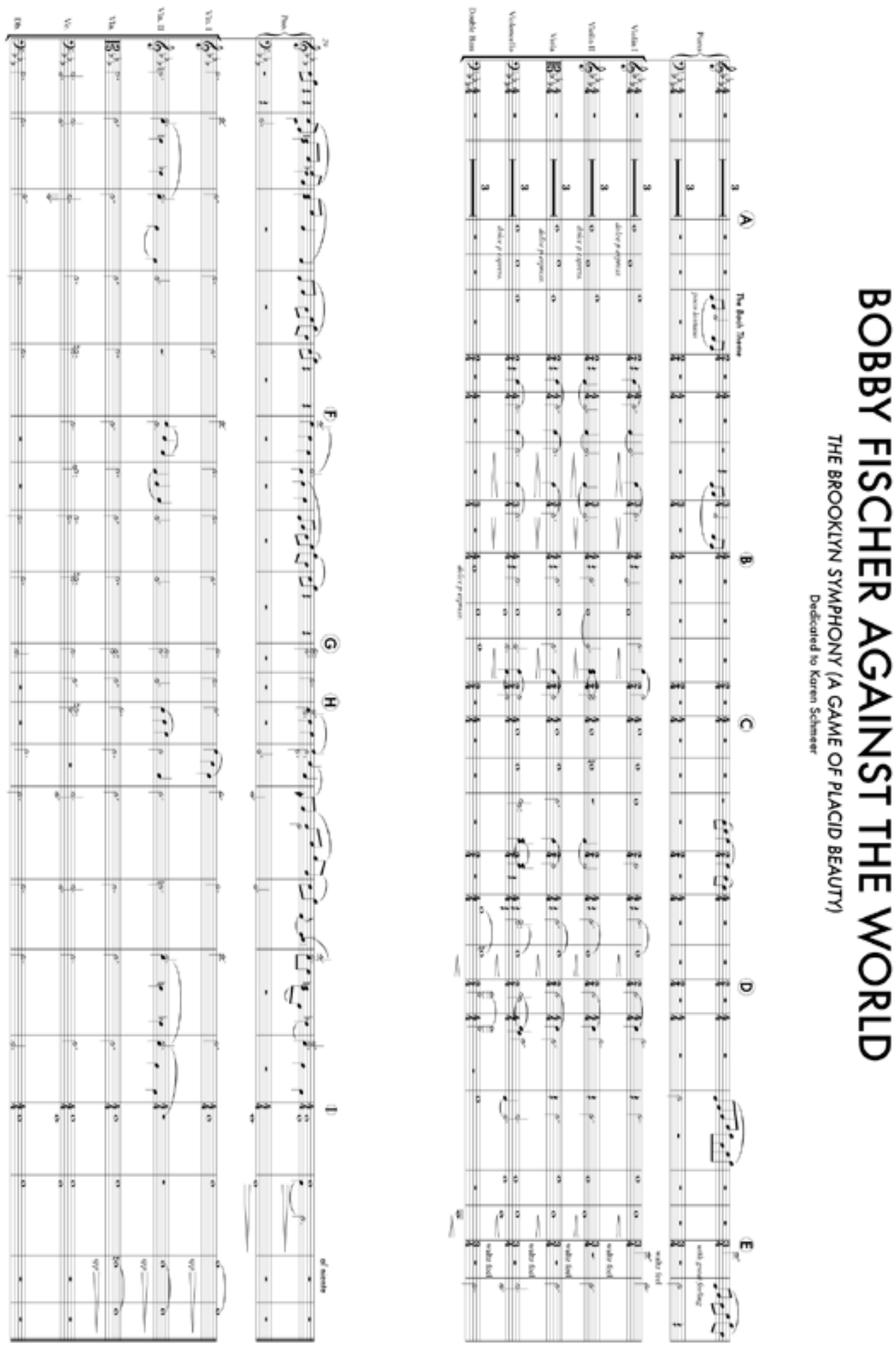

AMERICA
The Preamble
to the U.S. Constitution
CollegeHumor.com


Jeff
(guitar)
=
Thomas
Jefferson
Adam
(Bass)
=
John
Adams
Lyn
(Producer)
=
Benjamin
Franklin
George
(Manager)
=
George
Washington
edoc
nothing to hide

Bobby Fischer Against the World - soundtrack for G4G friends

by Philip Sheppard | Radiomovies Limited

A free hi-res download of the soundtrack to my ‘Bobby Fischer Against the World’ film. Download from here:
http://philipsheppard.bandcamp.com/album/bobby-fischer-against-the-world



GAMES

Oddly Enough

by Simon Aronson

(from *Try the Impossible* (2001)

This is not a magic trick. It's a bet, or very simple game, to be played among three players. The odds are straightforward – each player apparently has an equal chance of winning and the payoff to whoever wins is suitably 2 to 1. Nevertheless, it is a scam because the odds are not what they seem.

One of the problems with many hustles and scam bets is that the proposed rules of the game often aren't quite "normal." The procedures may be roundabout, in that they get to their point somewhat indirectly. They may feel somewhat contrived, in the sense that they aren't as clean as what one would have expected had a few people really just sat down "merely to bet" among themselves. In short, hustles often lack naturalness. Indeed, these extra steps or additional twists frequently are the essential camouflage in making the apparent odds different from the true odds.

I've always been fascinated by how a hustler can manipulate the odds in his favor, but being of a skeptical frame of mind and being nurtured in the deceptive motives and schemes of magicians, I know I would have hesitated to play in most of the scams I've read about. I wondered whether one could devise an utterly simple game whose rules appear completely understandable, fair and above-board, and natural – and yet still manipulate the laws of chance. The following bet would have suckered me in, because of its minimalist trappings and its "obviously" logical, straightforward procedure. It's precisely what three innocent guys might do, perhaps in a bar, to see who gets a "pass" on his share of the bill.

The Game

The rules are minimal. The hustler and two innocent marks each place their bet, say $5 each, into the pot. One of the marks shuffles and cuts a deck of cards and then deals one card face up to each player. If all three cards happen to be of the same color (i.e., all red or all black) the result is a tie or a "push," and the dealer would deal another round. If the dealt cards are two of one color and one of the opposite color, whichever player receives the single "odd" color card wins the pot. It's that simple.

The players can either stop after the first round, or they could each toss another bet into the pot and deal the next three cards face up to play another round.

On any round, the odds against any individual player receiving the odd-color card are obviously 1 out of 3, and the winner gets paid 2 to 1, so chance should give each player an equal opportunity to win or lose. At no time does the hustler ever need to touch the cards, and the mark's shuffle and cuts are free and genuine. Indeed, if the mark wants, after the first round, he could give the deck another free cut so that the second round gets dealt from a different place in the shuffled pack. What could be more fair?

The Scam

The scam consists of three factors. First, there is an initial secret set-up of the deck. Second, the deck is given just one riffle shuffle. Third, the position of the hustler at the table as the cards are dealt is key. All three are easily controlled.

The deck is initially set up with the cards secretly arranged in alternating colors. The hustler positions himself so that he will receive the *second* card dealt in each round. All the hustler needs to do is to sit immediately to the *right* of the mark who will do the dealing. So long as the cards are dealt in traditional fashion, going around from left to right, this seating arrangement will insure that the hustler gets the second of the three cards dealt.

If the cards are handled according to the following procedure, the hustler will win significantly more money than he will lose – because the "true" odds are that the hustler will receive the odd-colored card a full 50% of the time (i.e., on average, once out of every two rounds).

The Procedure

Let's call the mark sitting on your immediate left Lefty and the other player Righty. Explain the simple rules and have each of the three players toss an equal amount into the pot.

You're going to want Lefty to be the dealer, so for "fairness" and to lull Righty into the action, we'll have him mix the cards as follows: with the deck secretly stacked (i.e., in alternating colors), cut the pack toward Righty and ask him to complete the cut. Then have him cut the deck approximately in half and riffle the halves together. Finally, have him cut the deck again, this time toward Lefty. Lefty now completes the cut, and he's ready to begin dealing. (These cutting instructions can be varied as you like, since the deck may be cut as many or as few times as you like, either before or after the shuffle. The important point is that the deck be given only one riffle shuffle.)

The deck has now been legitimately cut and shuffled, so we're ready to play. (And, if Lefty prefers, he could cut the deck and complete the cut again, so that the dealing would start from a random and unknown position; it makes no difference.) Lefty deals one round of three cards off the top, dealing each card face up, starting with Righty, then to you, and finally to himself. The seating arrangement assures that you receive the second, or middle, card dealt. Look at the results, and give the pot to the winner (i.e., whoever receives the odd color card). It's that simple. If you prefer, all three cards could be dealt face down, and then the players would then turn them over to see who wins. This may affect the "look" of the game, but obviously it has no effect on the outcome.

If you desire, you can immediately suggest an additional bet, and then have Lefty continue to deal one more round (i.e., three more cards). On such a second round you can offer Lefty the option of either dealing the three cards from the point he left off, or cutting into the middle and dealing from there. In either instance, it's to your advantage to make that bet.

What's Going on

If the deck had not been shuffled, it would be relatively easy to figure out how to win, because an *unshuffled* alternating color stack *guarantees* that, if three consecutive cards are dealt from anywhere, the middle card *must* be the opposite color from the surrounding two. It's the shuffle that disarms people. Indeed, even when you comprehend the underlying probabilities, it's still hard to believe it really works in practice.

The easiest way to convince yourself of the real odds is to go through the above procedure and deal out the full deck into 17 separate rounds; you'll find that, over time, the second card dealt in each round will win about 8 or 9 rounds, and the two other hands will win about 4 or 5 each. (Since this is based on probability, don't be surprised if you need to go through a few full decks, more than just 17 rounds, before these overall odds start to appear. But they will appear eventually).

The reason this occurs is because of a novel application of the Gilbreath principle. Magicians familiar with this seminal concept are, of course, aware that the above shuffling and cutting procedure results in the deck's being in successive pairs, with each pair containing one of each color (let's refer to this as a "Mixed Pair"). The red card might come first in some pairs and second in other pairs, but the deck will, after one riffle shuffle, consist of consecutive Mixed Pairs. This means that you can *never* actually get three cards in a row of the same color (but your explanation of the rules is designed to let the marks think that such a "push" is a real possibility).

Second, and more importantly, it means that no matter where the cards are dealt from, every three consecutive cards dealt will always consist of one consecutive Mixed Pair, plus one more card (the "Non-pair" card). This Non-pair card must always be either the first or the third card dealt (never the "middle" card of the three), because the Mixed Pair is always of two *consecutive* cards. The Mixed Pair thus must comprise either the first two cards dealt or the last two cards, and the Non-pair card will fill the remaining space. Moreover, this Non-pair card can never be the winning card among the three because its color will always match one of the two Mixed Pair cards. This means that the winner in each round must always be one of the two cards in the Mixed Pair – namely, the one whose color is opposite to that of the Non-pair card.

The result of the above is that, in any given round, each of the two cards comprising the Mixed Pair has a 50/50 chance of winning that round, and the Non-pair card has no chance of winning. Since the middle card dealt is *always* one of the two cards comprising the Mixed Pair (regardless of whether the Mixed Pair falls either at positions 1 and 2 or positions 2 and 3), the chance of the second dealt card's winning any particular round is 50%.

One fascinating aspect of this analysis is that it is counter-intuitive. It would seem at first impression (to me, anyway) that if the second card dealt in each round has a "higher" chance of winning, then if the deck had been cut just one card deeper or shallower, then such an alternative cut should move the "increased likely winner" to fall into the first (or the third) position. But, in actual fact, this is not the case – because the cut simply determines whether the Mixed Pair will fall either to positions 1 and 2, or 2 and 3. In either case the card dealt second from that point, as determined by that cut, will still be one of the Mixed Pair. It's always comforting, both in scams and in magic, when an underlying principle is counter-intuitive, because there's less chance that it can be re-constructed later on.

Comments

(1) ***Odds And Ends.*** The above scam is perhaps the simplest thing that can be accomplished with such a set-up, but there are other facets which I've played with. For instance, as you watch the cards being dealt, as soon as you see a "double" (two cards dealt consecutively of the same color) you can know both the color of the next card before it's dealt, and also know the division points between each Mixed Pair from there on, as the deal continues. This allows you, in a more elaborate demonstration, to secretly count along for successive rounds and even predict the next winning color. If you were actually playing successive rounds, you might be able to vary the size of your bet, betting more when you knew you would win next and less when you saw you were going to lose.

Don't be tempted to deal too many successive rounds – because, frankly, it can look "too good." Unless Lefty is an erratic shuffler, it's quite possible for long runs of regularly alternating colors to appear over successive deals, and such a repeated pattern could enlighten Lefty and Righty that something's not quite random. It's better just to play a few rounds with the odds skewed in your favor, and call it quits. Tactically, if you know beforehand that someone habitually shuffles in odd clumps of twos and threes, select him to play Righty's shuffling role.

I have experimented with altering the set-up in minor ways, by varying the alternating color scheme in just a few areas of the deck. Such a modified set-up can creates the possibility of a round of three cards occasionally being all the same color. This minor variation does change the odds in a de minimus way, but heightens the sense that the outcome is actually determined by chance. It's something to consider if you're going to play more than a couple of rounds.

For analytic purposes, the text presents a "bare bones" procedure. You can easily make it more convincing if *you* start by giving the deck one or two simple false table riffle shuffles and then casually hand it to Righty, for his cuts and shuffle. Likewise I have omitted any discussion of the possibility of adding just a tiny bit of sleight-of-hand – but it's easy to see how the addition of "second dealing" can produce different results.

(2) ***Background and Credits.*** My starting point for "Oddly Enough" was Nick Trost's "Odd Man Wins" (Trost, *The Card Magic of Nick Trost*, 1997, p. 93). Trost's game procedure is quite different from the ideas outlined above since his purpose is to create an obviously controlled magic trick in which the "mark" *never* wins. To do this, he uses the Gilbreath principle in a fairly traditional manner, dealing out the entire deck into four piles. The victim then selects one pile, and depending on which pile the mark has chosen, the performer then discards one of the remaining three piles. Next the performer enlists the aid of a second spectator to play his confederate, and each takes one of the two remaining piles. Between the two of them, they always beat the mark because one or the other of them must always have the odd-color card.

Trost's procedure is acceptable in a magic context, but the illogicality of dealing out four piles just to discard one takes it out of the realm of creating a realistic betting situation. Likewise, Trost's magical goal of overtly demonstrating complete control despite a shuffle (so that the performer or his confederate always wins) is antithetical to the psychology of a valid scam, which generally lets the mark win just enough so that he doesn't ever realize he's being taken.

I wondered whether, by eliminating these two "magic trick" elements, I could make the conditions more closely resemble an uncontrolled, legitimate game. I didn't want to use a confederate, and I wanted my hand to be the winner (as opposed to Trost's procedure in which one single mark is the loser). So I explored the possibility of using the Gilbreath principle to secretly skew the odds if only three hands were dealt, and was delighted when my initial trials proved successful. I immediately went out dining with some of my lawyer friends, and the ensuing bets paid for the evening. The result is "Oddly Enough."

(3) ***Exculpation.*** At the risk of being repetitive, please understand that this scam does skew the odds in your favor – but on any one or two rounds it doesn't *guarantee* a win. The bet is certainly worth making if you can afford the possible loss, but please don't bet your house on it.

This scam and the odds-skewing principles set forth are published for purposes of entertainment and amusement only. So is the preceding sentence.

Tiebreaker Dice

Rod Bogart
bogart@gmail.com

Background

In 2012, Eric Harshbarger developed an intriguing set of dice that help players decide who should go first in a game with two to four players. Rather than 1-12 on each die, the numbers 1-48 are distributed across a set of four dice. Largest number wins, and the distribution is determined so that each die has an equal chance of being the winner when all possible outcomes are considered. For more info, see http://www.ericharshbarger.org/dice/#gofirst_4d12

Proposal

Given the idea of Go First Dice, I'd like to introduce the concept of Tiebreaker Dice. The goal is to return the 12-sided dice to their normal numbering, with some additional information to break ties.

The current set of Go First dice has the following numbers on the faces of the dice:

Die A:	1	8	11	14	19	22	27	30	35	38	41	48
Die B:	2	7	10	15	18	23	26	31	34	39	42	47
Die C:	3	6	9	16	20	21	28	29	33	40	44	45
Die D:	4	5	12	13	17	24	25	32	36	37	43	46

Table 1

By performing ((N-1) modulo 4) on Table 1, it produces the following data:

Die A:	0	3	2	1	2	1	2	1	2	1	0	3
Die B:	1	2	1	2	1	2	1	2	1	2	1	2
Die C:	2	1	0	3	3	0	3	0	0	3	3	0
Die D:	3	0	3	0	0	3	0	3	3	0	2	1

Table 2

This code could be used to indicate which die wins in a tie. If the dice are colored Red, Green, Blue, and Yellow, each die could have the numbers one through twelve as normal, with additional dots of the other dice colors. For a given 1-12 value, one die will have zero colored dots, one die will have a single dot, one die will have two dots, and the last die will have three colored dots.

In the event of a tie, the winner is the die that has the other die's dot color. A die face with zero dots loses all ties, and a face with three dots face wins all ties.

This proposal would work just as well as the 1-48 numbering scheme, but has the added benefit of allowing the dice to be used as regular 1-12 dice.

Repairing an apparent bias

However, with the Go First dice numbering, the Tiebreaker dice notation system appears to be biased. It is not, but this is not visually obvious. The die associated with the second row of Table 2 never wins a four-way tie.

It is possible to rearrange the codes in Table 2 to get the same Go First behavior with dice that also appear to be fair with Tiebreaker notation.

Die A:	0	3	2	1	2	1	3	0	2	1	0	3
Die B:	1	2	3	0	1	2	0	3	1	2	3	0
Die C:	2	1	0	3	3	0	2	1	0	3	2	1
Die D:	3	0	1	2	0	3	1	2	3	0	1	2

Table 3

With the above encoding, each Tiebreaker die has three chances of being first (or second, etc) if all four players roll the same number. This repairs the apparent bias of the previous dice. Of course, these dice have all the other desirable properties of Go First dice; when any two or three of them are used, each die will win an equal amount of ties.

Table 4 provides the dot patterns for each face of a set of Tiebreaker Dice that are Red, Green, Blue, and Yellow (R G B Y).

	R	G	B	Y
1		R	RG	RGB
2	YBG	YB	Y	
3	BY	BYR		B
4	G		GRY	GR
5	YG	Y	YGR	
6	B	BR		BRG
7	GYB		GY	G
8		RBY	R	RB

9	BG	B		BGR
10	Y	YR	YRG	
11		RYB	RY	R
12	GBY		G	GB

Table 4

A Strategy for **Borders**
a variant of Dots & Boxes

Jason Colwell
Associate Professor of Mathematics
Department of Mathematics and Computer Science
Northwest Nazarene University
Nampa, ID

---

introduction In this article we describe a winning strategy for the game of Borders, a game proposed by the author in his article for the G4GX gift exchange.

First we recall the description of the game, and two differences between it and Dots & Boxes.

description The game is played on an $a \times b$ grid (the cells being 1×1). The sides of the cells we call edges. (Each edge borders one or two cells.) Their initial state is "empty". A move consists of building a fence segment along an empty edge, after which the edge is no longer empty. If the fence built completes an enclosure of one or more cells, all cells within the enclosure are claimed by that player, except for any cells already claimed by the other player.

differences from Dots & Boxes The game Borders is different from Dots & Boxes in two ways:

- The building of a single fence segment may result in the player claiming multiple cells.
- Claiming a cell does not permit the player to build another fence segment.

The following conjecture was made by the author in the original article:
conjecture Suppose the game of Borders is played on an $a \times b$ board. The game is a win for the first player if $a + b$ is odd, and for the second player if $a + b$ is even.

In this article, we prove the conjecture by describing a winning strategy.

notation The grid is positioned in $\mathbb{R}^2$ with corners at $(0,0)$, $(a,0)$, $(0,b)$, and (a,b). For $1 \le i \le a$ and $0 \le j \le b$ we denote by

$$\overline{\boldsymbol{(i,j)}\mathbf{W}}$$

the fence segment from (i,j) "west" to $(i-1,j)$ (or the move consisting of building that fence). For $0 \le i \le a$ and $1 \le j \le b$ we denote by

$$\overline{\boldsymbol{(i,j)}\mathbf{S}}$$

the fence segment from (i,j) "south" to $(i,j-1)$ (or the move consisting of building that fence). We denote, for $1 \le i \le a,\ 1 \le j \le b$, by

$$\boxed{\boldsymbol{i,j}}$$

the cell with corners $(i-1,j-1)$, $(i,j-1)$, $(i-1,j)$, and (i,j).

"accessible" We say that one cell is "accessible" from another if they are adjacent and their common border is an empty edge.

"k-canal" By "k-canal", we mean a sequence of k cells bordered each by exactly two fence segments, each cell accessible from the previous one, and the first and last cells of the k-sequence not accessible from any cell outside the k-sequence having more than one fence segment bordering it. (Note the difference between the "canal" in Borders and what is termed a "chain" in Dots & Boxes.) Playing on a k-canal replaces it with a k-dead-end (if the new fence is built at one end) or with two canals whose lengths sum to k (if any other fence is built).

"k-dead-end" By "k-dead-end", we mean a sequence of k cells, each one accessible from the previous one, each cell except the last bordered by exactly two fence segments, the last bordered by three fence segments, and the first not accessible from any cell having more that one fence segment bordering it. To put it more intuitively, a k-dead-end is a k-canal closed off at one end. The k cells of a k-chain can be claimed in one move, and has a value of k to the player who plays on it.

"k-corral" By "k-corral", we mean a collection of k cells that can be claimed in one move. Note that a k-dead-end is a k-corral where any move will result in claiming some of the cells. Contrapositively, on any k-corral that is not a k-dead-end, a move can be made that will not result in the claiming of cells.

the strategy The strategy consists of attempting to perform the following Actions, listed in order of decreasing priority.

1. Do not create any dead-ends.
2. If there is a single corral, claim it.
3. If there are two corrals of equal size, claim one of them.
4. If there are two dead-ends of unequal length, claim enough squares from the closed end of the longer one to reduce its length to that of the shorter one.
5. if there is an edge where building a fence segment would create two corrals, one of which is not a dead-end, the two corrals containing numbers of empty edges of different parity, build a fence in the non-dead-end corral, so as not to claim any cells. This will cause the two potential corrals to contain numbers of empty edges of equal parity.
6. Play on a $2k$-canal, building a fence so as to split it into two k-dead-ends.

We shall call the "protagonist" the first player if $a + b$ is odd, or the second player if $a + b$ is even. The other player shall be called the "antagonist". The above strategy, when followed by the protagonist, will prevent him from losing. The game may end in a draw if the board has an even number of cells. In that case, an additional bit of strategy is necessary to ensure a win for the protagonist.

assumption If the protagonist follows the strategy just described, both players lose nothing by avoiding creating corrals as long as possible. We shall assume that they do this. The board's cells will then be eventually partitioned into canals (as well as, possibly, cells accessible from their ends that have no additional empty edges). If there are canals only of even length, the game will be a draw. On the other hand, as playing on an odd-length canal results in a loss of at least 1, both players will avoid this as long as possible, and play will continue until all even-length canals are split equally between the players, and only odd-length canals remain. The antagonist will have to play first on each of these, losing 1 point each time. Therefore, the game will be at best a draw for the antagonist, and a loss if there are any odd-length canals. If the protagonist can, while still following the strategy, ensure that there is at least one odd-length canal created, he can win the game. This additional bit of strategy will complete the proof that the protagonist has a win. The antagonist's possible moves are checked exhaustively. The first few moves by both players are described in the following discussion, and the subsequent moves left to the reader.

proof of strategy First note that as no more than two dead-ends can be created in a single move, Actions 1 and one of Actions 2, 3, and 4, can be performed, whenever one's opponent's move creates a dead-end. Any of these possibilities is of non-negative value to the player moving. Two corrals containing numbers of unbuilt edges of unequal parity will never simultaneously appear, due to Action 5 performed on a previous turn.

If the opponent's move does not create a dead end, then Action 1 may be performed except in the case where every place available to build a fence, belongs to a canal (and it will be to only one canal). The combination of Actions 1 and 5 will be of no loss to the player moving.

In the exceptional case, the player moving will try to perform Action 6, which has a value of 0. Only if Action 6 is impossible can there be loss to the moving player.

Note that at this point, if both players have played rationally, all squares claimed have been from even-length canals, split evenly between the two players. Each such canal now contains an even number of empty edges. Thus the number of empty edges contained in territory already claimed on the board is even. All unbuilt fence segments in unclaimed territory belong to odd-length canals, each of which has an even number of empty edges. This means that the number of fence segments is equal in parity to the total number of edges on the board, which is $2ab + a + b$. But this has the same parity as $a + b$, which means it would have to be the antagonist's, not the protagonist's, turn to move. That is, the protagonist will never have to deal with this situation where both Actions 1 and 6 are impossible.

It follows that the protagonist has at least a draw. □

additional bit of strategy The strategy described so far is enough to prevent a loss, but a draw may still result – only if all the canals into which the board is partitioned are of even length. Otherwise, as described above, the antagonist – and him only – will be forced at least once to play on a canal of odd length, resulting in a strict advantage to the protagonist.

We now describe a sub-strategy, constituting "Action 7" in the overall strategy until the odd-length canal has been constructed. Effectiveness of this sub-strategy is seen by exhaustive search. Most of the essentially different possibilities are treated below. Additional possibilities are handled similarly, or are of obvious cost to the antagonist.

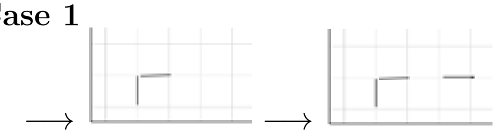
In the case where this sub-strategy is to be used, the board has an even number of cells; equivalently, at least one of a and b is even, say a. If the protagonist is the second player, there will be one fence already built when the protagonist begins using this strategy. In this case, the protagonist should choose the top row or the bottom row as necessary to avoid contact with the fence segment already built.

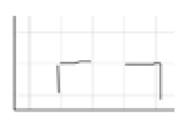
(The only way this last thing could be impossible is if $b = 2$ and the first move is some $\overline{(i,1)\mathrm{W}}$, but in this case it is relatively easy for the protagonist to make his first move $\overline{(1,1)\mathrm{W}}$ or $\overline{(a,1)\mathrm{W}}$ and ensure that a corner cell is bordered by two fences, which causes it to become a 1-canal, as desired.)

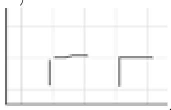
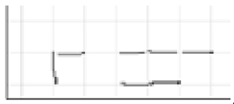
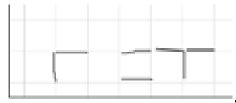
Say, without loss of generality, that the row chosen by the protagonist on which to begin play is the bottom row. The protagonist wishes to ensure that an odd canal is created in the bottom rom. The protagonist's first move is $\overline{(1,1)\mathrm{S}}$. If possible, his next move is to turn $\boxed{1,1}$ into a 1-canal by building $\overline{(1,1)W}$. But if the antagonist plays on $\boxed{1,1}$, making this impossible, then the protagonist's second move is $\overline{(2,1)\mathrm{W}}$.

The 10 essentially different states of the board after the next move by the antagonist, are depicted below. The response of the protagonist is shown in each case. Notes on the progress of the game after that, are given, with verification left to the reader.

Case 1

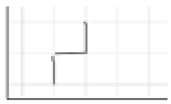
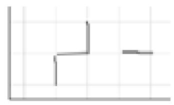


The protagonist should try to ensure that at least the fences  are built, which will force an 1-canal. Otherwise, a position containing one of

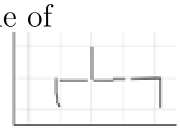
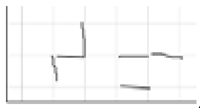
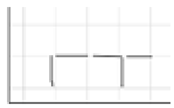
, , ,

can be achieved. All three involve shifting the play by an even number of cells to the right (2,4, and 4, respectively), where this same sub-strategy may be applied again. (If this occurs repeatedly, eventually the play will reach the right side of the board and result in an odd-length canal.)

Case 2

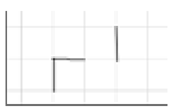
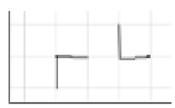
$\longrightarrow$  $\longrightarrow$ 

One of

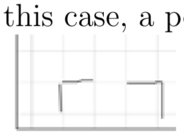
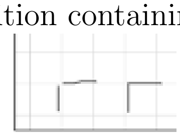
, , ,

can be achieved. In the first one, the objective has been accomplished. In the second and third, the play has been shifted two cells to the right, where this same sub-strategy may be applied again.

Case 3

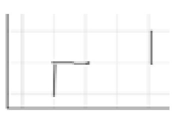
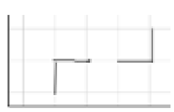
$\longrightarrow$  $\longrightarrow$ 

In this case, a position containing

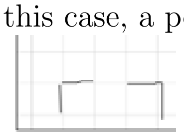
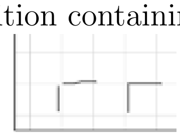
 or 

can be achieved, the first ensuring an 1-canal and the second shifting the play two cells to the right.

Case 4

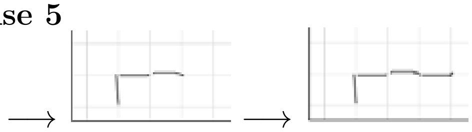
$\longrightarrow$  $\longrightarrow$ 

In this case, a position containing

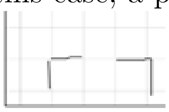
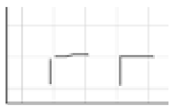
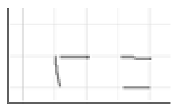
 or 

can be achieved, the first ensuring a 1-canal and the second shifting the play two cells to the right.

Case 5

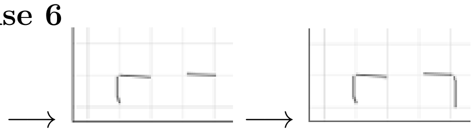


In this case, a position containing

 or  or 

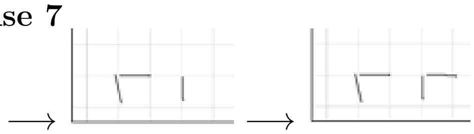
can be achieved, the first ensuring an 1-canal and the other two shifting the play two cells to the right.

Case 6



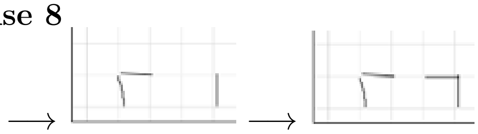
Here, a 1-canal is already assured.

Case 7



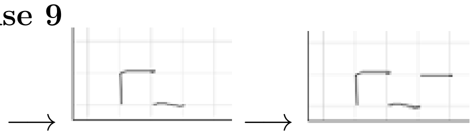
Here, the play is shifted two cells to the right.

Case 8

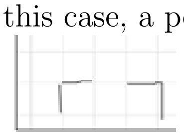
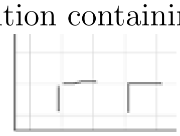
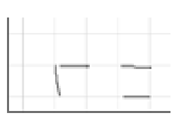


Here, a 1-canal is already assured.

Case 9

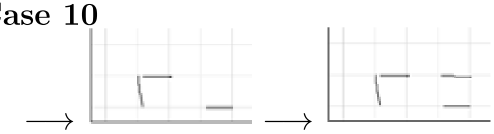


In this case, a position containing

 or  or 

can be achieved, the first ensuring a 1-canal and the other two shifting the play two cells to the right.

Case 10



Here, the play is shifted two cells to the right.

---

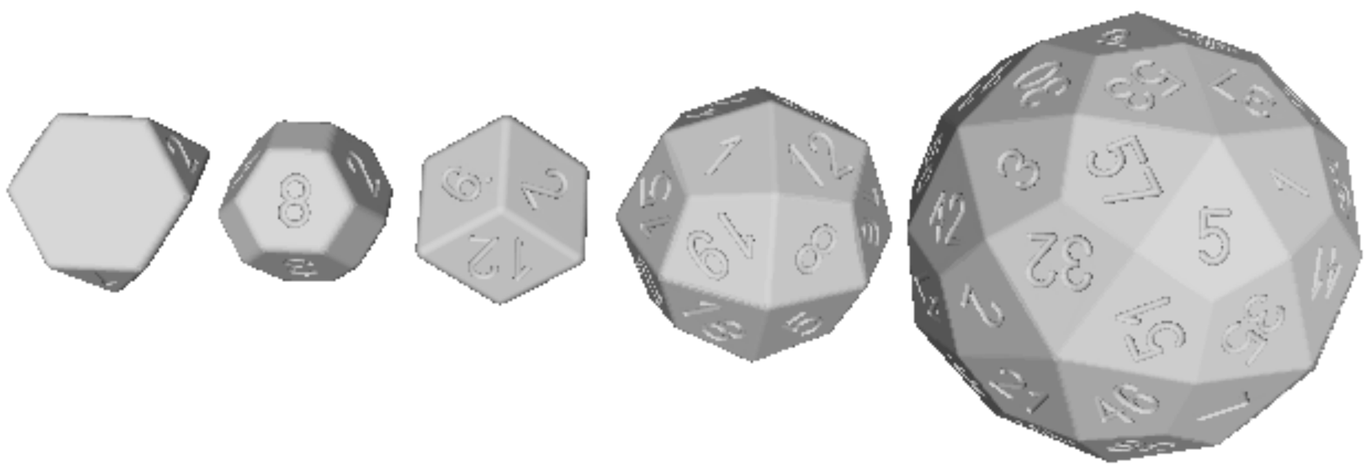
plans for future research The author plans to write a web program for the game of Borders, implement the strategy described in this article for the server's play against a human, and see how difficult it is, on an even-size board, for the human antagonist to force a draw (by making sure all canals created are even) if the protagonist server does not use the "additional bit of strategy". It would be nice to replace the exhaustive checking of cases by a more elegant formulation.

Unique Polyhedra Dice

by Robert Fathauer | Tessellations

Five new and unique polyhedral dice.

Three-dimensional rendering of new polyhedral dice:



A Cluster Analysis of Richard's PBEM Server

Lyman Hurd[1]

For G4G11, March 2014.

Introduction

The goals of this paper are twofold. First, it is my intention to share with the larger community knowledge of one of my personal favorite places on the Internet, Richard's PBEM Server ("Richard's PBeM Server"), a play-by-email server founded by Richard Rognlie in 1994! Secondly, I would like to make use of the openness of the platform to apply machine learning techniques to see what patterns can be discerned from user behavior, since Richard's server keeps records on games played going back to its foundation.

History of the server

Richard's server was founded in 1994 as means to play games by email in an automated fashion. Richard had previously been able to play Trax on a server at UC Berkeley, but the availability was not complete and the set-up did not automate tasks such as maintaining the user database. Gamerz was launched with the ability to play Trax (TraxGame) and Twixt. Before the first public announcement, the list also included Hex, Abalone and C++Robots (since relegated to its own implementation).

In 1998 Richard wrote: "[PBMServ] currently supports 50+ games, 1200+ users and 3500 requests/day" ("PBMServ History"). It is apparent from the experiments below that it has only continued to build. My personal history with the site started in 2001 when I became addicted to playing the game Zèrtz, which I had recently bought (Zèrtz).

Games

The range of games is broad but the emphasis is on abstract board games such as backgammon, chess, go, etc. The distinguishing feature of the server is that each of the games has its board drawn purely in ASCII characters and therefore is suitable for playing with any email reader in existence. In the figure below, a game of Quoridor is shown in progress.

[1] Google Inc., 1600 Amphitheatre Pkwy, Mountain View, CA 94043

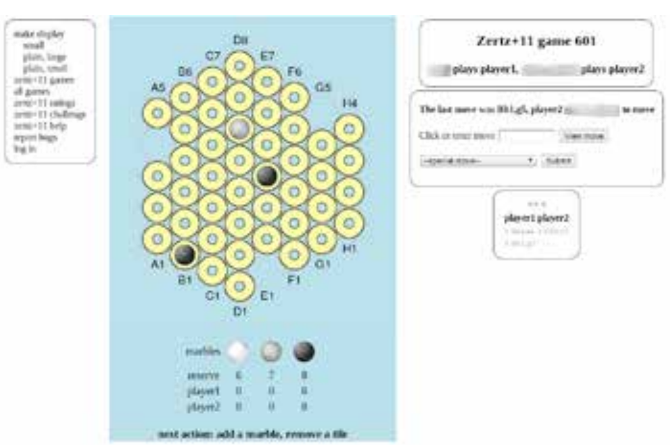
```
                         [N]

            A   B   C   D   E   F   G   H
            |   |   |   |   |   |   |   |
        +-----------------------------------+
        |                                   |
     8--|   +   +   +   +   +   +   +-------|--8              (X)
        |                                 O |           Fences - 3
     7--|   +   +   +   +   +   +-------+   |--7        ===========
        |                       |           |            | | |
     6--|   +   +   +   +   +   |   +-------|--6
        |       |               |           |
     5--|   +   |   +   +-------+-------+   |--5
[W]     |       |       |                   |     [E]
     4--|   +   +-------|   +   +   +   +   |--4
        |               |                   |
     3--|-------+   +-------+-------+-------|--3              (O)
        |           |                       |           Fences - 1
     2--|   +-------|   +   +   +   +   +   |--2        ===========
        |         X |                       |            |
     1--|-------+   +   +   +   +   +   +   |--1
        |                                   |
        +-----------------------------------+
            |   |   |   |   |   |   |   |
            A   B   C   D   E   F   G   H

                         [S]
```

Rules at: http://www.gamerz.net/pbmserv/quoridor.html
Graphics: http://www.gamerz.net/pbmserv/Quoridor/Quoridor.php?731&html

There are variants of traditional games, such as a large number of variants on the game of Backgammon (also called Hit), such as Moultezim and Plakoto, which are traditional variants on the game popular in other countries (Tzannes).

There are more modern games such as the games of Project GIPF by Kris Burm. In addition to abstract games there are card games such as Spades and Hearts, and some less traditional games such as Werewolf (also known as Mafia).

The server has been set up as a platform and allows for the addition of new games. Therefore a dedicated group of developers has sprung up to add new games to the server on a volunteer basis (if not previously mentioned, everything on the site is without cost). A relatively recent (in server time) addition to the site has been a graphical front end to allow games to be played via the web. This front end is completely integrated with the more traditional mail server and it is perfectly possible to mix and match with one player using the graphical interface and one email or even to have one player alternate between the two methods.



An exciting trend has been having game developers bring their games to Gamerz as a means of exposing them to an audience of fans. The prolific designer Cameron Browne, author of Connection Games and Hex Strategy, has added over forty games currently only playable online on the server. Most of the games he designed himself, but Yavalath has the distinction of having been designed by a computer (Evolutionary Game Design).

The games he has added include Druid[2], Mambo, Dragons, and a two-player game, Conway, based on John Conway's Game of Life (Mathematical Games)[3]

To continue a non-exhaustive list, Doug Zander has contributed Power Drain, Luke Pebody contributed Cooper Young and Mark Ballinger and I brainstormed Haggis as an adaptation of the Sid Sackson game Haggle (Sackson).

The Experiment

The experiment was conducted by downloading from the server ratings for 255 different games (there were actually even more data files than this, but I immediately excluded a handful of games, largely experimental, which had no complete games). Later, after working with the data I made the decision to cut off games with fewer than 50 completed games, which trimmed the list by 75 to leave 180 games to cluster. The excluded list included many interesting games (including two I had coded myself), but I was worried that the statistics would be too noisy. Ironically, the last game to make the cut (i.e., have over 50 games completed) was called Borderline.

Then there were the users whose behavior I tracked. There were in all 2362 distinct users. At first I weeded out all users who had only played one type of game, but later decided that this pruning would lead to unexpected effects in the distance metric used and so I reversed that decision, i.e., every game was represented by a vector with 2362 dimensions. The most prolific user on the site (in terms of breadth) had completed games in 176 different categories.

The statistics yielded for each user, games won, lost, drawn and the user's rating (discussed more below). To simplify matters, the only data kept was the total number of games completed (wins + losses + draws). Using a technique similar to the "bag of words" technique for document classification ("Bag-of-words Model"). Each game was then associated with an N-dimensional vector where each dimension represented a distinct user, i.e., each game was expressed as a bag of users.

Some users are huge fans of certain games and I wanted to make sure that these superfans did not drown out the contributions of their peers, so I stepped in and made a further arbitrary decision to cap the number of games played by a single user at 20 (the number of games necessary to be an "established player" on the server).

---

[2] To the insistent commenters on Boardgamegeek, everyone realizes that the druids did not actually build Stonehenge and they certainly did not put down stones willy-nilly for the purpose of walking over them.

[3] **Achievement unlocked:** obligatory John Conway reference.

The reason for this cap was that the next step was to normalize the vectors to unit length. Without taking that step, it would quickly be the case that some games were essentially only represented by a few dedicated fans. Before I made this adjustment, there were some games deemed close that seemed implausible given the nature of the game (not that there were no surprising correlations in the final results).

Next, the problem was to consider how close games were to each other. I used the “cosine distance” which is simply:

$$1 - x \cdot y$$

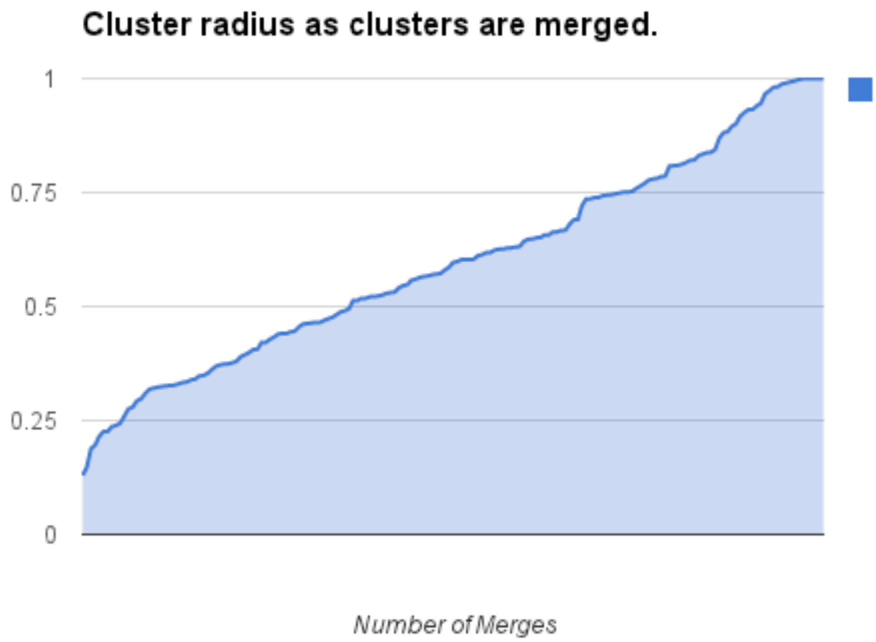
for normalized game vectors x and y. Since the components of every game vector are guaranteed to be non-negative, the result is always a number between 0 and 1. To be equivalent (distance zero) two games would have to be played by exactly the same group of players in the same proportion.

Two sets of games were considered to have distance equal to the diameter of the combined set, i.e., the furthest distance achievable by a pair of points in the combined set. By using this criterion instead of trying to compute centroids of clusters, the program never actually had to perform the normalization or dot product calculation after the initial run and all the clustering operations were able to operate based on a lookup table.

Clustering was done by agglomeration. The starting point was a list of 180 clusters with each containing one game, and then the program recursively removed the closest pair of clusters from the list and replaced them with a new cluster formed from their union. This new cluster required expanding the lookup table of distances to account for the distance of this new cluster from all existing clusters, however in this setup, this could be obtained by setting the distance of a cluster to the new cluster to be the maximum of the distances from that cluster to its two subcomponents.

Results

This graph marks the progress of the algorithm as it successively combined clusters to form new clusters. Therefore, the left edge of the graph corresponds to the situation of having a set for every game and the rightmost extreme corresponds to the result of combining all 181 games into a single cluster. At each point in the progression, the program kept track of the diameter of the newest formed cluster. The hope was that there would be a logical point of inflection in this graph, which would be true if there were really discrete clusters such that merging to that point would entail relatively little gain in error and combining beyond that point would mark a sharp trend upwards. The actual data, however, showed a much smoother transition, which meant there was not necessarily an obvious stopping point or ideal number of clusters.



Here are the results of the clustering algorithm when it was allowed 20 distinct clusters. The choice to stop at 20 was arbitrary, but it was the first level of aggregation that put the Project GIPF games into their own cluster, and 20 was approximately the number of different hand-curated categories used by Richard Rognlie in the game descriptions on the server's front page.

Cluster (number of plays)	Notes
Cluster 1 Total Plays (147460) Backgammon (87554) Scramble (16291) Plakoto (13795) Moultezim (10445) Hypergammon (9447) Nackgammon (4269) Deadgammon (2325) Lingo (1729) Grandgammon (1605)	Backgammon variants and word games. I found it surprising that the algorithm lumped these two together...
Cluster 2 Total Plays (28694) Othello (4258) Go (3895) Pente (3417) TwixT (2490) LoA (2476)	Most of these are established abstract games with a long history.

Hex (2156) Amazons (2063) Wari (1560) Havannah (890) Gomoku (790) Phutball (645) Draughts (615) Hexxagon (606) Hexade (533) Dama (432) Epaminondas (426) NMM (389) Connect4x4 (383) FireAndIce (355) Emergo (315)	
Cluster 3 Total Plays (23617) Renju (22461) Ninuki (1039) S-Pente (117)	Renju (RenjuNet)] has a significant fan base all its own. It is the professional version of the game gomoku (five in a row).
Cluster 4 Total Plays (14965) Trax (7864) Conhex (1597) Quoridor (1366) Gonnect (1270) Druid (966) Akron (622) Unlur (574) Stymie (321) Oust (126) Crossway (94) Jungle (83) Batalo (82)	
Cluster 5 Total Plays (8491) Chess (4003) DarkChess (1366) OmegaChess (973) DoubleChess (646) ProgressiveChess (342) QuickChess (327) RennChess (237) AvalancheChess (129) Chex (123) Amoeba (100) Bughouse (90)	Chess and its variants...

GrandChess (87) Capablanca (68)	
Cluster 6 Total Plays (6852) Zertz (2629) Zertz+11 (971) Yinsh (817) Tzaar (716) Dvonn (711) Punct (543) Gipf (393) Tamsk (72)	Project GIPF
Cluster 7 Total Plays (6544) Mambo (850) Yavalath (595) Y (481) Margo (472) Mutton (419) Halves (364) Holo (355) Lambo (243) Dragons (218) Chameleon (180) Palago (171) Forms (164) Chroma (150) Che (149) Limit (134) Osbo (130) Thoughtwave (127) Antipod (110) Boche (109) Jade (106) Gates (105) Mono (105) Orbit (86) Sonar (81) Dna (80) Cross (73) Trichet (68) Star (64) Ndengrod (63) Blobs (62) Visavis (61) Trilbert (61) TimeVectors (56)	Many games by Cameron Browne.

Malaka (52)	
Cluster 8 Total Plays (6040) ToW (1203) Abalone (660) ROthello (420) ChineseCheckers (395) Ataxx (391) K-Pente (380) Neutron (366) Tanbo (352) Connect4 (349) Gravity (261) Checkers (244) Susan (239) Rings (195) Plotto (193) Hexbo (178) Connectris (125) MaxCheckers (89)	
Cluster 9 Total Plays (5439) Yacht (2670) Warship (1385) CooperYoung (785) Toot (402) Sudoku (115) Chaos (82)	
Cluster 10 Total Plays (4609) Wizard (1682) Perudo (1177) MHearts (524) Cathedral (489) Spades (312) Blackout (169) WarpAndWeft (91) Conway (83) Hive (82)	This includes many of the card games although it mixes in pure abstracts such as Cathedral.
Cluster 11 Total Plays (3731) Shogi (1696) ChuShogi (1036) TenjikuShogi (566) Xiangqi (433)	Shogi and most of its variants (although Minishogi has a cluster of its own).
Cluster 12 Total Plays (3204)	

CoNeutron (1948) Robots (1197) Koan (59)	
Cluster 13 Total Plays (2452) Fanorona (376) Powerdrain (332) Alak (246) Soccolot (228) Blackbox (224) Breakthrough (219) Psycho (209) NCBackgammon (162) Stack (111) Tumble (107) Pitch (102) Survival (76) Focus (60)	Many of these are the games of Doug Zander
Cluster 14 Total Plays (2434) Dots (816) Spangles (657) Andantino (387) Oddthello (129) Quadrature (122) Plotto5 (108) Entropy (82) Connexxions (77) Qubic (56)	
Cluster 15 Total Plays (1517) Octi (615) OctiLite (330) RazzledazzleX (311) Octi3base (183) Razzledazzle (78)	
Cluster 16 Total Plays (911) Stratego (261) Hnefatafl (213) DoubleMoveChess (209) Trinim (176) Borderline (52)	
Cluster 17 Total Plays (416) Reversi (144) Majorities (86)	

Projex (66) Quadrex (62) LoopTrax (58)	
Cluster 18 Total Plays (223) Onyx (147) Accasta (76)	
Cluster 19 Total Plays (152) Terrace6x6 (89) Terrace (63)	
Cluster 20 Total Plays (104) MiniShogi (104)	

Further Exploration

There were several facets of gameplay that were not explored by this analysis. For one thing, the popularity measure was only conducted using the metric of games completed. The data contained in addition the ratings (see http://www.gamerz.net/pbmserv/ratings.html). This information could potentially be used as a different source of information, although it would probably be necessary to restrict to "established" players, i.e., those with at least twenty games played, as ratings tend to be highly noisy before that point.

One different problem that has occurred to me is whether one can determine from a ratings distribution the relative degree of skill in a game or decide whether there are tiers of players. Elwyn Berlekamp distinguishes levels of knowledge that characterize played of Dots and Boxes based on specific revelations. It has been my experience that Zèrtz exhibits a similar phenomenon, especially since on the smaller board it is possible to lose the game on the first move against a knowledgeable opponent.

Another factor not taken into consideration was the length of time games had been available on the server. This obviously gave a bias towards games that had been around longer. For newer games there is a recently implemented "history" command, which is more granular than the "ratings" command used here. It would have helped take that into account.

Lastly, games take differing lengths of time, which means that number of games is not necessarily the best metric to use. Time spent playing may be a better metric or number of moves played.

Conclusion

This concludes the exploration of Gamerz. The list of games is large and grows larger and larger over time as long as there are individuals willing to extend it.

Acknowledgements

The author would like to thank the developers who participate in building and extending Gamerz and in particular: Cameron Browne, Jonathan Bush, John Williams, and Doug Zander, whose comments led to great improvements in this paper. Most of all, the author would like to thank Richard Rognlie for giving me in his words many happy opportunities to reduce my productivity.

Works Cited

"Bag-of-words Model." Wikipedia. Wikimedia Foundation, 30 Mar. 2014. Web. 30 Mar. 2014. <http://en.wikipedia.org/wiki/Bag-of-words_model>.

Berlekamp, E. R. Dots and Boxes. 1971. Print.

Browne, Cameron. Connection Games: Variations on a Theme. Wellesley, MA: A.K. Peters, 2005. Print.

Browne, Cameron. Evolutionary Game Design. London: Springer, 2011. Print.

Browne, Cameron. Hex Strategy: Making the Right Connections. Natick, Mass: A.K. Peters, 2000. Print.

"Cameron's Games Page." Cameron's Games Page. Web. 30 Mar. 2014. <http://www.cameronius.com/games/>.

"Cosine Similarity." Wikipedia. Wikimedia Foundation, 26 Mar. 2014. Web. 30 Mar. 2014. <http://en.wikipedia.org/wiki/Cosine_similarity>.

"Games Project: Gipf, Tamsk, Zertz, Dvonn and Yinsh." Games Project: Gipf, Tamsk, Zertz, Dvonn and Yinsh. Web. 30 Mar. 2014. <http://www.gipf.com/>.

"MATHEMATICAL GAMES The Fantastic Combinations of John Conway's New Solitaire Game "life"" Conway's Game of Life: Scientific American, October 1970. Web. 30 Mar. 2014. <http://www.ibiblio.org/lifepatterns/october1970.html>.

"PBMServ History." PBMServ History. Web. 30 Mar. 2014. <http://www.gamerz.net/pbmserv/history>.

"Ratings." Ratings. Web. 30 Mar. 2014. <http://www.gamerz.net/pbmserv/ratings>.

"The Renju International Federation Portal - RenjuNet." The Renju International Federation Portal - RenjuNet. Web. 30 Mar. 2014. <http://www.renju.net/>.

"Richard's PBeM Server." Richard's PBeM Server. Web. 30 Mar. 2014. <http://www.gamerz.net/pbmserv>.

Sackson, Sid. A Gamut of Games. New York: Random House, 1969. Print.

Tzannes, N. S., and Basil Tzannes. Backgammon Games and Strategies. South Brunswick, NJ: A.S. Barnes, 1977. Print.

"Welcome to the Official TRAX Website!" Trax Resource Index. Web. 30 Mar. 2014.

<http://www.traxgame.com/index.php>.

"ZÈRTZ (1999)." ZÈRTZ. Web. 30 Mar. 2014.

<http://www.boardgamegeek.com/boardgame/528/zertz>.

Mixed Hands Blackjack
A New Blackjack Problem

Bill Liles
lilesw@gmail.com

June 25, 2016

1 Statement of the Problem

This problem follows standard Blackjack rules except for one exception. In standard Blackjack, a player may play more than one hand but the player is not allowed to move cards from any other there hands to any other of their hands.

In this version of the game, the player is allowed to pool all the cards in all of their hands and redistributed them into the same number of hands consisting of two cards each. For example if the player receives two hands with hand one containing an Ace and a three and hand two contains a ten and a seven, the player could rearrange the cards so that one hand now contains an Ace and a ten (Blackjack) and the other had contains a three and a seven for a total of ten. But once the hands have be rearranged, they cannot be rearranged again.

The problem is what is the best strategy for rearranging the cards to maximize winnings.

2 Variations

There are a number of variations to the above problem, some are which are listed here. For example, the player could rearrange the cards so that some of the new hands can have more than two cards. Or there could be multiple players and they could interchange cards between themselves, which include buying and selling cards. Or the dealer is also allowed to have multiple hands, in which case, which hand do the players play against? Or hands that have not busted are allowed to be rearranged multiple times. They are many such variants.

3 Relationship to casino games

In the problem formulated above, the player plays multiple hands and can rearrange the cards amongst those hands. There is a version of this played in casinos, called Blackjack Switch or Blackjack Swap. In this version of the game, the player plays two hands and is allowed to interchange the second card dealt to each hand.

In our version, the player can play more than two hands and all cards can be rearranged.

4 Terminology

In this paper, the following terminology will be used.

(1) A,2,3,4,5,6,7,8,9,X represents the numeric value of a card, where A represents Ace and X represents a 10 count (i.e. a ten, Jack, Queen or King).

(2) M, N, P, Q, R, S represent arbitrary cards

(3) H represents the number of hands

(4) C represents the number of cards in all hands $C = 2 * H$

(5) ! represents factorial $n! = n * (n - 1) * ... * 1$

(6) !! represents double factorial $n!! = n * (n - 2) * ... * 1$ for n odd

(7) ? represents summation of integers from 1 to n expressed as $n?$

(8) $R(n)$ represents the number of possible hands arrangements for n cards.

(9) $P(n)$ represents the number of distinct hands

5 Number of possible hands orderings

Assuming one could not rearrange the cards between hands but could order the hands will be played in, how many such orderings are there? For example assume the first hand dealt contained A2 and the second hand contained 34. By reordering the hands, the first hand will now contain 34 and the second hand contain A2.

This is the same question as how many ordering are there of n objects. In this case the number of objects is H so the number of orderings in $H!$.

In the actual play of a game, if the player is card counting, this could have a slight theoretical advantage.

6 Number of possible rearrangements

Given C cards and that each hand must have two cards in the hand, how many arrangements are possible. Examples will be given for 1,2 and 3 hands.

Remember that the order of the cards in a Blackjack hand are irrelevant.

In the following examples, the cards within a hand are listed vertically and the hands are spaced apart horizontally. A small horizontal gap means that the hands are one arrangement. A larger horizontal gap signifies a new arrangement.

Example for one hand using cards M and N.

M
N

For one hand there is only one arrangement. $R(2) = 1$

Example for two hands using cards M, N, P and Q.

M	P		M	N		M	N
N	Q		P	Q		Q	P

For two hands there are three arrangements. $R(4) = 3$

For three hands we can generalize from two hands using cards M,N,P,Q,R and S. Notice that card M is paired with cards N, P and Q in the different arrangements. So given the initial hands as being:

M	P	R
N	Q	S

Card M will be paired with cards N,P,Q,R and S. That is five different pairings. For each pairing, there are four other cards to be arrange. From the example of two hands, we know that there can be three hand arrangements for four cards. Thus the total number of arrangements for three hands is $5 * 3 = 15$ or $R(6) = 15$.

This can be express as $R(C) = (C-1) * R(C-2)$. This can be re-expressed in non-recursive terms as $R(C) = (C-1)!!$. Remember that C is the number of cards. This can be also expressed in terms of the number of hands, H, as $R(H) = (2*H-1)!!$.

The above was all done assuming that the order of the hands does not matter. If the order of the hands does matter than we need to include that into the possible arrangements. From Section 5 we know there are $H!$ orderings so the total number of rearrangements when ordering is important is $H! * R(H) = H! * (2*H-1)!!$.

7 Number of possible different distinct hands

This is slightly different than the prior section. The question is what is the number of possible distinct hands. If we refer again to the example of two hands we have the different rearrangements as:

M	P		M	N			N
N	Q		P	Q		Q	P

The distinct hands are MN, MP, MQ, NP, NQ, PQ. Thus there are six distinct possible hands. The formula for calculating this is straight forward. Pick any card, in this case M. It must be able to be paired with all possible other cards. If there is a total of C cards, then the number of possible pairings is $C-1$. Now pick any other card, in this case N, and it must be able to be paired with all other cards, but it has already been paired with the first card picked, so there are $C-2$ possible new pairings. Continue for the next card, which in this case is P. But it has been already paired with the prior two selections so there are only $C-3$ pairings possible.

So the total number of possible distinct hands, $P(H)$ is simply the summation of all of these numbers. This can be expressed as shown below.

$$P(H) = \sum_{i=1}^{2*H-1} i = (2*H-1)?$$

8 Minimum number of unique hands and arrangements

So far, all the calculations have been based upon all the cards in all the hands are distinct. That does not have to be the case.

For example, with two hands, all four cards could be the same such as a four. Is this case the number of distinct hands is one and the number of rearrangements is one.

For three hands from a single deck, there could be four cards of the same value, and the remaining two cards with the same value as each other, and different from the four cards of like value. In this case, the number of distinct hands is two and the number of arrangements is also two.

For four hands from a single deck, there could be four cards of the same value, and the remaining four cards having the same value as each other, and different from the first four. For example, there could be four fives and four sevens. In this case, the number of distinct hands is three and the number of rearrangements is three.

I am still working on the general formula for the minimum number of unique hands and arrangements when duplicate cards are allowed.

9 Conclusion

In this paper, a new version of Blackjack was presented that raises different mathematical problems than the traditional version. In future work, the general formula to determine the minimum number of unique hands and arrangements will be presented, and also a strategy for selecting an optimal arrangement of the cards.

THE PREHISTORY OF NIM GAME

The aim of this paper is to recount the ancestors of the Nim game, taking as a reference the one introduced by Charles Leonard Bouton in his article, 1901.[1] We will see that the Nim already existed in a different form, which is called nowadays the additive version or the one pile game, and our purpose will be to follow its evolution over the centuries.

I. The Origin of the Name "Nim", Quite a Nimstory!

In a first paragraph, we will see that the word "Nim" appeared with Bouton. Yet, it would be naïve to believe that the Nim was completely invented by Bouton and that it did not exist in other forms or under other names before 1901. Games as simple as Nim, which does not require any board game nor perennial stand, are handed over verbally; therefore this orality is difficult to track and rewriting the story and the evolution of these games is the most often a laborious task.

A. The Germanic Origin

After Bouton's article was published, a lot has been extensively written about the name "Nim" and its possible origins. A link was pointed out between the Nim and the Fan-Tan,[2] a game of Chinese origin; but it was wrong, as chance is part of the Fan-Tan. Richard Epstein[3] explains the Eastern link by the simplicity of the Nim's structure and the strategically subtle moves when it comes to a mathematical point of view. Fan-Tan also seemed to have been played by American students as early as the end of the 19th century.[4]

---

[1] Charles, Bouton, " Nim, A Game with a Complete Mathematical Theory ", *The Annals of Mathematics*, 2nd

[2] Martin, Gardner, *Hexaflexagons and other mathematical diversions The First Scientific American Book of Puzzles and Games*, 1958, Chicago and London, Revised Edition, The University of Chicago Press, 1988. p. 151.
Raymond, Archibald, "The Binary Scale of Notation, A Russian Pleasant Method of Multiplication, The Game of Nim and Cardan's Rings", *The American Mathematical Monthly*, 25.3 (March 1918): 132-142. p.141.
Claude, Berge, "Théorie générale des jeux à *n* personnes", *Mémorial des sciences mathématiques*, Fascicule 138, 1957, pp. 1-114, Paris, Gauthier-Villars, 1957. p. 28.

[3] Richard, Epstein, *The Theory of Gambling and Statistical Logic*, Elsevier, 2nd Edition, 2009. p. 334.

[4] James, Glimm, John, Impagliazzo, Isidore, Singer (Editors), *The Legacy of John von Neumann*, *Proceedings of symposia in pure mathematics*, Vol. 50, American Mathematical Society, 1990. p. 283.

It was not before May 1953 that Alan Ross[5] denied the Chinese origin of Nim in a short note published in the *Mathematical Gazette*, in which he explained that Chinese names never end with a –m. Nevertheless, this note did not prevent from the confusion being made between the two games. According to Ross, the most likely explanation would be that the word Nim comes from the imperative of the German verb *nehmen*, which means to take (the imperative being *nimm*) and that Bouton would have chosen this word remembering his academic years spent in Leipzig where he received his PhD.[6]

This short note of May 1953 was not to go unnoticed, and in December of the same year,[7] Joseph Leonard Walsh, a student then a colleague of Bouton at Harvard University, sent a message to the editor of the *Mathematical Gazette* in order to confirm the German origin of the word Nim, explaining that *nimm* was frequently used during a game. And there was more to come. Eight years later, still in the *Mathematical Gazette*, in October 1961, N. L. Haddock[8] published a note about Alan Ross's note and suggested a link between the Nim game and the Mancala,[9] taking as a basis Harold James Ruthven Murray's work, *A history of board-games other than chess* (1952). Haddock stated that according to Murray, the Mancala game would be of Egyptian or Arab origin, and Nim would be a derived form[10] for the European and American continents. In short, there were many various speculations about the origins of the Nim game and its name:

> Did [Bouton] have in mind the German nimm (the imperative of nehmen, "to take") or the archaic English "nim"[11] ("take"), which became a slang word for "steal"? A letter to The New Scientist pointed out that John Gay's Beggar's Opera of 1727 speaks of a snuffbox "nimm'd by Filch", and that Shakespeare probably had "nim" in mind when he named one of Falstaff's thieving attendants Corporal Nym. Others have noticed that NIM becomes WIN when it is inverted.[12]

---

[5] Alan, Ross, "Mathematical Note No. 2334 : The Name of the Game of Nim", *The Mathematical Gazette*, 37.322 (May 1953): 119-120.

[6] Martin Gardner also considered this possibility in a work in 1983 in which he devoted a chapter to the Nim game and the Hackenbush. See Martin, Gardner, *Wheels, Life and Other Mathematical Amusements*, New York, W. H. Freeman Company, 1983. p. 143.

[7] Joseph, Walsh, "Correspondence: The Name of the Game of Nim", *The Mathematical Gazette*, 37.322 (December 1953): 290.

[8] N. L., Haddock "Mathematical Note No. 2973 : A note on the game of Nim", *The Mathematical Gazette*, 45.353 (October 1961): 245-246.

[9] Mancala comes from the Arabic *mâkala* and refers to the same game than the *Awélé* (with the exceptions of some variations) or the *Wuri*. Actually, the name depends on the geographical place we are. The game board consists of two parallel rows of six holes, each hole containing the same number of seeds. The aim of the game is "to sow" the seeds, in turns, redistributing them in other holes and taking them when they move from one row.

[10] Indeed, conversely to Mancala, the Nim game does not need any board with holes and its aim is not to win as many points as possible.

[11] Richard Epstein uses the word *niman* (and not *nim*) as the verb from the archaic English meaning "to take" or "to steal". Richard, Epstein, *The Theory of Gambling and Statistical Logic*. p. 335.

[12] Martin, Gardner, *Wheels, Life and Other Mathematical Amusements*. p. 143.

The true story still remains nebulous and will remain so, unless new information about what Bouton had in mind when he wrote his article emerged.

B. The Game "la luette ", "les luettes", or "l'aluette" [13]

The game "l'aluette" or "jeu de la vache" is a French card game mainly played in rural and coastal areas between Gironde and the Loire estuary, in a geographical area that is under the influence of the dialects of the Poitou and of Saintongeais. It is a trick-taking game played by four players (two teams of two) with forty-eight Spanish-looking cards, one of which displaying a cow, hence the name of the game "jeu de la vache". This game was still played in cafés around 1960 but is given up today. It seems it was introduced in France during the 16th century. Indeed we can find the game "les luettes" in François Rabelais's *Pantagruel* (1532), *Gargantua* (1534) and the *Cinquième livre* (1564). Yet, the rules of the game are not precisely defined and consequently, nothing enables us to state that the game "les luettes" in Rabelais's works corresponds with the game "l'aluette" as it is known today. But in one of the first French-English dictionaries, published in 1611 by Randle Cotgrave (born 16th, dead in 1634), *A Dictionarie of the French and English Tongues*,[14] we find at the headword "*luettes*" the following definition: "Luettes. Little bundles of peeces of Ivoirie cast loose upon a table; the play is to take up one without shaking the rest, or else the taker looseth."[15] Studies were conducted in order to highlight the influence of Rabelais on the interpreters, readers and emulators of the French language, more especially on Cotgrave. In a work dated 1930, Lazare Sainéan explained: "At the beginning of the 17th century, Randle Cotgrave, an English lexicographer, undertook to explain all the specificities of Rabelais' language, and this significant work (1611) remains nowadays one of the main sources for the comprehension of Rabelais' works."[16] "Above all, Cotgrave is Rabelais' glossary, his first and sole interpreter in the lexicography area."[17] This suggests that the game "les luettes" in Rabelais' work consisted of chip stacks and that the initial idea was to take away some chips without moving the others.

---

[13] http://fr.wikipedia.org/w/index.php?title=Aluette&oldid=89760744 [12.11.2013]

[14] Randle, Cotgrave, *A Dictionarie of the French and English Tongues*, London, 1611.

[15] Idem. http://gallica.bnf.fr/ark:/12148/bpt6k50509g/f588.image.r=cotgrave%20randle.langFR

[16] Lazare, Sainéan, *L'influence et la réputation de Rabelais*, Paris, Librairie Universitaire J. Gamber, 1930. p. 34: " Au commencement du 17ème siècle, un lexicographe anglais, Randle Cotgrave, s'attacha à expliquer toutes les particularités de la langue de Rabelais, et ce travail considérable (1611) reste encore aujourd'hui une des principales ressources pour l'intelligence de l'œuvre." My translation.

[17] Idem. p. 82 : "Cotgrave est avant tout le glossateur de Rabelais, son premier et unique interprète dans le domaine de la lexicographie." My translation.

These rules are closer of those of a game like the Mikado but Bouton may have drawn his inspiration from this ancient game and abandoned the *physical skill* side in favour of a more strategic angle. This kind of transformation can also be observed in the game of Kayles by the famous puzzlist Henry Dudeney: originally, it was a simple bowling game of skill, then it turned into a more mathematical parlour version.

II. Mathematical Recreations in Europe

A. The Birth of Mathematical Recreations

Mathematical recreations can be considered as a new genre, halfway between pure recreation, the full educational tool, and the launch of challenges between scientists. These "marginal"[18] mathematics have not had the same aims over time:

> The first recreational works date back to the 1620s, with *Les problèmes plaisants et délectables sur les nombres* by Claude Gaspard Bachet de Meriziac and *Les Récréations mathématiques composées de plusieurs problèmes plaisants et facétieux* by Henry von Etten, while the famous Jacques Ozanam's *Récréations mathématiques et physiques* were published in 1692. The aim of the ancient recreations was above all to "pique one's curiosity", whereas those that appeared at the end of the 19th century and the beginning of the 20th had three other purposes.
> The first was to teach mathematics […] The second was to circulate new mathematics [...] The third aim was to educate with sharing recent historical researches [...] Moreover, these recreations sometimes explicitly compensate for the weaknesses that, at the end of this century [19th], France often acknowledged in the field of its mathematic research as well as in public education.[19]

We will see that this new genre significantly developed thanks to Bachet's and Ozanam's works during the 17th century, but yet that Luca Pacioli introduced it at the end of the 15th century. Our study mainly deals with the first books of mathematical recreations, those that

---

[18] "à la marge" : Evelyne, Barbin, "Les *Récréations* : des mathématiques à la marge ", *Pour la Science* (February-March 2007): 22-25. p. 22. My translation.

[19] Idem. "Les premiers ouvrages de récréations datent des années 1620, avec *Les problèmes plaisants et délectables sur les nombres* de Claude Gaspard Bachet de Méziriac et les Récréations mathématiques composées de plusieurs problèmes plaisants et facétieux d'Henry von Etten, tandis que les fameuses *Récréations mathématiques et physiques* de Jacques Ozanam sont publiées en 1692. Ces récréations anciennes ont surtout pour but de « piquer la curiosité », tandis que celles qui paraissent au tournant des XIXe et XXe siècles ont trois autres motifs.
Le premier motif est d'instruire aux mathématiques […]. Le deuxième est de diffuser des mathématiques nouvelles […]. Le troisième est de cultiver en faisant connaître les recherches historiques récentes. […] De plus, ces récréations suppléent parfois explicitement aux déficiences que la France de cette fin de siècle [le 19ème] se reconnaît souvent, aussi bien dans sa recherche mathématique que dans son instruction publique." My translation.

were aimed to "pique one's curiosity". We will focus on recreational puzzles presented as a game opposing two players who take turn trying to reach a given number, n, with adding up numbers in the range from 1 to k. This recreational problem is a former version of Bouton's Nim and the player who knew the solution could easily impress his opponent by displaying his mastery during every game; it was a way to be a social success, even to have a hold over the people who did not have the key. This tendency can be observed too, as we will see further on, in the practice of Tiouk-Tiouk in West Africa.

B. Luca Pacioli (Italy, 1508)

A simplified variation of the Nim game appeared in Europe for the first time during the Renaissance with the Italian mathematician Fra Luca Bartolomeo de Pacioli (1445-1517) and his treatise *De Viribus Quantitatis* written between 1496 and 1508. Pacioli was one of the most famous mathematicians of his time.[20] According to David Singmaster,[21] the *De Viribus Quantitatis* can be considered as one of the first works entirely devoted to mathematical recreations. Pacioli began to write this manuscript in Milan where he was teaching between 1496 and 1499.[22] He was then at the height of his career, a prominent member in the intellectual circle of the Duke of Milan, Ludovico il Moro (1452-1500).[23] The opisthographic manuscript, kept at the Bologna University Library, is written in Italian[24] and consists in 309 sheets, 24 by 16,5 cm (9.4"x6.5) organised into three parts. The first includes 120 arithmetical recreations (*Delle forze naturali cioé de Arithmetica,* only 81 problems are indexed), the second part consists of 139 problems dealing with geometry and topology (*Delle forze naturali cioé de Arithmetica*, 134 problems are indexed) and the third part contains a few hundred proverbs, poems, riddles and puzzles (*De documenti morali utilissimi*, 85 are indexed). One of Pacioli's key aims was to reveal the power of numbers and to demonstrate that they could be understood in a concrete way with the card games, dice, Tarot and board games he proposed. One problem of the first part (*XXXIIII*, see Fig. 1), considered as the first

[20] In 1496, Pacioli published the most important mathematics work after Fibonacci (1202), the *Summa de Arithmetica, Geometrica, Proportioni et Proportionalità*; it was his major work. See David, Singmaster, "*De Viribus Quantitatis* by Luca Pacioli: The First Recreational Mathematics Book", in Erik Demaine and Martin Demaine and RODGERS Tom, (Editors), *A Lifetime of Puzzles*, A K Peters, Ltd, Wellesley, 2008. pp. 77-122.
[21] *Ibid*. p. 77.
[22] *Ibid*. pp. 81-82. Pacioli left Milan in 1499 and went to Florence following Sforza's overthrow by a French invasion. He taught at the universities of Florence and Pisa between 1499 and 1507.
[23] Idem.
[24] Pacioli preferred Italian to Latin.

one pile game,[25] is the following one: “finish any number before the opponent, without taking more than a certain finite number”.[26]

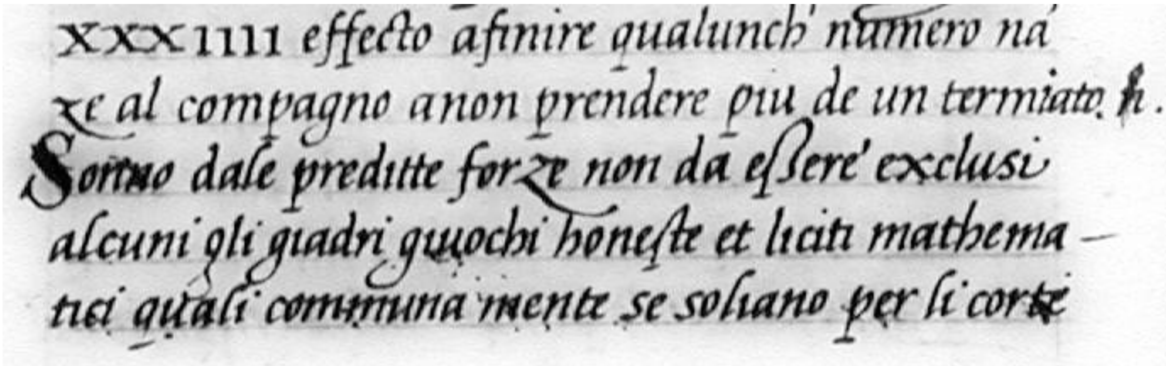

XXXIIII effecto afinire qualunch' numero na
ze al compagno anon prendere piu de un termiato .n.
Sortno dale predutte forze non da essere exclusi
alcuni gli giadri giuochi honeste et liciti mathema-
tici quali communa mente se soliano per li corte

Fig. 1: XXXIIII proposed by Pacioli in the first arithmetical part of *De Viribus Quantitatis.*
Source: **www.uriland.it/matematica/DeViribus/Pagine/175.JPG**[27], page 073v

This rather unclear wording makes sense when Pacioli gives the solution: he suggests that two players reach the number 30 with adding in turns numbers ranging from 1 to 6. Pacioli justifies the number 6 as the higher number of points that can be reached with a dice:

> It is customary to say among two [persons], that when you take points from a dice *alternatim* [in turn], you can take any number that you wish as long as it does not exceed 6, because in dice there is no greater [number of] points than 6; and the first person takes it on himself to get to 30 before his companion.[28]

In fact, it is a simplified version of Bouton’s Nim game, the one that is played with only one pile consisting of 30 objects and in which each removal is limited to a maximum of 6 objects. This version is called “additive” because one adds numbers instead of reducing piles, which was often the case in the first versions of Nim. Yet, the solution remains based on the same principle as Bouton’s Nim: there exist “safe combinations” that secure the win, provided that one plays correctly. Before revealing the method to find the safe combinations, Pacioli

[25] *Ibid*. pp. 91-92.
[26] Luca Pacioli, *De Viribus Quantitatis*, 1508. p. 073v: “ effecto afinire qualunch’ numero na’ze al compagno anon prendere piu de un termiato .n.” My translation.
[27] The whole manuscript was digitized and can be found on the following website: www.uriland.it/de-viribus-quantitatis-pages
[28] This translation of Pacioli’s Effect XXXIIII was kindly given to me by David Singmaster during the Board Game Studies Colloquium XVI, held in April, 3rd-6th 2013. It is part of David Singmaster’s translation (draft) of De Viribus Quantitatis. I warmly thank him for this help.

explains that this game is part of recreational, honest and legitimate mathematical games, which are fully entitled to be part of mathematics lessons and which everybody can go in as a mathematical recreation. Then he wonders if there is any advantage for the first or the second player to begin the game; finally, he quickly gives the winning strategy to be applied: the four safety steps, 2, 9, 16 and 23 have to be reached. At this stage, Pacioli does not explain how he has determined these safety steps, but it seems obvious that he succeeded in doing so using the following backward induction: if I don't want my opponent to win, I must put forward the highest number such as if my opponent added 1, 2, 3, 4, 5 or 6, he cannot reach 30. This number is 23; indeed, whatever number is added to 23, the resulting number is higher than or equal to 24, but remains less than or equal to 29. So, I will be able to complete up to 30 at the following turn. Reasoning in the same way with 23, the safety step before 23 is 16, then 9 and finally 2. So, I must manage to be the first to reach one of these safety steps, then the others up to 30. Pacioli does not explain safety steps by this method but he suggest a general method to find the safety steps of any game: "always divide the number that you wish to arrive at by one more than has been taken and the remainder of the said division will always be the first [step of the] progression [...]."[29] If the division comes out exactly and the remainder is zero, Pacioli considers that this case is more difficult and he clearly explains the backward induction he applies in a precise example. He takes the case when the number to be reached is 35 by adding numbers between 1 and 6:

> [...] For 35, take away 7, and 28 remains for the [step]; the other [step] takes away 7, and 21 remains; the other [step] 14; the other [step] 7. Therefore, he takes whatever he wants up to 6, and you will take, or actually you will make 7 the first degree, and then 14, 21, 28, and 35, and so on [...][30]

For Pacioli, the two situations are to be differentiated, whereas they actually need the same reasoning; they could come down to the following result: if the remainder of the division of n by k+1 (n being the number to reach and k the maximal number that can be added) is equal to zero, then we subtract k+1 to find the last safe combination, and so on until the first safety step is reached. If the remainder of the division of n by k+1 is different from 0, therefore it is the first safe combination and the others are determined by adding this remainder to the former safe combination. It must be noticed that Pacioli never gives this explanation by generalizing numbers and by using n and k. Indeed, it was not before 1901,

[29] Idem.
[30] Idem.

with Bouton's article, that the solution of a combinatorial game was to be formalized and studied as a general case with any number of piles and any number of chips in each pile. This is due to the belated development of algebra and the lack of an appropriate symbolism, which is required to represent unknown quantities and write equations. For centuries, men used clever arithmetical methods in order to solve problems that we would tackle nowadays with algebra. Consequently, it was then impossible to obtain general conclusions; each case was studied independently, and no formula, for which specific data to each example had only to be replaced, was developed.[31] René Taton confirms that fact:

> Consequently, the Renaissance algebra never presents *formulas*, but gives *rules* and offers *examples*. It is exactly what grammar does too, giving us *rules* that we have to follow, and *examples* that we have to comply with by declining names and conjugating verbs.[32]

Actually, this aspect can be found in all the cases of additive Nim game we analyse in this work.

The *De Viribus Quantitatis* manuscript is a collection of *mathematici ludi*, i.e. recreational mathematics including games or problems through which the author wished to teach mathematics avoiding the boredom due to the repetition of exercises frequently asked.[33] Before Pacioli, other authors had the same idea, such as Fibonacci, from whom Pacioli recognized frequent borrowings, Francesco and Pier Maria Calandri, but in the other arithmetic treatises (*trattati d'abbaco*), recreational problems are simply placed here and there in the text, in order to provide a rest to the reader who is learning; consequently, Pacioli's manuscripts can be considered as the first true treatise on the subject.[34] Vanni Bossi points out a *magical* nature, easily intelligible, in most of Pacioli's assertions.[35] It would seem that Pacioli did want the secret of the method to be well kept, in order to surprise the audience; indeed secret is the fundamental principle of any magic trick. Keeping this secret is the

---

[31] Vera, Sanford, *The History and Significance of Certain Standard Problems in Algebra*, Teachers College, Columbia University, New York City, 1927. p. 17.
[32] René, Taton, *La science moderne de 1450 à 1800*, Presses universitaires de France, Paris, 1958 (1ère édition Quadrige, avril 1995) p. 52: "En conséquence, l'algèbre de la Renaissance ne nous présente jamais de *formules*, mais nous donne des *règles* et nous offre des *exemples*. Exactement comme le fait la grammaire qui, elle aussi, nous donne des *règles* que nous devons suivre, et des *exemples* auxquels nous devons nous conformer en déclinant les noms et en conjuguant les verbes." My translation.
[33] Vanni, Bossi, "Magic and Card Tricks in Luca Pacioli's *De Viribus Quantitatis*", in Erik, Demaine and Martin, Demaine and Tom, Rodgers (Editors), *A Lifetime of Puzzles*, A K Peters, Ltd, Wellesley, 2008. pp. 123-130.
[34] *Ibid*. p. 123.
[35] *Ibid*. pp. 123-130.

essential condition for being able to amaze one's friends, especially women, "maxime donne", and those who do not know mathematic principles because they had no access to arithmetic knowledge.[36] This dimension of intellectual domination thanks to the control of the game is also to be found in the West African game Tiouk-Tiouk that we will talk about further.

Pacioli did not claim his originality loud and clear; some of the problems he suggested came from more ancient works or were studied or talked about in public schools at that time and handed out orally.[37] Some problems were even invented by his students and Pacioli encouraged them to do so. For instance, in one chapter, he mentions his disciple Carlo de Sansone from Perugia and in another one he names Catano de Aniballe Catani from Borgo, who would have played one of the problems in Naples, in 1486.[38] This date allows us to suppose that most of the problems in Pacioli's manuscript were in fact invented during the last quarter of the 15th century.

At the same time as *De Viribus Quantitatis*, the *Triparty en la science des nombres* by Nicolas Chuquet,[39] a Parisian doctor, came out in France without yet being published. This work "achieves a much higher level than the former works, even than Luca Pacioli's *Summa*."[40] Very few information is available on Chuquet, except that he was from Lyon and that he had a deep knowledge in arithmetic and algebra. Recently, it has been proved that Chuquet maintained contacts with the Italian tradition through provincial intermediaries.[41] The notebook of Francesco Bartoli, an Italian entrepreneur who regularly travelled between Italy and the South of France, gives some rare proofs of the hand-over of recreational problems through Europe:

> In addition to arithmetic tools such as exchange and multiplication tables, it contains a collection of thirty problems of the recreational sort. We can assume that Bartoli was only one of the many links in the trade routes by which the tradition of recreational mathematics passed from Italy to France and the Low Countries.[42]

The *Triparty* consists of two distinct parts: the first contains the *Triparty en la science des nombres* and the second deals with the *Applications des Règles du Triparty* with some

---

[36] *Ibid*. p. 124.
[37] Albrecht, Heeffer, *Récréations Mathématiques* (1624), *A Study on its Authorship, Sources and Influences*, October 2004. pp. 1-37. p. 18. Available on: http://logica.ugent.be/albrecht/thesis/Etten-intro.pdf [20.12.2011]
[38] Vanni, Bossi, *Magic and card tricks*... p. 125.
[39] (born between 1445-55- dead 1487-88)
[40] René, Taton, *La science moderne*… p. 19.
[41] Albrecht, Heeffer, p. 16.
[42] Idem.

sheets devoted to *Jeux et esbatements qui par la science des nombres se font.*[43] Unfortunately, we have not found games that could resemble the problem suggested by Pacioli, "yet, if we cannot assert that one of the works had directly influenced the others, their similarities prove that they are all part of a same tradition."[44] René Taton regrets that the *Triparty* had not the influence it should have had upon the development of algebra, "and it was Pacioli's *Summa* that, during the following century, was used as a starting point, and as a secondary source, for the theoretical and practical mathematical knowledge."[45]

Determining the (hi)story of a recreational game is extremely difficult because authors did not precise whether they borrowed the idea from someone else or not. This is particularly true in the case of the French Claude-Gaspard Bachet de Méziriac.

C. Claude-Gaspard Bachet (France, 1612)

Luca Pacioli's manuscript was kept in the archives of the University of Bologna for nearly five hundred years without being published. Maria Garlaschi Peirani[46] transcribed it into modern Italian in 1997, and in 2007, according to *The Guardian*, a translation into English had been undertaken.[47] The original book had rarely been consulted since the Middle Ages but it was undoubtedly a reference for later works. Actually, the additive version of Nim crossed centuries and frontiers and it was to appear again in various books of recreational mathematics as early as the 17th century. It is the case of *Problemes plaisans et delectables, qui se font par les nombres*,[48] the famous book by Claude-Gaspard Bachet known as de Méziriac (1581-1638), which is often considered as the first work on mathematical recreations (Fig. 2).

---

[43] Aristide, Marre, *Le triparty en la science des nombres par maistre Nicolas Chuquet Parisien publié d'après le manuscrit Fonds Français N°. 1346 de la Bibliothèque Nationale de Paris et précédé d'une notice par M. Aristide Marre*, Rome, 1881. p.32: "games and recreations, which are played thanks to the science of numbers." My translation.

[44] René, Taton, *La science moderne*… p. 19: "toutefois, si l'on ne peut affirmer que l'un de ces ouvrages ait influencé directement sur les autres, leurs similitudes démontrent qu'ils appartiennent à une même tradition." My translation.

[45] *Ibid*. pp. 21-22: "et ce fut la *Summa* de Pacioli qui, pour le siècle à venir, servit de point de départ, et de source secondaire, du savoir mathématique théorique et pratique." My translation.

[46] Luca, Pacioli, *De viribus quantitatis*, trascrizione di Maria Garlaschi Peirani dal Codice n. 250 della Biblioteca universitaria di Bologna ; prefazione e direzione di Augusto Marinoni, 1997.

[47] http://www.theguardian.com/world/2007/apr/10/italy.books
Lucy, McDonald, "And that's Renaissance magic…" *The Guardian*, 10 April, 2007.
As far as I know, no official translation has been published yet, conversely to what was announced in the article.

[48] Claude-Gaspard, Bachet, *Problemes plaisans et delectables, qui se font par les nombres*, Lyon, 1st edition, 1612.

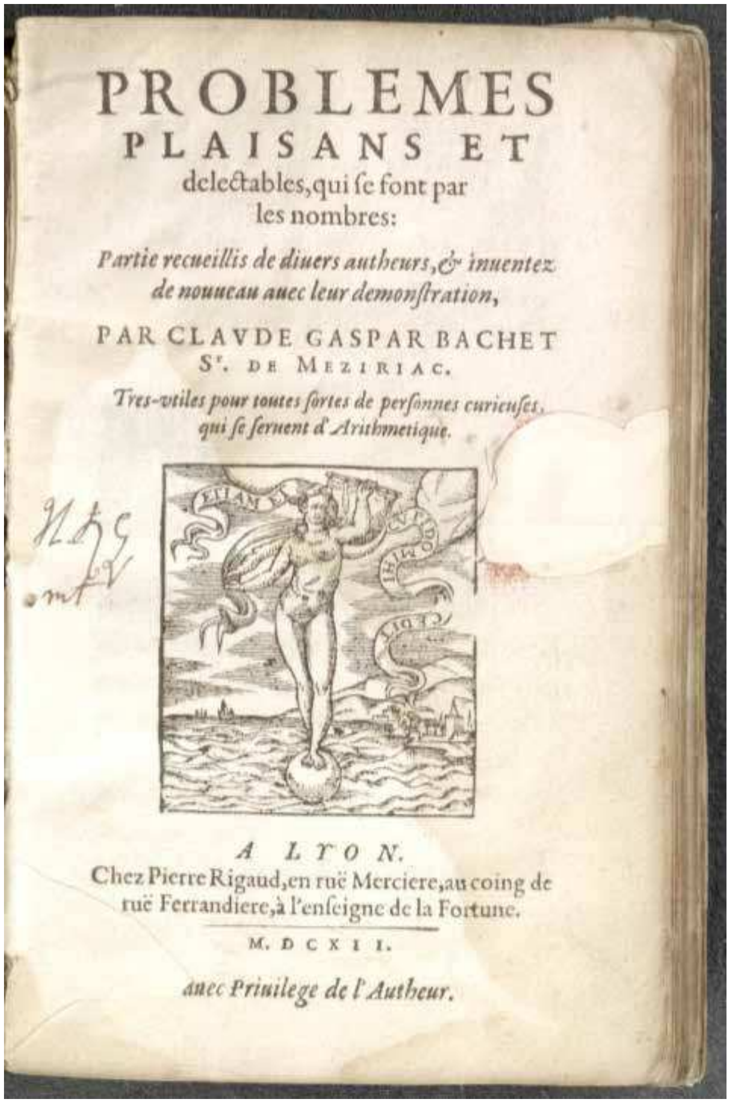


Fig. 2: Frontispiece of *Problemes plaisans et delectables, qui se font par les nombres* de Claude-Gaspard Bachet, first edition, 1612.

Source: Claude-Gaspard, Bachet (1612)

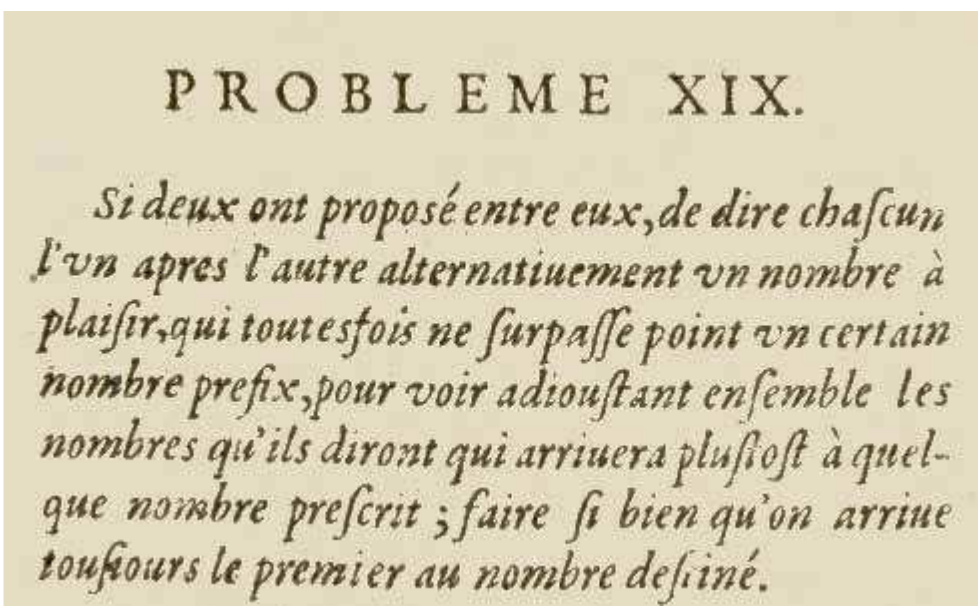
It was indeed the first book on the subject to be published, but the original idea of a collection of recreational mathematic problems was due to Pacioli.[49] Bachet was a mathematician and a French translator (Fig. 3), known for publishing the Greek text of Diophante's *Arithmétique* (3rd century) with a Latin translation added.

[49] Vanni, Bossi, *Magic and Card Tricks*… p. 124.

Fig. 3: Claude-Gaspard Bachet de Méziriac.

Source: Wikipedia,

http://commons.wikimedia.org/wiki/File:WP_Claude_Gaspard_Bachet_de_M%C3%A9ziriac.jpg



PROBLEME XIX.

Si deux ont proposé entre eux, de dire chaſcun l'vn apres l'autre alternatiuement vn nombre à plaiſir, qui toutesfois ne ſurpaſſe point vn certain nombre prefix, pour voir adiouſtant enſemble les nombres qu'ils diront qui arriuera pluſtoſt à quelque nombre preſcrit; faire ſi bien qu'on arriue touſiours le premier au nombre deſtiné.

Fig. 4: Problem XIX presented in Bachet's 1612 work

Source: Claude-Gaspard, Bachet, p. 99

Problem XIX (Fig. 4) in the first edition of Bachet's 1612 book about mathematic recreations proposes a version of Nim that is similar to the one formulated by Pacioli.[50] Bachet studies the case when two players must reach 100 by adding numbers ranging from 1 to 10, "or any smaller number [...] such as the player who will say the number that achieves 100 is recognized the winner."[51] Then he explains the strategy that ensures the win: "But to win unerringly, add 1 to the number that cannot be exceeded, here 10; you obtain 11 and then, always remove 11 of the number to be reached, 100; you will obtain these numbers 89, 78, 67, 56, 45, 34, 23, 12, 1."[52] These are the equivalents of Pacioli's safety steps, applied to a different configuration of numbers. It can be noticed that Bachet gives the solution, without proving it, of the given example; yet he makes a start on a generalisation when he advises the reader to add 1 to the number that must not be exceeded. Indeed, in any case, when the number to be reached is n and the added number not to be exceeded is k $(k \leq n)$, the steps to reach are the numbers m such as $m \equiv 1 \bmod(k)$, that is to say 1, $k+1$, $2k+1$, $3k+1$... Bachet rightly points out: "if the two players know the trick, the one who begins will inevitably win."[53] Then he suggests a demonstration to "formulate the general rule",[54] which is actually not that general, because the author's proof is limited to one example with 100 and 10 and Bachet only explains why it is clever to reach the steps 89, 78, etc. Backward induction, which is essential to this kind of games, is used here: Bachet starts at the end of the game, demonstrating that the opponent cannot reach 100 from 89 and goes back to the beginning of the game to determine the starting situation. Yet, he concludes: "the rule is infallible and perfectly demonstrated."[55] A notice meant to "bring diversity in the game practice"[56] proposes to play with other numbers – reach 120 without adding numbers higher than 10, or reach 100 with numbers not exceeding 8 or 9. Next, Bachet advises to choose an opponent who does not know anything about the strategy, and nevertheless to remain clever:

> So, if your opponent ignores the subtlety of the game, you must not take the same remarkable numbers necessary to win unfailingly, because doing so, you will highlight

[50] The fact that Pacioli was a major source of inspiration for Bachet has been studied in detail: about a third of the problems suggested by Bachet are directly linked to Pacioli. See Albrecht Heeffer, p. 18.

[51] Claude-Gaspard, Bachet, 1612. p.100: " ou tout nombre moindre. […] et que celui qui dira le nombre accomplissant 100, soit réputé pour vainqueur." My translation.

[52] Idem. "Or pour vaincre infailliblement, ajoute 1 au nombre qu'on ne peut passer, qu'est ici 10, tu auras 11, et ôte continuellement 11, du nombre destiné 100, tu auras ces nombres 89, 78, 67, 56, 45, 34, 23, 12, 1." My translation.

[53] Idem. "si les deux qui jouent à ce jeu savent tous deux la finesse infailliblement celui qui commence remporte la victoire." My translation.

[54] *Ibid*. p.101: "pour former la règle générale." My translation.

[55] Idem. "la règle est infaillible et parfaitement démontrée." My translation.

[56] Idem. "pour apporter de la diversité dans la pratique de ce jeu." My translation.

> the trick, and if he is a man of good intelligence, he will immediately notice these numbers, as he will see you are always choosing the same: but at the beginning, you can say other numbers on the fly, until you came nearer the wanted number, because then you will be able to add some of the necessary numbers for fear of being surprised.[57]

Better being safe than sorry…

This problem and its resolution appear in the second edition in 1624 (Fig. 5), under the same statement (Fig. 6); only the numbering is changing.

PROBLEMES PLAISANS ET DELECTABLES, QVI se font par les nombres.

Partie recueillis de diuers Autheurs, partie inuentez de nouueau auec leur demonstration.

Par CLAVDE GASPAR BACHET, Sieur de Meziriac.

Seconde Edition, reueuë, corrigée, & augmentée de plusieurs propositions, & de plusieurs Problemes, par le mesme Autheur.

Tres-vtile pour toutes sortes de personnes curieuses, qui se seruent de l'Arithmetique, & Mathematique.

A LYON,

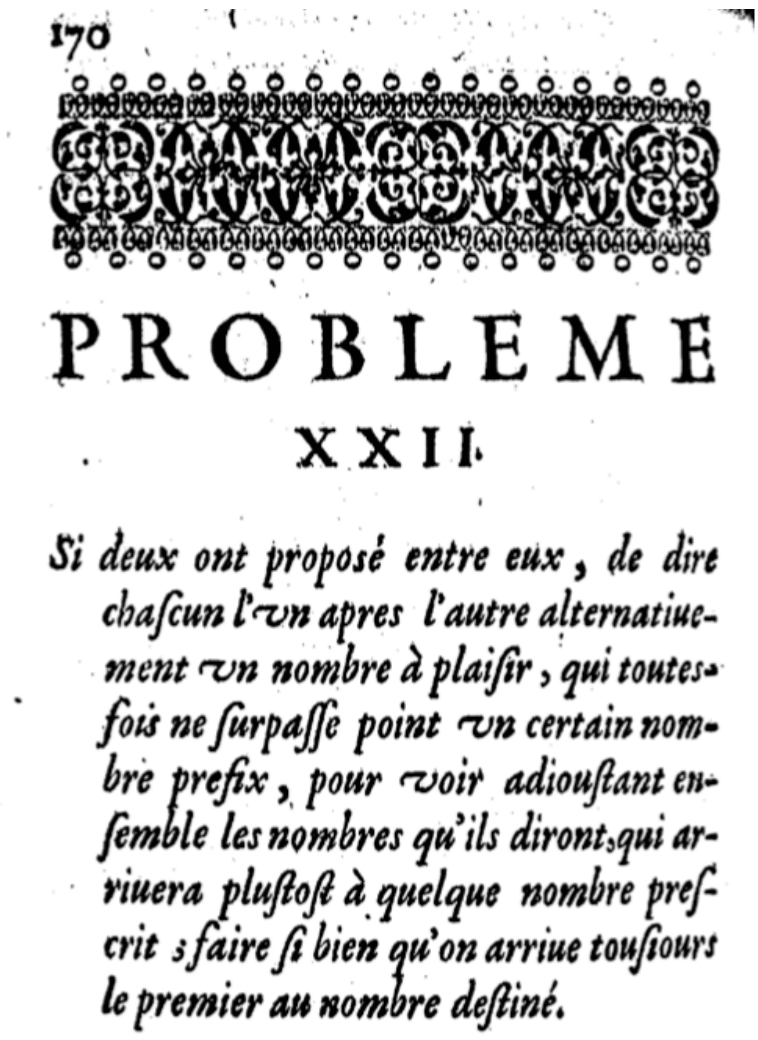
Chez PIERRE RIGAVD & ASSOCIEZ, ruë Merciere, au coing de ruë Ferrandiere, à l'Enseigne de la Fortune.

M. DC. XXIIII.

Fig. 5: Frontispiece of *Problemes plaisans et delectables, qui se font par les nombres* by Claude-Gaspard Bachet, second edition 1624.

Source: Claude-Gaspard, Bachet (1624)

[57] *Ibid*. p.102: "Partant, si ton adversaire ne sait pas la finesse du jeu, tu ne dois pas prendre toujours les nombres remarquables et nécessaires, pour gagner infailliblement, car faisant ainsi, tu découvriras trop l'artifice, et s'il est homme de bon esprit il remarquera tout incontinent ces nombres là, voyant que tu choisis toujours les mêmes : mais au commencement tu peux dire à la volée des autres nombres, jusqu'à ce que tu approches du nombre destiné, car alors tu pourras facilement accrocher quelque un des nombres nécessaires de peur d'être surpris." My translation.


170

PROBLEME

XXII.

Si deux ont proposé entre eux, de dire chaſcun l'vn apres l'autre alternatiuement vn nombre à plaiſir, qui toutesfois ne ſurpaſſe point vn certain nombre prefix, pour voir adiouſtant enſemble les nombres qu'ils diront, qui arriuera pluſtoſt à quelque nombre preſcrit; faire ſi bien qu'on arriue touſiours le premier au nombre deſtiné.

Fig. 6: Problem XXII as stated in Bachet's book 1624.

Source: Claude-Gaspard, Bachet, p. 170

In the same year, at the French university of Pont-à-Mousson, an octavo volume entitled *Récréation mathématique, composée de plusieurs problèmes plaisants et facétieux, En fait d'Arithmétique, Géométrie, Mécanique, Optique, et d'autres parties de ces belles sciences* was published. It was the first appearance of the words "mathematical recreations" in the title of a book.[58] The numerous revised and corrected editions that followed this publication do not make it possible to confidently state the paternity of this work, as the frontispiece does not mention any author. Opinions differ on this point and three names are brought forward: Henry

[58] Albrecht, Heeffer.

Van Etten, who signed the dedication, Jean Leurechon (1591-1670), a Jesuit whose name is used in almost every library to register the book, and Jean Appier Hanzelet (1596-1647), an engraver and printer at the university of Pont-à-Mousson, who published the book. Albrecht Heeffer led a survey to discover the true author of *Recréations mathématiques*. This survey also lists the different sources of the mathematical problems found in the volume. It would seem that the thirty-one arithmetical and combinatory problems, about one third of the collection, directly come from *Problèmes plaisans* by Bachet who is quoted in the preface and in some of the notes.[59] It is amusing to notice that the authors of mathematical recreation books often referred to Bachet,[60] simply because this latter forgot to give the origin of his recreations, whereas he collected them from Pacioli, Alcuin, Tartaglia, Cardan,[61] probably Chuquet,[62] and certainly from oral tradition too. Nevertheless, it must be admitted that as Bachet was the author of the most important seventeenth-century translation of Diophante's *Arithmetica,* it is therefore not surprising if one of his main concerns was related to arithmetical problems. Bachet hardly gave references, but a long tradition of recreational mathematics throughout the Middle-Ages and the Renaissance has made it possible to find some sources, among them Luca Pacioli.[63] Let us now consider another source of Nim, in additive version, which appeared in Germany shortly after the publishing of Bachet's work.

D. Daniel Schwenter (Germany, 1636)

One of the first German occurrences of the one-pile Nim can be found in *Deliciae Physico-Mathematicae* by Daniel Schwenter (1585-1636), a mathematician, inventor, poet and bookseller. Problem XLV is stated as follows: "you both must count to 30. The winner is the one who first reaches 30. But it is not allowed to add more than 6 at each turn."[64]

---

[59] Albrecht, Heeffer, p. 13.

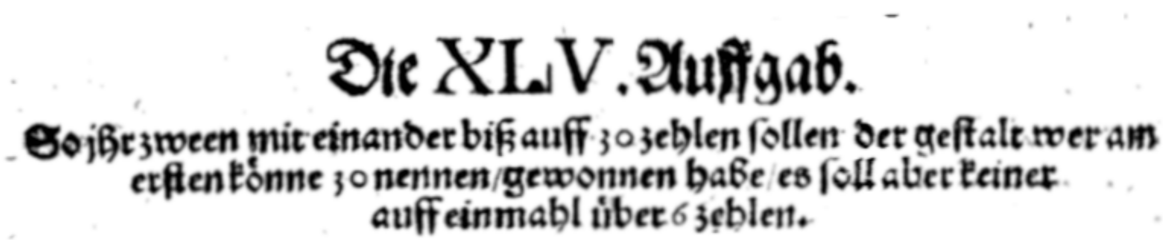
[60] For example, it is the case of Jacques Ozanam (1640-1718), a French mathematician, author of *Récréations mathématiques et physiques, Qui contiennent les Problèmes et les Questions les plus remarquables, et les plus propres à piquer la curiosité, tant des Mathématiques que de la Physique ; le tout traité d'une manière à la portée des Lecteurs qui ont seulement quelques connaissances légères de ces Sciences*, Paris, 1778. His work will be detailed later.

[61] Walter, Rouse Ball, rev. by Harold Coxeter, *Mathematical Recreations and Essays*, New York, The Macmillan Company, American Edition, 1947. p. 2.

[62] Albrecht, Heeffer, p. 17.

[63] *Ibid*. pp. 15-22.

[64] Daniel, Schwenter, *Deliciae Physico-Mathematicae*, Nuremberg, 1636. p. 78: "So ihr zwei sollt miteinander bis 30 zählen. Wer als erstes auf 30 kommt, hat gewonnen. Es darf aber keiner auf einmal über 6 zählen" My translation with help from a German friend…



Die XLV. Auffgab.
So jhr zween mit einander biß auff 30 zehlen sollen der gestalt wer am ersten könne 30 nennen/ gewonnen habe/ es soll aber keiner auff einmahl über 6 zehlen.

Fig. 7: Problem XLV presented in Schwenter's book, ed.1636.

Source: Daniel, Schwenter, p. 78

It is not clear whether Schwenter knew Pacioli's *De Viribus Quantitatis*, because he mentions Professor Gustavus Selenus[65] and his work about cryptography[66] in the very first sentence. According to Schwenter, Selenus explains that the winner will be the one who will choose numbers 9, 16 and 23. Unfortunately, because of the complexity of the Latin text, we have not found this passage in Selenus' book yet…

Schwenter is an interesting example because it shows how difficult it is to trace the links between the authors of books containing arithmetical problems. Former sources are sometimes quoted explicitly but it is not sure that these former sources did not take inspiration from other works that would not be mentioned.

E. Jacques Ozanam (France, 1694)

Jacques Ozanam (1640–1718) was a French mathematician more particularly known for his writings about trigonometric and logarithmic tables. The first edition of his *Récréations mathématiques et physiques*[67] dates back to 1694; many republications, added with revisions and additions, were to follow such as Jean-Etienne Montucla's edition in 1778.[68] These numerous republications make William Schaff say: "Ozanam may be regarded as the forerunner of modern books on mathematical recreations."[69] Nevertheless, Schaaf

[65] Gustave Selenus was the pseudo used by August II von Brunswick-Wolfenbüttel (1579-1666).

[66] Selenus wrote two books : one about Chess, *Das Schach – öder Königsspiel* (1616) and the second about cryptography, *Cryptomenytices et Cryptographiae libri IX* (1624).

[67] Jacques, Ozanam, *Récréations mathématiques et physiques, Qui contiennent les Problèmes et les Questions les plus remarquables, et les plus propres à piquer la curiosité, tant des Mathématiques que de la Physique ; le tout traité d'une manière à la portée des Lecteurs qui ont seulement quelques connaissances légères de ces Sciences*, Paris, 1778.

[68] Jean-Etienne Montucla (1725-1799) was a French mathematician, the author of *Histoire des Mathématiques* (1758).

[69] William, Schaaf, *Recreational Mathematics A Guide to the Literature*, The National Council of Teachers of Mathematics, INC., 1963. p. 1.

admits that Ozanam drew inspiration from Bachet's, Mydorge's and Leurechon's works and that "[…] his own contributions were somewhat less significant".[70]

We have based our study on the posthumous edition of *Récréations mathématiques et physiques*, dated 1778 and published by Claude-Antoine Jombert (173?-1788). In the first tome, *Contenant l'Arithmétique et la Géométrie*, the additive version of Bachet can be found under its simplest formulation:

> Problem XIV: two players agree to take in turns numbers smaller than a given number, for example 11, and to add them until one of the two persons can reach, for instance, 100; how should we proceed to be the first without fail?[71]

It is worth noticing that Ozanam openly asks the question: which is the strategy to apply in order to win, whereas his predecessors put down the problem without any questioning. Ozanam gives an explanation of the solution that is not so different from the one found in Bachet's book. He only completes the strategy to use if, instead of adding numbers ranging from 1 to 10, it is decided to choose numbers between 1 and 9. But he does not bring any changes in the statement of the problem nor in the explanation of the solution. This "inertia" in the evolution of the content in recreational books, numerous problems are actually similar in various books, is linked to the inertia in the solving methods of the given problems. Changes are generally to be found in the way problems are stated more than in the way they are solved. This "stasis" is also due to the fact that authors tended to copy the problems they had found in other sources:

> In certain instances, authors have been careful to state the immediate origin of their questions, but it must be confessed that this is done when one writer wishes to correct the work of another rather than when he merely wishes to acknowledge his use of the other's book.[72]

We will see later that copying a problem and including a slight variation without noticing that this variation completely changes the solution can sometimes prove to be tricky for the authors!

---

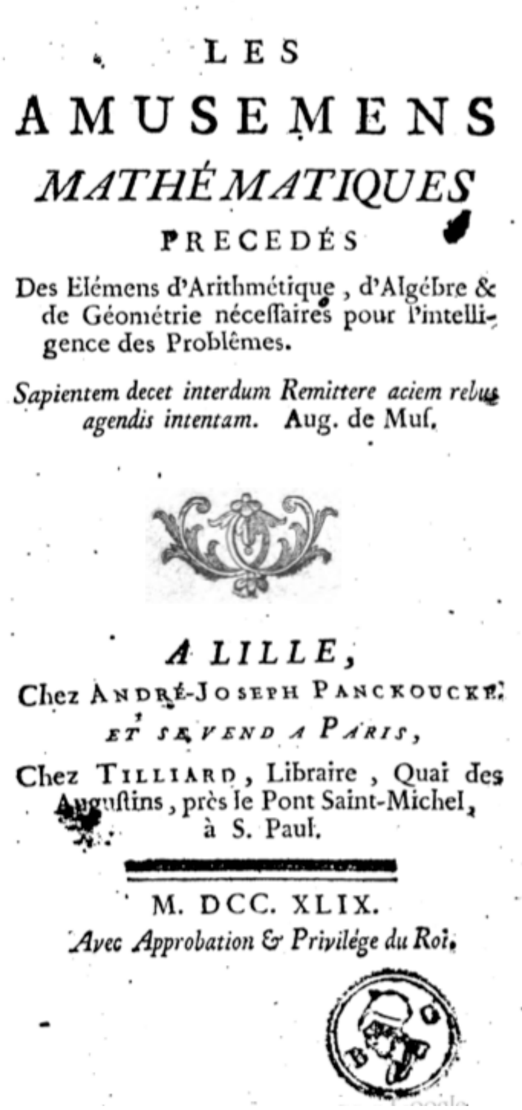
[70] Idem.
[71] Jacques, Ozanam, pp. 162-163. "PROBLÈME XIV : Deux personnes conviennent de prendre alternativement des nombres moindres qu'un nombre donné, par exemple 11, et de les ajouter ensemble jusqu'à ce l'un des deux puisse atteindre, par exemple, 100 ; comment doit-on faire pour y arriver infailliblement le premier ?" My translation.
[72] Vera, Sanford, *The History and Significance…* pp. 79-80.

F. André-Joseph Panckoucke (France, 1749)

In France, it is not before the end of the eighteenth century that an additive version of Nim was to be found in *Les amusemens mathématiques*[73] (Fig. 8) by André J. Panckoucke (1703-1753), a writer, bookseller and editor in Lille (1703-1753).[74] Panckoucke owned a bookshop situated Place Rihour[75] between 1728 and 1733 and he published a large number of works.[76] Among other books, he wrote the *Dictionnaire des proverbes françois et des façons de parler comiques, burlesques et familières...*(1748), *L'art de se désopiler la rate* (1754) and the *Amusemens mathématiques* (1749).[77] It seems that this editor from the North of France appreciated enjoyable pastime! But Panckoucke was also well read and an erudite who paid attention to scientific developments and to their practical applications.[78] His bookshop provided the intellectual elite of Lille with a large choice of books.

[73] André-Joseph, Panckoucke, *Les amusemens mathématiques, precedés Des Eléments d'Arithmétique, d'Algèbre & de Géométrie nécessaires pour l'intelligence des Problêmes*, Lille, 1749.
[74] His son, Charles-Joseph Panckoucke (1736 – 1798), was more famous than his father as he became the official editor and bookseller of the Imprimerie royale and of the Académie royale des sciences. He was a leading figure in the world of edition and diffusion of the encyclopaedic knowledge of the Enlightenment. He also corresponded with Voltaire and Rousseau.
[75] Place Rihour is a tourist place in Lille where the remains of a fifteenth-century castle can be seen. Palais Rihour was built by the Dukes of Burgundy of the Valois dynasty.
[76] See: André Joseph Panckoucke. (2013, November 2nd). *Wikipedia*. http://fr.wikipedia.org/w/index.php?title=Andr%C3%A9_Joseph_Panckoucke&oldid=97949812. [2.12.2013]
[77] For a better understanding, I have translated these three titles as follows:
- Dictionary of French Proverbs and of the Different Comical, Farcical and Colloquial Ways of Speaking.
- The Art of Killing Oneself Laughing.
- Mathematical Pastimes.

[78] Gilbert, Dalmasso, *Présence de la "chymie" dans la France du Nord, de la deuxième moitié du XVIIIe siècle au premier tiers du XIXe : sa diffusion et son enseignement public et privé, son application aux Arts*, PhD thesis submitted at the University of Lille 3 in 2005.



LES

AMUSEMENS

MATHÉMATIQUES

PRECEDÉS

Des Elémens d'Arithmétique, d'Algébre & de Géométrie néceſſaires pour l'intelligence des Problêmes.

Sapientem decet interdum Remittere aciem rebus agendis intentam. Aug. de Muſ.

A LILLE,

Chez André-Joseph Panckoucke.

ET SE VEND A PARIS,

Chez Tilliard, Libraire, Quai des Auguſtins, près le Pont Saint-Michel, à S. Paul.

M. DCC. XLIX.

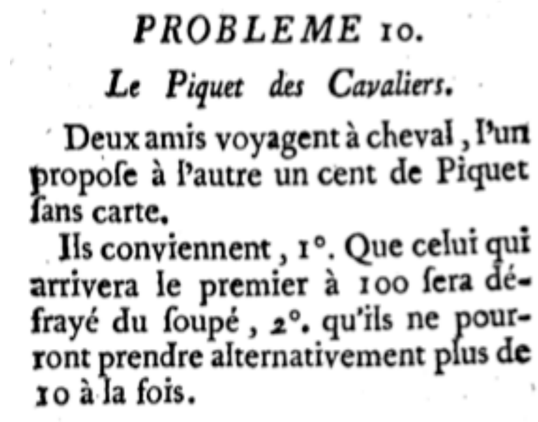
Avec Approbation & Privilége du Roi.

Fig. 8: Frontispiece of *Amusemens mathématiques* by André-Joseph Panckoucke, 1749.

Source: André-Joseph, Panckoucke (1749)

As the title indicates it, the book displays general results of arithmetic, algebra and geometry that are useful for resolving the 239 problems, which are proposed later with their solutions. Problem 10 is called "Le Piquet des Cavaliers"[79] and its wording is more fictionalized than in Pacioli's, Bachet's, Ozanam's or Schwenter's books. Below is a copy of this problem (Fig. 9):

[79] André-Joseph, Panckoucke, p.130.



PROBLEME 10.

Le Piquet des Cavaliers.

Deux amis voyagent à cheval, l'un propoſe à l'autre un cent de Piquet ſans carte.

Ils conviennent, 1°. Que celui qui arrivera le premier à 100 ſera défrayé du ſoupé, 2°. qu'ils ne pourront prendre alternativement plus de 10 à la fois.

Fig. 9: *Le Piquet des Cavaliers*, pastime proposed by Panckoucke.

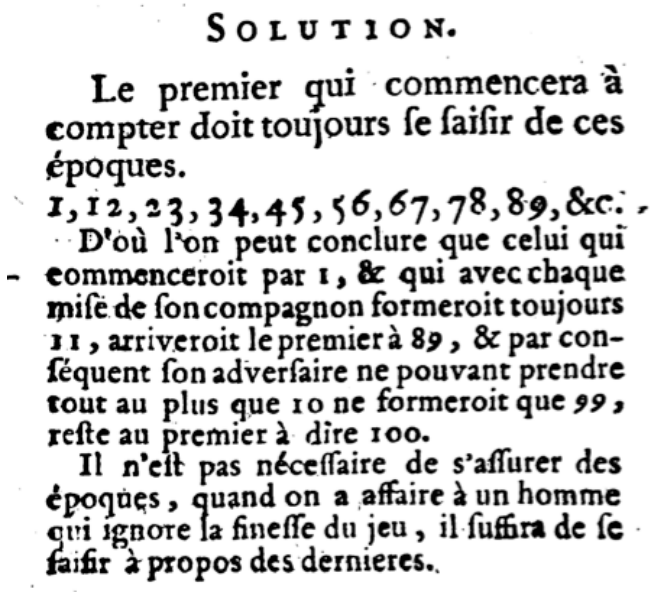
Source: André-Joseph, Panckoucke, p. 130[80]

Originally, *Piquet*[81] was a card game, which up to the middle of the nineteenth century, was one of the three games considered as the most dignified with Chess and Backgammon.[82] During the seventeenth century, it was played with 36 cards (the lowest being the 6). This game is described in the French comédie-ballet *Les Fâcheux* by Molière, performed in 1661. *Piquet* was also mentioned, spelled *Picquet*, in *Gargantua* by Rabelais (1534). The two players must take, in turns, a card from the pack and add its value to the sum already obtained with the former draws. The riders of our problem do not have any cards, which would not be very useful for riding, and they play orally, which is equivalent to one pile Nim. The given solution is very short; below is a copy (Fig. 10).

[80] My translation: "Two friends are riding; one of them suggests to play one "cent (one hundred) de Piquet" without cards.
Both agree that 1°. The first who will reach 100 will not pay the diner, 2°. They will not be allowed to take in turns a number higher than 10."

[81] Formerly, *Piquet* was called *Cent* (one hundred) because it was the number to reach in order to win a match.

[82] See: http://academiedesjeux.jeuxsoc.fr/piquet.htm



SOLUTION.

Le premier qui commencera à compter doit toujours ſe ſaiſir de ces époques.

1, 12, 23, 34, 45, 56, 67, 78, 89, &c.

D'où l'on peut conclure que celui qui commenceroit par 1, & qui avec chaque miſe de ſon compagnon formeroit toujours 11, arriveroit le premier à 89, & par conſéquent ſon adverſaire ne pouvant prendre tout au plus que 10 ne formeroit que 99, reſte au premier à dire 100.

Il n'eſt pas néceſſaire de s'aſſurer des époques, quand on a affaire à un homme qui ignore la fineſſe du jeu, il ſuffira de ſe ſaiſir à propos des dernieres.

Fig. 10: The solution of problem 10, *Le Piquet des Cavaliers* by Panckoucke.

Source: André-Joseph, Panckoucke, p. 130[83]

The steps to reach to ensure a win are called "époques", epochs, which is different from the words used by Bachet and Ozanam; this term was to be reused later by Guyot. The last sentence makes us think that Panckoucke read *Problèmes plaisants et délectables,* for he encourages the reader not to insist in reaching all the steps, "époques", but only those that are close to the number to be reached, just like Bachet did; a rather simple strategy, which is nevertheless efficient enough to eat out cheaply! Some forty years later, Henry Decremps (1746 – 1826) used this phrase: "*Principes mathématiques sur le piquet à cheval, ou l'art de gagner son dîner en se promenant*"[84] in his *Codicile de Jérôme Sharp.*[85] This assertion reinforces the idea that the player who knows the winning strategy can take advantage of his knowledge to obtain a favour from his fellow player. In this additive version of Nim, Panckoucke stresses the control you can exert over your opponent if you know the trick of the

[83] My translation:
"The first who will begin to count must always reach these steps: 1, 12, 23, 34, 45, 56, 67, 78, 89, etc…
Therefore, it can be concluded that the first who would begin with 1 and who would always make 11 when added with his friend's stake, would reach first 89; as his opponent cannot add more than 10, he could not reach but 99; consequently, it only remains to the first player to announce 100.
When playing with a man who is not aware of the finer points of the game, it is not necessary to ensure the first steps; it will be sufficient to ensure the last ones."

[84] For a better understanding, I give my translation of the title: "Mathematical Principles about the "Piquet à cheval", or the Art of Earning One's Diner while Riding".

[85] David, Singmaster, *Sources in Recreational Mathematics An Annotated Bibliography*, 9th preliminary edition, March 2004, available for consultation on: www.gotham-corp.com/sources.htm

game. It is no longer a matter of simply displaying an intellectual superiority but also a question of using it for material purposes. This reference highlights several new aspects of the game: first, the wording of the problem is formulated in a more fictionalized way, even if the given solution is the same as far as the solving method remains arithmetically identical; and secondly, the aim of the game and the interest of knowing the winning strategy are clearly emphasized.

G. Edmé-Gilles Guyot (France, 1769)

Edmé-Gilles Guyot (1706 – 1786) was a French physician and inventor, an author in the area of mathematics and physics, which he used to perform magic tricks such as optical illusions, projection of figures into smoke. Guyot worked on the development of magic lanterns used in phantasmagoria in order to show his experiments before a live audience and to popularize his discoveries. During the eighteenth century, this was indeed a common practice for teaching and disseminating sciences in France: "[...] mathematical exercises in public, which multiplied at that time with educational purposes [...] were meant to stress on applied mathematics, which were easier to understand by the common people who attended the meetings [...]."[86] Guyot's works were translated into English and German and were largely circulated in Europe.[87] In 1769, Guyot tackled the French edition of *Récréations Mathématiques*, which had been reissued more than twenty-five times between 1629 and 1680 by Claude Mydorge, Jacques Ozanam and Jean-Etienne Montucla who published the work in four volumes. Guyot titled the second volume *Nouvelles récréations physiques et mathématiques, Contenant, Toutes celles qui ont été découvertes et imaginées dans ces derniers temps, sur l'Aimant, les Nombres, l'Optique, la Chymie, etc. et quantité d'autres qui n'ont jamais été rendues publiques. Où l'on a joint leurs causes, leurs effets, la manière de les construire, et l'amusement qu'on peut en tirer pour étonner agréablement*.[88] This volume is devoted to recreations with numbers. Panckoucke's version of *Piquet à cheval*[89] can be found,

[86] René, Taton, *La science moderne*…, p.55: "[…] les exercices publics sur les mathématiques, qui se multiplient à l'époque avec des intentions pédagogiques […], sont portés à mettre fort l'accent sur les mathématiques appliquées, plus accessibles aux honnêtes gens qui y assistaient […] " My translation.
[87] See the English site of Wikipedia: http://en.wikipedia.org/wiki/Edm%C3%A9-Gilles_Guyot
[88] My translation: *New Physical and Mathematical Recreations including All of Those that have been Discovered or Created in Recent Times; Upon Magnet, Numbers, Optics, Chemistry, etc…and Many Other Things Never Made Public. To Which are Attached their Causes, their Effects, the Way to Construct Them and the Entertainment that can be Drawn in Order to Surprise Pleasantly*.
[89] Edmé-Gille, Guyot, *Nouvelles récréations physiques et mathématiques, Contenant, Toutes celles qui ont été découvertes et imaginées dans ces derniers temps, sur l'Aimant, les Nombres, l'Optique, la Chymie, etc. et*

in which "two riders who travel together, are bored when thinking of the distance that is still to be covered; they create a game that could help them to pass the time more pleasantly and agree to play a "Cent de Piquet" [...]"[90] What is at stake here is no longer the winning of a diner; it is simply a verbal recreation that helps passing the time. We can point out the use of the word "recreation" instead of "problem", which was used until then when referring to the object of the statement. The solution Guyot gives is similar to those given by the authors we studied previously; yet, the way the author presents it is rather different. Indeed, Guyot states:

> In order that the player who gives the first number can reach 100, and that his opponent cannot, he must remember the numbers 11, 22, 33 etc… of the problem mentioned above, and count such as there is always one unit more than these numbers; furthermore, he should give first the number 1, and because his opponent cannot take a number higher than 10, this latter will not be able to reach 12 that the first player will take, and then consequently, the numbers or "époques" (steps) 23, 34, 45, 56, 67, 78 and 89; reaching this last step, his opponent cannot prevent him from reaching 100 at the following turn, whatever number he himself could choose.[91]

Guyot does not clearly explain the backward induction necessary to find the "époques" and suggests right away that the first player should choose 1 so that his opponent could not reach the step 12. The solution is quite close to Panckoucke's in its wording as well as in its terminology, using "époque". Guyot also advises the player, "if his opponent does not know the trick and in order to better disguise this Recreation, to give indistinctly any numbers in the first turns, as far as around the end of the game, he takes the two or three last numbers necessary to win."[92] He adds that this recreation has no interest and is not pleasant if played by two people aware of the tricks, as far as "the first who gives the first number always wins."[93] Once again, it is better to play with someone who does not know the strategy. Guyot

---

quantité d'autres qui n'ont jamais été rendues publiques. Où l'on a joint leurs causes, leurs effets, la manière de les construire, et l'amusement qu'on peut en tirer pour étonner agréablement, Paris, 1769. pp. 27-29.

[90] *Ibid*. p. 27: "deux cavaliers qui voyagent ensemble, ennuyés du chemin qu'il leur reste encore à faire, inventent un jeu qui puisse leur faire passer le temps plus agréablement, et conviennent ensemble de jouer un Cent de Piquet." My translation.

[91] *Ibid*. pp. 27-28: "Afin que le premier qui nomme le nombre puisse arriver à 100, et que son adversaire n'y puisse y parvenir, il doit se souvenir des nombres 11, 22, 33, etc. du problème ci-dessus, et compter de façon qu'il se trouve toujours d'une unité au-dessus de ces nombres ; ayant en outre attention de nommer d'abord 1, attendu que son adversaire ne pouvant prendre un nombre plus grand que 10, ne pourra arriver au nombre 12, qu'il prendra alors lui-même et conséquemment ensuite les nombres ou époques 23, 34, 45, 56, 67, 78, et 89, à laquelle étant arrivé, quelques nombres que puisse choisir son adversaire, il ne peut l'empêcher de parvenir, le coup suivant, à 100." My translation.

[92] *Ibid*. p. 28: " ne connaît l'artifice de ce coup peut (pour mieux déguiser cette Récréation) prendre indistinctement toutes sortes de nombres dans les premiers coups, pourvu que vers la fin de Partie, il s'empare des deux ou trois derniers nombres qu'il faut avoir pour gagner." My translation.

[93] Idem. "attendu que celui qui nomme le premier a toujours gagné." My translation.

differentiates himself from his predecessors by insisting on the fact that the game is completely uninteresting if the two players know the solution. Furthermore, the aim of the game stated by Guyot differs from Panckoucke's; it is no longer a matter of using one's knowledge to achieve one's purpose. The author concludes with this interesting point:

> It (the recreation) can be played with any other numbers; and then, if the first wants to win, the number to be reached must not be equal to the one he can stake, because he could loose; but it is necessary to divide the higher number by the lower, and the remainder will be the number that the first player must say first to make sure he will win.[94]

Thus, the recreation is generalized for any numbers; Guyot gives an example that helps us to understand what he is trying to get at:

> If the number that we agree to reach is 30, and if the number we are allowed to say must be lower than 7, as 4 times 7 make 28, it remains 2 to reach 30 and this number is the one that the first player must pronounce in the first place; and so, whatever number his opponent will say, he must choose the number that will make 7 when added to the other, and he will necessary reach first the number 30.[95]

Guyot uses the division of n by k+1 and more particularly the remainder of this division, which sets the first number to say to ensure the win. This notion of division does not appear in Bachet's, Ozanam's and Panckoucke's works. Besides, Guyot gives an example, taking $n = 30$ and $k + 1 = 7$ to illustrate this generalization. This leads us to think that he got acquainted not only with Panckoucke's *Amusemens mathématiques* as far as the wording, the fictionalization and the terminology are concerned, but also with Pacioli's *De Viribus Quantitatis* or Schwenter's *Deliciae Physico-Mathematicae*; indeed, these two authors were the only ones who used the numbers 30 and 6 before 1769 in this additive version of one-pile Nim.

[94] Idem. "Elle [la récréation] peut se faire aussi avec tous autres nombres ; et alors si le premier veut gagner, il ne faut pas que le nombre où l'on doit arriver, mesure exactement celui jusqu'où on peut atteindre pour gagner, car alors on pourrait perdre ; mais il faut diviser le plus grand par le plus petit, et le reste de la division sera le nombre que le premier doit nommer d'abord pour être assurer du gain de la Partie." My translation.

[95] *Ibid.* pp. 28-29: "Si le nombre auquel on se propose d'atteindre est 30, et le nombre au-dessous duquel on doit nommer 7, 4 fois 7 faisant 28, il reste 2 pour aller à 30, et ce nombre est celui que le premier doit nommer d'abord ; alors quelque nombre que nomme l'adversaire, s'il y ajoute celui qui convient pour former avec lui celui de 7, il parviendra de nécessité le premier au nombre 30." My translation.

H. William Hooper (England, 1774)

The additive version of Nim crossed the Channel and landed on William's Hooper book entitled *Rational Recreations* in which "[…] the author has selected the principal part of the experiments from the writers on recreative philosophy of the last and present centuries".[96] The first edition of Hooper's book dates back to 1774 and the second, which we have worked on, dates back to 1783. Four volumes make up this work; the first mainly deals with "arithmetical and mechanical experiments"[97] and the others are related to optics, chromatic and acoustics (vol.2), pneumatic, hydrology and pyrotechnics (vol.4) and finally electrical and magnetical experiments in volume 3. The presentation of the eighth recreation of the first volume, in which we can find the additive version of Nim, is slightly different from the "problems", "recreation" or "effect" that were present in the former works. First, the recreation is entitled *The Magical Century*[98] and no longer *Cent de Piquet*. Then Hooper starts with an arithmetical reminder concerning the multiplication of the first nine digits by 11: "If the number 11 be multiplied by any one of the nine digits, the two figures of the product will always be familiar."[99] He illustrates this reminder with the following eleven-time table (Fig. 11):

11	11	11	11	11	11	11	11	11
1	2	3	4	5	6	7	8	9
11	22	33	44	55	66	77	88	99

Fig. 11: The multiplication table of the nine first non-zero natural numbers by 11.

Source: William, Hooper, p. 31

Next, Hooper chooses to add in turns counters piled up on a table, until he obtains 100, yet without adding more than 10 counters at the same time. It is worth noticing that Hooper opts for a more visual layout of the recreation, choosing the possibility to manipulate counters

[96] William, Hooper, *Rational Recreations. Volume the first. Containing arithmetical and mechanical experiments*, 2nd edition, London, 1783. p. i.
[97] *Ibid*. Frontispiece.
[98] *Ibid*. p. 31.
[99] Idem.

instead of abstract values that must be kept in mind. This recreation is really closer to a game including the handling of objects than an exercise of mental arithmetic. We can also point out a more direct link with the Nim of Bouton who considered piles of various-sized counters, which had to be manipulated too. Hooper states that, before starting to play, you must modestly (!) tell your opponent, "moreover, that if you stake first he shall never make the even century, but you will."[100] In order to achieve this, you must start with staking only one counter and then:

> [...] Remembering the order of the above series, 11, 22, 33, etc. you constantly add, to what he stakes, as many as will make one more than the numbers of that series, that is, as will make 12, 23, 34, etc. till you come 89, after which the other party cannot make the century himself, or prevent you from making it.[101]

In this last statement, we can easily understand the deadlock in which the opponent is, as he cannot win, nor make us lose. These properties can be directly compared to the ones of Bouton's safe combinations. Just like his predecessors did, Hooper suggests that "if the other party has no knowledge of numbers [...]",[102] you should choose any number in the first turns and then secure your win around the last steps such as 56, 67, 78 and 89. He specifies that "this Recreation may be performed with other numbers […]"[103] and that, in order to win, "[…] you must divide the number to be attained, by a number that has one digit more than you can stake each time, and the remainder will be the number you must first stake."[104] Solving the problem by using the division reinforces the idea that Hooper had knowledge of Guyot's *Nouvelles récréations physiques et mathématiques*. Yet, Hooper adds that in order to win, there must always be a remainder, which is true if we confine ourselves to playing the first stake and with an opponent who has the *knowledge of numbers*!

I. John Badcock (England, 1820)

The way John Badcock presents *A curious Recreation with a Hundred Numbers, usually called the Magical Century*[105] is exactly the same than Hooper's, at least in the first

[100] Idem.
[101] *Ibid.* pp. 31-32.
[102] *Ibid*. p. 32.
[103] Idem.
[104] *Ibid*. p. 31.
[105] John, Badcock, *Philosophical Recreations, or Winter Amusements: A Collection of Entertaining and Surprising Experiments in Mechanics, Arithmetic, Optics, Hydrostatics, Hydraulics, Pneumatics, Electricity,*

lines... Badcock echoes the explanation using the multiplication table of the nine first non-zero natural numbers by 11; he also assumes that players handle counters and have to reach 100 without adding more than 10 at each turn. Yet, at the start, each player has 50 counters. As a solution, Badcock does nothing else but copying Hooper's one, almost word for word, without noticing that the apparently light variation he introduced considerably changes the game and its solving! Indeed, if we start the game playing 1 counter, and if our opponent adds only one counter in each of his turns, we will have to use 10 counters in each turn to reach the safety steps (12, 23, 34, 45 etc.). Within five turns, we will already have taken 41 counters from our stock, compared with 5 used by our opponent, which makes a sum amounting to 46. Therefore, it is absolutely impossible for the first player to reach 100 first if he plays the way described above. Consequently, the game Badcock suggests is totally different from Hooper's *Magic Century,* because even if we agree that the player who has no counter left loses the game, the goal to reach at all costs is no longer the "époques" or the safety steps; we must also make sure that we have enough counters to keep on playing. This fact makes the solution considerably harder. Sometimes, it may be wiser to simply copy one's predecessors' works instead of introducing variations that are not mastered…

This change in the statement proves to be an interesting variation because it changes the solution. Unfortunately, the author did not take the chance to provide a brilliant contribution.

J. John Jackson (England, 1821)

The additive version of Nim can be found once again in an English work, dated 1821 and entitled *Rational amusement for winter evenings ; or, A collection of above 200 curious and interesting puzzles and paradoxes relating to arithmetic, geometry, geography, etc.*[106] John Jackson was a "Private Teacher of the Mathematics";[107] the preface precises that the author came across arithmetical and geometrical puzzles and that he regarded as relevant the idea to compile the most interesting riddles with their solutions into a small volume.[108] The

Chemistry, Magnetism, & Pyrotechny, Or Art of Making Fire Works, Together with the Wonders of the Air Pump, Magic Lanthorn, Camera Obscura, &c. &c. &. and A Variety of Tricks with Cards, London, Thomas Hughes ed., 1820. p. 33.

[106] John, Jackson, *Rational amusement for winter evenings ; or, A collection of above 200 curious and interesting puzzles and paradoxes relating to arithmetic, geometry, geography, etc. with their solutions, and four plates*, London, 1821.

[107] *Ibid*. Frontispiece.

[108] *Ibid.* p. i.

problem we are interested in is the forty-seventh of the chapter devoted to *arithmetical puzzles*[109] and is stated quite briefly:

> Two persons agree to take, alternately, numbers less than a given number; suppose less than 11, and add them together till one of them has reached a certain sum; suppose 100. By what means can one of them infallibly attain to that number before the other?[110]

The solution Jackson gives is as succinct as the statement; it suggests to choose the numbers 1, 12, 23, 34 etc. "[…] a series in Arithmetical Progression, the first term of which is 1, the common difference 11, and the last term 100",[111] in order to reach 89 and hence to win unfailingly. What is new here is the use of the words "Arithmetical Progression", written in capital letters in the book; this expression had never been used before. The idea of "progression", whether geometrical or arithmetical, goes back to the most ancient mathematical recreations: "This fact may be due to an innate fondness for rhythm and repetition, a trait that seems to be universal, or it may be due to the mystery of a series of numbers whose values increase so rapidly,"[112] [concerning the geometrical progressions].[113] Vera Sanford gives examples of riddles that appeal to geometrical and arithmetical progressions in Rhind Mathematical Papyrus, in Fibonacci, Tartaglia or Cardan.[114] Yet, none of these authors realized that these progressions, which they did not name so, were to be more than simple mathematical curiosities, "the reason may be found in the lack of symbolism and lack of scientific knowledge which operated previously […]".[115] They talked about problems in a purely arithmetical way, because they had no other mathematical tools, such as logarithms in some cases, to solve them. It was no longer the case in the nineteenth century and John Jackson, as a mathematics teacher, had undoubtedly knowledge of arithmetical and geometrical series, as well as their properties.

At this point, we stop our study on the ancestors of Nim related to the one-pile additive version. From the beginning of the nineteenth century, this version could be found in

[109] *Ibid*. p. 1.
[110] *Ibid*. p. 11.
[111] *Ibid*. p. 64.
[112] Vera, Sanford, *The History and Significance…* p. 55.
[113] My clarification.
[114] *Ibid*. pp. 55-57.
[115] *Ibid*. pp. 57-58.

numerous works of mathematical riddles and recreations in France, England and Germany,[116] sometimes presented as "Piquet sans cartes" or simply as a two-player game. Some authors set out the game under a subtractive version,[117] and a version "misère" appeared, in which the one who is the last to play is the loser. The range of works we have analysed represents the very first sources in which the ancestral version of Nim can be found and we have seen how difficult, sometimes impossible it is to create strong links between the different authors. We have also understood that it is rather simple to win when playing these additive versions, as far as we know addition and multiplication tables.

On the other hand, things become more laborious when it comes to the Bouton's Nim of 1901: it is still possible to do some mental arithmetic, but a more important mathematical knowledge, abstract thinking and a longer time to reflect are required to reach the win. This is also the case for combinatorial games that came after Nim, such as Wythoff's Nim or Moore's Nim_k, the solutions of which appeal to even deeper mathematics knowledge. The initial aim of the first Nim games that were displayed as mathematical recreations has disappeared through the ages; it is no longer a matter of creating puzzles in order to impress the fairer sex during high-society evenings, but a matter of discovering interesting mathematical properties, even discovering new ones that could lead to theories still undeveloped.

The next part of our study will be devoted to Tiouk-Tiouk, an African game, which could seem far from Bouton's Nim as it requires a board and offers the possibility to block a piece, but yet is similar to Bouton's Nim when we consider its solving. Additionally, Tiouk-Tiouk might be an ancestor of Nim game, but once again tracing its sources and its first appearances has proved to be difficult, especially because boards have been made of sand…

III. Tiouk-Tiouk in Western Africa

In 1955, Charles Béart, a school principal in tropical Africa published a work in two volumes, in which he made an inventory of games and toys in Western Africa.[118] One chapter is devoted to two-player combinatorial games without chance, such as board games (Chess,

[116] David, Singmaster, *Sources in Recreational Mathematics An Annotated Bibliography*, www.gotham-corp.com/sources.htm
[117] We start with number *n* and we can take off at the most *k* in each turn.
[118] Charles, Béart, *Jeux et jouets de l'Ouest africain*, Mémoires de l'Institut Français d'Afrique Noire n°42, IFAN DAKAR, Tome I et II, 1955.

Draughts, Tic-Tac-Toe), twelve-box games (Awélé) and other versions. According to the author,

> In Africa, there exist some games with very complicated grid patterns that were for a long time the privilege of only some upper classes, who kept secret the traditional methods that allow to defeat the opponent as early as the first moves as long as he does not know the ancient traditional methods of defence and the means to counterattack.[119]

This is the case for Tiouk-Tiouk, for which an optimal strategy exists as for the Bouton's Nim. A single page out of the 850 of Béart's work displays Tiouk-Tiouk; yet it stands out from other grid pattern games, since the aim is not to take the opponent's counters but to block them. On a grid consisting of 6, 8, 10 or 12 rows – the number of rows must be even – the first row filled with seeds is allocated to one player and the last row filled with sticks is attributed to the other one (Fig. 12).

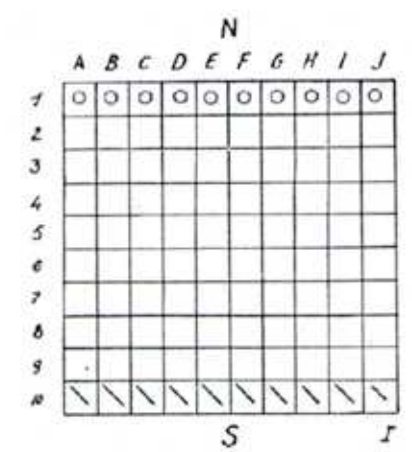


Fig. 12: Tiouk-Tiouk initial position.

Source: Charles, Béart, Tome II, p. 470

The two players alternate turns, and "each counter can be moved forwards or backwards as often as wanted, and as many squares as wanted too, but cannot jump over opponent piece.

[119] *Ibid.* Tome I, p. 53: "il existe en Afrique des formes de jeux à quadrillages très difficiles, qui furent longtemps permises seulement à certaines classes privilégiées et pour lesquelles les familles conservent, secrètes, des méthodes traditionnelles permettant d'écraser l'adversaire dès les premiers coups s'il ne possède pas lui même les vieilles méthodes traditionnelles de défense, et des moyens de reprendre l'offensive." My translation.

The player who succeeds in blocking all the counters of his partner will win."[120] If we compare this version with Bouton's Nim, we have in this case 6, 8, 10 or 12 piles and each contains the same number of objects; the gap between the seeds and the sticks is the same for each row. Transcribed into binary system, the starting position is a safe combination and the winner is the one who plays in second turn.[121] We do find this notion of safety step in the explanation given by Béart who names it "balance of distances":[122]

> At the start, the opposite counters are equidistant. The one who will play first and who, therefore, will break this balance, will loose. The second player will only have to restore the balance of distances with moving his counter forwards in the second column in such a way that intervals are equal in the two columns. He will keep this strategy until the end of the game, with always balancing the smallest interval proposed by the first player who will be finally stuck and who will loose.[123]

For example, in Fig. 13, the player S who has sticks, only needs to move his piece in A3 to rebalance the distances. On the other hand, in the configuration seen on picture Fig. 14, the player who must play will loose because he will inevitably break the balance of distances.

[120] *Ibid.* Tome II, p.471: "chaque pion peut se déplacer en avant ou en arrière à volonté, et d'autant de cases qu'il lui plaît, mais ne peut pas sauter par-dessus le pion du partenaire. A gagné qui arrive à bloquer tous les pions du partenaire." My translation.
[121] Providing that both players know the optimal strategy.
[122] *Ibid.* Tome II, p.471: "équilibre de distances." My translation.
[123] Idem. "Au départ, les pièces opposées sont à égale distance. Celui qui joue le premier et qui, par conséquent rompt cet équilibre, perdra. Il suffira au second de rétablir l'égalité des distances en avançant sa pièce dans la deuxième colonne de telle sorte que les intervalles soient les mêmes dans les deux colonnes. Le second joueur continuera cette tactique jusqu'à la fin de la partie en égalisant chaque fois sur le plus petit intervalle proposé par le premier joueur, qui, finalement, sera bloqué et perdra." My translation.

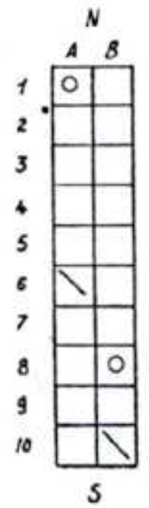


Fig. 13: Forward position of a Tiouk-Tiouk game: the player S has to move forward his stick in A3 to rebalance the distances.

Source: Charles, Béart, Tome II, p. 470

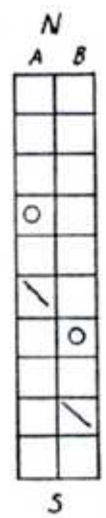


Fig. 14: Forward position of a Tiouk-Tiouk game: the player whose turn it is will loose for he will inevitably disturb the balance of distances.

Source: Charles, Béart, Tome II, p. 470

Béart specifies that Tiouk-Tiouk is generally proposed to a shepherd by a griot[124] who "generously offers him to draw the board and take the first-move advantage."[125] But we know now how disadvantageous it is to start the game! Yet, the griot does not cheat because "it is

[124] A griot is a traditional storyteller in West Africa. His origins go back to a time when writing did not exist. The griot is the keeper of oral tradition. The status of griot is passed within a cast. Griot families are specializes in divers areas: history, genealogy, storytelling and music.

[125] *Ibid*. p. 470: "généreusement il lui accorde le trait" My translation.

unnecessary, he is sure to win, when he wants if he does not start, and quite sure to win if he starts."[126] However there is a lot at stake with the outcome of this game.

Tiouk-Tiouk was listed in 1955 but it is impossible to date this game precisely and even to find its geographical origin. Béart classifies it within the grid games that have a peculiar status: "these are serious game par excellence, adult games; they are or were mainly the privilege of adults, of men, even of leaders; women and children could only imitate these games [...]."[127] This elitist aspect of Tiouk-Tiouk and other grid pattern games could be compared to the sixteenth-seventeenth-century mathematical recreations that were listed in works intended for educated classes who could afford to buy these books. "The most simple explanation for this singular situation lies in the fact that these games were introduced by a dominant society setting among a subjugated population."[128] Indeed, we can understand the intellectual pressure one can exert on somebody else who would not know the strategy and who would loose each game.

In 1988, Harry Eiss, author of *Dictionary of Mathematical Games, Puzzles and Amusements,*[129] listed another African version, and an Asian one, of Nim game and noticed that "whatever its origin, the game seems to have a universal appeal."[130] He added:

> A form of it known as Pebbles or Odds has been played in Africa and Asia for centuries. In this version, an odd number of pebbles, seeds, or whatever is placed in a pile, and players take turns selecting one, two, or three until all have been drawn. The player with an odd number in possession wins.[131]

Unfortunately, Eiss gives no other references and it is therefore difficult to trace a game that only requires pebbles or seeds and that does not leave lasting prints.

Games and mathematical riddles have been puzzling people for centuries, "[…] the human nature has changed but little, and problems that whet the imagination prove more fascinating than do the prosaic ones, whether a person lives in the sixteenth century or in the

---

[126] *Ibid.* pp. 470-471: "ce n'est pas nécessaire, il est sûr de gagner, quand il voudra, s'il n'a pas le trait, et à peu près sûr de gagner s'il l'a." My translation.

[127] *Ibid.* p. 451: "Ce sont là par excellence des jeux sérieux, des jeux d'adultes, et ils sont, ou ils furent, assez réservés aux adultes, et aux hommes, souvent même aux chefs ; les femmes et les enfants ne pouvaient qu'imiter ces jeux […]." My translation.

[128] Idem. : "L'explication de cette position singulière, la plus simple, est que ces jeux ont été introduits par une société dominante s'installant au sein d'une population subjuguée." My translation.

[129] Harry, Eiss**,** *Dictionary of Mathematical Games, Puzzles and Amusements*, Westport, Greenwood Press, 1988.

[130] *Ibid.* p. 188.

[131] Idem.

twentieth."[132] By the way, mathematical recreations were far from being outfashioned at the beginning of the twentieth century. Famous puzzlists Samuel Loyd's or Henry Dudeney's contributions are among the most popular; human infatuation for mysterious problems is timeless. This interest for puzzles and riddles is also marked by the growing complexity of the solutions, for example Piet Hein's *Superellipse*, Solomon Golomb's *Polyominoes*, Penrose's *Tilings* or Conway's *Surreal Numbers*. This complexity has enabled some mathematical theories to develop. Recreations have been considered as challenges to be taken up, yet within an entertaining frame. This tendency can be found as early as the first versions of Nim, of which solutions are rather easy to find out, provided that we take some time to do so. On the other hand, some no-trivial solutions were to appear later with Bouton's Nim and its variations. From that time on, the real mathematical history of Nim and its theorization have begun, when a sufficient keenness, even a genuine mathematical skill have been necessary to discover winning strategies. But this is another story…

[132] Vera, Sanford, *The History and Significance…* p. 62.

E1leven ... a Game.

by Carolyn “Cary” Staples | University of Tennessee .. School of Art

Growing up with cuisenaire rods, I understood math physically and visually as opposed to conceptually. I understood organization, sorting, collecting and comparing. I never made the transition to understanding how math abstractly describes the world. I would be stumped when I did not apply the rules correctly.

I was told I was wrong.

I was limited by what I could see and make.

In exploring the number “eleven” for this conference, I found the word “even” in the number.

But eleven is not even, but it is 1 and 1, it is symmetrical, it is a palindrome. 11 is a reflection of itself. It relates to a number of things that are associated with even. So is it possible to create an experience to allow users to explore the nature of eleven and not be wrong?

Each player has six dice. One for each letter of the word eleven. The “L” has been substituted with the numeral “1”. Players roll the cubes until they can generate a version of eleven; it could be all of the letters of the word or two cubes with the numeral “1”. Which option will you choose?

COLOR ADDITION ACROSS THE SPECTRUM OF MATHEMATICS

RON TAYLOR

Abstract. In this paper we introduce two sequential games whose rules are mathematical in nature, though no explicit mathematics is necessary during game play. Both games are based on color mixing rules which can admit a variety of mathematical interpretations. We discuss several of these realizations with an eye toward the novelty of the interpretation and from the perspective of using game play as a pedagogical strategy.

1. Introduction

> "*A feeling of adventure is an element of games. We compete against the uncertainty of fate, and experience how we grab hold of it through our own efforts.*" – Alex Randolph, game author

There are many games which can be analyzed using mathematics. Some of these involve some notion of chance like poker or games played with dice. Games which do not involve a notion of chance include checkers, tic-tac-toe and the ancient Chinese game of Go. While all of these games can be studied using mathematics, no explicit mathematics is necessary during game play. The commonalities between all such games, as with other parlour games, are the following:

(1) the *game* starts at a given state
(2) players play in succession according to a prescribed order
(3) each player has a choice of moves during his or her turn.

In between moves, or as a part of each move, there may be other elements to the game like the roll of a die or shuffling of cards.

These elements are characteristics of so-called sequential combinatorial games. The sequential nature arises from the fact that each player makes a

Date: October 16, 2013.

move before other players are allowed to make a move. This requirement, really a combination of (2) and (3), allows the remaining players to have this information available to them to incorporate into their own choices. The combinatorial nature is due to the the complexity of a game that is due to a – potentially large – number of possible moves for each player at each stage of the game in (3). The potential for a large number of legal moves may make it difficult for a player to nail down a winning strategy. According to Bewersdorff in [4]:

> "*There is no unified theory for the combinatorial elements in games. Nonetheless, a variety of mathematical methods can be used for answering general questions as well as solving particular problems.*"

While many games have analogs to serious real world problems which have a meaningful payoff at the end, and much of game theory is focused on such concerns, we will consider the payoff of these games here to be the amusement mentioned in the following definition of **game** found in [13]:

> *Game (n) – a competitive activity involving skill, chance or endurance on the part of two or more persons who play according to a set of rules, usually for their own amusement or that of spectators.*

There may be additional benefits in a pedagogical situation, and we do make a case for the pedagogical benefit of game play, but we purposefully exclude the notion of an outcome with any monetary value or increase in prestige. Meanwhile there are aspects of mathematical games that not only assist in preserving the enjoyment derived from playing a game over and over, but also enable the games to yield mathematical interpretations. These include the aforementioned chance element and the number of legal moves that each player can choose between at each stage of the game. An additional mitigating factor is the potential for different states of information among the individual players. That is, at each state of the game, when a player makes a move he may know more or less about the progress of the game. In chess, for example, each player is aware of every move that has been made so far. This also provides information about the possible remaining moves because of the positions of the pieces on the board. This is an example of a *perfect information* game. On the other hand, in a card game like bridge, the players may know what moves have been made so far, but players ordinarily do not know the cards being held by the other players. This is an example of an *imperfect information game.*

In the following section we introduce a pair of sequential games that involve some aspect of chance, but not to the same degree as poker or dice games. These games also have a strategic component like checkers or Go, but the chance element mitigates the pure strategy. One game, played with colored stones similar to those used in Mancala, is a perfect information game where the set of available moves and the current position of each player is known to all players at every stage of the game after the initial draw. The other game, played with special dominoes, is an imperfect information game where the set of available moves is known to all players, but each player's set of potential moves is hidden from all other players. After a brief introduction to the game play of both games, we turn to the structure of the common rules underlying both games and see that they can be interpreted in a variety of ways as different mathematical structures. Our goal is not to present a body of research about paths to an endgame, but rather we consider the structure of the rules of the games with an eye toward the pedagogical implementation of certain related mathematical concepts via a relatively simple method for playing with colors.

2. The games

The first game, called *Al-Jabar*, is played with a collection of colored stones. This game was initially developed by a father to be played with his son.[10] The game, as it currently exists, has evolved from this humble beginning to a sequential game with a robust mathematical structure. The game pieces are comprised of 10 stones each of the colors red, orange, yellow, green, blue, purple, and white along with 30 black stones.[1] Throughout the subsequent sections we denote black with the symbol K when we begin thinking about the mathematical structure of the games, and reserve B to denote the color blue.

Game play of *Al-Jabar* goes as follows:

(1) The black stones are placed in the middle of the playing surface.
(2) From the bag of the remaining stones, each player is dealt 13 stones at random.
(3) One colored stone is placed in the middle along with the black pieces as a starter.
(4) Players take turns exchanging pieces from their hands for pieces in the middle – up to 3 at a time – with the goal of reducing the number

[1]Note that the official rules describe these pieces as "clear or black" rather than just black. Here we will use the color black to describe the pieces, in part because this choice translates more easily to the discussion of the second game and the visual structure of the rules.

of stones in their hand. These exchanges are made using the rules of color addition.

(5) The final round of the game is indicated when a player ends his turn with exactly one piece in his hand or begins his turn with 2 or 3 pieces and no move that will reduce the number of pieces in his hand.

For a full description of the rules, along with information about rules variations, please see [2].

The second game, called *Spectrominoes*, is played with special dominoes according to the rules of the familiar domino game *Mexican Train*, see [9], as follows:

(1) Each player is dealt a number of dominoes, depending on the number of players.
(2) Players take turns playing their dominoes in a row by matching the colors along the ends or sides of the pieces according to the rules of color addition.
(3) A round ends when a player has run out of tiles to play.

The distinctive characteristic of this game is that the dominoes, instead of having a number of pips on each end, have one block of color on each end where the colors are chosen from the collection: red, orange, yellow, green, blue, purple, white and black. Sample moves are shown below.

FIGURE 1. Sample moves in the game of *Spectrominoes*.

In each case, the question remains: *What are the rules of color addition?*

3. Stipulation for pigmentation combination

The two games briefly described in the previous section are based on a set of straightforward color addition rules that are a mixture of additive and subtractive color mixing. Subtractive color mixing is based on a classical RYB color model that can be thought of as mixing paint colors. This is something that anyone who has played with finger paints would intuitively understand. An example of this is shown in Figure 2.

FIGURE 2. A fingerpainted masterpiece.

Additive color mixing is a more modern interpretation of color mixing where new colors are created by mixing different colors of light. One of the canonical models for this is the RGB standard for color on a computer monitor or television screen. Another example is the overlapping of colored lights in settings such as theaters as shown in Figure 3.

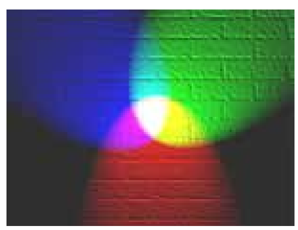


FIGURE 3. Color mixing via colored lights.

In the color addition rules for *Al-Jabar* and *Spectrominoes*, we mix the color mixing metaphors and default to having red, yellow and blue as the primary colors in both models. Additionally, using the metaphor of color mixing via colors of light, we include a notion of flipping a switch to turn on the light. This final idea gives the color addition rules a binary flavor which leads to a variety of mathematical interpretations.

There are five different color addition rules, with one of them being the same for each of the different colors. We describe them below in terms of their reliance on subtractive and additive color mixing fundamentals beginning with three rules based on subtractive color mixing. These are the more or less obvious rules that follow directly from the idea of mixing different colors of paint. They are:

$$\begin{aligned} \text{RED} + \text{YELLOW} &= \text{ORANGE} \\ \text{RED} + \text{BLUE} &= \text{PURPLE} \\ \text{YELLOW} + \text{BLUE} &= \text{GREEN} \end{aligned}$$

The other two rules are based on additive color mixing, with the extra notion of flipping a switch to turn on a specific light. The first of these is:

$$\text{RED} + \text{YELLOW} + \text{BLUE} = \text{WHITE}$$

We can interpret this as getting white light by mixing all of the colors of the spectrum which we can, in turn, create from the three primary colors red, yellow and blue. We will call this the **spectrum rule**. The final rule is a binary notion that arises when we think of getting a color by flipping a switch.

$$\text{RED} + \text{RED} = \text{BLACK}$$

In this case we can imagine that if we toggle the red light switch and then toggle it again, we get dark – symbolized by the color black – instead of a red light as a combination of red and red. There are, of course, eight versions of this rule, one for each color. These rules give rise to a variety of other unusual rules for color addition, which we will discuss as we consider the various ways we can interpret the rules using mathematical structures.

Now that we have the rules in place, we can have the means to experience abstract mathematical concepts somewhat tangibly through visual perception. This also allows us to investigate connections between these concepts and see the richness of the mathematical landscape. In [14] Wells gives us a nice motivation for making these connections.

> *"Exploration leads – as it does in natural history and geography – to important structures and features being identified, named and classified, so that the game develops its own language. These structures make abstract games playable and mathematics manageable."*

One of the pedagogical benefits of using the games as an approach to teaching mathematical concepts is that the rules are essentially simple, albeit somewhat counterintuitive, but this is where part of the benefit lies. Based on very straightforward ideas, surprising things can happen. That these things can illustrate a variety of mathematical principles can be a useful pedagogical tool since recreational mathematics can be though of as a mechanism for making serious mathematics more approachable to those in the process of developing expertise. We begin with the interpretation that uses, perhaps, the most obvious mathematical structure and continue to other less apparent connections.

4. The game is called Al-Jabar for a reason

The initial interpretation of the color addition rules, once the robust structure was in place, was as a group. In particular, we can think of the rules as a colored version of the structure of the group $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ where we have the colors corresponding to the elements of the group as follows.

RED	$\longleftrightarrow$	$(1,0,0)$	ORANGE	$\longleftrightarrow$	$(1,1,0)$
YELLOW	$\longleftrightarrow$	$(0,1,0)$	GREEN	$\longleftrightarrow$	$(0,1,1)$
BLUE	$\longleftrightarrow$	$(0,0,1)$	PURPLE	$\longleftrightarrow$	$(1,0,1)$
WHITE	$\longleftrightarrow$	$(1,1,1)$	BLACK	$\longleftrightarrow$	$(0,0,0)$

This allows us to rewrite the color addition rules from the previous section in $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ as follows:

$$R + Y = (1,0,0) + (0,1,0) = (1,1,0) = O$$
$$R + B = (1,0,0) + (0,0,1) = (1,0,1) = P$$
$$Y + B = (0,1,0) + (0,0,1) = (0,1,1) = G$$

$$R + Y + B = (1,0,0) + (0,1,0) + (0,0,1) = (1,1,1) = W$$

$$R + R = (1,0,0) + (1,0,0) = (0,0,0) = K$$

This also allows us to establish the somewhat counterintuitive color addition result of $GREEN + PURPLE = ORANGE$ as follows:

$$G + P = (0,1,1) + (1,0,1) = (1,1,0) = O.$$

For the sake of completeness, we have the following group table for color addition on the set of eight colors.

⋆	B	R	O	Y	G	B	P	W
K	K	R	O	Y	G	B	P	W
R	R	K	Y	O	W	P	B	G
O	O	Y	K	R	P	W	G	B
Y	Y	O	R	K	B	G	W	P
G	G	W	P	B	K	Y	O	R
B	B	P	W	G	Y	K	R	O
P	P	B	G	W	O	R	K	Y
W	W	G	B	P	R	O	Y	K

FIGURE 4. The group table for color addition.

In a classroom setting, game play of *Al-Jabar* would give students the experience of interacting with the group structure in a visceral way. Assembling pieces to make an exchange amounts to calculating a sum in the group. In order to determine combinations that create a BLACK sum, we would look for the inverse of a collection of pieces with a particular sum.

If, instead of thinking of the colors as ordered triples, we think of them as vectors in 3-space, then we can realize a geometric structure of $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ as the vertices of a cube with the action of vector addition taken under mod 2 addition.

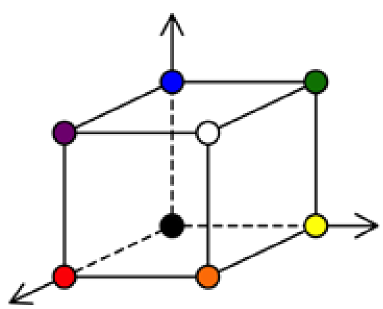


FIGURE 5. The rules of color addition as vectors in 3-space.

This may have applications in linear algebra or vector calculus as students begin to think about geometric and algebraic structures outside of the plane.

5. Finita geometricum

While the first geometric interpretation is closely related to the algebraic interpretation, being essentially the same thing with a slight change of notation, the second one is a bit more surprising. One caveat here is that we do not have an exact representation since one essential element is lacking in practice. However, we can overcome this with a bit of cleverness. In the meantime, this second interpretation gives a very nice visual representation of the rules at the same time as an introduction to a non-Euclidean geometry. The corresponding picture can also act as a visual aid during game play. This new geometric interpretation is as a coloring of the Fano plane. The Fano plane is a finite geometry with 7 points and 7 lines. The general structure looks like the picture below where the black dots are the seven points and the line segments, rather than being actual lines in the geometry them selves, indicate which points are on the lines. Note that the *circle*

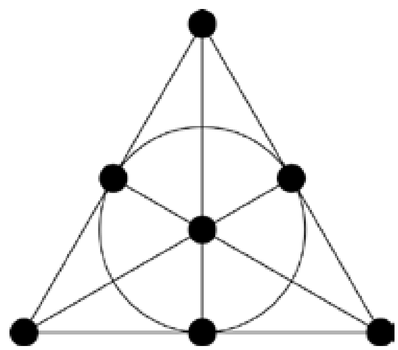


FIGURE 6. The Fano plane.

connecting the three midpoints indicates that these three points form a line in this geometry.

According to the Pólya Enumeration Theorem, there are

$$\frac{1}{168}\left(n^7 + 21n^5 + 98n^3 + 48n\right)$$

colorings of the Fano plane with n colors. For $n = 7$, we can choose one of the 7205 colorings as shown below to create a visual guide to the addition rules – with the exception of the binary structure given by black being the sum of any color with itself.[2] Notice that the midpoint of each side of the triangle is the secondary color corresponding to the sum of the colors of the two corresponding vertices, as in the metaphor of subtractive color mixing. But since the geometry does not have an inherent notion of betweenness, we can consider any point on any line to be the one *in the middle* so that we get RED=PURPLE+BLUE as well. To recover the spectrum rule, we can

[2]We make the following convention: we call such a coloring, using the 7 colors $\{R, O, Y, G, B, P, W, K\}$, a *spectral coloring* – or refer to the figure as being *spectrally colored* – as a nod to the colors of the spectrum.

perform the following calculation,

$$\begin{aligned} RED + YELLOW + BLUE &= RED + (YELLOW + BLUE) \\ &= RED + GREEN \\ &= WHITE \end{aligned}$$

by calculating the sums along individual lines. But this essentially forces the sum of the colors on a line to be black so that we recover, in some sense, the notion of the binary structure.

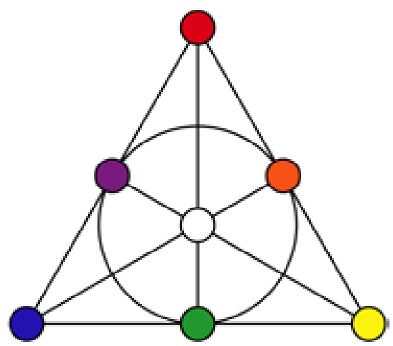


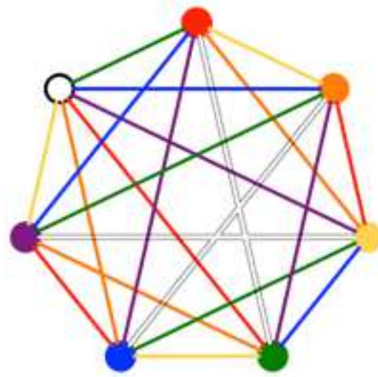
FIGURE 7. A *spectral* coloring of the Fano plane.

As an extension of this idea, we see from the work of Baez that we can recover the structure of the octonians from our coloring of the Fano plane.[4] In order to mitigate the appearance of the additive inverses, we can restrict the field of coefficients to $\mathbb{Z}_2$. We leave it as an exercise for the reader to determine the correspondence between the colors and the non-identity elements of the octonians.

This representation of finite geometries and division algebras through color addition seems to be a means by which students can be introduced gently to concepts that are far outside of their experiences from their secondary mathematics classes.

6. Totally colored complete graphs

In this section we consider a spectral coloring of K_7 in order to see the effects of various moves and the interdependencies between the colors. This graph is undirected and the color of the edge indicates the color that would be added to an individual vertex to create the color of the adjacent vertex. Similarly, the sum of the colors of the vertices at the ends of a given edge would be the corresponding color of the edge.



Notice that this is a proper coloring of the vertex set of K_7. It is also a proper coloring of the edge set of K_7. This makes the coloring a proper total coloring of K_7. This final feature is a variation on graph coloring that is both a proper vertex coloring and a proper edge coloring and also a coloring such that no edge shares a color with either of its incident vertices.

Moreover, this particular coloring is a minimal example of such a coloring since one fundamental property of the total coloring number $\chi''(G)$ of a graph G is that $\chi''(G) \geq \Delta G + 1$ where $\Delta(G)$ is the maximum degree of G. In this case we get equality, which is a well known result for complete graphs on an odd number of vertices.

With this particular coloring in place we have the following two results.

Proposition. *The edge sum of a cycle in a spectrally colored K_7 is black.*

Proof. Let S be a spectrally colored K_7 and let $C = e_1 e_2 \cdots e_k$ be a cycle in S. Then we can write the edge sum as

$$\begin{aligned} e_1 + e_2 + \cdots + e_k &= (v_1 + v_2) + (v_2 + v_3) + \cdots (v_k + v_1) \\ &= (v_1 + v_1) + (v_2 + v_2) + \cdots + (v_k + v_k) \\ &= K. \end{aligned}$$

Therefore the edge sum of a cycle is black. □

Proposition. *The vertex sum of a cycle in a spectrally colored K_7 is equal to the vertex sum of the remaining vertices.*

Proof. Let S be a spectrally colored K_7 and let $C = v_1 v_2 \cdots v_k$ be a cycle in S. Then the vertex sum in S is

$$\begin{aligned} c_{v_1} + c_{v_2} + \cdots + c_{v_k} &= [c_{v_1} + c_{v_2} + \cdots + c_{v_k}] + \sum_{v \in S} c_v \\ &= [c_{v_1} + c_{v_2} + \cdots + c_{v_k}] + (c_{v_1} + c_{v_2} + \cdots + c_{v_k}) + \sum_{v \notin C} c_v \\ &= (c_{v_1} + c_{v_1}) + \cdots + (c_{v_k} + c_{v_k}) + \sum v \notin C c_v \\ &= K + \cdots K + \sum_{v \notin C} c_v \\ &= \sum_{v \notin C} c_v \end{aligned}$$

as desired. □

The second proposition allows us to think about the exchanges made in the games in terms of different sets of pieces. It also provides an alternate reason for the unusual color addition rule

$$\mathsf{O} + \mathsf{G} + \mathsf{P} = \mathsf{K}$$

since the colors not in the cycle O–G–P are the colors in the spectrum move, whose sum is black.

Graph theory is an interesting visual branch of mathematics which is easily accessible to undergraduate students without requiring them to have an extensive background in order to understand interesting questions and even to begin doing research. Graph coloring is an active area in this discipline and the picture above is a nice entry point into this domain.

7. A knotty interpretation

The basic study of knot theory can be thought of as the search for ways to distinguish between knots. When this search is successful, the methods for distinguishing between knots are measurements of the complexity of a knot – so-called **knot invariants**. That is, the measurement given does not depend on the particular picture of the knot, or the way the knot is sitting in space, but rather the essential structure of the knot. One particular invariant is related to colorings of knot projections, those pictures of knots where a break in the string indicates that a strand is passing behind another strand. For a given prime number p, a knot is said to be p-colorable if there exists a projection of the knot which can be colored using the *colors* $\{0, 1, 2, \ldots, p-1\}$ so that at each crossing the sum of the two understands is equal to twice the over strand, mod p. That is, at a given crossing with the indicated coloring we have $x + y \equiv 2z \pmod{p}$.

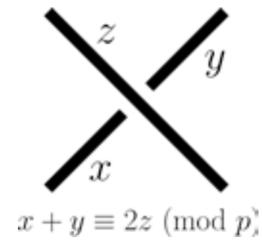


FIGURE 8. The p-colorability relation.

We can make a connection to knot theory by considering the following spectrally colored picture of the 7_1 knot:

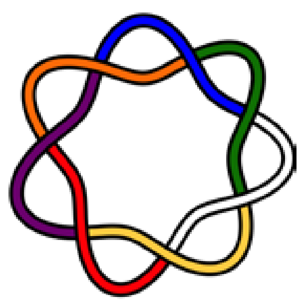


FIGURE 9. A colored projection of the 7_1 knot.

If we define a mapping $\bowtie$: $\{R, O, Y, G, B, P, W\} \to \{0, 1, 2, 3, 4, 5, 6\}$ between the usual colors in the picture and the mod 7 natural numbers as shown below, then we have a colored picture that satisfies the modular relationship described above, and which also follows the rules of color addition.

$$\bowtie = \{(\mathsf{W}, 0), (\mathsf{Y}, 1)(\mathsf{R}, 2), (\mathsf{P}, 3), (\mathsf{O}, 4), (\mathsf{B}, 5), (\mathsf{G}, 6)\}$$

Then we see that at the PURPLE-ORANGE-BLUE crossing the understrand sum would be $P + B \bowtie 3 + 5 \equiv 1 \pmod{7}$ while $2O \bowtie 2(4) \equiv 1 \pmod{7}$. Moreover, the sum of the three colors, as we have considered before is given by $P+O+B = Y \bowtie 1$. So we have a relationship that resembles the knot theory calculation for p-colorability and also retains aspects of color addition. We note here that we have not used the color BLACK because this notion of colorability uses a prime number of colors. So this correspondence would not account for a crossing in a knot in which all of the strands were the same color, which is an allowable coloring since $x + x \equiv 2x \pmod{p}$ for any p and for any $0 \leq x < p$.

Knot theory is another very visual entry point into higher mathematics for students, and colorability is one of the nice properties that is straightforward to check while also having deep connections to other knot theory invariants.

8. The difference is it

We have saved what is perhaps the most interesting interpretation for last, and it is in some ways the most abstract. As usual we consider the set of colors used in the game.

$$S = \{\mathsf{R}, \mathsf{O}, \mathsf{Y}, \mathsf{G}, \mathsf{B}, \mathsf{P}, \mathsf{W}, \mathsf{K}\}$$

From this set we can create the power set $\mathcal{P}(S)$ which will have $2^8 = 256$ elements. Then we can define an equivalence relation $\blacktriangle$ on the power set by saying that sets X and Y are equivalent, or $X \blacktriangle Y$ if the sum of the colors in the sets are equal. For example if we have

$$C_1 = \{\mathsf{O}, \mathsf{G}\} \quad \text{and} \quad C_2 = \{\mathsf{R}, \mathsf{B}\}$$

then we know $C_1 \blacktriangle C_2$ since the sums of the colors in both sets is purple. There is still the question of how to define the *color* of the empty set. For various reasons, it is logical to define the color of the empty set to be black. We see that if we define the color of the empty set to be clear, that is $\varnothing \blacktriangle \{\mathsf{K}\}$, then

Using this equivalence relation we can consider the function

$$\triangle : \mathcal{P}(S) \times \mathcal{P}(S) \to \mathcal{P}(S)$$

where $\triangle$ is the familiar symmetric difference operator. That is

$$A \triangle B = (A \cup B) \setminus (A \cap B).$$

Under this binary relation we can recover the structure of $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ by the operation $[A] \triangle [B] = [C]$ where $[X]$ is the singleton set which we choose as the representative of the associated equivalence class.

If we define the set $\mathcal{A}$ to be the elements of the power set of S with at most three elements along with the set $\{R, Y, B, W\}$ then we have a list of the possible moves in the game *Al-Jabar*. For example, there would be eight subsets in $\mathcal{A}$ which are equivalent to RED:

$$[\mathsf{RED}] = \left\{ \{\mathsf{R}\}, \{\mathsf{Y}, \mathsf{O}\}, \{\mathsf{G}, \mathsf{W}\}, \{\mathsf{B}, \mathsf{P}\}, \{\mathsf{O}, \mathsf{G}, \mathsf{B}\}, \{\mathsf{O}, \mathsf{P}, \mathsf{W}\}, \{\mathsf{Y}, \mathsf{G}, \mathsf{P}\}, \{\mathsf{Y}, \mathsf{B}, \mathsf{W}\} \right\}$$

Finally, if we allow for multisets with multiplicity two, then we can recover the idea of the sum of two colors being equal to black.

References

[1] C. Adams, *The knot book: An elementary introduction to the mathematical theory of knots*, Amer. Math. Soc., Providence, RI, 2004.

[2] *Al-Jabar: A Mathematical Game of Strategy*, `http://al-jabargames.com`, last accessed October 16, 2013.

[3] J. Baez, *The octonians*, Bull. Amer. Math. Soc. **39** (2002), 145-205.

[4] J. Bewersdorff, *Luck, logic and white lies: The mathematics of games*, translated by D. Kramer, AK Peters, Wellesley, MA, 2005.

[5] G. Chartrand & L. Lesniak, *Graphs and digraphs*, Chapman & Hall, London, 1996.

[6] M. Davis, *Game theory: A nontechnical introduction*, Basic Books, New York, 1983.

[7] T. Huffman, *Total colorings of graphs*, Master's Thesis, San Jose State University, 1989.

[8] G. Owen, *Game theory*, W.B. Saunders, Philadelphia, 1968.

[9] Rules of *Mexican Train*, `http://en.wikipedia.org/wiki/Mexican_train`, last accessed October 16, 2013.

[10] R. Schneider, personal communication.

[11] *Spectrominoes*, `https://www.facebook.com/Spectrominoes`, last accessed October 16, 2013.

[12] S. Vajda, *Mathematical games and how to play them*, Ellis Horwood, West Sussex, 1992.

[13] *Websters Encyclopedic Unabridged Dictionary of the English Language*, Dilithium Press, Ltd. New Jersey, 1994.

[14] D. Wells, *Games and mathematics: Subtle connections*, Cambridge University Press, 2012.

1	2	4	4	5	6
8	9	5	6	7	3
2	5	3	8	9	4
8	7	9	1	6	8

MAGIC

6	6	7	7	8	8
7	6	2	6	4	6
8	6	8	3	6	7
4	5	8	5	9	4

2	2	3	4	5	5
7	9	9	5	4	7
4	5	4	3	2	9
9	7	6	9	9	6

5	7	7	7	8	9
9	2	6	6	6	7
6	8	3	8	7	8
4	9	5	5	4	2

A Simple MatheMagics Trick
for G4G11

Print the previous page on cardstock and cut out the strips along the lines.

Tell someone that they can arrange as many of the strips as they want in whatever order they want (there are billions of possibilities) and you will add up the four multi-digit numbers instantly.

For example, if they pick the top six strips and place them in the order given above, then the first of the four six-digit numbers to be added up is 223,455.

The sum – and note that I am not even slightly pausing as I write this :-) – is 2,454,327. They can check your arithmetic with a calculator of course.

If you want to know how to perform this trick, I have put the method on page 4. That way you can choose to think about it first.

My variation of this trick uses the set of cards on the next page instead. They're purple to distinguish them. It is slightly more difficult to perform and much less likely for an audience to pick up on even if you perform the trick multiple times for them. I have not supplied the method here but you can email me if you can't figure it out.

Skona Brittain
SB Family School & SB Crafts
skona@sbfamilyschool.com or skona@sbcrafts.net

1	2	4	4	5	6
8	9	5	6	7	3
2	5	3	8	9	4
8	7	9	1	6	8

6	6	7	7	8	8
7	6	2	6	4	6
8	6	8	3	6	7
4	5	8	5	9	4

Spoilers

The Method:
The sum is the number in the 3rd row with a 2 prepended to it and 2 subtracted from its unit digit. So for the example above, 454,329 gives 2,454,327.

Why does this work?

Hint #1

It has something to do with the mathemagical number of 9.

Hint #2

Examine the sum of the three numbers in the other rows on each card strip.

Source

The original card strips, but not the explanation, are from a kit called Magic Science, which I purchased a couple of decades ago but can no longer find.

Kristine Hjulstad: How Magicians Fool Our Brain

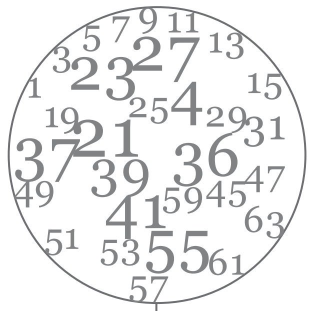
The didactic use of conjuring in teaching how the brain works and how it can be fooled.

Medical student and performer of *Theatrical Magic*, Kristine Hjulstad, explains how she uses magic tricks to illustrate the strengths and weaknesses of human perception. For three years Kristine has travelled with her performance ”Hjerne det”! (”The Brain”) entertaining and teaching pupils age 8-12 principles of neuro physiology and conjuring.

Contact:
www.kristinehjulstad.no
kristine.magi@hotmail.com

If you wish to view the complete talk og to:

http://urplay.se/Produkter/178895-UR-Samtiden-Skeptikerkongressen-2013-Hur-magiker-lurar-vara-hjarnor



MindReading Math

This is an ancient puzzle.

- A secret number is chosen by your friend. (between 1 and 63)
- Your friend looks at all 6 number cards.
- Your friend then places all the cards with their secret number on the table.
- You guess the secret number.

*The Secret is adding together the upper left hand corner number of the cards on the table.

1	3	5	7	9	11	13	15
17	19	21	23	25	27	29	31
33	35	37	39	41	43	45	47
49	51	53	55	57	59	61	63

2	3	6	7	10	11	14	15
18	19	22	23	26	27	30	31
34	35	38	39	42	43	46	47
50	51	54	55	58	59	62	63

4	5	6	7	12	13	14	15
20	21	22	23	28	29	30	31
36	37	38	39	44	45	46	47
52	53	54	55	60	61	62	63

8	9	10	11	12	13	14	15
24	25	26	27	28	29	30	31
40	41	42	43	44	45	46	47
56	57	58	59	60	61	62	63

16	17	18	19	20	21	22	23
24	25	26	27	28	29	30	31
48	49	50	51	52	53	54	55
56	57	58	59	60	61	62	63

32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47
48	49	50	51	52	53	54	55
56	57	58	59	60	61	62	63

… is it magic or math?

Magician **Lisa Menna** | menna.com
Visit: CAUSETOWONDER.ORG

An Introduction to Gilbreath Numbers

Robert W. Vallin
Department of Mathematics
Lamar University
Beaumont, TX 77710 USA
robert.vallin@lamar.edu

1 August 2016

Abstract

We begin with a "magic trick." The trick works as a result of the Gilbreath Principle, which can be proven using mathematical induction. In this paper we use the second version of the Principle to develop a new classification of number, the Gilbreath Continued Fraction. Once they are defined we then go on to describe Gilbreath Numbers and their place in the unit interval. Lastly, we look at some generalizations.

1 Introduction

1.1 A Magic Trick and Gilbreath Permutations

In 1958 an undergraduate math major at UCLA named Norman Gilbreath published a note in *The Linking Ring*, the official publication of the International Brotherhood of Magicians [6], in which he described a card trick. Stated succinctly, this trick can be performed by handing an audience member a deck of cards, letting him or her cut the deck several times, and then dealing N cards from the top into a pile. The audience member takes the two piles (the cards in hand and the set now piled on the table) and riffle-shuffles (in a riffle shuffle, the deck is split into two halves, one in each, hand, and the cards are released by the thumbs so that they fall on the table interwoven) them. The magician now hides the pile of cards (under a cloth, behind the back) and proceeds to produce pairs of cards where one card is black and the other red, claiming this is proof of the magician's powers.

The key to this trick is that the cards are pre-arranged in black/red order before the deck is handed out. Cutting the deck does not change this arrangement. When the top N cards are dealt into a pile, the black/red pattern is still there, but the order is reversed. It is then a mathematical induction argument to show that however the riffle-shuffle is performed, consecutive pairs of cards still maintain opposite colors. Let us look at a small example of this with eight cards and the subscripts just referring to first, second, third, etc. card in the original deck with R for red, B for black. The process is illustrated below.

Original order	After cutting	Two piles	Shuffled
B_1	B_5	B_1	R_8
R_2	R_6	R_2	B_7
B_3	B_7	B_3	B_1
R_4	R_8	R_4	R_6
B_5	B_1		R_2
R_6	R_2	R_8	B_5
B_7	B_3	B_7	B_3
R_8	R_4	R_6	R_4
		B_5	

If cards are taken two at a time from either the top or bottom of the deck, then we get pairs consisting of one of each color.

As is somewhat obvious, the colors do not matter as much as we have two types of cards. In fact, having a pattern of *two* types of characteristics does not matter. This trick also works if the cards are arranged by suit (e.g., Clubs, Hearts, Spades, Diamonds) and, after shuffling, dealt off four at a time.

Now let us turn this mathematical. For any nonempty set, S, a permutation on S is a bijection from S onto itself. We are only concerned with permutations on sets of numbers, beginning with $\{1, 2, 3, \ldots, N\}$, and eventually moving onward to $\{1, 2, 3, \ldots\}$, with a particular property (this is seen in the Ultimate Gilbreath Principle below). Notationally, we will denote our permutation using π and for any j, $\pi(j)$ refers to the number in the jth place of the permutation, not the placement of the number j in the permutation. So with $\{3, 4, 5, 2, 1\}$, $\pi(1) = 3$ and $\pi(5) = 1$.

Let us repeat the figure above, but this time with numbers rather than cards. Note that now that we are not performing a trick, the cutting, which gives the audience the idea of some randomness, is unnecessary.

Original order	Two "piles"	"Shuffled"
1	5	5
2	6	4
3	7	6
4	8	7
5	9	3
6	10	8
7		2
8	4	9
9	3	1
10	2	10
	1	

Now we write this as a permutation on $\{1, 2, 3, \ldots, N\}$. The permutation for the example above is

$$\{5, 4, 6, 7, 3, 8, 2, 9, 1, 10\}.$$

This is what is known as a *Gilbreath Permutation*.

Not every permutation is a Gilbreath Permutation. In fact, while there are $N!$ permutations on $\{1, 2, 3, \ldots, N\}$, there are only 2^{N-1} Gilbreath Permutations. This fact and the *Ultimate Gilbreath Principle* below, which tells us which permutations are Gilbreath Permutations, are from [2]. It is there the reader can find the proof of this.

Theorem 1 (The Ultimate Gilbreath Principle)
For a permutation π of $\{1, 2, 3, \ldots, N\}$ the following are equivalent:

1. *π is a Gilbreath Permutation.*
2. *For each j, the first j values*
   $$\{\pi(1), \pi(2), \ldots, \pi(j)\}$$
   are distinct modulo j.
3. *For each j and k with $jk \leq N$ the values*
   $$\{\pi((k-1)j + 1), \pi((k-1)j + 2), \ldots, \pi(kj)\}$$
   are distinct modulo j.
4. *For each j, the first j values are consecutive in $1, 2, 3, \ldots, N$; that is, if you take the first j values from the permutation, they can be rearranged as j consecutive numbers less than or equal to N.*

Just for the sake of pointing things out, it is Part Three that makes the magic trick work. It says whether you are taking cards from the deck in groups of two (red/black) or four (Clubs, Hearts, Spades, Diamonds), if the deck is set up correctly, then the magician gets one card of each type. Our interest is in the fourth part. This says even if the j numbers are not currently written in order, they appear in order in $\{1, 2, 3, \ldots, N\}$. So, for example, the permutations $\{2, 3, 1, 4\}$ and $\{3, 4, 2, 1\}$ are Gilbreath Permutations of $\{1, 2, 3, 4\}$, while $\{3, 4, 1, 2\}$ is not.

1.2 A Quick Look at Continued Fractions

Continued fractions have a long history with many interesting results. Even a quick look must contain lots of ideas. We present this subsection's theorems without proof, as they can be found in any text on continued fractions, such as [11]. To keep our investigations easier, we will keep with the simple continued fractions.

Definition 1 *Let* a_i*,* $i = 0, 1, 2, 3, \ldots$ *denote a collection (possibly finite) of integers*[1] *with* a_i *is a positive for* $i \geq 1$*. A* simple continued fraction *is an expression of the form*

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}}$$

To simplify our notation we will write this number as

$$[a_0; a_1, a_2, a_3, \ldots].$$

The individual $a_0, a_1, a_2, \ldots$ *are referred to as the* partial quotients *of the continued fraction expansion.*

Since our work on this is only concerned with numbers in the unit interval, we will refer the interested reader to [11] to learn more about all types of continued fractions.

Example 1 *The finite simple continued fraction* $[2; 3, 6]$ *represents*

$$2 + \frac{1}{3 + \frac{1}{6}} = 2 + \frac{1}{\frac{19}{6}} = 2 + \frac{6}{19} = \frac{44}{19}.$$

In the other direction,

$$\frac{32}{15} = 2 + \frac{2}{15} = 2 + \frac{1}{\frac{15}{2}} = 2 + \frac{1}{7 + \frac{1}{2}}.$$

So $32/15 = [2; 7, 2]$*.*

The technique above for turning a rational number into a continued fraction is guaranteed to terminate. Notice that in this dividing process ($32 \div 15$, then $15 \div 2$) the remainders, which in the next step become the denominators, are strictly decreasing. Thus the remainder must at some point become the number 1 finishing our continued fraction expansion in the rational numbers.

If x is a positive *irrational number*, then there exists a largest integer a_0 such that $x = a_0 + \frac{1}{x_1}$ where $0 < (x_1)^{-1} < 1$. Note

$$x_1 = \frac{1}{x - a_0} > 1$$

and is irrational (after all, x is irrational and a_0 is an integer).

1. Some references use complex numbers with integer coefficients.

We repeat this process, starting with x_1, find a_1, the greatest integer such that

$$x_1 = a_1 + \frac{1}{x_2}$$

where $0 < (x_2)^{-1} < 1$. As we continue we generate the continued fraction for x

$$[a_0; a_1, a_2, a_3, \ldots]$$

where this time the sequence of a_i does not terminate.

Example 2 *An example of this is $\sqrt{3}$ as*

$$\sqrt{3} = [1; 1, 2, 1, 2, 1, 2, \ldots] \text{ or, more conveniently, } [1; \overline{1, 2}].$$

This continued fraction expansion can be verified by solving the following equation

$$x = 1 + \cfrac{1}{1 + \cfrac{1}{2 + (x - 1)}},$$

where $(x - 1)$ is the repeating part of our continued fraction.

A quadratic irrational is a number of the form

$$\frac{P \pm \sqrt{D}}{Q},$$

where P, Q, and D are integers, $Q \neq 0$, $D > 1$ and not a perfect square. Lagrange, in 1779, proved that the continued fraction expansion of any quadratic irrational will eventually become periodic. It also goes the other direction. Euler showed that if a continued fraction is eventually periodic, then the value can be expressed as a quadratic irrational number.

There are also non-repeating, infinite, simple continued fractions. As an example,

$$\pi = [3; 7, 15, 1, 292, 1, 1, \ldots]$$

and we know this cannot end or repeat as π is a transcendental number (not the solution to a polynomial with integer coefficients). We will revisit transcendental numbers later.

We now turn our attention to *convergents.*

Definition 2 *Let x have the simple continued fraction expansion (finite or infinite) of $[a_0; a_1, a_2, a_3, \ldots]$. The* convergents *are the sequence of finite simple continued fractions*

$$c_0 = a_0, \quad c_1 = a_0 + \frac{1}{a_1}, \quad c_2 = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2}}, \quad \ldots.$$

and, in general, $c_n = [a_0; a_1, a_2, a_3, \ldots, a_n]$.

Typically, the next step is to represent these convergents as the rational numbers that they are. For example, $c_0 = a_0 = \frac{p_0}{q_0}$, $c_1 = a_0 + \frac{1}{a_1} = \frac{p_1}{q_1}$, $c_2 = a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = \frac{p_2}{q_2}$, and so on.

These convergents approach a limit. If we start with two convergents c_i and c_{i+1}, the literature on continued fractions then shows that

$$c_{i+1} - c_i = \frac{p_{i+1}}{q_{i+1}} - \frac{p_i}{q_i} = \frac{p_{i+1}q_i - p_i q_{i+1}}{q_{i+1}q_i} = \frac{(-1)^{i+1}}{q_{i+1}q_i}$$

From the definition of convergent it can be seen that q_i is always positive and increasing. Thus we have

Theorem 2 *For any simple continued fraction* $[a_0; a_1, a_2, a_3, \ldots]$ *the convergents,* c_i *form a sequence of real numbers where for all* i

- $c_{2i-1} < c_{2i+1} < c_{2i}$, *and*
- $c_{2i+1} < c_{2i+2} < c_{2i}$.

Thus the sequence of convergents has the property

$$c_1 < c_3 < c_5 < \cdots < c_{2i-1} < \cdots < c_{2i} < \cdots < c_4 < c_2 < c_0.$$

This brings us, finally, to a theorem that says any infinite simple continued fraction has meaning as a unique point on the real number line.

Theorem 3 *Let* $[a_0; a_1, a_2, a_3, \ldots]$ *represent an infinite, simple continued fraction. Then there is a point* x *on the real line such that* $x = [a_0; a_1, a_2, a_3, \ldots]$

2 Gilbreath Numbers

2.1 Creating Gilbreath Continued Fractions

We now wish to take a Gilbreath Permutation such as $\{3, 4, 2, 1\}$ and turn it into a continued fraction which then represents a real number in the unit interval. Our method is to write the entries in the permutation as a *Gilbreath Continued Fraction* whence $\{3, 4, 2, 1\}$ will become

$$0 + \cfrac{1}{3 + \cfrac{1}{4 + \cfrac{1}{2 + \cfrac{1}{1}}}}.$$

We can check that the simplified form of this rational number is 13/42. Now rather than write out the continued fraction we will take advantage of the fact that all the numerators are 1 and write the number in bracket notation noting the whole number part (0) and the denominator values in order, thusly

$$0+\cfrac{1}{3+\cfrac{1}{4+\cfrac{1}{2+\cfrac{1}{1}}}}=[0;3,4,2,1].$$

In the practice of continued fractions, we do not allow a representation to end with a 1. That is because

$$[0;3,4,2,1] \text{ is the same as } [0;3,4,3].$$

If the terminal digit is not allowed to be 1 then we get the property that continued fraction representations are unique, unlike decimal representations where $1.000\ldots=0.999\ldots$. This is one of the perquisites in dealing with continued fractions rather than decimals. However, not ending in 1 creates issues with Gilbreath Permutations, so we will allow the expansion to terminate with a 1.

If we begin with a Gilbreath Permutation of length N,

$$\{\pi(1),\pi(2),\ldots,\pi(N)\},$$

we may increase its length ad infinitum by inserting the numbers $N+1, N+2, N+3, \ldots$ to make a permutation on $\mathbb{N}$ and keep the property of being a Gilbreath Permutation. This must be done at the end, and with the numbers in order. Placing a new number anywhere else would mean we no longer have a Gilbreath Permutation. This is easiest to see with Part 4 of Theorem 1. Thus we can extend $\{4,3,5,2,1,6\}$ making the sequence

$$\{4,3,5,2,1,6,7\},\quad \{4,3,5,2,1,6,7,8\},\quad \{4,3,5,2,1,6,7,8,9\},\quad \ldots$$

which are each consecutive in $\mathbb{N}$ and the goal is the infinite ordered set of numbers

$$\{4,3,5,2,1,6,7,8,9,10,11,12,13,14,\ldots\}.$$

Another method of making an infinite string using Gilbreath Permutations is in Section 3. We will refer to the part of the expansion where $a_k=k$ as where the permutation straightens out.

Of course, each of these finite strings can be the terms in a finite simple continued fraction (representing a rational number) and this sequence of numbers converges to an irrational number that is the infinite simple continued fraction

$$[0;4,3,5,2,1,6,7,8,9,10,11,12,\ldots].$$

2.2 An Analysis Point of View for Gilbreath Continued Fractions

To make the notation easier to read, we will refer to our continued fractions using the typical notation $[0; a_1, a_2, a_3, \ldots, a_n]$, keeping in mind each $a_j = \pi(j)$. Let $\mathcal{G}$ represent the set of finite and infinite Gilbreath Continued Fractions. Moreover, let $\mathcal{G}_F$ be the set of Gilbreath Numbers whose representation is a finite string. That makes them the set of rational numbers in $\mathcal{G}$, and $\mathcal{G}_I$ the set of Gilbreath numbers whose representation is an infinite string, the irrationals in $\mathcal{G}$. These form a subset of the unit interval $[0, 1]$ and the natural question to ask is, "How much of the unit interval is taken up by these numbers?" We will show that $\mathcal{G}$ is a countably infinite set and a very sparse set in terms of category.

Theorem 4 *The cardinality of the set of Gilbreath Continued Fractions, $\mathcal{G}$, is $\aleph_0$, the cardinality of the natural numbers.*

Proof: For any fixed N, there are 2^{N-1} possible Gilbreath Permutations of $\{1, 2, 3, \ldots, N\}$. Thus the set of numbers in $\mathcal{G}_F$ that look like $[0; \pi(1), \pi(2), \ldots, \pi(N)]$ for each fixed natural number N is finite and the cardinality of $\mathcal{G}_F$ is $\aleph_0$. Now if $x \in \mathcal{G}_I$ there is a $k \in \mathbb{N}$ such that if we write $x = [0; a_1, a_2, \ldots, a_j, \ldots]$, then for $j \geq k$ we have $a_j = j$. In fact, this k is where $\pi(k-1) = 1$. So if we fix $N \in \mathbb{N}$ and insist $a_j = j$ for all $j > N$, how many prefixes are there for this continued fraction that are still Gilbreath? The answer is, of course another 2^{N-1}. Thus there are finitely many sequences for each place where the continued fraction straightens out. So $\mathcal{G}_I$ is countable, too, which means $\mathcal{G}$ is a countable set.

An immediate consequence of this is that the set $\mathcal{G}$ must be a first category set, the countable union of nowhere dense sets. However, we claim there is even more. The set $\mathcal{G}$ is in fact a *scattered* set. For the definition of this, we go to Freiling and Thomson [3].

Definition 3 *Let $S \subset \mathbb{R}$. We say S is scattered if every nonempty subset of S contains an isolated point.*

This idea of scattered has appeared in papers by Cantor, Young & Young, Denjoy, Hausdorff, and others, but without such colorful nomenclature. Scattered sets are different from countable sets and nowhere dense sets. The Cantor Set is nowhere dense (and uncountable), but not scattered The rational numbers are countable, but not scattered. Of course a scattered set cannot be dense, but could be first category. In [3], the authors prove that any countable G–delta set of real numbers is scattered.

Theorem 5 *The set of Gilbreath Continued Fractions, $\mathcal{G}$, is a scattered set in $\mathbb{R}$.*

Proof: For any $x = [0; a_1, a_2, a_3, \ldots, a_n] \in \mathcal{G}_F$ where n is fixed, this number must be isolated. For any $t \in \mathcal{G}_F$, let us assume t has length at most n. There are finitely many of these, so there exists a $\varepsilon_1 > 0$ so that the open ball (interval) centered at x with radius ε_1 does not intersect $\mathcal{G}_F$. If we were to append more numbers onto x, creating

$$y_k = [0; a_1, a_2, a_3, \ldots, a_n, n+1, n+2, \ldots, n+k]$$

these y_k are converging toward some irrational number $y = [0; a_1, a_2, a_3, \ldots, a_n, n+1, n+2, \ldots, n+k, \ldots]$. Thus there is an $\varepsilon_2 > 0$ such that the open ball centered at x with radius ε_2 intersects only finitely many y_k. This argument also explains why there is an ε_3 so that the open ball around x of that radius must miss $\mathcal{G}_I$. Thus the point is isolated.

If $x \in \mathcal{G}_I$, then x is obviously not an isolated point in $\mathcal{G}$ as there is a sequence of points in $\mathcal{G}_F$ that converge to x. However, there is an $\varepsilon > 0$ so that

$$B(x, \varepsilon) \cap \mathcal{G}_I = \emptyset,$$

where $B(x, \varepsilon)$ is the open ball with center x and radius ε. This argument has to do with the place k where the continued fraction expansion "straightens out"; that is, for $j \geq k$ we have $a_j = j$.

This result is powerful as it implies previous results. In [1] it is stated that any scattered set is necessarily both countable and nowhere dense.

Topologically, it is obvious that $\mathcal{G}$ is not open (no scattered set in $\mathbb{R}$ can be). It is, in fact, a closed set in the unit interval.

Theorem 6 *The set of Gilbreath Numbers is a closed set in* $[0, 1]$.

Proof: Pick an x in the complement of $\mathcal{G}$. This x has a continued fraction expanion that is *not* Gilbreath. Let N represent the least index where we see it is not Gilbreath; that is, $a_1, a_2, \ldots, a_{N-1}$ is ordered in $\{1, 2, 3, \ldots, N-1\}$. We create the following sequence of continued fractions

$$\begin{aligned} y_N &= [0; a_1, a_2, \ldots, a_{N-1}, b_N] \\ y_{N+1} &= [0; a_1, a_2, \ldots, a_{N-1}, b_N, b_{N+1}] \\ y_{N+2} &= [0; a_1, a_2, \ldots, a_{N-1}, b_N, b_{N+1}, b_{N+2}] \\ &\vdots \qquad\qquad\qquad \vdots \end{aligned}$$

where each y_n is a Gilbreath Continued Fraction.

Let

$$\varepsilon = \frac{1}{2}\left(\inf_{n \geq N}\{|x - y_n|\}\right).$$

Such an infimum must be positive since there is a unique limit for the y_n (which is not x). Then the open ball with center x and radius ε is contained in the complement of $\mathcal{G}$. Thus $\mathcal{G}$ is a closed set.

3 Generalizations

A Generalized Gilbreath Permutation is a permutation of $\{k, k+1, k+2, \ldots, k+(N-1)\}$ that has the form

$$\{\pi(1)+(k-1), \pi(2)+(k-1), \ldots, \pi(N)+(k-1)\},$$

where $1 < k$ and π is a Gilbreath Permutation. For example, the sequence $\{5, 6, 4, 7, 8, 3\}$ is the Gilbreath Permutation $\{3, 4, 2, 5, 6, 1\}$ with 2 added to all of the entries. These can be turned into simple continued fractions just as easily as with the ordinary Gilbreath Permutation. There are, however, some differences. The second way to turn a Generalized Gilbreath Permutation into an infinite string of numbers is to append the as yet unused natural numbers in such a way that the "Gilbreath-iness" of the string is maintained (that is; Theorem 1 is not violated). Given the Gilbreath Permutation $\{\pi(1), \pi(2), \ldots, \pi(N)\}$ either the number

$$\min\{\pi(1), \pi(2), \ldots, \pi(N)\} - 1$$

or

$$\max\{\pi(1), \pi(2), \ldots, \pi(N)\} + 1$$

can be put at the end of the string to maintain a Gilbreath Permutation.

When $k = 1$ the extension of a finite Gilbreath Permutation to an infinite Gilbreath Permutation has only one solution. This is not true for a Generalized Gilbreath Permutation since any or all of the numbers $\{1, 2, 3, \ldots, k-1\}$ can be placed in the permutation after the Nth entry (or course, not in just any order). As a quick example, the finite permutation $\{4, 3, 5, 2, 6\}$ can be extended to $\{4, 3, 5, 2, 6, 7, 8, 9, \ldots\}$ or $\{4, 3, 5, 2, 6, 1, 7, 8, 9, \ldots\}$ or $\{4, 3, 5, 2, 6, 7, 1, 8, 9, \ldots\}$ or $\{4, 3, 5, 2, 6, 7, 8, 1, 9, \ldots\}$ and many others.

Let us look at how this changes things with generalized, infinite Gilbreath Continued Fractions. Suppose x is the number given by

$$x = [0; 4, 3, 5, 2, 6, 7, 8, 9, \ldots] = 0 + \cfrac{1}{4 + \cfrac{1}{3 + \cfrac{1}{5 + \ddots}}}.$$

Since we can slip in a 1 at any space after the 2 and NOT lose the "Gilbreath-iness" of the situation we get a sequence of numbers

$$\begin{aligned} y_1 &= [0; 4, 3, 5, 2, 6, 1, 7, 8, 9, \ldots], \\ y_2 &= [0; 4, 3, 5, 2, 6, 7, 1, 8, 9, \ldots], \\ y_3 &= [0; 4, 3, 5, 2, 6, 7, 8, 1, 9, \ldots], \\ &\vdots \end{aligned}$$

of generalized, infinite Gilbreath Continued Fractions that converge to x. Hence unlike their counterpart, these points are not isolated points, but limit points. Although in this example the y_n's use all the natural numbers in their continued fraction form, this is not necessary. If x had been missing the values 1 and 2, then the 2 could have be snuck in and the 1 left out and still we would have a converging sequence.

This can become more complicated if even more terms, 1, 2, ..., $k-1$, are missing. Let

$$x = [0; 5, 6, 4, 7, 3, 2, 1, 8, 9, 10, 11, \ldots].$$

Create the new value

$$y = [0; 5, 6, 4, 7, 8, 9, 10 \ldots].$$

This is x with S, the set of missing numbers, equal to $\{3, 2, 1\}$, the values between the 8 (the k where where $a_j = j$ for all $j \geq k$) and 7 (that is $k-1$). Some subsets of S, with order intact, can be slipped in to the right of the 4 (because this now straightens out from the 7 onward to create Gilbreath Continued Fractions that converge to y). These subsets are $\{3\}$, $\{3, 2\}$, and $\{3, 2, 1\}$.

4 Other Gilbreath Numbers

We can create numbers another way by starting with the Gilbreath Permutation $\{3, 4, 2, 1\}$ and placing the numbers, in order, as the digits in the (finite) decimal expansion, 0.3421. We shall call such a number a *finite Gilbreath Decimal*.

There is the "cut-and-paste" method. This way $\{3, 4, 2, 1\}$ can become the decimal

$$0.3421342134213421\ldots.$$

This, by virtue of being a repeated decimal, is really a rational number which is equal to

$$\frac{3421}{9999} = \frac{311}{999}.$$

Admittedly, there seems to be nothing exciting there. Cut-and-paste with simple continued fractions gives us

$$[0; 3, 4, 2, 1, 3, 4, 2, 1, 3, 4, 2, 1, \ldots].$$

This gives us something a little more interesting than the rational number version. Recall that continued fractions with repeated patterns are quadratic irrationals.

Example 3 *Turning our attention to our repeating Gilbreath Continued Fraction* $x = [0; 3, 4, 2, 1, 3, 4, 2, 1, 3, 4, \ldots]$ *becomes the equations*

$$x = 0 + \cfrac{1}{3 + \cfrac{1}{4 + \cfrac{1}{2 + \cfrac{1}{1 + x}}}}.$$

This simplifies to the quadratic equation

$$29x^2 + 33x - 13 = 0$$

whose (positive) solution is

$$x = \frac{-33 + \sqrt{2597}}{58} \approx 0.309668434913 \ldots.$$

It is easy to generalize the repeating decimal representation:

$$x = 0.\pi(1)\pi(2) \ldots \pi(N)\pi(1)\pi(2) \ldots \pi(N) \ldots$$

has rational number representation

$$x = \frac{\pi(1)\pi(2) \ldots \pi(N)}{10^N - 1}.$$

Things are much uglier with repeating Gilbreath Continued Fractions. As a general example, $x = [0; a, b, c, d]$ leads us to the equation

$$x = \frac{bcd + bcx + b + d + x}{abcd + abcx + ab + ad + cd + ax + cx + 1}$$

and an even worse looking solution.

In decimal form, the "limit" of this repeating continued fraction is the irrational number $0.435216789101112 \ldots$, which is reminiscent of the better-known number

$$0.12345678910111213 \ldots.$$

Lastly, in lots of places (e.g. [12]) one finds that $[0; 1, 2, 3, 4, 5, \ldots]$ is a *transcendental* number. Hence the infinite Gilbreath Continued Fractions (regular and generalized) must also be transcendental. The number $[0; 1, 2, 3, 4, 5, \ldots]$ is related to Bessel Functions[2]. The nth Bessel Function $J_n(x)$ is the solution to the differential equation

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - n^2) = 0.$$

2. Thanks to Stephen Lucas of James Madison University for pointing this out to us.

One of the myriad of relations involving Bessel Functions is that

$$[0; 1, 2, 3, 4, \ldots] = -i\,\frac{J_1(2i)}{J_0(2i)} = 0.697774657964\ldots,$$

where $i = \sqrt{-1}$.

References

[1] A. Bruckner, J. Bruckner, B. Thomson, *Real Analysis*, Prentice-Hall, 1997.

[2] P. Diaconis and R. Graham, *Magical Mathematics: The Mathematical Ideas That Animate Great Magic Tricks*, Princeton U. Press, 2011.

[3] C. Freiling and B.S. Thomson, *Scattered sets and gauges*, Real Anal. Exchange, 21(1995/96), no. 2, p. 701-707.

[4] M. Gardner, *Mathematical Magic Show*, MAA, 1989.

[5] M. Gardner, *New Mathematical Diversions*, MAA, 2009.

[6] N. Gilbreath, *Magnetic colors*, The Linking Ring, 38(1958), no. 5, p. 60.

[7] C. Mulcahy, *The first Norman invasion*, Card Colm, MAA Online, August 2005.

[8] C. Mulcahy, *The second Norman invasion*, Card Colm, MAA Online, August 2006.

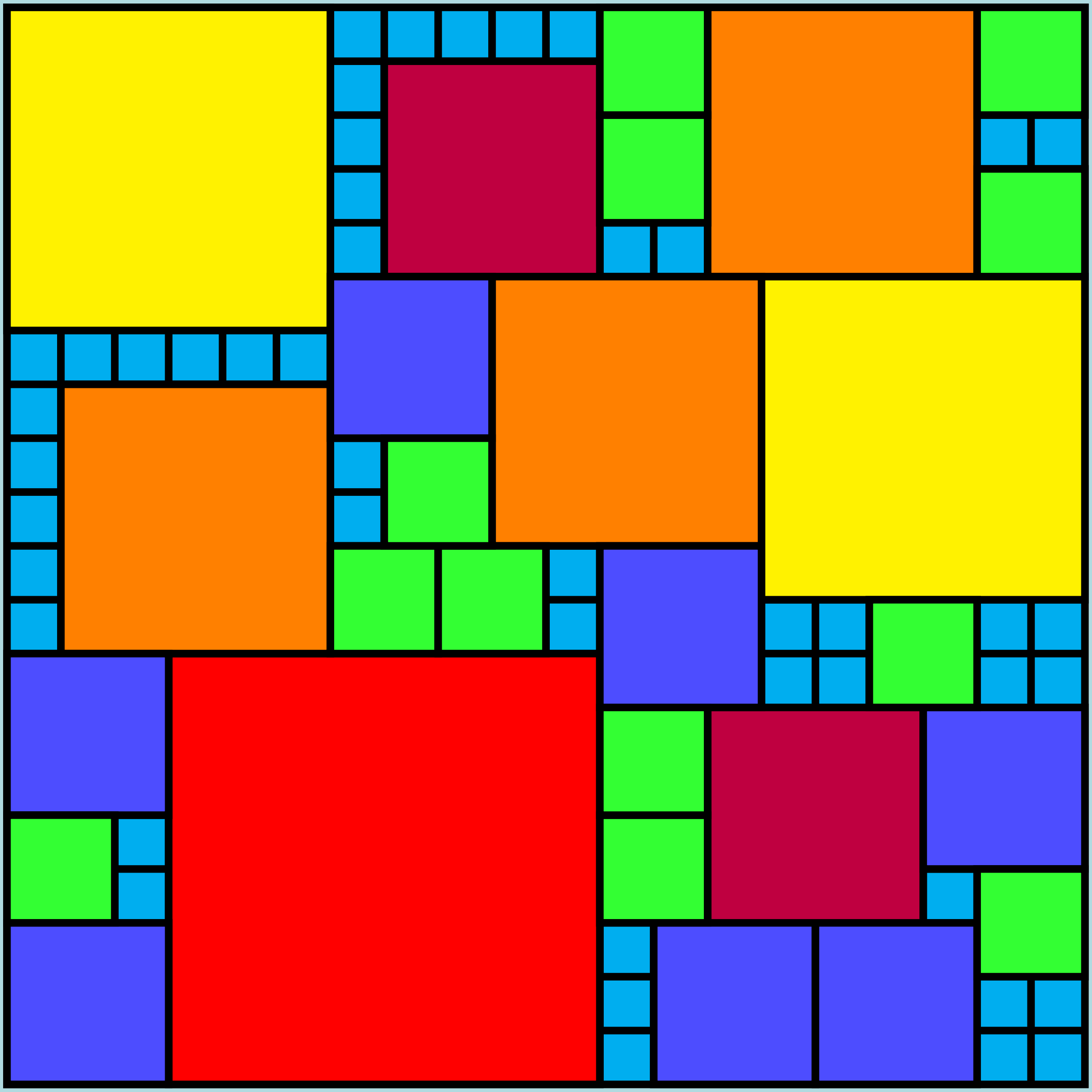
[9] C. Mulcahy, *The bilgreath principle*, Card Colm, MAA Online, August 2009.

[10] C. Mulcahy, *Gilbreath shuffling*, Card Colm, MAA Online, August 2012.

[11] C.D. Olds, *Continued Fractions*, MAA New Mathematical Library, 1963.

[12] D. Wells, *The Penguin Dictionary of Curious and Interesting Numbers*, Penguin Books, 1986.

[13] M. Wolf, *Continued fractions constructed from prime numbers*, `arxiv.org/pdf/1003.4015.pdf`.



MATH

Tiling Tetris Boards | Steve Butler, Jason Ekstrand & Steven Osborne | Page 190

Being Paid to Eat Pizza: Negative Prices in Supermarkets

Adam Atkinson (ghira@mistral.co.uk), G4G11

Some supermarkets in the UK (specifically Waitrose and Somerfield that I've noticed) occasionally have items with negative prices. I'd be interested to hear of other places where this happens and of more dramatic examples than I have seen. I'd also be interested to hear of places where this has been prevented, and how, and of possible amusing side-effects of the prevention.

Here is an extract from a receipt from my local branch of Waitrose (on 28/11/13):

WR PIRI PIRI HOUMOUS 29p
WR TOMATO HOUMOUS 29p
** WR HOUMOUS 2 for £3 ** -0.70

We see that I am charged 29p each for two tubs of houmous, then 70p is taken off because there is a "2 for £3" offer on houmous. As a result, I am paid 12p to take 2 tubs of houmous away. I spell it this way because the receipt does.

The first time I noticed this I bought 2 pizzas for 10p each and since they were "2 for 1" the receipt had 1 pound off. Result: I was paid 80p to take two frozen mushroom pizzas away.

It may seem that the above calculations make no sense. The "because" and "since" don't work. What is actually happening is that the items in question were discounted because they had reached their "sell-by" date, and they were also in "2 for 1" or "2 for £x" type offers. The way the offers are processed seems to be that if pizza is £1 and there is a 2-for-1 offer on it, you get £1 subtracted from your bill if you buy two pizzas. Similarly, if houmous is £1.85 but there is a "2 for £3" offer, when you buy two tubs 70p is taken off because the normal price of two tubs of houmous would have been £3.70. These amounts are subtracted even if the base price of the items has changed.

I have never dared to purchase only negatively priced items on a trip to the supermarket, but I have received email from someone who says he tried this and he was given his goods and some money at the checkout. I shall try this with an automated checkout rather than a staffed one the next chance I get and see what it does.

A possible new source of confusion is loyalty cards. The Waitrose card, for example, gives 10% off some items. I am keen to see how this combines with combo offers, discounts due to "sell by date" arrival or indeed both together. It could be that I get 10% off a negative price, in which case I would be best advised to go to the checkout once for positively priced items and once for negatively priced items, not using my loyalty card on the second visit. Or it could be that the 10% discount will apply to the undiscounted price. Considering how discounts seem to work in general, I think the last possibility is the most likely. Between getting my card and G4G11 no research opportunity has presented itself.

It seems unlikely that any of this matters. The quantities involved are always very small, and it may well not be worth a programmer spending any time at all "fixing" this. It may even be that it's cheaper for the supermarket to pay me 40p to remove a pizza instead of having staff throw it away 30 minutes later. It could be that I am being allowed to think I am getting away with something to encourage me to visit the supermarket and buy other things. I'd love to hear from someone with solid info on this.

Coupons less simple than "40p off" might also interact in surprising ways with all the above.

Telephone calls and the Brontosaurus

Adam Atkinson (ghira@mistral.co.uk)

This article provides more detail than my talk at G4G11 with the same title.

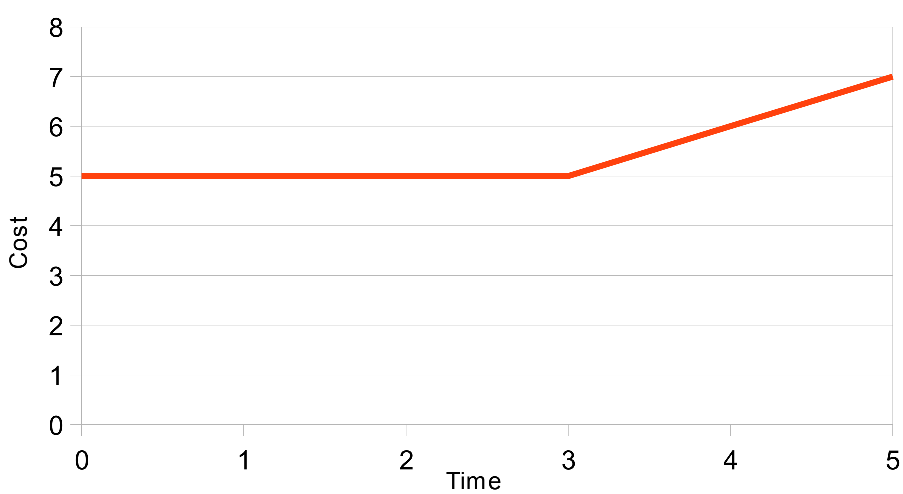
I am occasionally asked questions along the lines of "When do you ever use any of this stuff in real life?" or "What is the hardest mathematics you have ever used in real life?" I imagine other G4G people have had similar experiences.

Recently, I've found myself answering both these questions with an example which uses high-school level maths at most and thus should be relatively accessible.

I was working for a company which installed/maintained internal telephone systems for organizations of various sizes, including the links between these systems and the outside world. Note that what I did at this company did not involve telephone systems, so during the events of this story the whole situation was new to me. What I did may not necessarily reflect best practice on the part of people who really do this sort of thing for a living, but as with spherical cattle or drunks and street lights, we might be willing to sacrifice some accuracy/plausibility for the sake of creating a more accessible exercise. Also, and principally, I don't want readers to think that anything I say here reflects that company's actual approach to a problem like this.

I was approached by a manager and told that one of our customers felt that we were charging too much for phone calls. The costs worked like this: incoming calls and internal calls were free. Only calls to the outside world cost money, so from now on we shall only considering outgoing calls. And we shall only only consider standard outgoing calls (to national numbers, not international ones for example). Each call had an initial cost (some fixed amount of money) as soon as it was connected, and if it lasted at most a certain duration there was no extra charge. If the call lasted longer than that, there would be an additional charge (at some fixed rate) per unit of time over this length.

For the sake of argument, let's say that our price was 5p for any call up to 3 minutes, then 1p a minute after that.



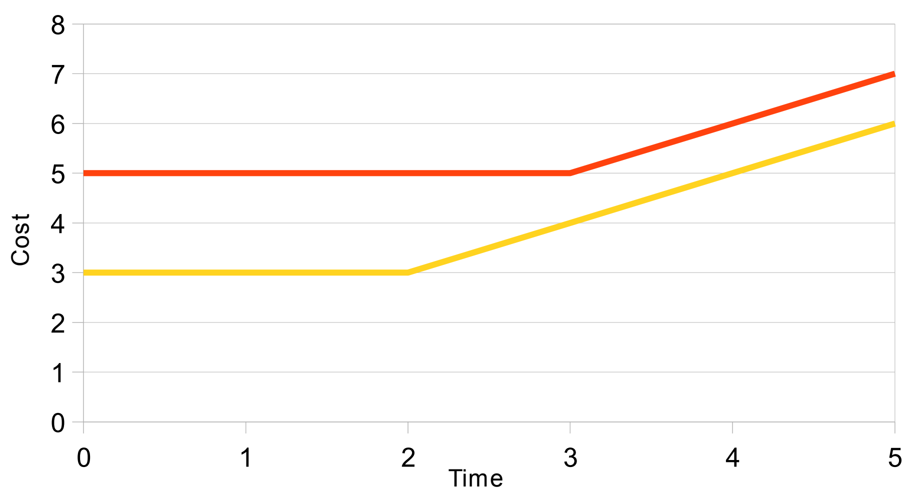
The customer had 5 quotes from our competitors, all expressed similarly: A for the first B minutes, then C per minute after that. The customer felt that our tariff was "clearly" too expensive. It was not "clear" to us why it was, so we asked, and were told that our value of C was too high. I was asked by my side if this was reasonable, and how I would compare such tariffs. As it happens, our value of A was smaller than some of the others, so I asked if the customer made enough "long" calls for our high value of C to cause a problem. It turns out neither side really knew, but since we ran the phone system it was, of course, perfectly possible to get a log of all calls going back months to look at this kind of thing.

With a list of call lengths, we can compare tariffs by seeing what each tariff would expect one to pay for that set of calls. However, one might not have such a log, or it might be so short that it might not be considered to be representative. Or the customer might fear that over time the length of calls might change.

Can we make any attempt at all to compare tariffs without a huge log of calls? I am asking this rhetorically, so clearly the answer must be "yes".

For starters, if C is too large then the cheapest way to make a very long call would be to hang up and re-dial every B seconds. This would be an annoying thing to have to do but one could imagine some people going to this much trouble. Certainly if modems were still a thing and costs were like this I would expect people to arrange for their modems to behave in this manner. Let's assume that even if some of our tariffs are like this, real people are not going to bother to redial all the time and will pay the tariff rate.

Any easy comparison is one that looks like this:

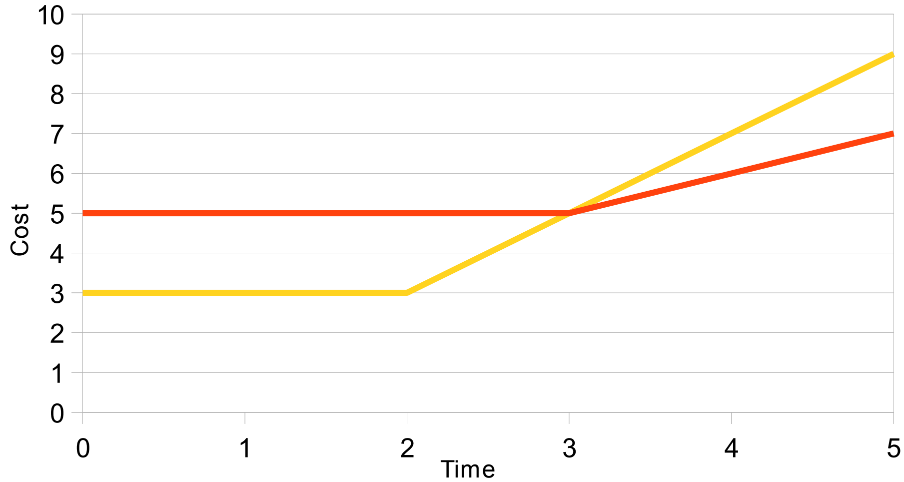


Clearly as the customer we would choose the yellow tariff here.

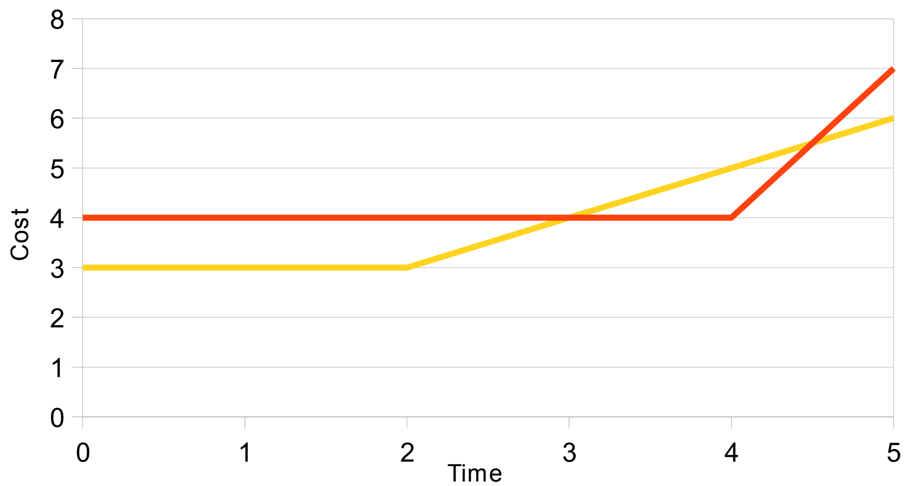
Let's assume any tariff which, like the red one here, is totally undercut by some other tariff is removed from further consideration.

Two other things can happen although the difference between them probably doesn't matter much.

In the first case one tariff could be cheaper for short calls and the other for longer calls:



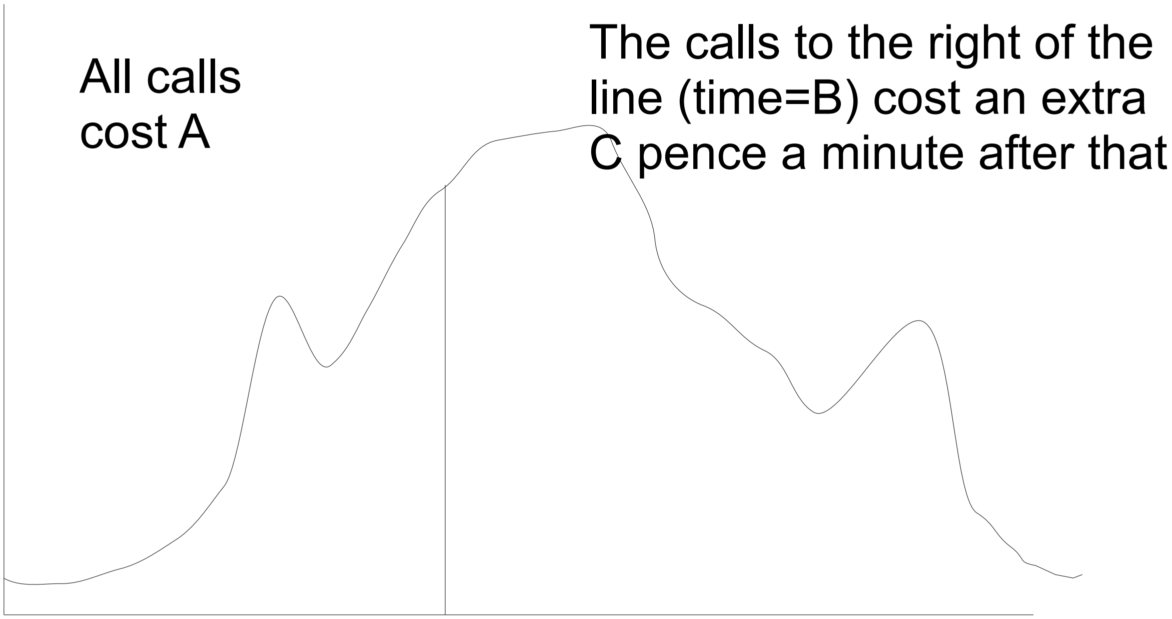
In the second case, one tariff could be better for medium length calls only.



In these cases, you need to know how many calls of each length the customer makes.

While in principle the probability distribution of call lengths could look like almost anything, perhaps in real life it can be treated as coming from some family with a small number of parameters.

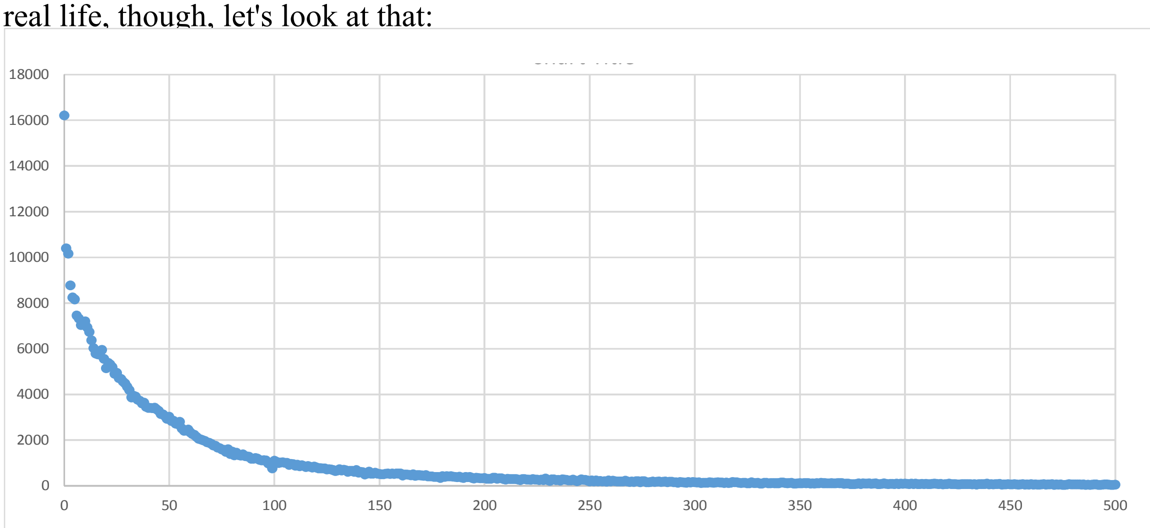
One might suppose that the probability distribution of phone calls might be brontosaurus-like. As A. Elk put it, the brontosaurus is thin at one end, much much thicker in the middle, and thin again at the far end:



Or perhaps more of a stegosaurus (as seen in the extra humps above) to make it less "nice" than a standard bell shape. I think we can all agree that telephone calls can't have negative length, but it might seem plausible that there's some common medium length and calls longer and shorter than that are less common.

If we know what this graph looks like then we can calculate the average cost of a call as:
A + C* Prob(Call lasts at least B) * (Mean additional duration of calls which last at least B),
since all calls get charged A immediately, then some get charged more. If we have a nice formula for our probability distribution we can turn this into something with integrals in it.

Calculating the mean additional duration is going to be possibly quite annoying. Wouldn't it be nice if it weren't annoying? The "spherical cows" assumption at this point is that the distribution of call lengths is exponential, because then the average additional duration of calls of length at least B is the same as the mean of the distribution as a whole. Since I have information about hundreds of thousands of calls made in real life, though, let's look at that:



Which is nothing like a brontosaurus at all. At first glace, this *does* look more like an exponential distribution. Indeed, the distribution sometimes used in exercises about this kind of thing is the exponential, which has only one parameter, so if you know the mean you know everything you need to. I already knew this when I started this exercise but had often wondered if this was mainly because it was the easiest distribution to do calculations with. Of course, to find out a customer's mean call length you need some call logs and if you have those you could run the calculations based on those as mentioned earlier.

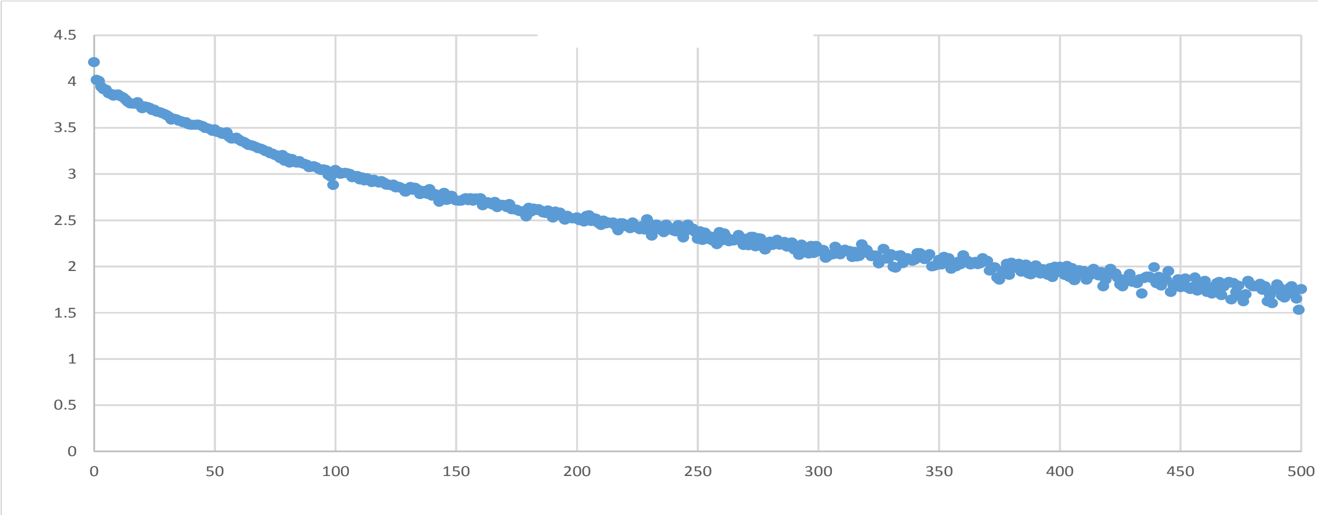
This graph is taken partway through cleaning up the call logs, and it seems possible that some information about the "cleaning" might be of interest.

A similar histogram made from the raw logs had a peak at 1 second rather than 0 seconds, which would have ruined the "it's exponential!" impression. I thought maybe this was because calls from 0.5 to 1.5 seconds were being called 1 second calls, so perhaps the 0 second calls had a very narrow time range. Actually it was stranger than that. In the call logs, if a call started in one calendar second and finished in the next one, it was called a one second call, even if it was actually, say, 0.1 seconds long, and this handling of "calendar seconds" clearly pushed many calls into seeming a substantial fraction of a second longer than they really were. Of course, for most purposes an error of under 1 second wouldn't matter, but we're worrying pointlessly about the relative heights of 0 and 1 second columns on a histogram here so let's try to fix this.

Fortunately, the call logs in the raw data used for this graph also contained information from which the length of the call in 50ths of a second could be deduced, and it is using that rather than the "duration" column that the above histogram was produced. Incomplete seconds are rounded down, so 0 to 49 50ths of a second count as 0 seconds, etc.

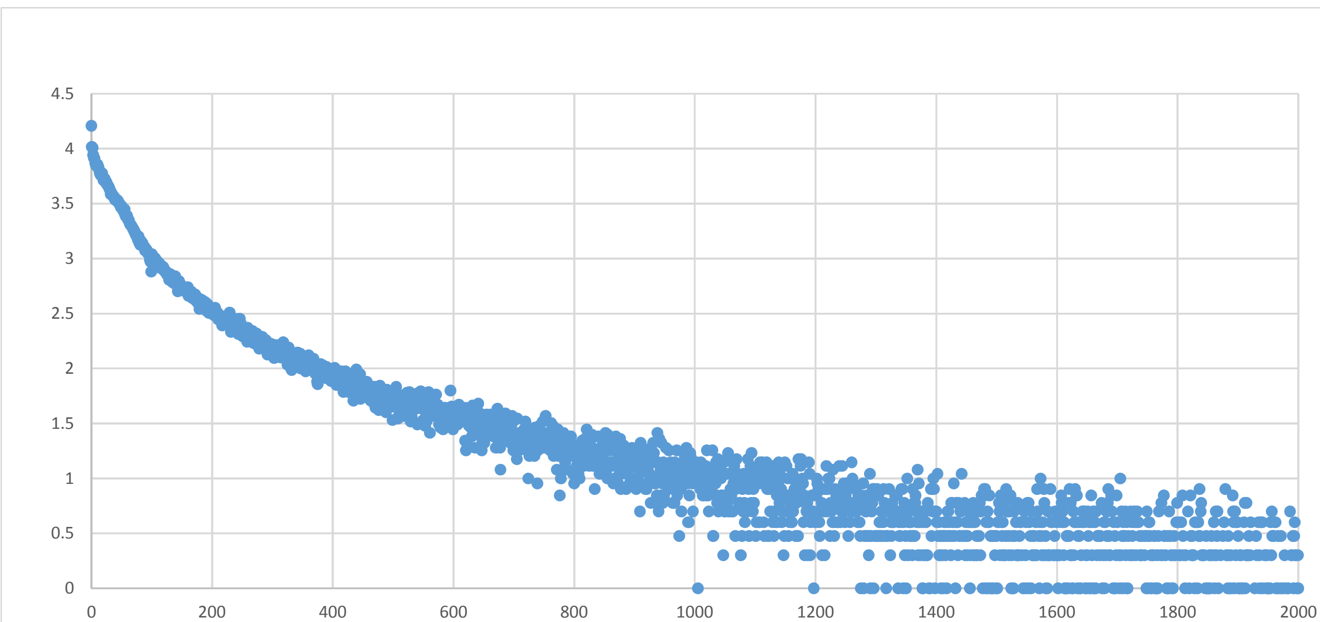
Unfortunately, the 0 second column is now incredibly tall. Can we find an excuse for making it shorter again somehow? Well, yes. In the call logs from the original story I was given a log only of outgoing calls, but I used a log of all calls to make this graph. I ought to remove internal and incoming calls from it, and one particular class of incoming call that shows up as being 0 seconds long is a call which is diverted automatically to voicemail. For some reason, such calls show up as a 0 second call to the phone followed by a real call to the voicemail system. Since these calls are incoming, they should be eliminated along with internal calls, international calls, calls to freephone numbers and so on.

Of course, I have not shown that the graph above really is exponential, merely that it looks closer to an exponential than to a brontosaurus. Using a log scale on the y axis would be the sensible thing to do here. As one might fear, this doesn't look like a straight line.



Actually I only cut off the x values at 500 seconds because the graph becomes quite noisy after that. It really continues to 45000 seconds. (The y axis here is a log scale even though it doesn't explicitly say so)

Here's a slightly wider version:



(Log scale on the y axis again) So it would seem that "Not really exponential, but closer to that than to a brontosaurus" is about the best we can do. I think the original pricing query, with an assumed exponential length distribution, could be used as an exercise in some context or other. And it is at least understandable why the exponential is found in exercises on this kind of topic.

I met Nancy while helping out at one of these festivals. They're an exciting way to play with math. —SV

The Julia Robinson Mathematics Festival

by Nancy Blachman

It was March, 1972. My tenth-grade math teacher, Mr. Forthoffer, was handing out another problem set from Saint Mary's College. We didn't have to do them, so some kids just left them on their desks. I had so much fun with the ones we'd gotten earlier in the year that I couldn't wait to start working on these. The first few problems were usually pretty straightforward, and solving them would boost my confidence for tackling others. I never could solve all of them, but I enjoyed trying.

I liked experimenting—cutting a 10x10 square into two pieces to make a rectangle, plugging values into equations, learning more about the problems as I worked. Even after I solved a problem, I liked thinking about whether there was an easier way to solve it.

Students who scored high on that year's qualifying problems would be invited to the Saint Mary's Math Contest at the end of the school year. Schools all over the San Francisco Bay Area sent busloads of students. I didn't actually care much about going. For me, the fun was in the problems I was doing at home.

I went to the contest that year and the next two, but it was a disappointment. Sitting alone in that room all day, without being able to discuss my ideas with my father—it wasn't nearly as much fun as the problems I got from school. To this day I remember them with deep fondness.

When I teamed up with Josh Zucker and the Mathematical Sciences Research Institute (MSRI) in 2006 to create something new, I was determined to bring back the best of the now-defunct Saint Mary's Math Contest, and leave behind what didn't work. We decided to emphasize fun, creating a mathematics festival instead of another competition. The festival would have dozens of tables with mathematical problems,

"We decided to emphasize fun, creating a mathematics festival instead of another competition."

puzzles, games, and activities, each with a facilitator to help students stay connected with the math.

We wanted the festival to nurture students, so we let them work individually or in groups. We hoped to attract students at a wide range of abilities with math at all levels, so we chose problems and activities that would connect to one another. We started each set with simple problems everyone could work out, leading to progressively more difficult questions (we even included unsolved research problems). We hoped to have so many problems that not even a mathematical genius could solve all of them during the festival.

The festival needed a name to match its spirit. Julia Robinson was a great mathematician who, along with two other mathematicians, was renowned for solving Hilbert's tenth problem. She lived not far from us in the San Francisco Bay area, and was a distinguished mathematics professor at the University of California at Berkeley for many years, until her death in 1985. It felt perfect to honor her legacy with this festival.

In March 2007, the first year we ran the festival, we were concerned that we might not get many students to sign up, but within a few weeks the festival was oversubscribed. With more registrants than space, we asked our sponsor, Google, for a tent to accommodate more students. They came through, and the day of the festival started out sunny and chaotic.

It all fell into place, with hundreds of students eagerly approaching the problems we'd devised. There were thirty tables with activities, puzzles, games, and problems. When we announced that sandwiches were available for lunch, many of the kids would not stop working. We may not have managed to feed their bodies, but we surely fed their minds! The prizes from Google were icing on the cake.

The response was so enthusiastic that we've been able to make the festival an annual tradition. And we've grown, offering festivals in over a dozen locations—California (eight different locations), Connecticut, Washington D.C., Michigan, Texas, North Carolina, Arizona, Virginia, Washington (state), and Wyoming.

My goal was to inspire, delight, and challenge children, as the Saint Mary's Math Contest did for me, but with more collaboration and less competition. Thanks to its many sponsors and volunteers, the Julia Robinson Mathematics Festival is a success, and I've seen my dream come true.*

*You can find more information at: **jrmf.org**

PUZZLE

St. Mary's Math Contest Sampler

Adapted from problems by Brother Alfred Brousseau

1. A company named JULIA has an advertising display with just the five letters of its name, lit up in various colors. On a certain day the colors might be red, green, green, blue, red. The company wishes to have a different color scheme for each of the 365 days of the year. What is the minimum number of colors that can be used for this purpose?

2. How would you decide whether a number in base 7 is even, based on its digits?

3. Given the sequence 1, 2, 4, 5, 7, 8, 10, ... where every third integer is missing, find the sum of the first hundred terms in the sequence.

4. Find the sum of the cubes of the numbers from 1 to 13. Now find the sum of the cubes of the numbers from 1 to n.

5. Using exactly five 5's, and the operations +, −, ×, ÷, and factorial (!), represent each of the numbers up to 30.

EXPLORATIONS

Candy Conundrum*

by Joshua Zucker

Julia Robinson Mathematics Festival

Colors

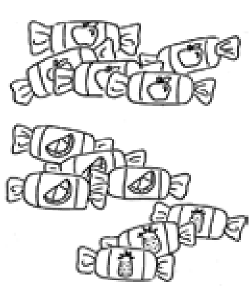
1. Some years ago, a candy company advertised the large number of flavors that could be made by mixing their candies in your mouth. How many are there really?

2. You have 5 red apple candies. How many different nonempty sets of candies could you put in your mouth?

3. You have 5 red apple and 4 green lime candies. How many different nonempty sets of candies could you put in your mouth?

4. You have 5 red apple, 4 green lime, and 3 yellow pineapple candies. How many different nonempty sets of candies could you put in your mouth?

Flavors



We'll consider two sets of candies to be the same flavor if the ratio of candies of each color in one set is the same as the ratio in the other. For instance, 2 green and 1 yellow is the same flavor as 4 green and 2 yellow; it's 2/3 green, or a 2:1 green:yellow ratio. Similarly 3 red is the same flavor as 2 red: pure red!

5. You have 5 red candies. How many flavors could you make?

6. You have 5 red and 4 green candies. How many flavors could you make?

7. You have 5 red, 4 green, and 3 yellow candies. How many flavors could you make?

---

*Find more problems at: **jrmf.org**

Geometry

There's a geometric interpretation of all of the above. For instance, with 5 red and 4 green candies, the possible combinations are ordered pairs. (2,1) and (4,2) are the same flavor, for example.

8. Expressed geometrically, what does it mean for two different sets of candies to be the same flavor? Assume for now that there are only two colors.

9. Describe a geometric way of understanding how many different flavors there are. Compare it to the numeric approach.

10. What does symmetry tell you about the number of flavors with k red candies and k green candies?

Generalizing

Now let's try for some bigger patterns.

11. If you have 1 candy of each of n colors, how many different flavors are possible? If you have k candies of 1 color, how many different flavors are possible? OK, sorry, that was too easy.

12. If you have 2 candies of each of n colors, how many different flavors are possible? If you have k candies of each of 2 colors, how many different flavors are possible?

13. Generalize as much as you can!

14. How do the previous answers change if the candies are large, with an upper limit to how many fit in your mouth at once?

15. What can you say about the relative probability of various flavors if you pick a random handful of size n out of a set of candies? Start by considering some easy cases, where n is small, and there aren't too many different flavors, and plenty of candies of each flavor (since n will limit you, it gets more complicated if you also have limits due to running out of candies).

16. What other questions can you think of about how to count combinations of candies?

Tiling Tetris Boards

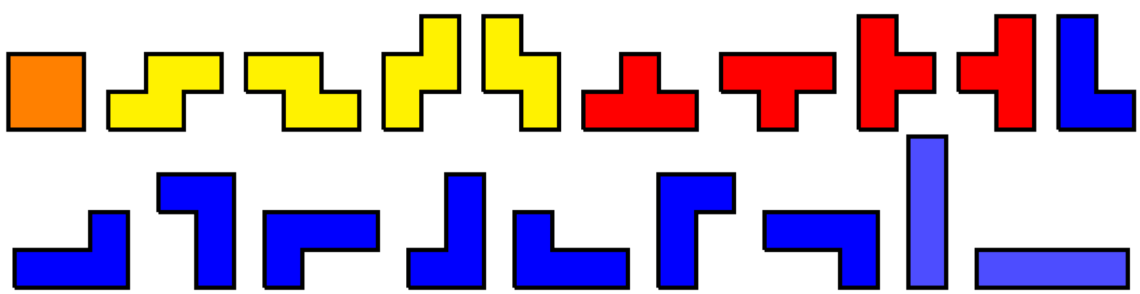
Steve Butler[*] Jason Ekstrand Steven Osborne

Introduction

One of the most popular topics in recreational mathematics are questions related to tiling. These come in various shapes and forms from Penrose tiles, to Golomb's polyominoes, to Marjorie Rice's discovery of pentagons that tile the plane (which she was motivated to find after reading Martin's column in Scientific American).

One of the simplest classes of polyominoes are the tetronimoes, i.e., made from four squares. These have become well known through the game Tetris, wherein these pieces continue to fall from the top into a 10×20 field and the players must arrange them so that whole rows are filled up (thus freeing up space as pieces continue to fall). While the game Tetris has been well studied, for example Erik Demaine has recently shown that the game is hard, the question of how many ways there are to tile the 10×20 board using Tetris pieces has not previously been considered.

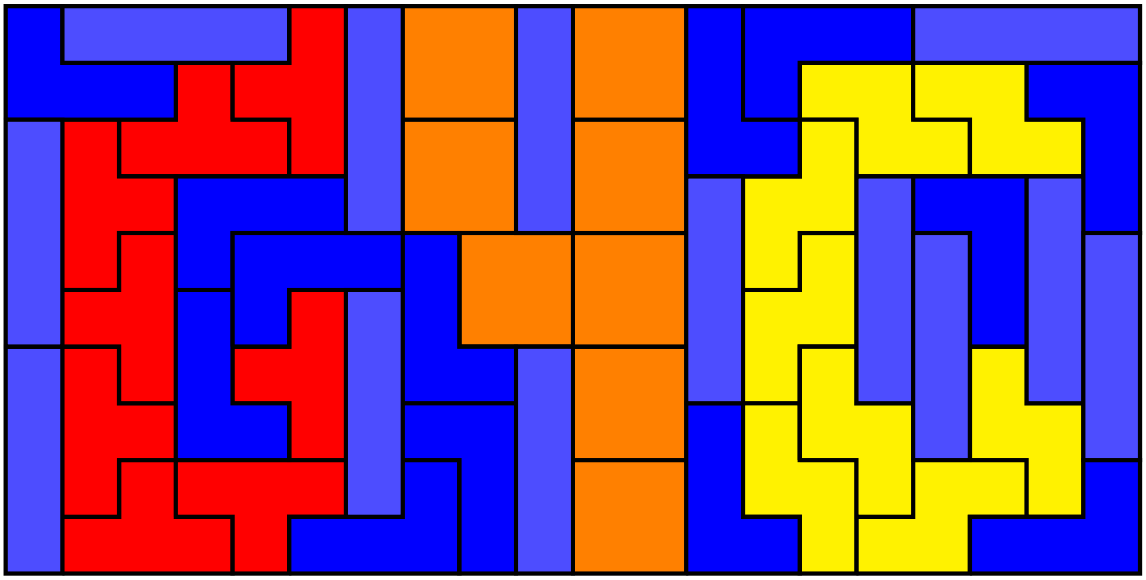
To be more precise we want to consider how many different ways that there are to tile a 10×20 board using the following possible sets of pieces (i.e., all possible Tetris pieces in all possible orientations):



We want to cover the board using some or all of these pieces. For instance we can easily cover the board using 50 of the 2×2 Tetris piece, so we know the number of possible ways to tile is at least one. Below we give another tiling that bears

[*]Iowa State University, Ames, IA 50010. butler@iastate.edu

some passing connection to this conference, showing that the number of ways is at least two.



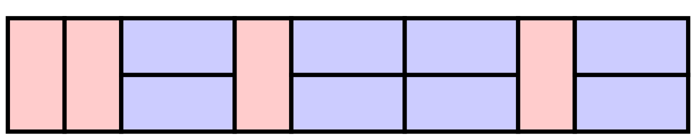
In fact the number of ways to tile is much more than 2. To be precise the number of ways to tile is

$$291,053,238,120,184,913,211,835,376,456,587,574.$$

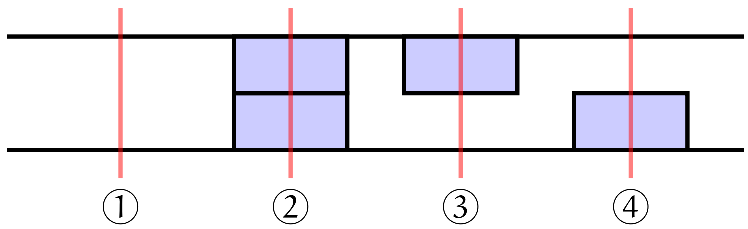
In other words, a little bit more than 291 decillion. This is a ridiculously large number, so large that all of the world's supercomputers working together for a billion years could barely begin to make a complete list of all of these. Nevertheless, we know exactly how many ways there are to tile, and this uses some simple ideas and a little bit of work with linear algebra (matrices). We will outline the basic approach by a simpler example involving tiling the $2 \times n$ board with dominoes.

Counting tilings on a $2 \times n$ board with dominoes

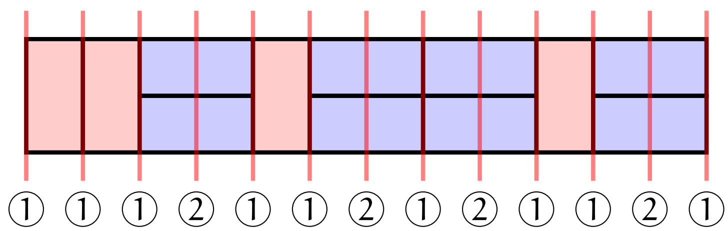
We consider a classical problem. Namely, the number of ways to tile the $2 \times n$ board using dominoes which can be either horizontal ▭ or vertical ▯. An example of this is shown below for a 2×12 board.



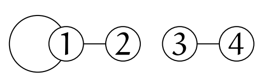
The usual way that people approach counting these is to find a recurrence relationship and then work with that recurrence relationship, ultimately leading to the Fibonacci numbers. We will take a slightly different approach and start by focusing on what happens in the transitions between columns (this can be thought of as sort of the zen approach to counting tiling, wherein it is not the tiles that are important but what happens between the tiles). To be precise there are four possible ways we can cover a transition between columns and these are drawn and labelled below.



If we take these labellings and apply them to the tiling given above we get the following:



The next thing to note is that there are limitations on what order things can occur. For instance a ① can either be followed by another ① (i.e., a vertical domino in between the two columns) or by a ②; but it can't be followed by a ③ or a ④ since that would result in uncovered parts of the board. We can represent our situation by drawing a graph wherein we connect any two crossings which can occur consecutively.



Finally, note that we always start and end our tiling with a ①, the tiling then corresponds to moving in this graph from vertex to vertex along edges connecting legal crossings. In fact, for every such way there is to start at ① then take n steps and have returned back to ① there is a tiling and vice-verse. Such a sequence of moves in the graph is called a walk that starts and stops at ①. So we now have the following.

Observation. *The number of tilings of our board is equal to the number of walks of appropriate length that start and stop at* ① *in this associated graph.*

We have transformed our problem from counting tilings to counting walks. The great news is that there is a mathematical tool which is built for counting walks. Namely, matrices and matrix multiplication. So first we can observe that ③ and ④ can never be reached so we only need to worry about ① and ②. Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

We are indexing this matrix by letting the first row and column correspond to ① and the second row and column corresponding to ②. This matrix keeps track of where we are allowed to move, i.e., we can move from ② to ① but we cannot move from ② back to ②. This is known as the *adjacency matrix*. Another way to view this matrix, is that it records how many ways we can get from one node to another node in the graph by taking one step.

Suppose

$$B_k = \begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} \end{pmatrix}$$

is such that $b_{ij}^{(k)}$ corresponds to the number of walks of length k that start at i and stops at j (so in terms of our problem we want the b_{11} entry in the matrix B_n). We have already noted that $B_1 = A$.

The next thing to observe is that the walks of length $k+1$ are closely connected to the walks of length k. For example if we want to end at ② then at the previous step we had to be at ①; on the other hand if we wanted to end at ① then at the previous step we could have been at either ① or ②. So this gives us the following:

$$B_{k+1} = \begin{pmatrix} b_{11}^{(k+1)} & b_{12}^{(k+1)} \\ b_{21}^{(k+1)} & b_{22}^{(k+1)} \end{pmatrix} = \begin{pmatrix} b_{11}^{(k)} + b_{12}^{(k)} & b_{11}^{(k)} \\ b_{21}^{(k)} + b_{22}^{(k)} & b_{21}^{(k)} \end{pmatrix} = \underbrace{\begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}}_{\text{Matrix multiplication}} = B_k A$$

In particular, we can now see us that $B_k = A^k$. So we have the following.

Observation. *The number of tilings of our board is equal to an entry in* A^n *where* A *is the adjacency matrix of our associated graph.*

So this means that counting walks in the end comes down to some simple matrix algebra.

The advantage of this approach is that as we multiply matrices together we don't have to keep track of the past history of all of our walks, i.e., this allows us to enumerate without having to individually generate each one. This is what allows for these ridiculously huge numbers to be computed.

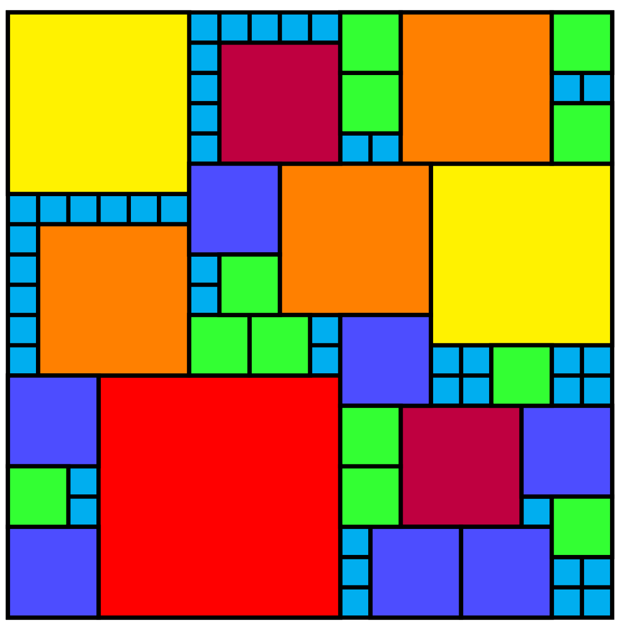
Counting tilings on the Tetris board with Tetris pieces

The count for tilings on the Tetris board is based on the same principles of turning it into a matrix multiplication problem. Namely, we started by looking at all the possible ways that Tetris pieces could cross column transitions of height 10. It turns out that there are about half a billion such crossings, and then we figured out which can occur consecutively. This gave us the graph which in turn gave us the matrix A. This is a **huge** matrix, but it is also fairly sparse in that most column transitions cannot occur consecutively. This allows us to easily pull an entry out of the matrix A^{20}, which is how we derived our count.

This technique works for tilings of the board with any collection of polyominoes, i.e., dominoes, triominoes, tetronimoes, and so on. This has also been extended to other problems, including the counting the number of ways to subdivide a 20×20 square into smaller squares of whole length. There are a little over five sexdecillion ways to subdivide the square, or to be more precise:

$$5,107,719,132,342,188,248,003,492,889,317,730,068,733,587,913,765,138$$

An example of one of these is shown below.



This goes to show that a little bit of linear algebra can go a long way!

/dev/joe Crescents and Vortices

by Joseph DeVincentis

Introduction

Erich Friedman posed this problem on his Math Magic web site in March 2012[1]: Find the shape of largest area such that N of them, in total, fit into two different shapes. In one variation of the puzzle, the shapes were circles of diameters 2 and 3, and once N was sufficiently large, the solutions that people (including Maurizio Morandi, Jeremy Galvagni, Andrew Bayly, and myself) were finding tended to be concave shapes that nested together in two different ways to fit inside the two circles. I discovered an interesting series of shapes of this sort which I named */dev/joe crescents*. It appears that Bayly's solution for N=9 is in fact one type of /dev/joe crescent, but I generalized the type and built many more solutions on this model.

Since they are meant to fit into circles, it is natural that their edges consist entirely circular arcs. This makes them relatives of Reuleaux polygons. In Reuleaux polygons, these arcs are centered at the vertices, a property which makes them have constant width. In /dev/joe crescents, the arcs are not centered at the vertices, but are determined in quite a different way, which leads to some very odd looking shapes.

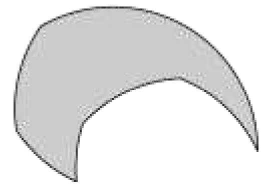


Fig. 1: A single /dev/joe crescent of type 1 for S=3, L=7

The governing rule that determines this shape is that they fit together, rotated with respect to one another, inside a circle. Although the shape above has 6 distinct arcs, I think of it as having 3 edges: the outside (the single arc which is much longer than the rest), the leading edge (the other 2 arcs on the convex side), and the trailing edge (the 3 arcs on the concave side).

In this shape, which I call a type 1 crescent, all of the arcs are from circles of radius 1. The three edges which appear to be the same size are in fact the same size. Four of the angles between arcs which (two inner and two outer) are the same. S=3 indicates that three of these fit inside a circle of diameter 2, so the outside edge is 1/3 of a circle.

One arc on the leading edge is the same length as two of the arcs on the trailing edge. That one arc can be aligned against either of the other two. One way, it takes S=3 pieces to go around a small circle with all of the outside edges forming that circle. The other way, L=7 indicates that it takes 7 of these wrap around and meet up with the first shape to make a circular pattern, with each outside edge tangent to the circle.

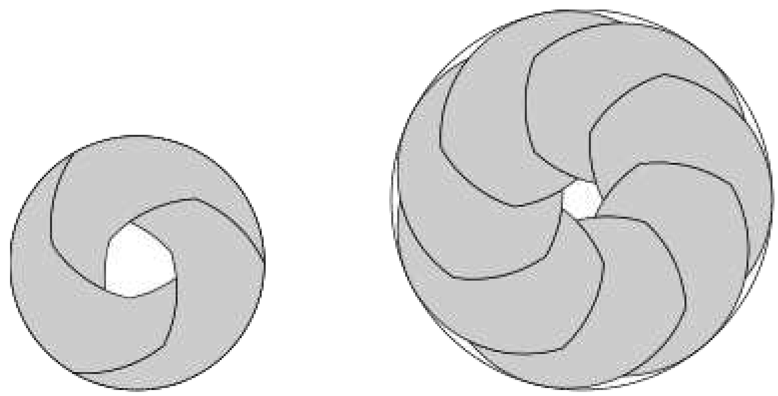


Fig. 2: /dev/joe vortex of type 1 for S=3, L=7

This creates two striking (and strikingly different) patterns. The small circle resembles a sort of vortex of fluid swirling into the central hole, and for this reason I gave the name */dev/joe vortex* to these circular patterns. There is also an optical illusion here which makes it possible to view it as a 3-dimensional shape, a bar of triangular cross-section which has been bent into a loop and twisted like a Möbius strip, although my impression of it is that it has a full twist, and so does not have the one-sidedness that comes with a Möbius strip. This illusion works better for some S,L pairs than for others.

In the large circle, the pattern looks like a set of rather oval dominoes which have been set up in a tight circle and toppled into one another. The way that each piece completely fills the concavity in the next piece creates this illusion. It's not always that way, but when the ratio of S:L is appropriate for filling a large portion of the area of the circles, it commonly is. Despite the very different appearance, it's the same tile used in both patterns.

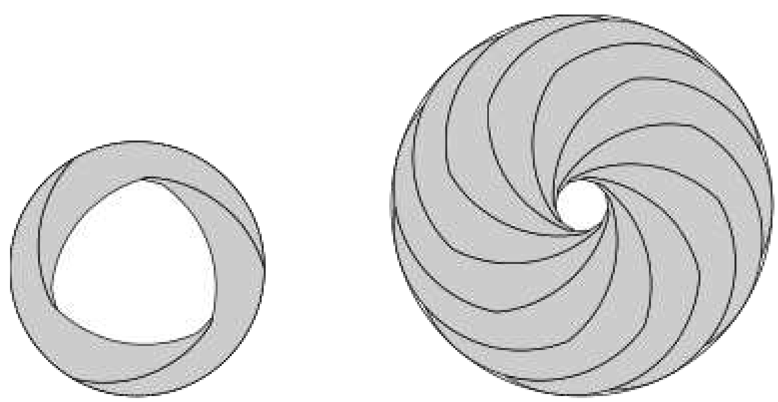


Fig. 3: /dev/joe vortex of type 1 for S=3, L=11

Since the areas of the two circles are in the ratio 4:9, and the goal was to fill as large a portion of the two circles as possible, the number of pieces used in the small and large circles should be close to this ratio in order to fill a large portion of the circle. However, there is nothing preventing the creation of vortices using numbers quite different from this ratio; it just creates a large hole in the middle of one of the circles, as in the case for S=3, L=11, shown in Figure 3.

1 http://www2.stetson.edu/~efriedma/mathmagic/0312.html

Determination of the Shapes

The shape of a crescent is determined by the outside (which is always 1/S of a circle), the length of the common arc, and the angle between two arcs. In the larger circle, only a part of the outside arc is actually on the outside of the overall shape. Part of the outside arc lies against one of the common arcs from the trailing edge of another crescent. The portion of outside arc which can actually fit on the outside is determined by the overlap of L equally spaced circles of diameter 2, each internally tangent to the circle of diameter 3. The common arc is the difference between these two arc lengths, and can be determined purely as a function of S and L. A detailed derivation appears below.

The preceding construction also determines the sharp angle between the outside arc and the trailing edge. In the small circle, this sharp angle forms a smooth edge with the angle between the outside arc and the leading edge of another piece at the edge of the circle, and that is one of the common angles, so this sharp angle determines the common angle.

The shapes I have shown so far all have 6 edges, but this is not always the case for /dev/joe crescents. A crescent ends when the leading edge and trailing edge intersect. The final arc on each edge is cut short at this point. Two additional parameters, A and B, are defined as the number of arcs making up the leading and trailing edges, respectively. They aren't always 2 and 3. In typical cases which are good candidates for the area maximization of the original problem, usually A=S-1 and B=S.

Special Cases

In some cases, as in the one shown in Figure 2, the hole in the center cannot be filled, because in the small circle it is bounded entirely by portions of the trailing edge, while in the large circle it is bounded entirely by portions of the leading edge. However, as seen above in Figure 3 for S=3, L=11, this is not always the case. When the last partial arc of trailing edge is longer than the last partial arc of leading edge, and B=A+1, both holes are bounded by trailing edge. When this happens, the solution can be improved by adding a small projection to each piece, such that the entire hole in the center of the large circle is filled. These bits will protrude into the much larger hole of the small circle, giving that hole a shape like a sawblade.

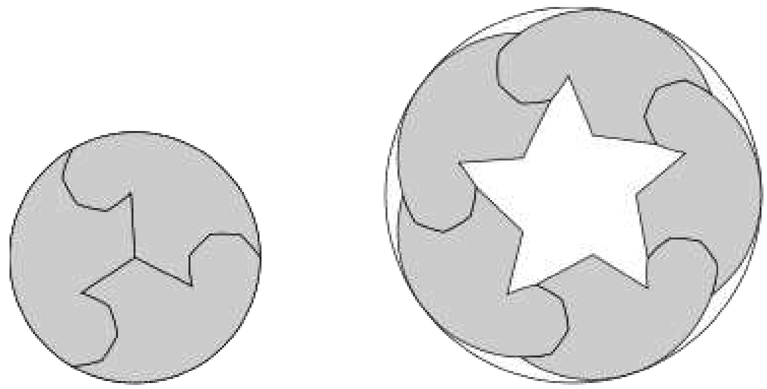


Fig. 4: /dev/joe vortex of type 1 for S=3, L=5 (with A=B=4 and then connected to the center)

At the other extreme in S:L, there are cases where the common edge is too short for the leading and trailing edges to intersect, and instead the edges curl back on themselves. In these cases, at an appropriate point the edges can be ended and simply connected straight to the middle of the small circle. This happens for S=3, L=5 as shown in Figure 4.

A very special case occurs when L=2S. This seems to cause A=B=S and for the leading and trailing edges to intersect in the center of the circle, with the final arcs bisecting one another, resulting in a completely filled small circle without any need to deviate from the standard design. When I tested the S=3, L=6 shape I got a shape which I think is identical to Bayly's solution for this case. Figure 5 shows this and the next two members of this family of solutions.

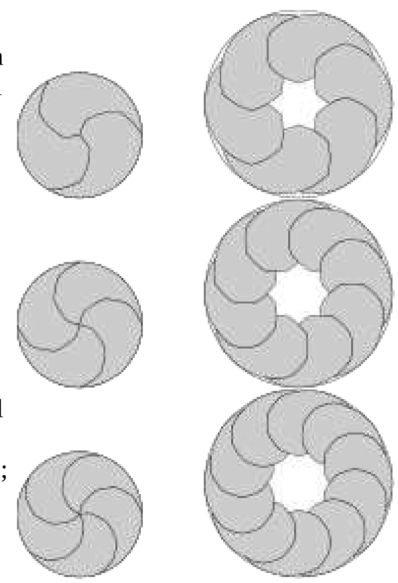


Fig. 5: /dev/joe vortices of type 1 for L=2S, S=3 to 5

Type 2 Vortices

In the original problem (whose goal, remember, was to maximize the area covered), only the total number of pieces N=S+L was specified, not the individual values of S and L. However, each S,L pair has a maximum theoretical area which is reached when one of the circles is completely filled; either $N\pi/S$ if the small circle is filled, or $9N\pi/4L$ if the large one is filled. By the time I developed this strategy, there were already solutions for most cases up to N=16 which covered enough area that only the S,L pair with the largest value of this area limit even had a hope of matching that score. And in some cases my solutions did not measure up. When more of the larger circle was filled, I noticed that the central hole was sometimes quite tiny, and the area would not be large enough to reach the goal even if it was filled in. The problem was that I had made the outside be made of arcs of radius 1, and this left some little slivers of area along the edge of the circle which could never be filled.

This led me to develop a second type of crescent, one where the outside was an arc of radius 1.5. This allowed the slivers along the edge to be captured, while losing slivers along the edge of the small circle, which had area to spare.

It turns out that inscribing S arcs of radius 1.5 and length equal to 1/L of the large circle along the inside of the small circle does not completely fill the small circle for typical S:L ratios. The natural thing to do then is to space them equally and let the remaining space along the edge of the small circle become another edge of the shape. So the outside now consists of two arcs, one of radius 1.5 and one of radius 1. Since in the large circle, the arcs of radius 1.5 completely fill the edge, those arcs are always on the outside and never nest against another arc. Instead, the arcs of radius 1 become the common arcs that are repeated along the leading and trailing edges. This leads to solutions like the one shown in Figure 6.

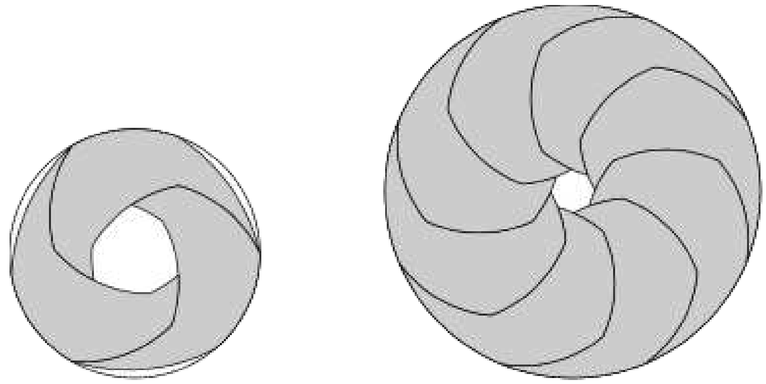


Fig. 6: /dev/joe vortex of type 2 for S=3, L=8

Type 3 Vortices

It is possible that for some systems very close to the S:L ratio of 4:9, a hybrid approach in between these two may provide a better area. In this shape, part of the border is an arc of radius 1.5 but it is less than 1/L of a circle. This is a subject for further research and will not be covered here.

Mathematical Basis of the Construction

I applied geometry and trigonometry to diagrams of the figures I was drawing to determine the placement of the various points necessary to draw these diagrams and determine their areas.

Begin by constructing a circle of radius 1.5 and a sector of that circle of angle $2\pi/L$ bisected by a horizontal radius of the circle. An angle π/L appears above the horizontal and another π/L below it. Now draw a sector of radius 1 and angle 2pi/S, tangent to the circle at its intersection with that same horizontal, and with its bottom point on the lower radius of the first sector. Necessarily, the distance between the centers is 1/2. The arc of radius 1 drawn here is the outside edge of the crescent. The part which lies inside the first sector is the part which lies on the outside when the crescents are placed in the larger circle; symmetry requires that this arc exactly fits in the sector of angle $2\pi/L$.

Let A be the intersection of this arc with the upper radius of the first sector. Draw another radius inside the second sector, to A. This divides the second sector into two pieces, an upper portion of angle φ and a lower portion of angle θ which is itself bisected by the horizontal radius; $\varphi+\theta=2\pi/S$. In addition, define $\beta = (\pi-\varphi)/2$ as one of the equal angles in this triangle with edges 1, 1, d.

Construct a vertical line through A, and define x to be the distance along this vertical line from A to the horizontal radius, and y the distance from the center of the sector of radius 1 to the right angle just created. Construct a chord across the sector of angle phi; define d to be its length. This produces Figure 7.

Now we have $\sin(\theta/2) = x$ and $\tan(\pi/L) = x/(y+0.5)$ and also $x^2+y^2 = 1$. Solve the last equation to get $y = (1-x^2)^{1/2}$, and substitute it into the second equation:

$\tan(\pi/L) = x/((1-x^2)^{1/2}+0.5)$

$x = \tan(\pi/L)((1-x^2)^{1/2}+0.5)$

$x = \tan(\pi/L)(1-x^2)^{1/2}+0.5 \tan(\pi/L)$

$x-0.5 \tan(\pi/L) = \tan(\pi/L)(1-x^2)^{1/2}$

Now square both sides to eliminate the square root:

$(x-0.5 \tan(\pi/L))^2 = \tan^2(\pi/L)(1-x^2)$

$x^2+0.25 \tan^2(\pi/L) - x \tan(\pi/L) = \tan^2(\pi/L) - x^2\tan^2(\pi/L)$

$x^2(1+\tan^2(\pi/L)) - x \tan(\pi/L) - 0.75 \tan^2(\pi/L) = 0$

Solve the quadratic for x:

$x = (\tan(\pi/L) +/- (\tan^2(\pi/L)+3\tan^2(\pi/L)(1+\tan^2(\pi/L)))^{1/2})/(2+2\tan^2(\pi/L))$

$x = (\tan(\pi/L) +/- (4\tan^2(\pi/L)+3\tan^4(\pi/L))^{1/2})/(2+2\tan^2(\pi/L))$

Pull a $\tan(\pi/L)$ out of the numerator:

$x = (1 +/- (4+3\tan^2(\pi/L))^{1/2})\tan(\pi/L)/(2+2\tan^2(\pi/L))$

As long as L is at least 3, the tangent has a positive value. For

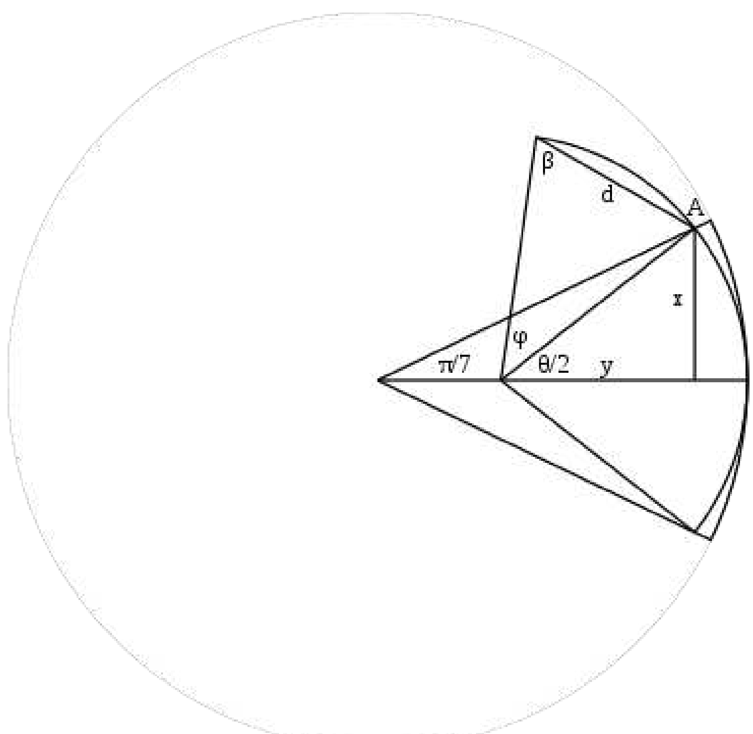


Fig. 7: Construction used to determine length of outside and common arcs for Type 1 vortices. S=3 and L=7 for this example.

x to be positive, we need the first term to be positive, and $(4+3\tan^2(\pi/L))^{1/2} > 2$, so we need the positive choice of the +/- sign.

$x = (1 + (4+3\tan^2(\pi/L))^{1/2})\tan(\pi/L)/(2+2\tan^2(\pi/L))$

It's possible to apply other identities here but it doesn't really get much simpler than this. This form is useful in that it only requires calculating one trig function to numerically calculate x from L. Since $\sin(\theta/2) = x$ means $\theta = 2\,\mathrm{asin}(x)$, we can calculate θ from L as well.

We also have $\varphi = 2\pi/S - \theta$ and $d = 2\sin(\varphi/2)$.

We don't need x and y in the diagram any more. They were merely a means to calculate θ from L. Drop them (and the vertical line which was of length x), and duplicate the remaining lines rotated counterclockwise by 2π/L.

Connect the centers of the two sectors of radius 1 with a line segment. If this was repeated all the way around the circle, the segments like this would form a regular L-gon with circumradius 1/2. Let n be the length of one of these lines.

$n = 1/2\,(2\sin(\pi/L)) = \sin(\pi/L)$

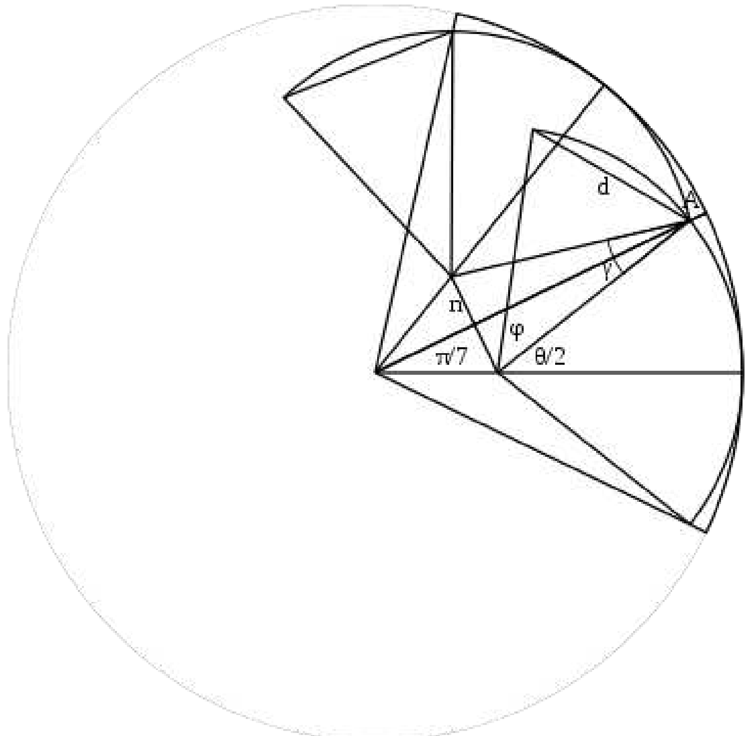


Fig. 8: More of the construction

Also, define γ as the angle at A in the triangle now formed with sides 1, 1, n. So $\gamma/2 = \mathrm{asin}(n/2)$ or $\gamma = 2\,\mathrm{asin}(n/2)$. This yields Figure 8.

The next goal is to add another arc to the shape to the left of the arc with chord d. It will be congruent to that arc, and share one endpoint with it. The angle at which this arc is drawn must be determined.

I used the angle between chord d and the corresponding chord for the next arc. In order to do this, we need to take the drawing here and place it on the small circle, centered at the center of the small sector from the upper tile. Some superfluous lines have been removed, but we need to keep the entire upper tile and the arc with chord d, since it forms part of the trailing edge of our tile. This produces Figure 9.

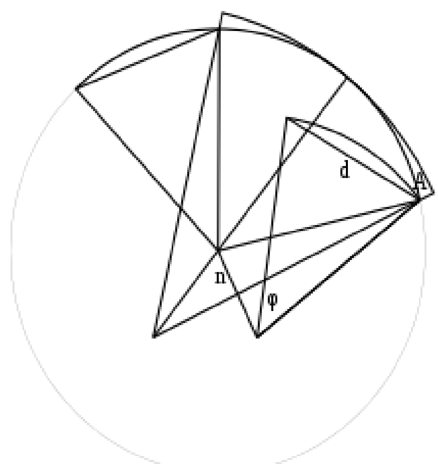


Fig. 9: Moving to the small circle

The next tile would be properly placed at the same center and rotated by 2π/S. See Figure 10. The arc with chord labeled d in the upper tile forms part of the leading edge of the lower tile. I used the angle between the two chords (d and its adjacent copy) to determine the placement of the succeeding arcs. This angle α is composed of two copies of the angle, β, between the chord and one of its radii minus the overlapping section, γ. The whole angle $\alpha = 2\beta - \gamma = \pi - \varphi - \gamma$.

Now we can dispense with the extra lines and just draw copies of this sector such that their chords form angles α until the leading and trailing edges meet. This is shown in Figure 11. It looks as if parts of three of these sectors and chords form a trapezoid, but it's a near miss and the two segments that appear to coincide to form the long edge actual intersect, form an angle of less than one degree.

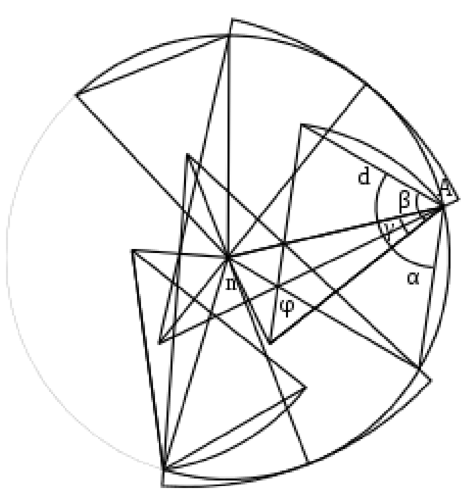


Fig. 10: Two copies of the partial tile on the small circle.

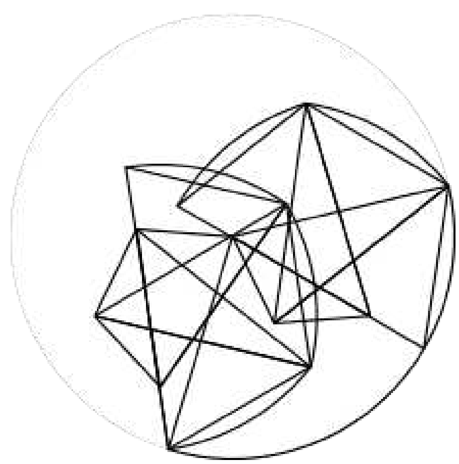


Fig. 11: Completing the tile.

The Surprising Coincidence

In the previous construction, for S=3 and L=7, θ is some irrational fraction of a circle, and so are φ and γ. But α works out to exactly 13π/21. So if I continued drawing these arcs and chords until I had drawn 21 of them, they would then start to coincide after I've drawn 21 chords. These chords form the perfect star polygon {21/4} (Figure 12), that is, a polygon with 21 equal sides and 21 equal angles which loops 4 times around the center, with each vertex is connected to the 4th nearest one on either side.

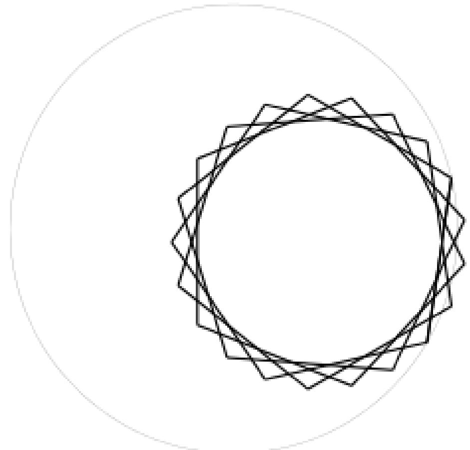


Fig. 12: Star polygon {21/4} formed by iterating chords across common arcs of the S=3, L=7 tile well beyond the point needed for its leading and trailing edges to meet.

Why is this? Why should these shapes lead to such an angle? It works for other cases, too. In general, the chords form a star polygon {SL/(L-S)} (with any common factors divided out). In the L=2S special case, the polygon is {SL/S} or a simple regular L-gon.

Proof of this angle: Go back to the diagram in the large circle which features the segment x, but extend it to length 2x all the way across the chord of angle θ. Repeat this all the way around the large circle so that these segments form a regular L-gon. See Figure 13.

Each angle of this L-gon is composed of two copies of the angle shown here as η which, due to the triangle it is in, is (π-θ)/2, and one copy of γ. Since it's the interior angle of a regular L-gon, it measures π - 2π/L. So π - 2π/L = π - θ + γ, or γ = θ - 2π/L. But α = π - φ - γ and φ = 2π/S - θ, so α = π - 2π/S + θ - θ + 2π/L = π - 2π(1/S-1/L) = π - 2π(L-S)/SL. So SL copies of this angle add to (SL - 2L + 2S)π, which is exactly the angle we need to make the star polygon {SL/(L-S)}.

It's easier to see this if you take the supplementary angle π-α, which equals 2π(L-S)/SL, which is just enough to bend the polygon around L-S full revolutions after SL copies of the angle.

The Arc-Through-the-Center Coincidence Proven

There is still the unexplained coincidence in the L=2S cases where the arc passes through the center of the small circle.

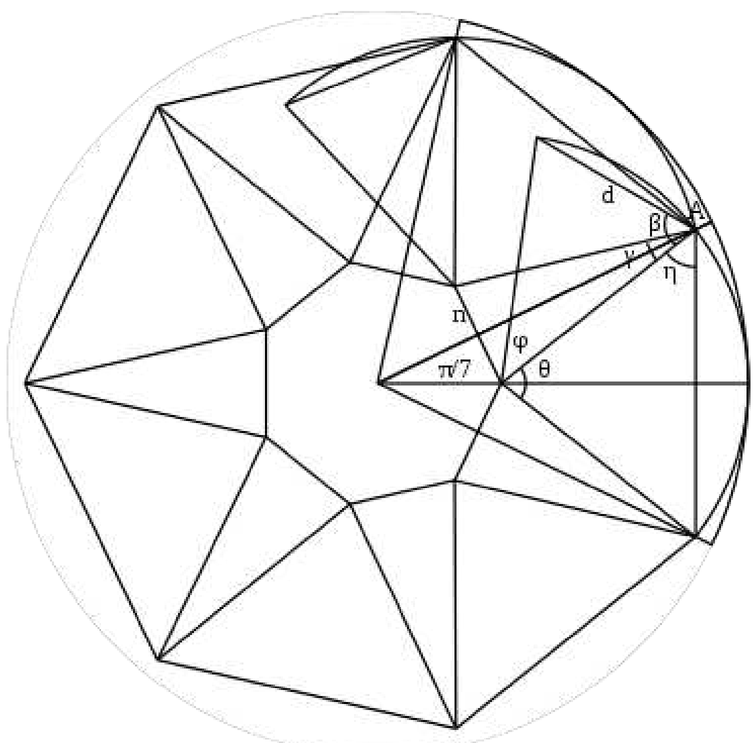


Fig. 13: The L-gon of chords of length 2x.

In these cases, the polygon has an even number of sides, and the distance of interest is the distance from the midpoint of one arc to the midpoint of the opposite arc. Aside from the conclusion I want to reach that this distance is 1, it's hard to measure this relative to any of the rest of the problem. But draw a segment bisecting each sector. These segments for two adjacent sectors intersect at some distance from the arc's midpoint (and another distance from the center about which the arc is drawn) which by symmetry is the same for both sectors. It's the same by symmetry for each pair of adjacent sectors, so it is the same for all of them, and all of them intersect in a single point, C. The distance from this C to the arc's midpoint B is half the distance of interest.

To find the length of segment BC, I calculate the distance from C to the center D about which the arc is drawn, and subtract that from the known radius, 1. CD and its counterpart in the adjacent sector form two legs of an isosceles triangle, where the third side is the segment identified earlier as n = sin(π/L). These segments of length n form another star polygon similar to the one formed by the chords, so the angle at C is (L-S)/SL of a full revolution, or 2π(L-S)/SL.

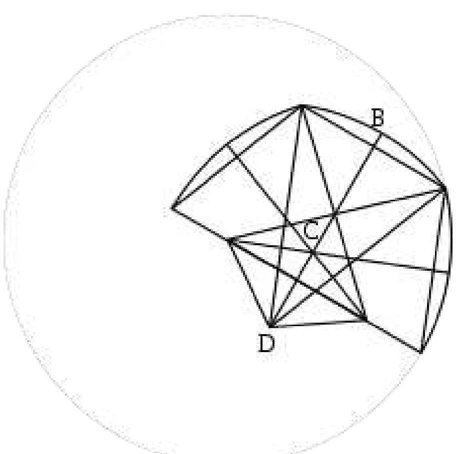


Fig. 14: Bisecting the sectors

So CD / (n/2) = sin(π(L-S)/SL) or CD = sin(π(L-S)/SL) / 2sin(π/L). In the special case where L=2S, sin(π(L-S)/SL) = sin(π/L) so CD = sin(π/L) / 2sin(π/L) = 1/2. That makes BC also 1/2, and this is why the distance from arc midpoint to opposite arc midpoint is 1 in these cases.

Conway and The 3x+1 Problem Continued

Gary Greenfield

Abstract

We present a conjecture concerning a family of iterative functions that includes the most well-known example ($3k+1 \to 4k+1$, $3k \to 2k$, $3k-1 \to 4k-1$) associated with Conway's generalization of The 3x+1 Problem.

I don't think Martin Gardner's Mathematical Games column devoted to The 3x+1 Problem [2] made any lasting impression on me. I only became interested in the problem after reading the "stopping times" paper by Riho Terras [8]. Dan Drucker and I then began kicking around some ideas and fumbling with it a bit, but never got anywhere. When we mentioned this to Erdös during a mid-70s visit he made to Wayne State University he repeated a phrase we had previously heard attributed to him:

"Mathematics may not yet be ready for such problems."

and followed that up with:

"We will have to find something else for you to do."

We did indeed find other things to do, but over the years thanks to well-publicized teasers such as Richard Guy's [4] students continually re-discover the problem and it has always remained at the fore.

Until Conway's related paper [1] was reprinted in Lagarias' book [7, pp. 219–224] it was relatively obscure. I can't remember how I first learned of it (perhaps from the references in the preprint of [6] Lagarias sent me) but in the late-90s I secured a copy of Conway's paper via Interlibrary Loan. Therefore, when prospective University of Richmond Honor's student, Robin Givens, expressed interest in The 3x+1 Problem, although reluctant, I gave her the go-ahead to look at Conway's paper and see what she could learn about generalizations of The 3x+1 Problem. Obtaining significant or noteworthy results was not possible — after all, the problem IS hard — but a conjecture we formulated concerning a generalization [3] may be of interest and I will briefly cover it below.

With $p > 1$ fixed, Conway considers the generalization

$$g(n) = a_i n + b_i$$

where $i = n \bmod p$, and $a_0, b_0, \ldots, a_{p-1}, b_{p-1}$ are *rational* constants chosen such that $g(n)$ is always *integral*. Conway proves that it is undecidable whether given a function g and positive integer n such that $g(n)/n$ is periodic there exists an integer k such that the k-fold iterate $g^k(n) = 1$.

Although it is not mentioned in Conway's paper, according to Guy [5] the motivating example was:

$$t(n) = \begin{cases} \frac{2}{3}n + 0 & \text{if } n \bmod 3 = 0 \\ \frac{4}{3}n - \frac{1}{3} & \text{if } n \bmod 3 = 1 \\ \frac{4}{3}n + \frac{1}{3} & \text{if } n \bmod 3 = 2 \end{cases}$$

and according to Lagarias [6] this example was found in a 1932 journal of Collatz. The function t is a permutation on the positive integers with known finite cycles (1), (2 3), (4 5 7 9 6) and (44 59 79 105 70 93 62 83 111 74 99 66). It is not known if the iterates of t starting with 8 form an "infinite" cycle or, assuming it does, if there is more than one infinite cycle.

Because t can be viewed as the mapping:

$$\begin{array}{lcl} 3k+1 & \to & 4k+1 \\ 3k & \to & 2k \\ 3k-1 & \to & 4k-1 \end{array}$$

the generalization of The 3x+1 Problem that Robin and I considered was the family of functions c_q defined for *odd* $q > 1$ by:

$$c_q(n) = \begin{cases} \frac{q+1}{2}k & \text{if } n = qk \\ (q+1)k + \ell & \text{if } n = qk + \ell, \text{ with } 0 < |\ell| \le \frac{q-1}{2}. \end{cases}$$

Note that t is the function c_3. There are several elementary facts about this family that the reader is invited to discover on his or her own (e.g., only c_3 has a 2-cycle). What I wish to highlight here is the following conjecture:

Remainder Conjecture: For any finite sequence $r_0, r_1, \ldots, r_m$ such that $0 \le |r_i| < (q+1)/2$, there exist an n such that the i-fold iterate ${c_q}^i(n)$ satisfies ${c_q}^i(n) \bmod q = r_i$.

The conjecture says there exist $k_0, k_1, \ldots, k_m$ such that if we set $n = qk_0 + r_0$, then ${c_q}^i(n) = qk_i + r_i$ for $0 < i \le m$. The conjecture is true when all the r_i are identically zero. Its significance is that it reveals there are effectively q^m cases one must consider in order to decide on a case by case basis whether or not c_q has a cycle of length m.

References

[1] J. H. Conway, Unpredictable iterations, Proc. 1972 Number Theory Conference, University of Colorado, Boulder, Colorado (1972) 49–52.

[2] M. Gardner, Mathematical Games, Scientific American, 226 (June 1972) 114–118.

[3] R. Givens, On Conway's Generalization of the 3x+1 Problem, Honors Thesis, University of Richmond, 2006.

[4] R. K. Guy, Don't try to solve these problems!, Amer. Math. Monthly, 90 (1983) 35–41.

[5] R. K. Guy, Private communication.

[6] J. Lagarias, The 3x+1 problem and its generalizations, Amer. Math. Monthly, 92 (1985) 3–23.

[7] J. Lagarias, The Ultimate Challenge: The 3x+1 Problem, Amercian Mathematical Society, Providence, RI, 2010.

[8] R. Terras, A stopping time problem on the positive integers, Acta Arith., 34 (1978) 241–252.

Loss Aversion and Stock Price

G4G14 Gift Exchange

Stephen D. Hassett

This exchange item consists of two papers that served as the basis for "The Risk Premium Factor" (Wiley 2011). They are in the realm of behavioral economics and establish the relationship between stock price and Loss Aversion/Prospect Theory (Kahneman and Tversky, 1979) for which Kahneman won the Nobel Prize in Economics.

The first paper was published in the Journal of Applied Corporate finance and describes a simple mathematical model that explains the level of the S&P 500 Index with considerable accuracy. The second was not published and instead was incorporated into the author's book. It expands the discussion to explore the relationship of stock price to loss aversion coefficients associated with Prospect Theory (Kahneman and Tversky, 1979) for which Kahneman won the Nobel Prize in Economics.

- The RPF Model for Calculating the Equity Market Risk Premium and Explaining the Value of the S&P with Two Variables (Journal of Applied Corporate Finance Spring 2010)
- How the Risk Premium Factor Model and Loss Aversion Solve the Equity Premium Puzzle (Unpublished)

<u>About the Author</u>

Steve Hassett is a technology executive and author of the "The Risk Premium Factor: A New Model for Understanding the Volatile Forces that Drive Stock Prices" (Wiley 2011) and has also published in the Journal of Applied Corporate Finance, Ad Age, CNBC.com and is a regular contributing author for the Seeking Alpha investment website. He holds an MBA from the Darden School of Business at the University of Virginia and a B.S. from Rensselaer Polytechnic Institute.

VOLUME 22 | NUMBER 2 | SPRING 2010

Journal of APPLIED CORPORATE FINANCE

A MORGAN STANLEY PUBLICATION

In This Issue: **Strategy and Valuation**

Morgan Stanley

The RPF Model for Calculating the Equity Market Risk Premium and Explaining the Value of the S&P with Two Variables

by Stephen D. Hassett, Hassett Advisors

While driving increases in shareholder value is one of the most important responsibilities of any business leader, many executives are handicapped by their limited understanding of what drives value. And they are not alone. Even prominent economists say that stock market valuation is not fully understood. For example, in a 1984 speech to the American Finance Association, Lawrence Summers said,

It would surely come as a surprise to a layman to learn that virtually no mainstream research in the field of finance in the past decade has attempted to account for the stock-market boom of the 1960s or the spectacular decline in real stock prices during the mid-1970s.[1]

Some people see the stock market as arbitrary and random in setting values. But despite occasional bouts of extreme volatility (including, of course, the recent crash), most academics (and many practitioners) would likely agree with the proposition that the market does a reasonably good job of incorporating available information in share prices. At the same time, however, certain factors can clearly cause the market to misprice assets. These include problems with liquidity, imperfect information, and unrealistic expectations that can knock valuations out of line for a period of time. But such limitations notwithstanding, over a longer horizon the market appears to be reasonably efficient in correcting these aberrations.

The RFP Valuation Model introduced in this article is intended to explain levels and changes in market values and, by so doing, to help identify periods of likely mispricing. As such, the model offers a general quantitative explanation for the booms, bubbles, and busts—that is, the series of multiple expansions and contractions—that we have experienced over the past 50 years. The model explains stock prices from 1960 through the present (March 2010), including the 2008/09 "market meltdown." And it does so using a surprisingly simple approach—one that combines generally accepted approaches to valuation with a simple way of estimating the Market or Equity Risk Premium (ERP) that produces remarkably good explanations of market P/E ratios and overall market levels. To show you what I mean, Figure 1 shows how the P/E ratio predicted by model, when applied to S&P Operating Earnings, explains levels of the S&P 500 over the past 50 years, the earliest date for which I had reliable earnings data.

My approach to estimating the Equity Risk Premium is the most original part of this overall hypothesis. Many if not most finance theorists have assumed that the Equity Risk Premium is a constant that reflects the historical difference between the average return on stocks and the average return on the risk-free rate (generally the return on the 10-year U.S. government bonds). But if we also assume that long-term real interest rates do not change and that real growth can be approximated by real long-term GDP growth (also generally assumed to be stable), then the market-wide P/E would also be absolutely constant over time.

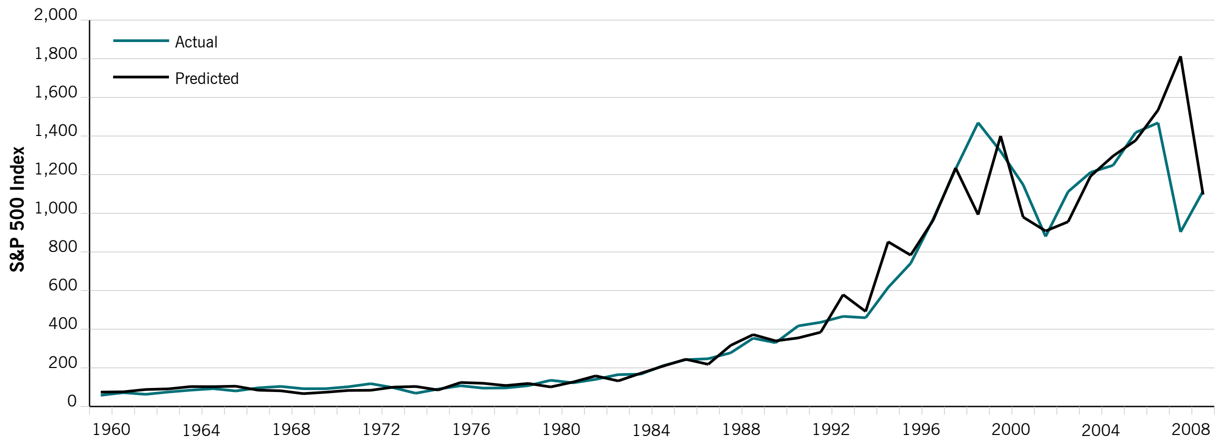
But, of course, the P/E multiple on the earnings of the S&P 500 is volatile, with year-end values ranging from 7.3 in 1974 to 29.5 in 2001. One possible objection to the idea of a constant risk premium is its implication that, when the risk-free rate increases, investors are satisfied with a premium that is smaller as a proportion of the risk-free rate. In this article, I suggest that the Equity Risk Premium is not a fixed number but a variable that fluctuates in direct proportion to the long-term risk-free rate as a fixed percentage, not a fixed premium. When used with the constant growth model, the cost of capital can be determined by the following formula:

$$\text{Equity Risk Premium} = \text{Risk-Free Long-Term Rate} \times \text{Risk Premium Factor} \quad (1)$$

This relationship can be used to explain why and how the risk premium varies over time; as interest rates vary, so does the risk premium. This Risk Premium Factor (RPF) appears to have held steady for long periods of time, changing just twice during the 50-year period from 1960 to the present (July 2009). Based on my calculations, the RPF was 1.24 from 1960-1980, 0.90 from 1981-June 2002, and 1.48 from July 2002 to the present. As we saw earlier in Figure 1, the model does a very good job of predicting market levels, even through the present financial crisis.

1. Quoted by Justin Fox, *The Myth of the Rational Market: History of Risk, Reward and Delusion on Wall Street*, p. 199. (Harper Collins, New York, 2009).

Figure 1 **S&P 500 Actual vs. Predicted—1960–2009**



This result is also consistent with investor "loss aversion," the well-documented (by Kahneman and Tversky) willingness of investors to sacrifice significant gains to avoid considerably smaller losses. One of their studies produced a loss aversion coefficient of 2.25,[2] which implies that participants, on average, would be indifferent to the outcome of a coin flip promising either an expected but uncertain $325 or a guaranteed $100. The analogous calculation for the RPF model suggests that if the risk-free rate were 4% and the RPF 1.48, investors contemplating a $1,000 investment would assign roughly equal value to a guaranteed (bond-like) $40 and equities with an expected return of $99.

Valuing Constant Growth

The place to start is with the simplest valuation model, the Constant Growth Equation. This model derives from, and represents a specific case of, the Discounted Cash Flow (DCF) model that is used to determine the net present value of a projected stream of future cash flows. In the case in question, it is a perpetual stream of cash flows with a constant rate of growth. Instead of assuming different levels of earnings in each period, it assumes a constant growth rate off the base year and a constant cost of capital.

The DCF model can be expressed as follows:

$$P = \sum E_1 / (1+C)^1 + E_2 / (1+C)^2 + \ldots + E_n / (1+C)^n \qquad (2)$$

where E is cash flow and C is cost of capital. If you assume that E grows at a constant rate (G),

$$P = \sum (E_0 \times (1+G)^1) / (1+C)^1 + (E_0 \times (1+G)^2) / (1+C)^2 + \ldots + (E_0 \times (1+G)^n) / (1+C)^n \qquad (3)$$

the result simplifies to:

$$P = E / (C - G) \qquad (4)$$

This equation, which is not so much a theory as an indisputable mathematical concept, is the expanded form of the core insight that the value of a perpetual stream is the amount of the payments divided by the required rate of return. In other words, the value of a guaranteed $100 perpetual annuity in a market where the long-run risk-free return is 10% is $1,000 ($100/.10).

The next step is to take the constant growth version of this model (equation 4) and apply it to market valuation by substituting S&P operating earnings for the variable E above.

P = Price (Value of S&P 500 Index)

E = Earnings (Reported operating earnings for the prior four quarters as reported by S&P) as a proxy for cash flow

G = Expected long term growth rate

C = Cost of equity capital

This formula can also be restated to predict the Price-Earning (P/E) ratio of the S&P 500 as follows:

$$P/E = 1 / (C - G) \qquad (5)$$

2. Daniel Kahneman and Amos Tversky, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5 (1992):297-323.

Table 1 **Growth Drives P/E**

Long-term Growth	Predicted P/E
0%	12.6
2%	16.7
4%	25
6%	50

These two equations, when used with the right assumptions (as discussed below) can be helpful in understanding the valuations of both individual companies and the overall market.

Some academics and practitioners argue that equity should be valued as the present value of not earnings or cash flows, but of the dividend payments actually made to shareholders—an argument that is embodied in the Gordon (or Dividend) Growth Model. Some proponents of this model advocate a modified approach that values all corporate distributions, share repurchases as well as dividends. One well-known advocate of this model is Nobel Laureate Paul Krugman, who wrote:

Now earnings are not the same as dividends, by a long shot; and what a stock is worth is the present discounted value of the dividends on that stock—period, end of story.[3]

I disagree, and for several reasons. For starters, Modigliani and Miller demonstrated in their famous 1961 article on the "irrelevance" of dividend policy, that it is the underlying expected earnings power of companies, not their dividend payouts, that determine corporate market values.[4] Dividend policy is as much a reflection of a company's capital structure and investment opportunity set as of its expected future profits—and decisions to pay out capital may often reflect a maturing of the business and a scarcity of profitable investment opportunities. What's more, most promising growth companies pay no or minimal dividends—and certainly for those companies, the current levels and changes in earnings are likely to be more reliable indicators than dividends of future profitability.

Why Growth Rate and Cost of Capital Matter—Lessons from the Constant Growth Equation

Assume you have an asset with a cost of capital of 12%, a growth rate of 2% and cash flow of $100. Using the Constant Growth model, the value can be calculated as follows: $100 / (12% - 2%) = $1,000. This might be called the "intrinsic value" of the asset and, as such, it offers the best guide to what it should trade for.

We can also apply this model to a share of stock to determine its intrinsic value. In place of cash flow, we use earnings per share (EPS) of $2.00 with the same cost of capital and growth rate, and the result is $2.00/(12% - 2%) = $20.00. Since EPS is $2.00 and price is $20.00, the Price to Earnings Ratio (P/E) is $20/$2 or a P/E of 10. While the market may value it differently, if these assumptions are true, this formula tell us its intrinsic value.

P/E ratios are often used to assess whether share prices are expensive or cheap. A P/E of 8 is considered very low, but when Google had a P/E of 60 or more, some thought it was very high. Is a company with a P/E of 10 a bargain compared to a company with a P/E of 20? We can explore this question using the constant growth equation.

Take the same company and now assume that its cost of capital drops to 8%, its growth rate increases to 3%, and its earnings stay the same. These might seem like small changes, but their impact is dramatic: $2.00/(8% - 3%) = $40.00, a doubling of value with the P/E rising to 20. If growth increases to 5% (in line with nominal long-term GDP growth), the share price rises to $66, and the P/E is 33. (For additional examples of how P/E varies based on growth for a company with an 8% cost of capital, see Table 1.)

The formula P = E / (C – G) shows that earnings relate directly to price. What many managers fail to realize is that investors don't look at earnings in a vacuum; they parse the information in earnings in order to estimate growth. And that's why the reporting of earnings often causes the P/E to change.

So, for all its simplicity, the Constant Growth model has some important lessons:

1. Small changes in growth make a big difference in value
2. Cost of capital is important, so we better get it right
3. Earnings drive value (stock price) but also contain information

While it may not be difficult to project current earnings, the big challenges are forecasting growth and getting the right cost of capital.

A Short Overview of Risk Premiums

The Capital Asset Pricing Model (CAPM) can be used to determine the cost of equity for an individual firm or the market overall. The model takes the form of the following equation: **Cost of Equity = R_f + β x (ERP),** where R_f = Risk-Free Rate (and we will use the yields on 10-year Treasuries as a proxy); β = Beta, which measures the sensitivity of the stock to market risk (which, by definition, is 1.0 for the entire

3. Krugman, Paul, "Dow 36,000: How Silly Is It?", *The Official Web Page of Paul Krugman*, accessed August 2009, http://web.mit.edu/krugman/www/dow36K.html.

4. Franco Modigliani, Merton H. Miller, "Dividend Policy, Growth, and the Valuation of Shares," *Journal of Business*. 1961, vol. 34, no. 4.

Table 2 **ERP Drives Valuation**

R_f	ERP	Cost of Equity	GDP + Inflation	Predicted P/E
5%	3%	8%	5%	33
5%	4%	9%	5%	25
5%	5%	10%	5%	20
5%	6%	11%	5%	17
5%	7%	12%	5%	14

market); and ERP = Equity Risk Premium (the calculation of which will be the main subject of this discussion). Given that the Beta of the broad market is 1.0, the Cost of Equity for the market as a whole can be expressed as $C = R_f + ERP$.

While the risk-free rate is easily determined, the risk premium is not. In fact, there is no clear consensus on how this should be done. The Equity Risk Premium (ERP) is the expected return an investor requires above the risk-free rate for investing in a portfolio of equities. It makes sense that if 10-year Treasury yields represent the safest (risk-free) long-term investment, then investors will require higher expected rates of return to buy riskier securities like corporate bonds or equities. My own considerable experience in valuing businesses has made it clear to me how sensitive valuations can be to one's estimate of the ERP (a topic I return to later).

The most common way of estimating the ERP is to measure the historical premiums that investors have received relative to Treasury yields and assume that investors will expect that rate of return in the future. Depending on method and time-period, this can range from 3% to 7% or more. Other methods include surveys and forward-looking estimates based on current stock market levels. There is a huge body of research on measuring equity risk premiums. Indeed, entire books have been written on the subject.

Many researchers have argued that the Equity Risk Premium changes over time—and that such fluctuations are a major source of stock price changes—and also that the ERP has experienced a "secular" decline during the past few decades. In their book *Dow 36,000*, for example, Kevin Hassett (no relation) and James Glassman pushed this argument to its reduction ad absurdum when suggesting that the risk premium could vanish entirely since, given a sufficient amount of time, stocks appeared virtually certain to outperform bonds.[5] In *The Myth of the Rational Market*, Justin Fox quotes Eugene Fama, one of the pioneers of the efficient market hypothesis, as saying, "My own view is that the risk premium has gone down over time basically because we've convinced people that it's there."[6] Roger Ibbotson, a well-known compiler of ERP statistics, has suggested that the recent decline in the risk premium should be viewed as a permanent, but non-repeating event, "We think of it as a windfall that you shouldn't get again," he said.[7]

The Effects of Risk Premium on Valuation

Table 2 shows the expected effects of differences in ERP (ranging from 3% to 7%) on valuations and P/E ratios. Using the constant growth model, P/E = 1 / (C – G), if we assume that the market will grow with long-term estimates of real GDP at 3% plus long-term inflation at 2%, our estimate of stock market P/E would have P/E = 1 / (C – 5%). (Note: Real GDP + Inflation is Nominal GDP). With Treasury yields at 5%, and ERPs ranging from 3%-7%, our range of cost of capital (R_f + ERP) is from 8% to 12%. Table 2 also shows the P/E implied for the overall market given this range of estimates of ERP and cost of capital. To provide some perspective on these numbers, if the S&P 500 were at 1,200 with its current P/E of 19, it would increase more than 25% to 1,593 with a P/E of 25 and the same level of earnings!

A New ERP Theory: The Risk Premium Factor (RPF) Model

Conventional theory says that if the Equity Risk Premium were 6.0% and 10-year Treasury yield was 4.0% then investors would expect equities to yield 10%. The theory also implies that if the 10-year Treasury was 10%, then investors would require a 16% return, which represents a proportionally smaller premium.

For reasons discussed below, I will argue that investors expect to earn a premium that is not fixed, as in the conventional CAPM, but varies directly with the level of the risk-free rate in accordance with a "Risk Premium Factor" (RPF). While this proportional RPF is fairly stable, it can and does change over longer periods of time.

To illustrate the concept, with an RPF of 1.48, equities are expected to yield 9.9% when Treasury yields are at 4.0%. But if Treasury yields suddenly rose to 10%, equities would have to return 24.8% (10 + 1.48 x 10 = 24.8) to provide investors with the same *proportional* compensation for risk. In this example, an increase in interest rates (and inflation) causes the risk premium to jump from about 6% to 15%, suggesting that interest rates have a greater impact on valuation and market price than is generally recognized.

To test this approach, we must determine not only the

5. James K. Glassman and Kevin A. Hassett, *Dow 36,000: The New Strategy for Profiting From the Coming Rise in the Stock Market*, (Times Business, New York, January 1, 1999).
6. Justin Fox, *The Myth of the Rational Market: History of Risk, Reward and Delusion on Wall Street*, p. 263. (Harper Collins, New York, 2009).
7. Ibid.

Figure 2 **10-Year Treasury Yields—1960–2009**[12]

Source: U.S. Treasury

Risk Premium Factor, but estimates for the other variables in the following equation:

$$P/E = 1 / (C - G) \qquad (11)$$

In the analysis that follows, I use the following variables and assumptions:

P = Price (Value of S&P 500)

E = Actual Earnings (Annualized operating earnings for the prior four quarters as reported by S&P). Earnings, while not ideal, are used as a proxy for cash flow and seem to work very well

G = Expected long-term projected growth rate, which is broken down into Real Growth and Inflation, so $G = G_R + I_{LT}$

G_R = Expected long-term real growth rate. Long-term expected real growth rate (G_R) is based on long-term GDP growth expectations on the basis that real earnings for a broad index of large-cap equities will grow with GDP over the long-term. A rate of 2.6% is used with the same rate applied historically.[8]

I_{LT} = Expected long-term inflation, as determined by subtracting long-term expected real interest rates (Int_R) from the 10-year Treasury, where Int_R is 2%; based on the average 10-year TIPs Yields from March 2003 to the present.[9]

C = Cost of Capital is derived using Capital Asset Pricing Model, where for the broad market, $C = R_f + ERP$

R_f = Risk-Free Rate as measured using 10-year Treasury yields

ERP = Risk Premium Factor (RPF) x R_f

RPF = 1.24 for 1960 – 1980; 0.90 for 1981 – 2001; and 1.48 for 2002 – present. The RPF for each period was arrived at using a linear regression to fit the assumptions above to actual PE.[10]

When using these assumptions for the present period—that is, with an RPF of 1.48—the formula reduces to:

$$P/E = 1/ (R_f \times (1+RPF) - (R_f - 2\%) - 2.6\%) \qquad (12)$$

Explanatory Value of the RPF Valuation Model

As can be seen in Figures 2-6, the actual values deviated significantly from the predicted values at the end of 2008 and the first quarter of 2009, but had returned to something like parity by June 2009. I believe that these deviations from the model were attributable mainly to the abnormally low yields for 10-year Treasuries that had been in effect since late 2008, when the "flight to quality," along with the Federal Reserve's purchase of notes beginning in March 2009, caused the 10-year Treasuries to be overpriced.[11] As shown in Figure 2, yields then fell to as low as 2.2%, as compared to a more "normal" range of 4.1% to 5.1% in 2006 and 2007 (and rarely

8. "Economic Projections and The Budget Outlook," Whitehouse.gov, Access Date March 15, 2009, http://www.whitehouse.gov/administration/eop/cea/Economic-Projections-and-the-Budge-Outlook/.

9. "H.15 Selected Interest Rates", The Federal Reserve Website, Accessed March-July 2009, http://www.federalreserve.gov/datadownload/Choose.aspx?rel=H.15.

10. All data used in the analysis is available for download at: http://sites.google.com/a/hassett-mail.com/marketriskandvaluation/Home.

11. "Fed in Bond-Buying Binge to Spur Growth," *The Wall Street Journal Online*, March 19, 2009, http://online.wsj.com/article/SB123739788518173569.html.

12. H.15 Selected Interest Rates", The Federal Reserve Website, accessed March-January 2010, http://www.federalreserve.gov/datadownload/Choose.aspx?rel=H.15.

Figure 3 **S&P 500 P/E Actual vs. Predicted—1960–2009 (Annual)**

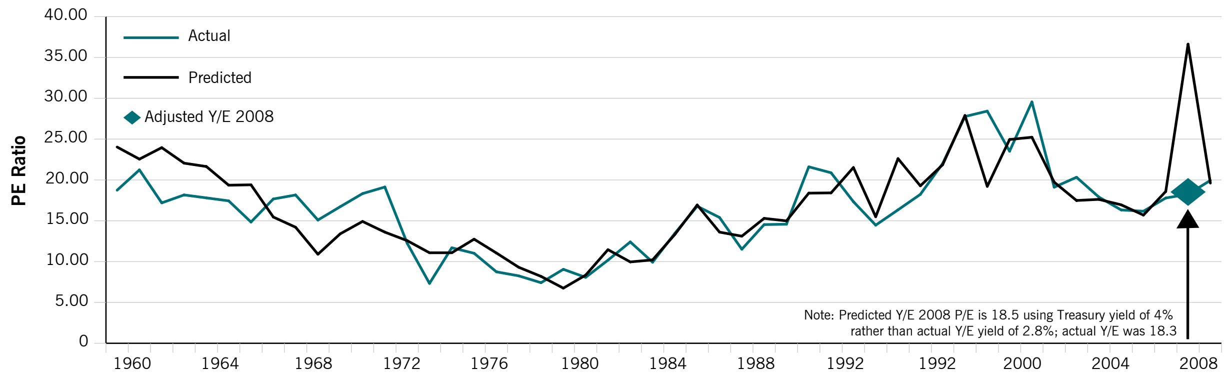
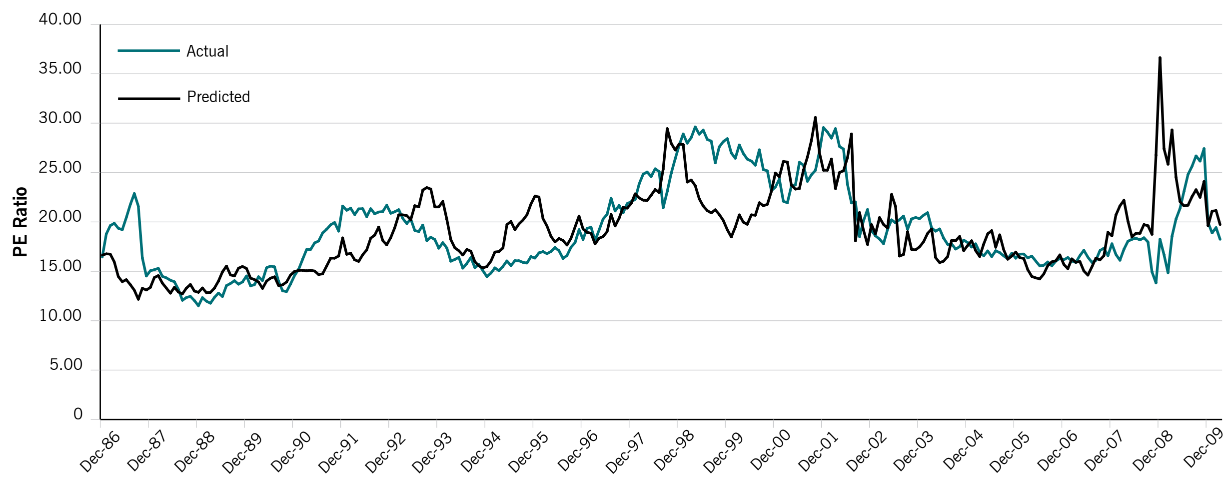


Figure 4 **S&P 500 P/E Actual vs. Predicted—1988–March 2010 (Monthly)**



less than 4% since 1960).

To compensate for these abnormally low Treasury yields Figure 3 shows the P/E ratios that would likely have prevailed if Treasury yields had remained at a still low, but more normal yield of 4%.[13] And as shown in each of Figures 3-5, when we normalize the 2008 R_f variable in this way, the actual year-end valuations correspond closely with the predicted values. One use of the model is to spot anomalies—and I believe that Treasury yields during the 2008/09 financial crisis were an anomaly.

Also plainly visible in Figure 3 is the decline in P/E ratios in the 1970s, reflecting the increase in interest rates during that period. It also shows the jump in P/Es during the 1980s, reflecting the drop in inflation and interest rates.

Figure 4 shows the application of the same model using monthly data from the end of 1986 through March 2010.[14] Like Figure 3, Figure 4 shows the return of values to parity by middle of 2009. And as can be seen in Figure 5, the RPF model explains overall market valuation levels when actual S&P operating earnings are applied to the P/E ratio during the period 1960–2009.[15] Using both year-end annual data for the past 50 years and monthly data for the past 20 years, then, the RPF model appears to do a very good job explaining valuations. And that in turn would suggest that, at any

13. While earnings are released quarterly, the model was extended to monthly and daily price data by using actual closing prices for S&P 500 and 10-Year Treasury yields along with S&P 500 operating earnings as a constant for each month in the quarter. The quarterly earnings were applied for the month preceding quarter end (i.e., Dec – Feb = Q1) under the assumption that market expectations would have incorporated earning expectations. Again, it assumed that as the end of quarter approaches earnings estimates should be within a reasonably close to those actual earnings ultimately reported and embodied in share prices. Earnings and S&P Averages 1960-1988 from Damodaran Online: Home Page for Aswath Damodaran (New York University) http://pages.stern.nyu.edu/~adamodar/; S&P Earnings and levels from 1988 – Present from Standard and Poors Website, http://www2.standardandpoors.com/portal/site/sp/en/us/page.topic/indices_500/2,3,2,2,0,0,0,0,0,1,5,0,0,0,0,0.html; Calculations and methodology by the Author.

14. See Note 13.

15. See Note 13.

Figure 5 **S&P 500 Actual vs. Predicted—1988–March 2010**

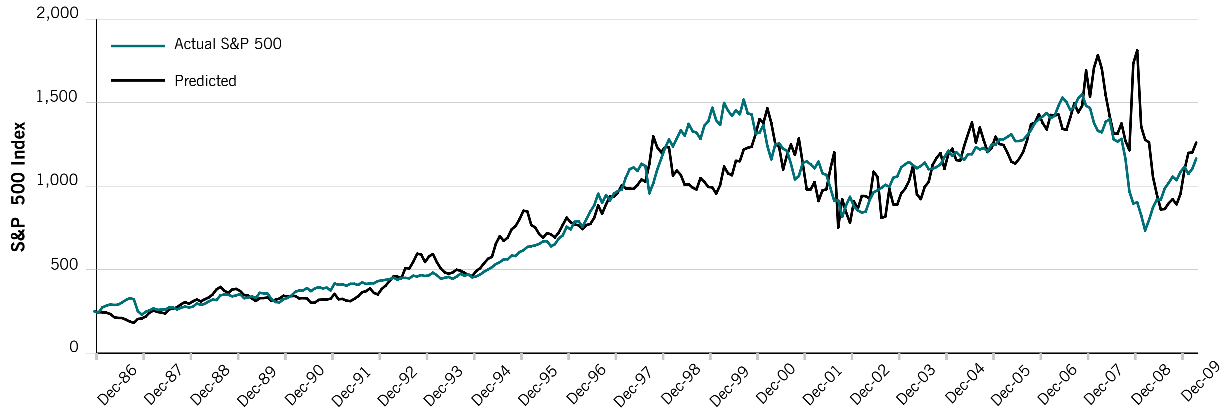


Table 3 **Estimated Risk Premium Factors**

Period	RPF
1960 – 1980	1.24
1981 – Q2 2002	0.90
Q3 2002 – Present	1.48
6%	50

Table 4 **RPF Valuation Model R Squared Results**

	R Squared	
Dataset	Full Dataset	Excluding Meltdown
1960 – 2008 (Annual)	89.5%	96.3%
1986 – September 2009 (Quarterly)	80.6%	88.0%
January 1986 – September 2009 (Monthly)	86.3%	90.8%
January 1986 – September 2009 (Daily)	86.5%	90.9%

point in time, the general level of market pricing and P/E ratios are driven mainly by just two factors: interest rates and expected earnings.

Estimating the Risk Premium Factor (RPF)

The RPF was estimated by fitting the model to actual levels of the S&P 500 over the period 1960 to the present. This analysis revealed two distinct shifts in the RPF since 1960. Table 3 shows the RFP factors that provide the best fit for each period.

The overall fit was assessed by calculating the R^2s of the regressions using the appropriate RPF for each time period. As previously discussed, the meltdown after September 2008 drove down the risk-free rate to an unsustainable level and left a trail of historical earnings that clearly did not reflect expectations. As also discussed previously, these factors are now back in line. To adjust for this recent anomaly, the R^2 was calculated excluding meltdown time period beginning September 2008.

As reported in Table 4, after excluding the meltdown period, the RPF Valuation Model explains a remarkably high 96% variation of stock prices over the past 50 years, as well as 91% of the daily variation.[16]

Consistency with Prospect Theory/Loss Aversion

As mentioned earlier, Daniel Kahneman and Amos Tversky first developed "prospect theory" in 1979, proposing that individuals have a sufficiently strong preference for avoiding losses that they are willing to pass up considerably larger gains. (Kahneman won the Nobel Prize in Economics in 2002 after Tversky passed away in 1996.) Such "loss aversion" in turn causes individuals to seek compensation for risk that is greater than what would be indicated by expected value of the outcomes. For example, if you were offered a certain $100 or $201 for correctly guessing a coin flip, you *should* prefer the coin flip. Not surprisingly, most people require higher levels of compensation to take the bet.

Numerous studies have been conducted to determine how much additional compensation is required; this is called the loss aversion coefficient. In a 1992 study, Kahneman and

16. For daily calculation, actual closing prices for S&P 500 and 10-Year Treasury are used; daily earnings were derived using same approach as monthly earnings as explained in Note 13.

Tversky reported finding a coefficient equal to 2.25.[17] In other words, people on average were indifferent to a coin flip for $325 versus a guaranteed $100. Other studies found coefficients of loss aversion in the range of 1.43 to 4.8.[18]

Such coefficients are consistent with my RPF findings, in which equities require premiums ranging from 90% to 148% over 10-year Treasury yields (roughly equivalent to loss aversion coefficients between 1.90 and 2.48). And the two concepts appear to have another important similarity. Stock market investors, like the subjects in these studies, appear to expect an incremental return for bearing risk that increases proportionally with the level of the risk-free interest rate. For example, if you were indifferent between $10 guaranteed and $30 on a coin flip, you probably would not accept that same fixed $20 premium over the expected value if the stakes were raised and you were offered a choice between a certain $100 and a contingent $220. Likewise, if the risk-free rate is 4% and the RPF is 1.48, a $1,000 investment in bonds would offer a guaranteed $40 and equities an expected return of $99, or a $59 premium. But if bonds instead yielded 10% and the guaranteed return rises to $100, a $59 premium would probably look much less attractive.

Potential Causes for Shifts in The Risk Premium Factor (RPF)

The RPF has shifted twice in the past 50 years, once in 1981 and again in July 2002. The period from 1960-1981 was characterized by increasing inflation expectations, rising from 1.8% in 1960 to 11.7% in 1981.[19] In 1981, the trend reversed and inflation expectations began to decline. The 1981 shift in RPF from 1.24 to 0.90 could have resulted from this change in inflation expectations driven by world events, with the decline in inflation resulting in higher real after-tax equity returns. Events during 1981 that could have contributed this change include:

• Resolution of the Iran hostage crisis. The reduction of tensions could have increased expectations of stability and a secure oil supply bringing with it lower inflation and less risk of an economic shock.[20]

• Inauguration of the Reagan era, with tax reduction leading to higher real after-tax returns.

At the same time, my analysis shows that the RPF increased from 0.90 to 1.48 in mid-2002. The decline of the rate of long-term inflation ended in 2002, with long-term inflation expectations having declined from a peak of 11.7% in 1981 to 2.0% in 2002. From 2002–2008, the rate of inflation has remained fairly stable, fluctuating in the 2% - 3% range. Other events that could have caused or contributed to the shift in 2002 include:

• Department of Justice investigation into Enron. Enron, Tyco and WorldCom's destruction of confidence in reported earnings may have led to increase risk premium factor.

• The enactment of Sarbanes Oxley in response to accounting scandals. The act faced severe criticism for imposing significant costs on public companies. Some suggested high compliance costs would cause capital to flee to less regulated markets, increasing the premium required for U.S. equities.

• Congressional authorization of war in Iraq. Expectations of a protracted war with Iraq could have increased expectations that increased borrowing to fund the war would lead to increased inflation and tax rates in the future.

Potential Weaknesses in RPF Theory and Methodology

Proper application of the model requires an understanding of its potential weaknesses:

• *All data points are current actual or historical.* While the market is forward looking, all data in the analysis are based on actual results. Even 10-year Treasury yields, which embody expectations about future real interest and inflation, were sampled at a single point in time, along with earnings that are not released until well after the quarter ends. Analysts' estimates are widely accepted as being embodied in current share price and would be expected to be reasonably close to actual before the end of each quarter.

• *Reasons for changes in Risk Premium Factor (RPF) are not fully explained.* The RPF has changed twice over the past 50 years and has historically held for long periods of time. While I have suggested a few possible reasons for the two changes in the RPF over the past 50 years, it is clear that further explanation and understanding is necessary.

• *The RPF may seem to be set arbitrarily to fit actual.* Given the good linear regression fit across a relatively large number of data points, the RPF seems to make sense and provide good result. Nevertheless, this remains a valid concern.

• *RPF cannot be projected.* Thus far it only seems possible to discern the RPF with hindsight. Still this would seem superior to other methods for determining risk premiums that produce less definitive results. For example, if the RPF changed just two times over 50 years, one might argue that in any given year there is a 96% chance (48 out of 50) that the RPF will remain constant over the next year.

17. Kahneman and Tversky. (1992), cited earlier.

18. Abdellaoui, Mohammed, Bleichrodt, Han and Paraschv, Corina, Loss Aversion Under Prospect Theory: a Parameter-Free Measurement (October 2007). Management Science, 10:1659-1674.

19. Calculation of inflation expectations based on difference between 10-Year Treasury yield and assumed 2% long-term real interest rate

20. "1981: Tehran frees US hostages after 444 days" BBC Website, Accessed March 15, 2009, http://news.bbc.co.uk/onthisday/hi/dates/stories/january/21/newsid_2506000/2506807.stm.

Declining Interest Rates Explain More than Half of S&P 500 Index Growth Since 1981

Interest rates are much more important than is generally recognized. Some contend that the effects of interest rates on corporate values are limited to the direct impact on corporate borrowing and consumer spending. Such observers tend to argue that although the cost of capital rises with inflation, for the market as a whole, the negative effect of this increase is directly offset by the positive effects of inflation on earnings. In other words, in the equation $V = E / (C – G)$, since C and G increase by the same amount (inflation), the expected impact of inflation is zero.

By contrast, the RFP Model suggests that since the ERP increases proportionally with the risk-free rate, it rises faster than the growth in earnings, causing a decline in valuations. So, in addition to the direct negative impact of interest rates on earnings, higher rates also have a large impact on P/E multiples.

The highest monthly finish of the S&P 500 was October 2007, when it closed at 1549. The highest annual finish of the risk-free rate was 1981, when the 10-year Treasury yield ended the year at 13.7%. Between these two mileposts, the S&P 500 Index increased 1264%, from 122 to 1549. During the same period, S&P Operating Earnings increased only 588%, rising from 15.2 to 89.3. Thus, earnings accounted for only 47% (588%/1264%) of the growth of the S&P 500 during this period.

And since the increase in S&P earnings account for less than half of the increase in its value, much of the remaining increase can be attributed to decreases in the risk-free rate—and with the 10-year Treasury yields falling to 4.47% in October 2007, the cost of capital dropped from over 26% at the end of 1981 to about 11% in 2007. And according to the RPF model, over 50% of the appreciation over the past 29 years is explained by reductions in both the RPF and risk-free rate. More specifically, the model provides a way of explaining the remarkable increases in corporate P/E multiples since the 1960s—one that relies largely on changes in interest rates (which embody expected inflation) during that period.

The RPF Model and Market Efficiency: Exploring Major Market Events From 1986–2009

The RPF Model can help demystify valuation and also help explain major market vents over the past 20 or so years. The exploration of these events may also serve to shed some light on the efficient market hypothesis.

The Efficient Market Hypothesis (EMH) was first fully proposed by Eugene Fama in his doctoral thesis at the University of Chicago in the 1960s. In short, it states that the markets are "informationally efficient" in the sense that all available information is incorporated in the current stock price. The implication is that since all information is embodied in the current price, it should be difficult for investors to beat the market year in and year out.

Over time it has been much debated and variations have emerged that allow exceptions for holders of private information (say, management) small stocks that are not heavily traded. The EMH has been much criticized, particularly by professional money managers who would be out of work if the market were perfectly efficient. After all, if the pros can't outperform the market, why not just buy index funds?

Many people take the EMH to mean that the markets are *always* right. Today even Fama admits the market makes mistakes: "In a period of high uncertainty, it's very difficult to figure out what the right prices are for stocks."[21] And Ken French, a frequent collaborator with Fama and Professor at the Tuck School of Business at Dartmouth, said in an interview jointly conducted with Fama that:

> *The efficient market hypothesis is just a model and, like all interesting models, it is not literally true. There are mistakes in prices even if one considers just publicly available information and, since people use financial prices to help decide how to allocate resources, those mistakes must affect the underlying reality. Of course, the existence of mistakes does not imply they are easy to find.*[22]

How the RPF Valuation Model Explains October 19, 1987 (Black Monday)

U.S. and global markets plunged on October 19, 1987, with the S&P 500 declining more than 20%. The cause of the decline has been much discussed, with program trading often cited as the main culprit along with portfolio insurance (derivatives).[23]

The application of the RPF Model to this period is revealing. As shown in Figure 6, which shows actual versus predicted S&P levels,[24] the market appears to have gotten "ahead of itself"—thereby creating a bubble of sorts—in anticipating an increase in earnings and values. As can be seen in Figure 7, interest rates began to climb in March 1987, rising from 7.25% in March to 9.25% in October, driving down the predicted P/E and the predicted level of the S&P 500.[25] Yet despite flat earnings, the market grew by 12% from February to September (and a total of 25%

21. "CBS Money Watch, http://moneywatch.bnet.com/investing/article/eugene-fama-why-you-cant-time-the-market/277142/.

22. "Fama/French Forum" http://www.dimensional.com/famafrench/2009/04/qa-bias-in-the-emh.html.

23. "Black Monday 10 Years Later: 1987 Timeline," *The Motley Fool Website*, accessed March 2009, http://www.fool.com/features/1997/sp971017crashanniversary-1987timeline.htm.

24. See Note 13.

25. See Note 14.

Figure 6 **Actual vs. Predicted During October 1987 Crash**[32]

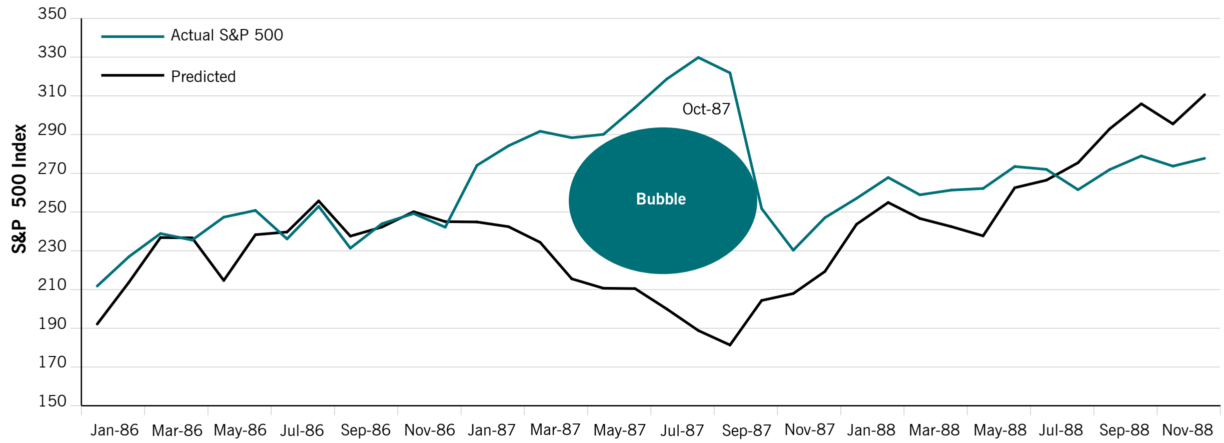
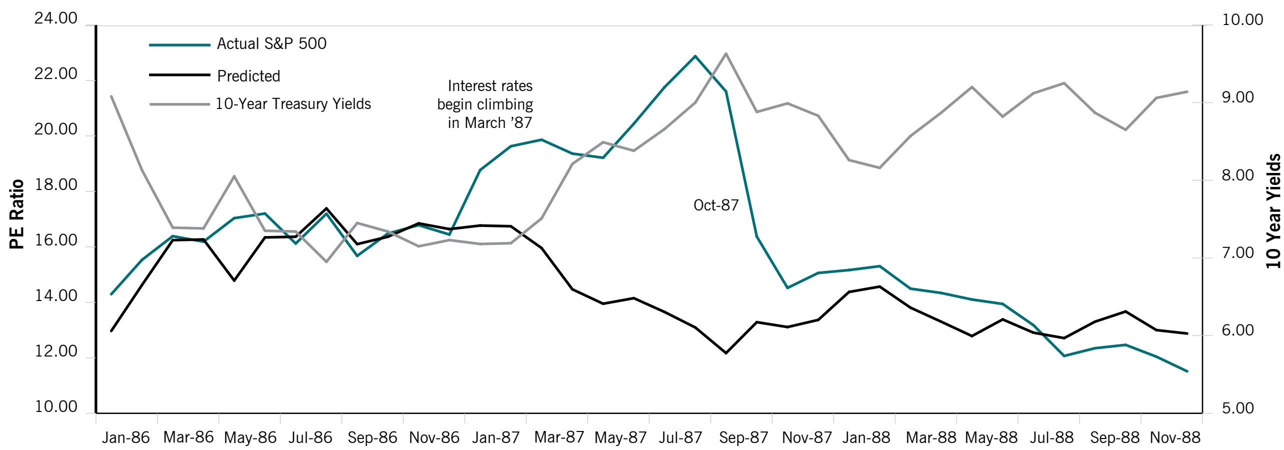


Figure 7 **Interest Rate Impact on October 1987 Crash, Actual S&P 500 Month-end data–10-Year Treasury Yields**



from December). With the market crash in October, the predicted and actual fell back into parity, with both figures suggesting the creation and bursting of a bubble.[26]

The suggestion offered by the RPF model in this case is that the underlying cause of the crash was excessive valuation relative to the sharp rise in interest rates. While actual and predicted levels often deviate, without a shift in the RPF, they tend to fall back in line.

But why did the market fall on October 19 and not November 19? The market began its decline in August. During the days before October 19, Iran had attacked a U.S flagged tanker, exacerbating fears that oil prices would continue to rise.[27] Perhaps this solidified the belief that earnings would not rise and inflation would stay high, keeping interest rates high. And this point of view was rapidly assimilated into the market. My own belief is that these developments were nothing more than the pinpricks that popped the balloon—actions that, while not particularly momentous in and of themselves, were enough to cause an unbalanced state to return to a more sustainable equilibrium. While derivatives and program trading may have aggravated the market decline once the decent began, they were not the fundamental cause, but rather part of the mechanism that helped to restore equilibrium.

26. See Note 14.

27. "Iranian Attacks on Kuwaiti Port Called Cause for U.S. to Retaliate," *The New York Times*, October 18, 1987, http://www.nytimes.com/1987/10/18/world/iranian-attacks-on-kuwaiti-port-called-cause-for-us-to-retaliate.html.

Figure 8 **Actual vs. Predicted during the 2000 dot.com Bubble, S&P 500 Month-end data–10-Year Treasury Yields**

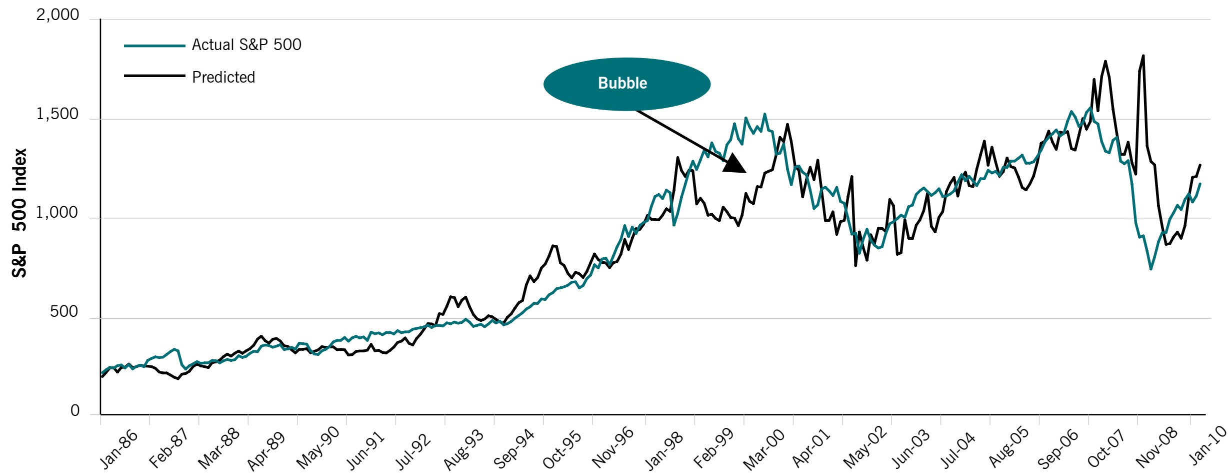


Figure 9 **NASDAQ January 1999–May 2002**

Source: Yahoo! Finance

2000 "Dot Com" Bubble: RPF Model Suggests Significant Bubble for the S&P 500

The NASDAQ peaked on March 10, 2000, at 5,132 in what is widely considered to be a bubble driven by excessive valuations of the Internet and other technology companies. Many economists such as Robert Schiller, author of *Irrational Exuberance*, argued that the entire market was embroiled in a speculative bubble throughout this period.[28]

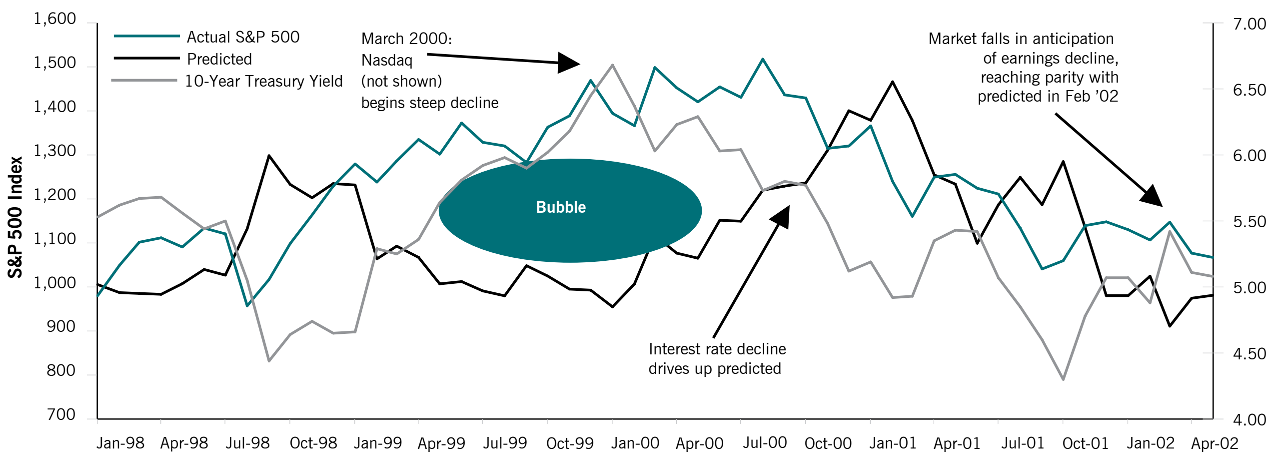
Application of the RPF Model to the S&P 500, strongly suggests that a significant bubble did exist. Indeed, Figure 8 suggests that the dot.com bubble of the late 90s was by far the largest during the period 1986 through 2009.

The model was not applied to the NASDAQ because it would be inappropriate to assume that the long-term growth of the smaller cap and technology heavy NASDAQ would equal long-term GDP growth and that volatility (Beta) would be the same as the S&P 500. As shown in Figure 9, the NASDAQ had declined by 32% in mid-April 2000 from its March 10 high, and by 51% by the end of 2000.

What explains this plunge in prices? From November 1998 until March 2000, 10-year Treasury yields increased from 4.6% to 6.2%. While the NASDAQ began to run up in late 1999, as can be seen in Figure 10, the S&P 500 Index began to diverge from RPF Model predictions in January

28. Robert J. Schiller, *Irrational Exuberance*, (Princeton University Press).

Figure 10 **Dot.com Bubble Close Up, Actual S&P 500 Month-end data–10-Year Treasury Yields** 32



1999. As also shown in the figure, the S&P 500 Index did not begin its decline until August 2000. (Remember the model is applied using actual reported operating earnings, so predicted levels at any point are backward looking and do not reflect expectations.) However, the market began to anticipate that the NASDAQ meltdown would have a negative impact on earnings and the index followed.[29] And since S&P earnings fell by 27% from March 2000 to December 2001, the RFP Model appears to have "signaled" that earnings would fall well in advance of the actual reported drop.

The implication, then, is that the bubble was created by the combination of inflated earnings levels with rising 10-year Treasury yields that the market was somehow slow to recognize. To the extent the increases in interest rates were orchestrated by the Fed to cool an overheating economy, investors may have misread the signal and expected the increase in interest rates to be temporary. But, as the rate increases began to affect earnings, the market began a sharp repricing as the new point of view was assimilated.

How the RPF Valuation Model Explains 2008–2009 Meltdown and Recovery

The bursting housing bubble and mortgage crisis ultimately led to the meltdown that began September 2008. By August 2008, the S&P 500 had already fallen by 16% from its May 2007 peak. During this period, 10-year Treasury yields declined from around 5% to less than 4%. As illustrated in Figure 11, this led to an increase in predicted levels of the S&P 500 index.

According to the Case-Schiller Home Price Index, home prices fell more than 10% from second quarter of 2006 to the fourth quarter of 2007 and a total of 18% by the second quarter of 2008.[30] This historically large decline led to (well-founded) concerns about financial instability and the elimination of an important source of disposable income. Once again, in anticipation of a decline in earnings, the S&P 500 index fell while the RPF Model (using reported operating earnings) showed an increase in predicted levels as interest rates declined. The lines for expected and actual S&P values in Figure 11 begin to converge in August 2008, just before the worst of meltdown began in September and October. Investors were unable to absorb the seriousness of the pending crisis, so while the market fell in anticipation of an earnings decline, the expectations did not come close to reflecting the magnitude of the situation.

As can be seen in Figure 11, the flight to quality and resulting drop in Treasury rates clearly drove up the predicted levels to abnormal highs. But, as interest rates returned to a more normal level by June 2009, the predicted and actual levels returned to parity.

RPF Model implications for efficient markets?

• Over a longer period of time, the market is efficient if one allows for oscillations around true value, but is also subject to making mistakes. These mistakes can create bubbles.

• Over time the bubbles are deflated and the market returns to predicted levels as new long-term views are assimilated.

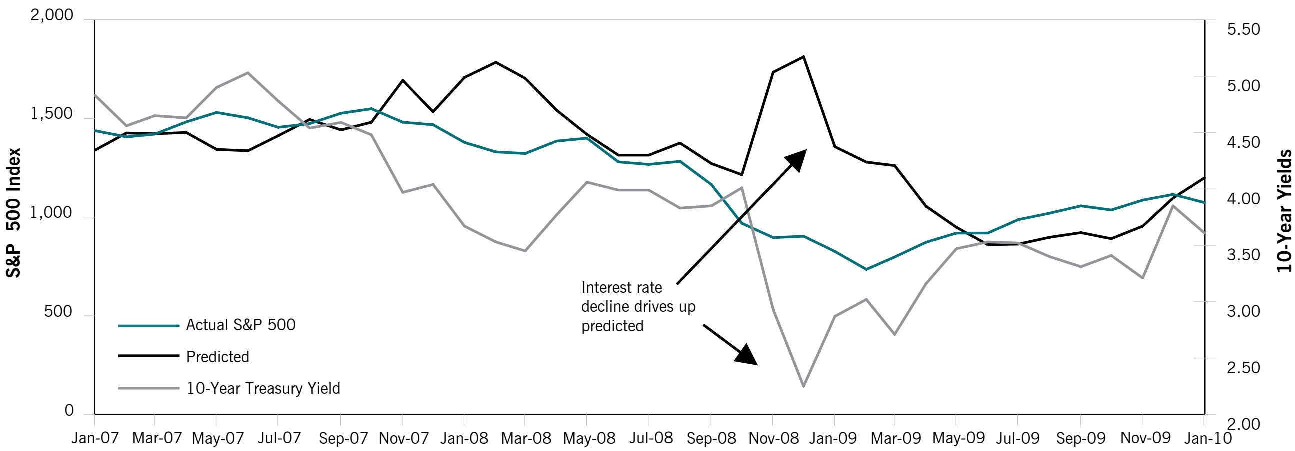
• The RPF Valuation model has shown to be useful in identifying bubbles before they pop.

This pattern supports the contention that the valuation model would have worked well during this period with a

29. See Note 13.

30. "S&P/Case-Schiller Home Price Indices," *Standard and Poors Website*, accessed March to April 2009, http://www2.standardandpoors.com/spf/pdf/index/csnational_values_022445.xls.

Figure 11 **Actual vs. Predicted During 2008–2009 Meltdown, S&P 500 Month-end data–10-Year Treasury Yields**



normalized interest rate. It also shows how the market led predicted levels as it incorporated expected rather than actual historical operating earnings.

In sum, analysis of these major market events with the RPF Model supports the contention that markets make mistakes in processing information. It also suggests that market prices oscillate around a true fair value price. But, as highlighted throughout this discussion of three major market events, these deviations can be very large.

2010 Outlook

As of this writing, on April 14, 2010, the S&P 500 Index closed at 1,211, as compared to a predicted level of 1,260—still 4% below the predicted level. In addition to looking at the market today, the model can help inform an opinion about the future. S&P estimates 2010 operating earnings of $75.27. If we also assume the 10-year Treasury remains unchanged at 3.83%, the S&P 500 Index would be predicted to end the year at 1,485—a gain of another 23%. But if the bond rate rises to 5%, even with the growth in earnings, the S&P's predicted value at year end is 1,107—a drop of 9% from the current level.

Conclusions

Many people view the market valuation process as a black-box driven by emotion, leaving many managers unsure what strategies they can pursue to increase shareholder value. Using two main variables, the RPF Valuation model highlights a number of important principles that can be used to inform the valuation of all companies in most (though not all) circumstances:

1. The Equity Risk Premium is not a constant, but a relatively stable Risk Premium Factor (RPF) that is applied to the risk-free rate (10-year Treasury yields).

2. The Risk Premium Factor is consistent with the loss aversion coefficient associated with the prospect theory (of Kahneman and Tversky).

3. The Risk Premium Factor Valuation Model [$P = E / (R_f \times (1+RPF) - (R_f - Int_R + G_R))$] effectively explains both P/E and S&P 500 Index levels using readily available information and simplifying assumptions.

4. Growth is a critical component of valuation, and the impact of growth on value is easily quantified using the RPF model.

5. Interest rates drive market value—and the fair value of the market (P/E Ratio) cannot be estimated without considering interest rates.

6. Interest rates have a greater impact on market price and valuation than is generally recognized, with low rates more beneficial and high rates more punishing.

7. Declining interest rates were a major factor in the long bull market from 1980 through 2007.

8. The RPF model suggests that if Treasury yields remain in the low 4%–5% range and earnings recover to 2006/07 levels, the market could stage a rally and recover to record levels, with the S&P 500 Index rising to the range of 1,300–1,700.

9. Though efficient and rational over longer time periods, the market is prone to occasional, generally short-lived oscillations and pricing errors.

STEVE HASSETT is president of Hassett Advisors based in Atlanta, Georgia, which specializes in corporate development and growth strategies. Previously, he was VP-international and emerging businesses at the Weather Channel, founder of a Web and mobile software company, and a corporate finance consultant with Stern Stewart & Co.

Journal of Applied Corporate Finance (ISSN 1078-1196 [print], ISSN 1745-6622 [online]) is published quarterly, on behalf of Morgan Stanley by Wiley Subscription Services, Inc., a Wiley Company, 111 River St., Hoboken, NJ 07030-5774. Postmaster: Send all address changes to JOURNAL OF APPLIED CORPORATE FINANCE Journal Customer Services, John Wiley & Sons Inc., 350 Main St., Malden, MA 02148-5020.

Information for Subscribers *Journal of Applied Corporate Finance* is published in four issues per year. Institutional subscription prices for 2010 are:

Print & Online: US$416 (US), US$499 (Rest of World), €323 (Europe), £255 (UK). Commercial subscription prices for 2010 are: Print & Online: US$556 (US), US$663 (Rest of World), €429 (Europe), £338 (UK). Individual subscription prices for 2010 are: Print & Online: US$105 (US), £59 (Rest of World), €88 (Europe), £59 (UK). Student subscription prices for 2010 are: Print & Online: US$37 (US), £21 (Rest of World), €32 (Europe), £21 (UK).

Prices are exclusive of tax. Asia-Pacific GST, Canadian GST and European VAT will be applied at the appropriate rates. For more information on current tax rates, please go to www3.interscience.wiley.com/about us/journal_ordering_and_payment.html#Tax. The price includes online access to the current and all online back files to January 1997, where available. For other pricing options, including access information and terms and conditions, please visit www.interscience.wiley.com/journal-info.

Journal Customer Services:For ordering information, claims and any enquiry concerning your journal subscription please go to interscience.wiley.com/support or contact your nearest office.
Americas: Email: cs-journals@wiley.com; Tel: +1 781 388 8598 or +1 800 835 6770 (toll free in the USA & Canada).
Europe, Middle East and Africa: Email: cs-journals@wiley.com; Tel: +44 (0) 1865 778315.
Asia Pacific: Email: cs-journals@wiley.com; Tel: +65 6511 8000.
Japan: For Japanese speaking support, Email: cs-japan@wiley.com; Tel: +65 6511 8010 or Tel (toll-free): 005 316 50 480. Further Japanese customer support is also available at www.interscience.wiley.com/support
Visit our Online Customer Self-Help available in 6 languages at www.interscience.wiley.com/support

Production Editor: Joshua Gannon (email:jacf@wiley.com).
Delivery Terms and Legal Title Prices include delivery of print journals to the recipient's address. Delivery terms are Delivered Duty Unpaid (DDU); the recipient is responsible for paying any import duty or taxes. Legal title passes to the customer on despatch by our distributors.

Back Issues Single issues from current and recent volumes are available at the current single issue price from cs-journals@wiley.com. Earlier issues may be obtained from Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. Tel: +1 518 537 4700, Fax: +1 518 537 5899, Email: psc@periodicals.com

This journal is available online at Wiley InterScience. Visit www.interscience.wiley.com to search the articles and register for table of contents e-mail alerts.

Access to this journal is available free online within institutions in the developing world through the AGORA initiative with the FAO, the HINARI initiative with the WHO and the OARE initiative with UNEP. For information, visit www.aginternetwork.org, www.healthinternetwork.org, www.healthinternetwork.org, www.oarescience.org, www.oarescience.org

Abstracting and Indexing Services
The Journal is indexed by Accounting and Tax Index, Emerald Management Reviews (Online Edition), Environmental Science and Pollution Management, Risk Abstracts (Online Edition), and Banking Information Index.

How the Risk Premium Factor Model and Loss Aversion Solve the Equity Premium Puzzle

Stephen D. Hassett

Hassett Advisors
www.HassettAdvisors.com
SHassett@HassettAdvisors.com

This Draft: September 22, 2010
First Draft: September 20, 2010

Abstract

The term “equity premium puzzle” was coined in 1985 by economists Rajnish Mehra and Edward C. Prescott. The equity premium puzzle in considered one of the most significant questions in finance. A number of papers have explored the fundamental questions of why the premium exists and has not been arbitraged away over time. This paper expands upon the findings implicit in the Risk Premium Valuation Model (Hassett 2010) that the equity risk premium is a function of risk free rates. Since 1960 the equity risk premium has been 1.9 – 2.48 times the risk free rate. The long term consistency of this relationship with loss aversion coefficients associated with Prospect Theory (Kahneman and Tversky, 1979) suggest it as a solution to the equity premium puzzle and support the experimental findings of Myopic Loss Aversion (Thaler, Tverseky, Kahneman and Schwartz, 1997).

Introduction

The equity premium puzzle in considered one of the most significant questions in finance. The term "equity premium puzzle" was coined by Mehera and Prescott in their 1985 paper, "The Equity Premium, A Puzzle,"[1] referring to the inability to reconcile the observed equity risk premium with financial models.

In the analysis, they use short-term treasuries as the risk free rate to calculate the real return on equities over numerous historical periods. They conclude that on average short-term treasuries have produced a real return of about 1% over the long-term, while equities have yielded 7%, implying a premium of about 6% or seven times the risk free return. Unable to reconcile a 7 x premium with financial models, they term it a puzzle.

Since then numerous papers have also attempted to explain the difference, including Shlomo Benartzi; Richard H. Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle"[2] which attempts to explain it in relation of loss aversion as first described in a paper by Daniel Kahneman and Amos Tversky in 1979.[3] They state:

> "The second behavioral concept we employ is mental accounting [Kahneman and Tversky 1984; Thaler 1985]. Mental accounting refers to the implicit methods individuals use to code and evaluate financial outcomes: transactions, investments, gambles, etc. The aspect of mental accounting that plays a particularly important role in this research is the dynamic aggregation rules people follow. Because of the presence of loss aversion, these aggregation rules are not neutral."

Our mental accounting for gains and losses determines how we perceive them.

Loss Aversion

Loss aversion refers to the fact that people are more sensitive to decreases in wealth than increases. Empirical estimates find that losses are weighted about twice as strongly as gains (e.g., Tversky and Kahneman (1992)[4]; Kahneman, Knetsch, and Thaler (1991)[5], Thaler, Tversky, Kahneman, Schwartz (1997)[6]). The pain of losing $100 is roughly twice the perceived benefit of gaining $100, so on average their subjects required equal odds of winning $200 to compensate for the potential loss of $100. In other words, the average subject required a gain of twice the potential loss to take a gamble that had equal chance of loss or gain. This was in stark contrast to the belief that people, as rational beings, evaluated the expected value and would be indifferent to a chance of gaining $100 to losing $100 if the odds were 50/50; if the gain were tilted to be slightly favorable they should take the bet. In reality, losing hurts more; people on average do not find the prospect of gaining $101 along with an equal chance of losing $99 to be an attractive wager. In their experiments, they found that subjects required about $200 to be willing to accept the 50/50 proposition of losing $100. Kahneman won the Nobel Prize in Economics in 2002 after Tversky passed away in 1996. Of course all people do not behave this way all the time, otherwise Las Vegas would not exist!

Loss Aversion and Corporate Decision Making

Incorporating loss aversion into financial thinking is in many ways a significant departure from how finance is often taught and practiced. In business school, I was taught to rely on net present value and expected value. A project with positive net present values should be pursued and that when faced with a range of outcomes, the expected value can be calculated by assigning probabilities to each outcome. The mantra: Pursue all NPV positive projects.

My experience has been that the business world rarely works this way. Due to corporate as much as individual loss aversion, decision makers are often much more risk averse, viewing the consequence of failure much greater than the rewards for success. Investments that have only slightly positive NPV or expected value are usually not pursued. Even the more risk tolerant individuals would tend to avoid risk if the organization takes a very dim view of loss.

This is why it is so important for organizations to employ incentive structures that reward sustainable growth in value and prudent risk taking. My own experience is that organizations without such incentives tend to be very risk averse. When decisions come down the internal calculus that investing successfully results in no reward, while failure results in unemployment or at least limited advancement, investment and growth are sure to slow. I would also argue that this also explains risk taking for traders on Wall Street where outsized rewards are given for success compared to the stigmas and punishments for failure. It's not that traders have high tolerance for risk, it's that in using OPM (Other People's Money) the penalty for failure is small.

Attempts to Solve The Equity Premium Puzzle

As discussed above, Mehra and Prescott(1985) coined the phrase "Equity Premium Puzzle" because they estimated that investors would require a very high coefficient of relative risk aversion (of the order of 40 or 50) to justify the observed equity risk premium of 7%. Mehra and Prescott revisited the topic two decades later with their 2003 paper, "The Equity Premium in Retrospect" where they continued to try and solve the puzzle by comparing real returns and ask whether the equity premium is due to a premium for bearing non-diversifiable risk. They conclude the answer is no unless you assume the individual has an extreme aversion to risk; many times higher than the 2x return seen in the lab.

They approach the problem using a general equilibrium model and compared short-term real risk free rates to observed equity premium. While I am not in a position to opine on the use of these models in evaluating equity premium, for several reasons I will discuss shortly, I believe that the use of short-term real rates is mistaken. I am not surprised they could not explain the rational for investors to such a dramatic disparity, since in my opinion they are not making the right comparison. Rather than using short-term real rates, they should be using long-term nominal rates.

What they did was a bit like measuring the speed of one moving vehicle from another moving vehicle. If Car A is moving at 60 mph and Car B is behind it at 66 mph and car C is next traveling at 61 mph, car C will see itself gaining on car A at just 1 mph. From the perspective of car C, car B is gaining on car A at a rate of 6 mph or 6 x faster than itself. This is all fine unless we care about their speed relative to a neutral observer who is not moving. Relative to the neutral observer, Car B is only going 10% faster than Car A.

Mehra and Prescott did not pick the right relative observation point. By using real returns they are measuring the difference from a moving vehicle. If we look at this from the perspective of real returns then the relative premium looks huge. But if we look at from the perspective of nominal returns, the neutral observer, then the premium it is not unreasonable. This is consistent with both the way individuals have been shown to evaluate gains and losses and with financial theory.

The mental accounting of investors focuses on the nominal returns. It's what investors track and how money managers are compensated. So it makes sense that that proper basis for evaluating the risk premium relative to the risk free rate is long-term nominal returns. For example, let's assume inflation is 2%. If an investor is considering a $1,000 investment with Treasuries at 4%, the yield is guaranteed to be $40 per year with a full return of principal. While the investor is exposed to interim fluctuations in value, the coupon and return of principal are guaranteed. Alternatively, the same investor considering an investment in the S&P 500 Index, would be evaluating the expected return relative to the nominal long-term rate rather than the real short term rate. In this case, expected equity returns of 10% would look good, yielding on average $100 per year rather than $40. If we calculate real returns by subtracting the 2% inflation, the $80 return for equities dwarfs the $20 for treasuries.

Now let's assume that expected inflation rises to 6% and the risk free rate jumps to 8%, so a new $1,000 bond would yield $80. If you applied the same 6% premium for equities, you get an expected yield of $140. Sure the real returns are the same, but doesn't the risky $140 look less attractive compared to a guaranteed $80?

Is it the right thing to track? Maybe not, but it is the reality. If investors compare their returns on equities to the nominal return of other investments, any attempt to explain the premium must compare the relative return as perceived by investors. Nominal not real returns should be used.

Long-term Treasury rates are used in determining cost of capital since they embody the market's best guess on long-term inflation. Even though this means they are not truly risk free, it is the best market estimate of expected interest rate and inflation risk; it is the right reference point. While it's true that using real equity returns accounts for the actual inflation component, it does not account for interest rate risk. In order account for expected inflation, most practitioners use long-term treasuries as the risk free rate. In doing so, they also incorporate a risk factor for interest rates.

Required return can be thought of as follows:

Nominal Equity Return = Real Equity Return + Inflation **(1)**
= Short-term Risk Free Rate + Inflation + Interest Rate Risk Premium + Equity Risk Premium **(2)**

If you subtract inflation from both sides to derive the real required return, you are still left with interest rate risk, which includes risk of unexpected inflation. So by using real equity returns and short-term risk free rate, you still have to account for the interest rate risk premium.

Real Equity Return = Short-term Risk Free Rate + Interest Rate Risk Premium + Equity Risk Premium **(3)**

Essentially, what Mehra and Prescott were calling the equity risk premium, was really the equity risk premium plus the interest rate risk premium.

Some believe that interest rates do not have a material impact on equity returns since inflation will result in earnings growth and since equities are priced as a multiple of earnings, as earnings grow equity prices increase with inflation. As I will discuss later, inflation has a huge impact on equity prices.

In "Myopic Loss Aversion and The Equity Premium Puzzle," Benzarti and Thaler (1995) they posit that the high degree of loss aversion is due to "myopic loss aversion" in that investors are sensitive to interim losses as equity markets fluctuate. They suggest that investors look at nominal returns since that is what is reported, therefore that's what investors look at. They find that a loss aversion factor of 2.25 to 2.78 is consistent with observed risk premiums if investors evaluate their portfolios about once a year and overall results are very sensitive to frequency of evaluation. In "The Effect of Myopia and Loss Aversion on Risk," Thaler, Tversky, Kahneman, Schwarts (1995), looked at this question through lab experiments found that subjects were more loss averse when they evaluated their returns more frequently and that they viewed guaranteed outcomes as a reference point with an evaluation period of about one year (13 months). In other words, investors evaluate their portfolios annually and expect a premium proportionate to the nominal risk free rate. As we will see below the RPF Valuation Model provides real world support for these findings.

Determining the Equity Risk Premium

In introducing the Risk Premium Valuation Model[7] (Hassett 2010), I posited that rather than being a fixed premium, the Equity Risk Premium fluctuates with the risk free rate, maintaining a constant proportionate relationship. The Equity Risk Premium equaled the Risk Free Rate times a constant factor. That factor (Risk Premium Factor) ranged from 0.9 – 1.48 between 1960 and today. So substituting into the formula where Cost of Equity = Rf + ERP,

Cost of Equity = Risk Free Rate + Risk Free Rate x Risk Premium Factor (RPF) **(4)**

Simplifying to:

Cost of Equity = Risk Free Rate x (1 + RPF) **(5)**

The RPF does not change frequently. In fact it has shifted only twice since 1960:

Period	RPF
1960 – 1980	1.24
1981 – Q2 2002	0.90
Q3 2002 – Present	1.48

Table 1: Estimated Risk Premium Factors

A Risk Premium Factor of 0.9 – 1.48, means Cost of Equity equals the Risk Free Rate times 1.9 – 2.48, very close to the findings on loss aversion factors.

The factor was determined by applying a set of simplifying assumptions to the constant growth formula:

$$P = E / (C - G) \text{ or } P/E = 1 / (C - G) \qquad (6)$$

Variables and assumptions used are as follows:

P = Price (Value of S&P 500)

E = Actual Earnings (Annualize operating earnings for the prior four quarters as reported by S&P). Earnings, while not ideal, are used as a proxy for cash flow and seem to work very well

G = Expected long term projected growth rate, which is broken down into Real Growth and Inflation, so $G = G_R + I_{LT}$

G_R = Expected long-term real growth rate. Long-term expected real growth rate (G_R) is based on long-term GDP growth expectations on the basis that real earnings for a broad index of large-cap equities will grow with GDP over the long-term. A rate of 2.6% is used with the same rate applied historically.[8]

I_{LT} = Expected long-term inflation, as determined by subtracting long-term expected real interest rates (Int_R) from the 10 Year Treasury, where Int_R is 2%; based on the average 10 Year TIPs Yields from March 2003 – present.[9]

C = Cost of Capital is derived using Capital Asset Pricing Model, where for the broad market, $C = R_f + ERP$

R_f = Risk Free Rate as measured using 10 Year Treasury yields

ERP = Risk Premium Factor (RPF) x R_f

RPF = 1.24 for 1960 – 1980; 0.90 for 1981 – 2001; and 1.48 for 2002 – present. The RPF for each period was arrived at using a linear regression to fit the assumptions above to actual PE. All data used in the analysis is available for download at: http://sites.google.com/a/hassett-mail.com/marketriskandvaluation/Home

Including all assumptions, the formula reduces to:

$$P = E / (R_f \times (1+RPF) - (R_f - Int_R) - 2.6\%) \qquad (7)$$

$$\text{Or } P/E = 1/ (R_f \times (1+RPF) - (R_f - Int_R) - 2.6\%) \qquad (8)$$

The model explains stock prices from 1960 - 2009 with R Squared around 90%[10] to actual index levels from 1960 – 2009 as shown in graph below.

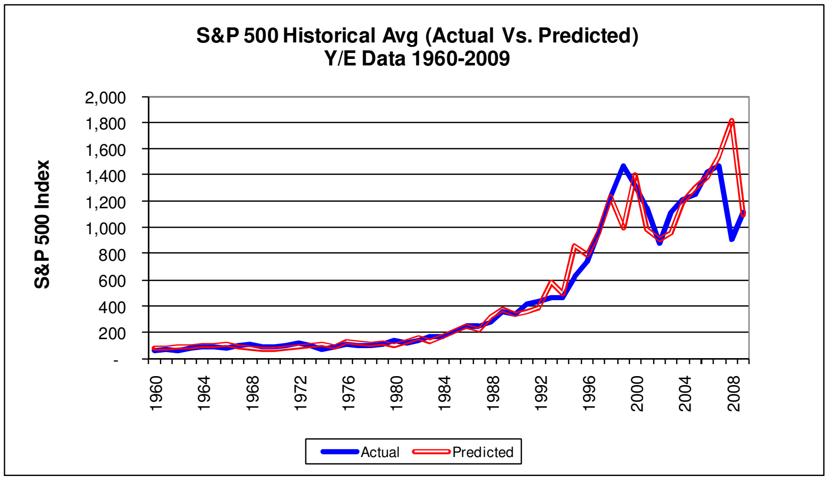


Figure 1: S&P 500 Actual vs. Predicted - 1960- 2009

The model only works if we assume that the Equity Risk Premium is conditioned on the Risk Free Rate, meaning that it gets bigger when the Treasury yields increase and smaller when they shrink. In fact one reason that I suspect many studies compared real returns, rather than nominal returns, may be the belief that inflation does not impact valuation. One common belief is that since profits will grow with inflation, inflation does not matter when discounted back. Another look at the constant growth equation can help understand this thinking:

P / E = 1 / (C – G), where **(9)**

C = Rf + ERP **(10)**

G = Real Growth + Expected Inflation **(11)**

Rf = Real Interest Rate + Expect Inflation **(12)**

We can restate the equation for P/E as:

P/E = 1/ ((Real Interest Rate + Expect Inflation) – (Real Growth + Expected Inflation), **(13)**
Expected Inflation is canceled out and:

P/E = 1/ (Real Interest Rate + Real Growth) **(14)**

Since we assume the Real Interest Rate and Real Growth are a constant over the long term, P/E is also a constant. And, this would be true if the Equity Risk Premium were a constant. But if we assume that the Equity Risk Premium moves with the Risk Free Rate, then we get the relationship charted above, which is a very good fit with historical data.

Impact of Inflation on Value

Some argue that inflation should not have an impact on equity values, since higher costs can be passed on in the form of higher prices, so on average, earnings growth should keep up with inflation. If you assume P/E ratios should be a constant, say, 19 then with earnings of $2.00 share a company would trade at $38.00. With 5% inflation, earnings would grow to $2.10 and the share price to $39.90 – a gain of 5% which just matches inflation.

We get the same result using a constant growth model and a fixed Equity Risk Premium. Let's assume the Equity Risk Premium is 6%, the Risk Free Rate is 7%, which embodies 5% inflation, and real long term growth rate of 2.6%. Using the formula P/E = 1 / (C-G) we get, P/E = 1 / ((7%+6%) – (5%+2.6%) for a P/E of 18.5. If we lower the inflation rate to 2% the risk free rate drops to 4% and we calculate P/E = ((4%+6%)-(2%+2.6%) = 18.5. As shown earlier, any change inflation cancels itself out.

However, if we derive the Equity Risk Premium using the RFP Model, then the Equity Risk Premium varies with inflation. More inflation results in a higher risk premium. Using a 2% real interest rate, Table 2 below demonstrates the impact of inflation on P/E:

Inflation	R_f	ERP	Cost of Equity	G	Predicted P/E
2.0%	4.0%	5.9%	9.9%	4.6%	18.8
3.0%	5.0%	7.4%	12.4%	5.6%	14.7
4.0%	6.0%	8.9%	14.9%	6.6%	12.1
5.0%	7.0%	10.4%	17.4%	7.6%	10.2
6.0%	8.0%	11.8%	19.8%	8.6%	8.9

Table 2: Inflation Drives Valuation

Since investors expect a proportionately higher return over risk free, as inflation rises they apply a greater discount to future earnings, resulting in a lower present value, resulting in a lower multiple.

Back to Loss Aversion

We know that individuals have different tolerances for risk. If the RPF is 1.48, that implies the market as a whole has a loss aversion coefficient of 2.48. That is the average of all investors, not every individual. We would expect some to have lower coefficients and others higher. Gambling addicts destroy their own lives, knowing the odds are not better than even, implying a loss aversion coefficient of less than 1.0. Likewise, some people are more risk averse than average. This is one of the factors that act to set price.

The prices for individual stocks are set at the margin. For example, Google closed today at $476 and traded about 2.5 million shares. But with 320 million shares outstanding, that is less than 1%. The price is set by the investors trading that 1%. The implication is that the owners of the remaining 99% think Google is worth more than the current $476 and some number of investors would be willing to buy Google at a lower price. Mechanically the way this works is that sellers offer to sell a number of shares at a certain price, called the Ask, and potential buyers offer to buy at a specified price, called the Bid. The Bid for Google might be 200 shares at $476.07 and the Ask 700 shares at $476.18. The difference, $0.11 in this case, is called the Bid-Ask spread. These are the current best offers to buy and sell. For high volume stocks like Google, the Bid-Ask spread is small, just 0.02% in this case. For lower volume equities the spread will generally be higher.

If an investor places a marker order to, say, buy 500 shares, the first 200 shares will be filled at the current Bid price for 200 shares at $476.17. The remaining 300 shares will be filled by the next best ask price, which will be $476.17 or higher. It is not the consensus or average estimate of value that determines the price, but the price at which investors at the margin are willing to buy or sell at any moment. So if I don't own shares of Google and I think it's worth just $400 or even $100, I am not a factor in setting the price. But if in the moment described above, I enter a bid for 200 shares at $476.18, the order is immediately filled and, for that moment, I am the price setter.

Similarly, investors with loss-aversion coefficients at the extremes should not be expected to have much market impact. An investor with a loss aversion coefficient well above 2.5 will be risk averse and have portfolio skewed towards government bonds, while and investor with a loss aversion coefficient near 1.0 will always have a portfolio that is mostly equities. Therefore neither will have much impact on price setting. On the other hand, investors with loss aversion coefficients around 2.5 will be more likely to be shifting their portfolios between bonds and equities and have a larger impact on pricing.

Conclusion

Loss aversion is hard wired into us and drives a number of decision processes that seems to include how investors set prices in the stock market. Thaler, Tversky, Kahneman, Schwarts (1995) found evidence of what they called Myopic Loss Aversion and demonstrated the expectations of risk premiums were consistent experimental findings for loss aversion if portfolios were evaluated annually. The Risk Premium Factor Valuation Model (Hassett 2010) provides real world evidence that the market actually behaves this way. Combing evidence that the risk premium varied with the risk free rate in a proportion consistent the findings in behavioral studies, suggests that Loss Aversion is the answer to the equity premium puzzle.

Endnotes

[1] Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economies*, vol. 15, no. 2 (March):145–161.

[2] Benartzi, S., and R. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics*, vol. 110, no. 1 (February):73–92.

[3] Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, vol. 46, no. 2 (March):171–185.

[4] Kahneman, D., A. Tversky. 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5:297-323.

[5] Kahneman, Daniel., Jack L. Knetsch and Richard.H. Thaler, 1991." The endowment effect, loss aversion, and status quo bias." *Journal of Economic Perspectives,* vol. 5, no. 1 (Winter): 193-206

[6] Thaler, R., A. Tversky, D. Kahneman, and A. Schwartz. 1997. "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test." *Quarterly Journal of Economics*, vol. 112, no. 2

[7] Hassett, S. D. (2010), The RPF Model for Calculating the Equity Market Risk Premium and Explaining the Value of the S&P with Two Variables. *Journal of Applied Corporate Finance*, 22: 118–130. http://bit.ly/d6b1Py

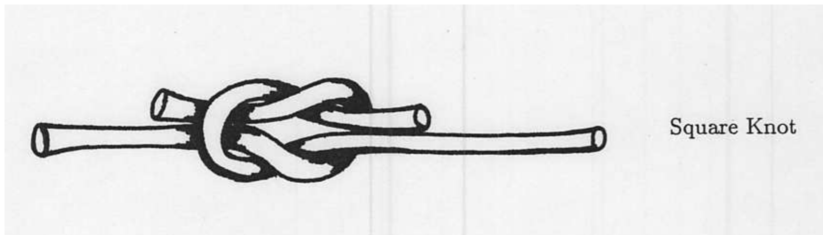
[8] "Economic Projections and The Budget Outlook." Whitehouse.gov, Access Date March 15, 2009, http://www.whitehouse.gov/administration/eop/cea/Economic-Projections-and-the-Budge-Outlook/

[9] "H.15 Selected Interest Rates", The Federal Reserve Website, Accessed March-July 2009, http://www.federalreserve.gov/datadownload/Choose.aspx?rel=H.15

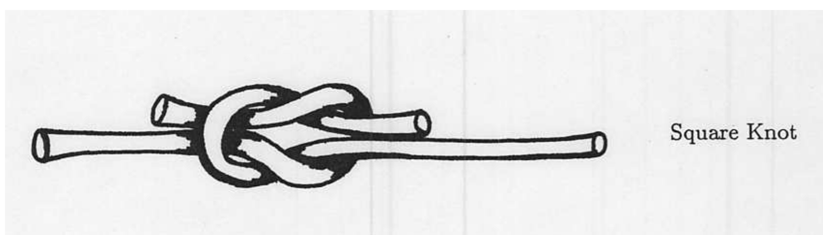
[10] See Hassett (2010)

The Granny Knot, the Square Knot and the Analysis of Hitches
by Louis H. Kauffman

Here is the square knot.

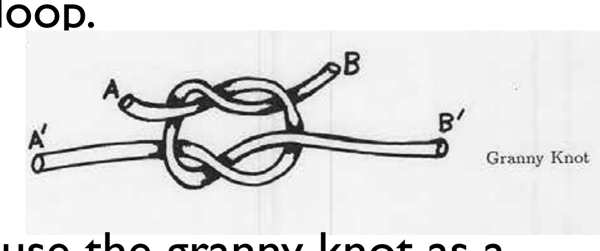


The square knot makes an excellent splice. That is, you can reliably join two lengths of rope by using the square knot as shown above.



The reason the square knot is so good as a splice is that forces applied to the two lengths of rope cause each of the two loops in the splice to constrict the base of the other loop.

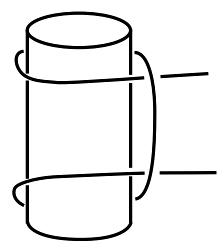
Here is the granny knot.



You might attempt to use the granny knot as a splice, but this is a dangerous idea. Forces applied at A' and B' will not constrict the bases of the loops, and the whole assemblage can slip.

These properties of the square knot and the granny knot are well known. What is not so well known is that each splice can be converted to a 'hitch' by making one of the lengths of rope the axis (post) for the hitch. When we make the conversion, it turns out that the granny knot becomes the well-known clove hitch, a very reliable hitch! The very dangerous granny knot converts to a very reliable hitch. Conversely, the very reliable square knot converts to a less reliable but still workable hitch that we call the square hitch.

The clove hitch.



The square knot hitch

We will give illustrations of the conversion of the square and granny into corresponding hitches after the next section. In this next section we give an excerpt from the Author's book "Knots and Physics" (World Scientific, 1991-2012) where we give an exposition of the analysis of hitches due to Bayman:

Benjamin Bayman, Theory of Hitches, Amer. J. Physics, Vol. 45, No. 2, Feb. 1977.

After the excerpt, we show the conversions and we analyse the square hitch in the same fashion as we shall have already analysed the granny hitch aka the clove hitch.

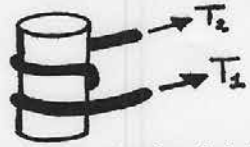
Excerpt from "Knots and Physics" by LK.

1°. Theory of Hitches.

This section is based on the article [BAY].

We give a mathematical analysis of the properties of **hitches**. A **hitch** is a mode of wrapping a rope around a post so that, with the help of a little friction, the rope holds to the post. And your horse does not get away.

First consider simple wrapping of the rope around the post in coil-form:



Assume that there are an integral number of windings. Let tensions T_1 and T_2 be applied at the ends of the rope. Depending upon the magnitudes (and relative magnitudes) of these tensions, the rope may slip against the post.

We assume that there is some friction between rope and post. It is worth experimenting with this aspect. Take a bit of cord and a wooden or plastic rod. Wind the cord one or two times around the rod. Observe how easily it slips, and how much tension is transmitted from one end of the rope to the other. Now wind the cord ten or more times and observe how little slippage is obtained – practically no counter-tension is required to keep the rope from slipping.

In general, there will be no slippage in the T_2-direction so long as

$$T_2 \leq \kappa T_1$$

for an appropriate constant κ. This constant κ will depend on the number of windings. The more windings, the larger the constant κ.

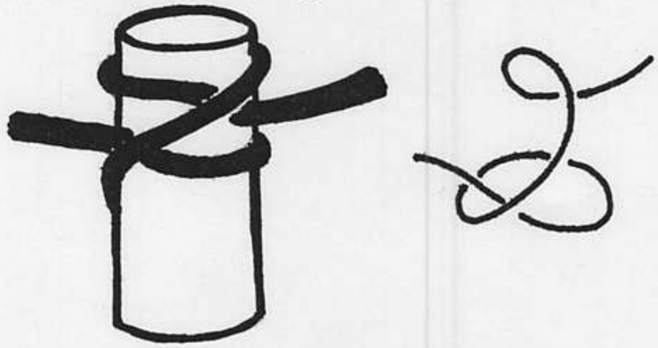
A good model is to take κ to be an exponential function of the angle (in radians) that the cord is wrapped around the rod, multiplied by the coefficient of friction between cord and rod. For simplicity, take the coefficient of friction to be unity so that

$$\kappa = e^{\theta/2\pi}$$

where θ is the total angle of rope-turn about the rod.

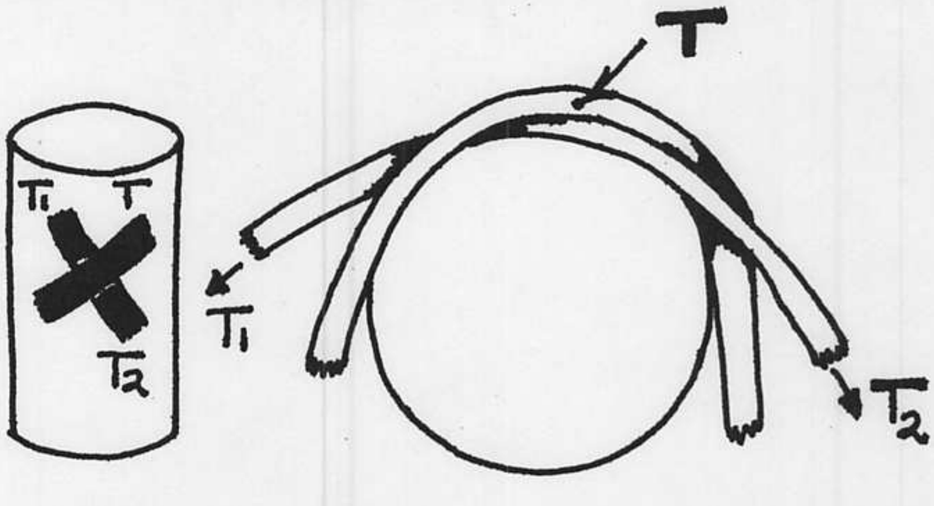
Thus, for a single revolution we need $T_2 \leq eT_1$ and for an integral number n of revolutions we need $T_2 \leq e^n T_1$ to avoid slippage.

A real hitch has "wrap-overs" as well as windings:



Clove Hitch

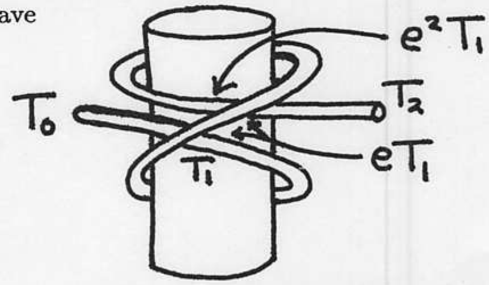
Here, for example, is the pattern of the clove hitch. In a wrap-over, under tension, the top part squeezes the bottom part against the rod.



Hold Fast
$T_2 \leq T_1 + uT$

This squeezing produces extra protection against slippage. If, at such a wrap-over point, the tension in the overcrossing cord is T, then the undercrossing cord will hold-fast so long as $T_2 \leq T_1 + uT$ where u is a certain constant involving the friction of rope-to-rope, and T_2 and T_1 are the tensions on the ends of the undercrossing rope at its ends.

With these points in mind, we can write down a series of inequalities related to the crossings and loopings of a hitch. For example, in the case of the clove hitch we have



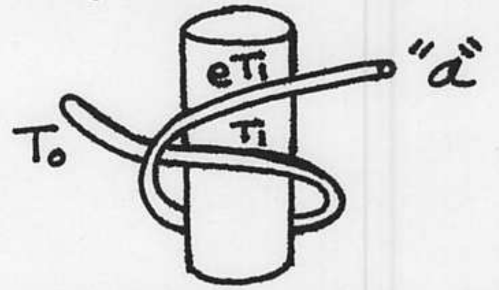
$$\begin{aligned} T_1 &\leq T_0 + ueT_1 \\ T_2 &\leq e^2T_1 + ueT_1 \end{aligned}$$

that the equations necessary to avoid slippage are:

$$T_1 \leq T_0 + ueT_1$$
$$T_2 \leq e^2T_1 + ueT_1.$$

Since **the first inequality holds whenever** $ue > 1$ or $u > 1/e$, we see that **the clove hitch will not slip** no matter how much tension occurs at T_2 just so long as the rope is sufficiently rough to allow $u > 1/e$.

Remark. Let's go back to the even simpler "hitch":



$\boxed{T_1 \leq T_0 + ueT_1}$

Our abstract analysis would suggest that this will hold if $ue > 1$. However, there is no stability here. A pull at "a" will cause the loop to rotate and then the "u-factor" disappears, and slippage happens. A pull on the clove hitch actually tightens the joint.

This shows that in analyzing a hitch, we are actually taking into account some properties of an already-determined-stable mechanical mechanism that happens to be made of rope. [See also Sci. Amer., Amateur Sci., Aug. 1983.]

There is obviously much to be done in understanding the frictional properties of knots and links. These properties go far beyond the hitch to the ways that ropes interplay with one another. The simplest and most fascinating examples are the **square knot** and the **granny knot**. The square knot pulls in under tension, each loop constricting itself and the other - providing good grip:

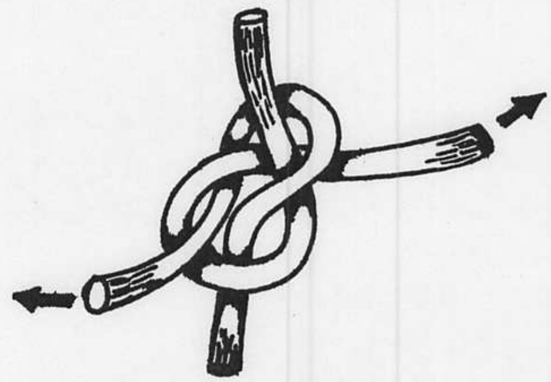
Square Knot

Construct this knot and watch how it grips itself.

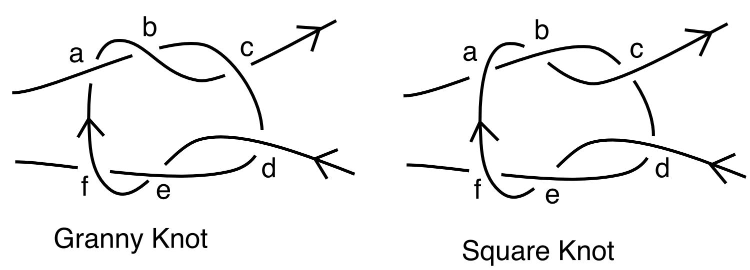
The **granny** should probably be called the **devil**, it just won't hold under tension:

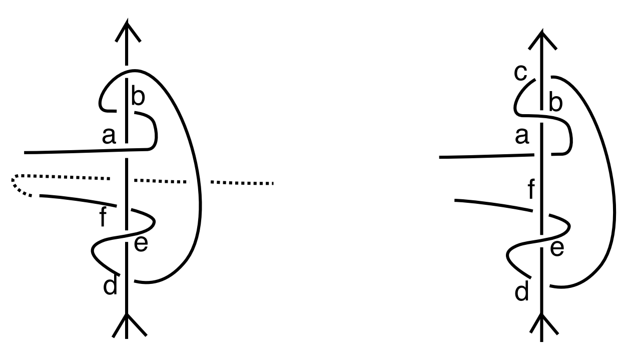
Granny Knot

Try it! Ends A and B, are twisted perpendicular to ends A' and B' and the rope will feed through this tangle if you supply a sufficient amount of tension.

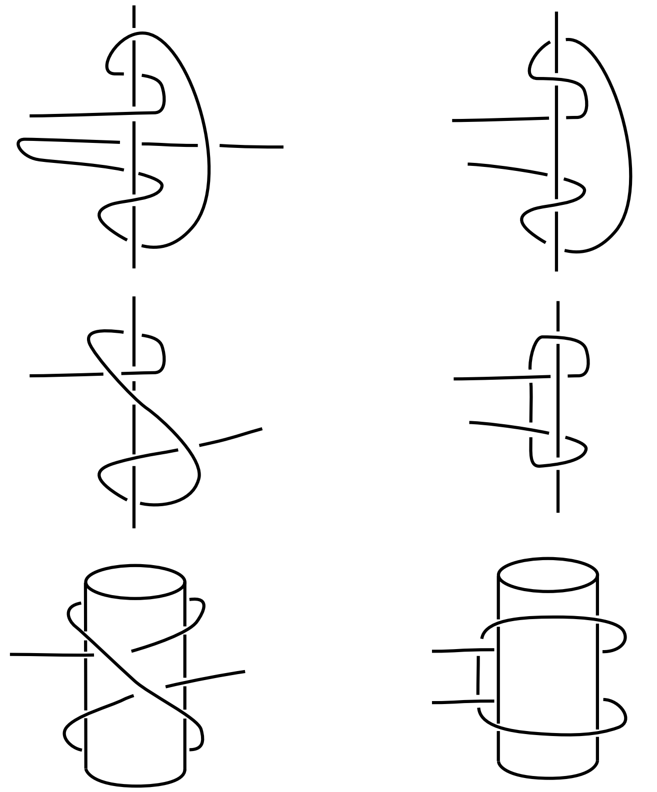


Interconversion of the Granny and the Square Knot to Corresponding Hitches.





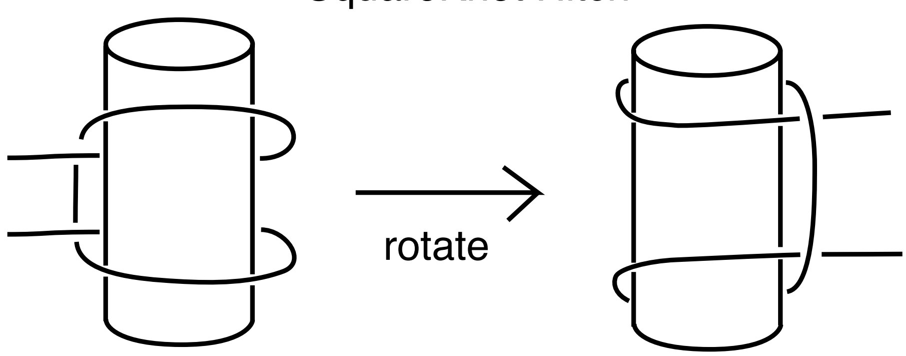
Straighten the Oriented Axis in order
to convert the knot (splice) to a hitch.

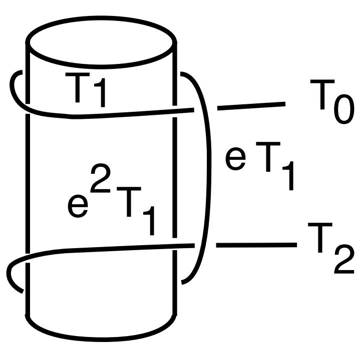


Granny Hitch same as the Clove Hitch

SquareKnot Hitch

SquareKnot Hitch





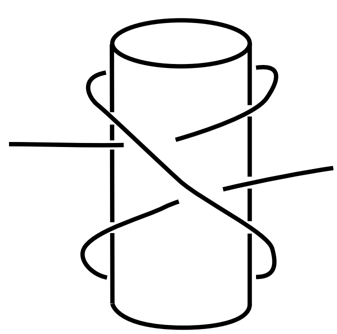
$$T_1 < T_0 + ueT_1$$

$$T_2 < e^2T_1 + ue\,T_1$$

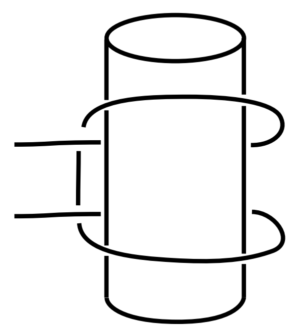
Need ue >1 and so hitch will hold for sufficiently rough rope.

In our Bayman-type analysis of the square knot hitch, we found that if the parameter $u > 1/e$ where e is post's friction parameter, then the hitch will hold. This has the same appearance as the end-result of our analysis of the clove hitch in the middle section of the paper, but if you make these hitches you will see that there is a world of difference between the square hitch and the granny hitch (aka clove hitch). The granny hitch makes good contact with the post, while the square hitch does not make good contact. The weaving points in the square hitch tend to be pulled away from the post. This means that much more depends upon the friction of the rope against itself than in the granny hitch.

The moral of our story is that there is much physical lore to be extracted from individual weaving patterns. A full mathematical analysis of knots and their behaviour under forces and friction is a subject for the future.



Granny Hitch
same as the
Clove Hitch



SquareKnot Hitch

When is the next Thankgivukkah?

Doron Levy[1]

In the October 8, 2013 issue, the Boston Globe stated that the next time the first day of the Jewish holiday of Hanukkah will coincide with the American observance of Thanksgiving is "79,043 years from now, by one calculation" [1]. In contrast, on November 22, 2013, National Geographic stated that we will have to wait "until the year 79,811" for the next occurrence of Thanksgivukkah [2]. How come two respectable news outlets managed to provide similar yet different answers to the same calculation?

In principle, calculating the next occurrence of Thanksgivukkah should be a rather straightforward task. For any given year, one has to (i) calculate the date of the American thanksgiving, which is defined as the fourth (and not the last) Thursday in November; (ii) calculate the date of the first day of Hannukkah; and (iii) check if these two dates match. Instead of resolving the discrepancy in this calculation, I prefer to leave it as a challenge. Accordingly, this note will discuss some issues that one may wish to take into account in carrying out this task.

The holiday of Hannukkah is observed 8 days starting on the 25th day of the month of Kislev according to the Hebrew calendar. It is important to note that the Jewish day does not begin on midnight, but on the sunset before it. This means that every Jewish day overlaps with two "standard" days. While the first day of Hannukkah of the Jewish year 5774 coincided with Thanksgiving 2013, the first candle was actually lit on Wednesday night - as this is the beginning of the day of Thursday according to the Jewish tradition.

A Hebrew year consists of 12 months in a common year and 13 in a leap year. In the Hebrew calendar, leap years occur in years 3, 6, 8, 11, 14, 17, and 19 of a 19-year cycle. In a leap year, the month of Adar is renamed as Adar II and has 29 days instead of 30 days. An additional 30-days month (Adar I) is added before Adar II. Another confusing aspect of the Hebrew calendar is that there are two months (Heshvan and Kislev) that may have 29 or 30 days, depending on various factors. In some rare cases, the leap year can be extended by an extra day by moving the New Year holiday of Rosh Hashana one day forward. A comprehensive discussion can be found in [8, Chapter 7].

Gauss derived a formula for the date of the holiday of Passover, which falls on the first day of the month of Nisan (see, e.g., [10]). Given Gauss formula for the date of Passover it is a reasonable adjustment to compute the first day of Hannukkah, as all that is required is to figure out whether it is a leap year, and what is the exact number of days in the month of Kislev. These factors determine the number of days that separate Hannukkah from Passover.

How do we know when is Thanksgiving? Clearly, for any given year, if we know the day of the week of November 1, we can easily find the date of the fourth Thursday in November. Many algorithms were devised to find the day of the week for a given day. A rather comprehensive list can be found at [4]. Here I would like to mention three examples of interest to the attendees of the G4G meetings:

1. Corinda on page 71 of "13 steps to mentalism" presents a method for "A day for any date". This method is described as a mentalism effect in which "The performer invites members of his audience to call out any date they like; upon hearing the date, the performer gives the exact day of the week that that date falls on and delivers his reply within seconds. Everything is achieved by a quick calculating mental system."

   The algorithm that Corinda provides is one of the easier algorithms to memorize and mentally compute, which makes it suitable for stage demonstrations. It can be summarized as follows: (i) Add a quarter to the value of the last two digits of the year. (ii) Add the code value for the month (Table 1). (iii) Add the day of the month. (iv) Divide the total by seven. (v) The remainder tells you the day according to Table 1. This algorithm gives the correct date for years 1801-1900. When the date is in the twentieth century you must deduct two from the final remainder and when the date occurs in the eighteenth century, two should be added to the final remainder.

[1]Department of Mathematics and Center for Scientific Computation and Mathematical Modeling (CSCAMM), University of Maryland, College Park, MD 20742.

June	0
September, December	1
January (leap year), April, July	2
January, October	3
May	4
February (leap year), August	5
February, March, November	6

Sunday	1
Monday	2
Tuesday	3
Wednesday	4
Thursday	5
Friday	6
Saturday	0

Table 1: Months and Days for Corinda's algorithm

2. Lewis Carroll published an algorithm "to find the day of the week for any given date" [6]. This algorithm was discussed by Martin Gardner in [9]. A copy of Carroll's paper is shown in Figure 1.

3. The Doomsday Rule. The doomsday is the day of the week of the last day of February. The Doomsday Rule is a method devised by John Conway for computing the day of the week of any given date. A nice explanation of this algorithm can be found in the notes by Graham [3].

All algorithms for calendar calculations must account for leap years. The Julian calendar, imposed by Julius Caesar in 45 B.C., following the advice of Sosigenes, the Alexandrian astronomer, was based on a 4-year leap year cycle – one day being added every four years. The current calendar, the Gregorian calendar, was instituted by the Pope Gregory XIII in 1582. According to the Gregorian calendar, a leap year is a year that is divisible by 4, unless it is divisible by 100, unless it is divisible by 400. Leap years were established in order to provide occasional corrections to the length of the year, so that the calendar follows the "astronomical" year.

The astronomical Almanac online [5] defines a tropical year as the period of time for the ecliptic longitude of the Sun to increase 360 degrees. Since the Sun's ecliptic longitude is measure with respect to the equinox, the tropical year comprises a complete cycle of seasons, and its length is approximated in the long term by the civil (Gregorian) calendar. The mean tropical year is approximately 365 days, 5 hours, 48 minutes, 45 seconds. This estimate corresponds to approximately 365.24219 days. This value is actually does not remain constant as it changes over time due to a variety of reasons.

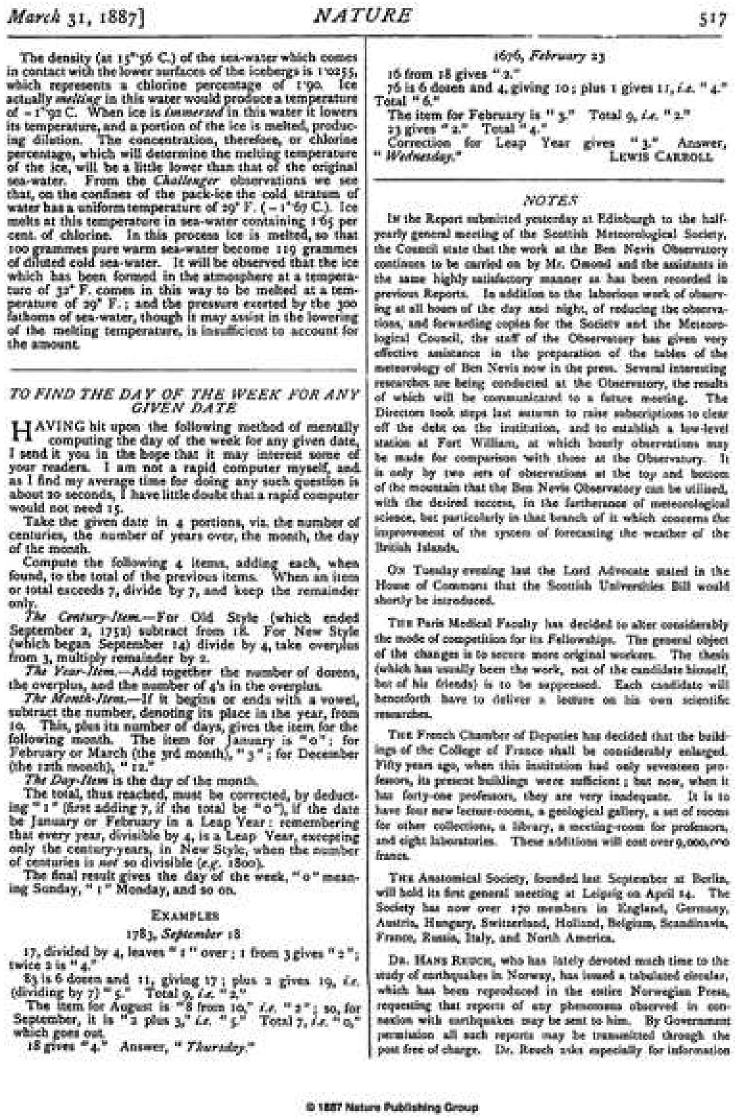
Leap years are nothing but a correction to the calendar in order to avoid too big of a drift between the calendar and the tropical year. In the Julian calendar, the average length of a year is 365.25. This is improved by the Gregorian calendar in which the average length of a year is $365 + 97/400 = 365.2425$, a better approximation of 365.2422... The Julian or Gregorian calendars are not the only way to correct approximate the length of the year by adding days. One interesting alternative, involves approximating 365.24219 as a continued fraction:

$$365.24219 \approx 365 + \cfrac{1}{4 + \cfrac{1}{7 + \cfrac{1}{1}}} = 365\frac{8}{33} = 365.2424...$$

In practice, such an approximation can be implemented by defining 8 years in a 33-year cycle as leap years. The specific choice of which 8 years is of no importance. This is exactly the correction proposed by the Persian mathematician, philosopher and poet, Omar Khayyam: a 33-year cycle where the years 4,8,12,16,20,24,28, and 33 are leap years [11]. An even more accurate approximation can be obtained, e.g., by adding one term to the continued fraction ($1/1$ will be replaced by $1/(1 + 1/3)$). A calendar based on the resulting fraction ($31/128$) will correspond to an average year of 365.2421875 days.

Is there any real need for more accurate corrections? The error in the Gregorian calendar is approximately $|365.24219 - 365.2425| = 0.00031$ days per year. In comparison, the error in the Khayyam calendar is $|365.24219 - 365.2424242...| = 0.00023424...$ days per year. This means that the Gregorian calendar shifts by one day in approximately 3225 years, while the Khayyam calendar shift by one day in approximately 4269 years, rendering it a better approximation.

If we have to wait about 80,000 years for the next occurrence of Thanksgivukkah, multiple corrections will have to be made to the calendar (and in fact also to the Hebrew calendar). This means that any *correct* answer to the challenge, is guaranteed to be *wrong*...



March 31, 1887] NATURE 517

The density (at 15°·56 C.) of the sea-water which comes in contact with the lower surfaces of the icebergs is 1·0255, which represents a chlorine percentage of 1·90. Ice actually *melting* in this water would produce a temperature of −1°·92 C. When ice is *immersed* in this water it lowers its temperature, and a portion of the ice is melted, producing dilution. The concentration, therefore, or chlorine percentage, which will determine the melting temperature of the ice, will be a little lower than that of the original sea-water. From the *Challenger* observations we see that, on the confines of the pack-ice the cold stratum of water has a uniform temperature of 29° F. (−1°·67 C.). Ice melts at this temperature in sea-water containing 1·65 per cent. of chlorine. In this process ice is melted, so that 100 grammes pure warm sea-water become 119 grammes of diluted cold sea-water. It will be observed that the ice which has been formed in the atmosphere at a temperature of 32° F. comes in this way to be melted at a temperature of 29° F.; and the pressure exerted by the 300 fathoms of sea-water, though it may assist in the lowering of the melting temperature, is insufficient to account for the amount.

TO FIND THE DAY OF THE WEEK FOR ANY GIVEN DATE

HAVING hit upon the following method of mentally computing the day of the week for any given date, I send it you in the hope that it may interest some of your readers. I am not a rapid computer myself, and as I find my average time for doing any such question is about 20 seconds, I have little doubt that a rapid computer would not need 15.

Take the given date in 4 portions, viz. the number of centuries, the number of years over, the month, the day of the month.

Compute the following 4 items, adding each, when found, to the total of the previous items. When an item or total exceeds 7, divide by 7, and keep the remainder only.

The Century-Item.—For Old Style (which ended September 2, 1752) subtract from 18. For New Style (which began September 14) divide by 4, take overplus from 3, multiply remainder by 2.

The Year-Item.—Add together the number of dozens, the overplus, and the number of 4's in the overplus.

The Month-Item.—If it begins or ends with a vowel, subtract the number, denoting its place in the year, from 10. This, plus its number of days, gives the item for the following month. The item for January is "0"; for February or March (the 3rd month), "3"; for December (the 12th month), "12."

The Day-Item is the day of the month.

The total, thus reached, must be corrected, by deducting "1" (first adding 7, if the total be "0"), if the date be January or February in a Leap Year: remembering that every year, divisible by 4, is a Leap Year, excepting only the century-years, in New Style, when the number of centuries is *not* so divisible (*e.g.* 1800).

The final result gives the day of the week, "0" meaning Sunday, "1" Monday, and so on.

EXAMPLES

1783, *September* 18

17, divided by 4, leaves "1" over; 1 from 3 gives "2"; twice 2 is "4."

83 is 6 dozen and 11, giving 17; plus 2 gives 19, *i.e.* (dividing by 7) "5." Total 9, *i.e.* "2."

The item for August is "8 from 10," *i.e.* "2"; so, for September, it is "2 plus 3," *i.e.* "5." Total 7, *i.e.* "0," which goes out.

18 gives "4." Answer, "*Thursday*."

1676, *February* 23

16 from 18 gives "2."

76 is 6 dozen and 4, giving 10; plus 1 gives 11, *i.e.* "4." Total "6."

The item for February is "3." Total 9, *i.e.* "2."

23 gives "2." Total "4."

Correction for Leap Year gives "3." Answer, "*Wednesday*."

LEWIS CARROLL

NOTES

IN the Report submitted yesterday at Edinburgh to the half-yearly general meeting of the Scottish Meteorological Society, the Council state that the work at the Ben Nevis Observatory continues to be carried on by Mr. Omond and the assistants in the same highly satisfactory manner as has been recorded in previous Reports. In addition to the laborious work of observing at all hours of the day and night, of reducing the observations, and forwarding copies for the Society and the Meteorological Council, the staff of the Observatory has given very effective assistance in the preparation of the tables of the meteorology of Ben Nevis now in the press. Several interesting researches are being conducted at the Observatory, the results of which will be communicated to a future meeting. The Directors took steps last autumn to raise subscriptions to clear off the debt on the institution, and to establish a low-level station at Fort William, at which hourly observations may be made for comparison with those at the Observatory. It is only by two sets of observations at the top and bottom of the mountain that the Ben Nevis Observatory can be utilised, with the desired success, in the furtherance of meteorological science, but particularly in that branch of it which concerns the improvement of the system of forecasting the weather of the British Islands.

ON Tuesday evening last the Lord Advocate stated in the House of Commons that the Scottish Universities Bill would shortly be introduced.

THE Paris Medical Faculty has decided to alter considerably the mode of competition for its Fellowships. The general object of the changes is to secure more original workers. The thesis (which has usually been the work, not of the candidate himself, but of his friends) is to be suppressed. Each candidate will henceforth have to deliver a lecture on his own scientific researches.

THE French Chamber of Deputies has decided that the buildings of the College of France shall be considerably enlarged. Fifty years ago, when this institution had only seventeen professors, its present buildings were sufficient; but now, when it has forty-one professors, they are very inadequate. It is to have four new lecture-rooms, a geological gallery, a set of rooms for other collections, a library, a meeting-room for professors, and eight laboratories. These additions will cost over 9,000,000 francs.

THE Anatomical Society, founded last September at Berlin, will hold its first general meeting at Leipzig on April 14. The Society has now over 170 members in England, Germany, Austria, Hungary, Switzerland, Holland, Belgium, Scandinavia, France, Russia, Italy, and North America.

DR. HANS REUSCH, who has lately devoted much time to the study of earthquakes in Norway, has issued a tabulated circular, which has been reproduced in the entire Norwegian Press, requesting that reports of any phenomena observed in connexion with earthquakes may be sent to him. By Government permission all such reports may be transmitted through the post free of charge. Dr. Reusch asks especially for information

Figure 1: Lewis Carroll's Nature paper on finding the day of the week for any given date

References

[1] http://www.bostonglobe.com/news/nation/2013/10/07/convergence-hanukkah-thanksgiving-unleashes-creativity/MxbqB0YtKA2b8AtR7XGwsL/comments.html

[2] http://news.nationalgeographic.com/news/2013/11/131124-thanksgivukkah-thanksgiving-hanukkah-calendar-turkey-potatoes/

[3] http://people.cst.cmich.edu/graha1sw/Pub/Doomsday/Doomsday.html

[4] http://en.wikipedia.org/wiki/Determination_of_the_day_of_the_week

[5] http://asa.usno.navy.mil/SecM/Glossary.html

[6] Lewis Carroll, To Find the Day of the Week for Any Given Date, Nature, March 31, 1887.

[7] Tony Corinda, 13 Steps to Mentalism, Robbins & Co.

[8] Nachum Dershowitz and Edward Reingold, Calendrical Calculations, 3rd Edition, Cambridge University Press, 2008.

[9] Martin Gardner, The Universe in a Handkerchief: Lewis Carroll's Mathematical Recreations, Games, Puzzles, and Word Plays, pp.24–26, Springer-Verlag, 1996

[10] Zvi Har'El, Gauss Formula for the Julian Date of Passover, http://www.math.technion.ac.il/S/rl/docs/gauss.pdf

[11] Rickey, V.F., Mathematics of the Gregorian Calendar. The mathematical Intelligencer, **7** (1985), pp.53–56.

Three Pairs of Problems

Andy Liu

We present three pairs of problems.

Problem 1A.
Kim and his friend Abdullah were going from Lahore to Benares along the scenic Grand Trunk Road, which would take them through Umballa, Delhi and Alighur. Excited about their adventure, they went through Umballa and reached Delhi in a week. Then Abdullah wanted to go back while Kim wanted to push on. Nevertheless, they stayed together. However, the dispute had slowed them down, so that they only made it to the next stop each week. Whenever they were in Umballa, Delhi or Alighur, the argument about which way to go would flare up again. Eventually, they reached Benares, but only after having changed directions ten times. How many different week-by-week itineraries could they have followed?

Problem 1B.
Kim, Abdullah and their friend Chota Lal were going from Lahore to Alighur. Kim and Abdullah went along the scenic Grand Trunk Road which would take them through Umballa and Delhi. Chota Lal left Lahore at the same time, but along a straight road directly to Delhi. At Umballa, Abdullah went on a striaght road directly to Alighur while Kim continued on the Grand Trunk Road. At Delhi, Chota Lal rejoined Kim as they both arrived at the same time, and they pushed on without stopping. Finally, all three of them arrived at Alighur simultaneously. Kim travelled at the same constant speed, whether alone or in company. Each of Abdullah and Chota Lal travelled at a constant speed when alone. There was a statue of Rudyard Kipling at the intersection of the two straight roads. Was it possible for Abdullah and Chota Lal to have met each other there?

Problem 2A.
Forty Thieves, ranked 1 to 40, were trying to cross a river in a boat which took two of them to row. However, if the ranks of two of them differ by more than 1, they would refuse to be in the boat together. This meant that only two could cross at a time, but the same two must bring he boat back. Their leader, whose rank was 1, appealed to Ali Baba for assistance. As it happened, Ali Baba also wanted to cross the river. “I can get you guys over,” he said, “if you make me an honorary thief with same rank as you.” After some hesitation, the leader accepted Ali Baba as his equal, and Ali Baba master-minded the operation. What was the minimum number of one-way crossings required to get everyone across

Problem 2B.
Forty Thieves, ranked 1 to 40, had some gold coins in their possession equal to their respective ranks. Someone else had stolen 41 gold coins from the Royal Mint. If any group of the Forty Thieves had exactly 41 gold coins among them, they would all be hanged. Ali Baba was ordered to round them up and put them in prison, entering two at a time. One of the two could surrender all his gold coins to the other, and then Ali Baba could release him immediately. No trading of gold coins could occur in prison as the thieves were held in individual cells. Naturally, no thief was willing to yield his gold coins, unless he was ordered to do so by Ali Baba. What was the maximum number of thieves Ali Baba could have in prison and yet save them all?

Problem 3A.

An Evil Witch had imprisoned Princess Anna. When Prince Boris got to the Evil Castle, he found fifteen veiled ladies being formed into a line by the Evil Witch. “One of these veiled ladies is your Princess Anna. Seven of the others were transformed by me from poisonous frogs. The remaining seven were poisonous toads. After I have opened the door to my castle, they will go inside and sit down in order. Then you can come in and kiss any of them. If she is indeed the Princess, the two of you are free to go. If not, you will find out whether you have just kissed a poisonous frog or a poisonous toad. This information will not be of much use to you,” she laughed, “because you will die from the kiss.” While the veiled ladies followed the Evil Witch into the Evil Castle, the Fairy Godmother materialized beside Prince Boris. “I will help you in two ways. First, here are three life-saving pills,” she whispered as she handed them to him. “Second, I will give you a divine revelation. The first seven veiled ladies were poisonous frogs and the next seven were poisonous toads. Princess Anna is last in line.” Then she vanished. Feeling that he did not even need the pills, Prince Boris stepped confidently into the Evil Castle. To his dismay, the veiled ladies were in clockwise order at a round table, with no indication of who was the first to sit down. Could he still rescue Princess Anna?

Problem 3B.

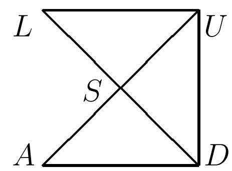
After he had imprisoned Prince Boris, an Evil Wizard granted Princess Anna an audience. “Here are thirty-five cards, each with a number on the back,” he said as he dealt them on the table so that they formed a five by seven array. For each card, the side with the number faced down, so that Princess Anna could not see it. Then he continued, “You may turn over one of the cards and read the number on it. You must then tell me the sum of all thirty-five numbers. If it is correct, you may take your Prince Boris with you. If not, you will suffer a horrifying death.” While he closed his eyes in glee as he planned Princess Anna’s demise, the Mafia Godfather materialized beside Princess Anna. “I will give you a divine revelation. In any two by three or three by two array, the sum of the numbers on the six cards is always 336.” He vanished just as the Evil Wizard opened his eyes and fixed his gaze at Princess Anna. Could she rescue Prince Boris?

Solution to Problem 1A.

Let k_n denote the number of possible itineraries for Kim and Abdullah if they changed directions n times. We claim that $k_n = k_{n-1} + k_{n-2}$. Consider first the case where n is even. This means that they would go onto Benares. Hence the last change in directions must occur in Delhi or Umballa. If it occured in Delhi, this means that he must have returned from Alighur, and had gone there earlier from Delhi. If we shorten the last segment “Delhi — Alighur — Delhi” to just “Delhi”, we will get a path counted in k_{n-2}. On the other hand, if the last change in directions occured in Umballa, this means that he could have gone back to Lahore had he not done so. If we replace the segment “Delhi — Alighur — Benares” at the very end by “Lahore”, we will get a path counted in k_{n-1}. The case where n is odd can be handled in an analogous manner. Just interchange “Lahore” with “Benares” and “Umballa” with “Alighur”. This justifies the claim. It follows that k_n are just the Fibonacci numbers, with $k_0 = 1$ and $k_1 = 2$. Hence $k_2 = 3$, $k_3 = 5$, $k_4 = 8$, $k_5 = 13$, $k_6 = 21$, $k_7 = 34$, $k_8 = 55$, $k_9 = 89$ and $k_{10} = 144$.

Solution to Problem 1B.

Let L, U, D, A and S be the locations of Lahore, Umballa, Delhi, Alighur and the statue of Rudyard Kipling. If Abdullah and Chota Lal do meet each other at S, this must occur after Abdullah left U and before Chota Lal arrived at D. Let the lengths of these two time intervals be x and y respectively. Let the constant speeds of Kim, Abdullah and Chota Lal be a, b and c respectively. Since $LU + UD > LD$, $a > c$. Since $UD + DA > UA$, $a > b$. Now $UD = a(x + y)$, $US = cx$ and $SA = by$. We have $US + SA = cx + by < a(x + y) = UA$, which is a contradiction. It follows that Abdullah and Chota Lal cannot have met each other at S.



Solution to Problem 2A.

We first show that 153 crossings are sufficient. In the first 8 crossings, we get thieves 39 and 40 over to the far shore, as shown in the chart below.

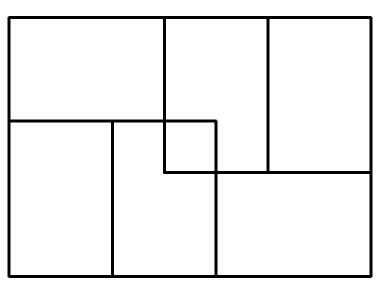
Crossing	Ranks of thieves		
Number	on Near Shore	in Boat	on Far Shore
First	3,4,...,40	1,1,2	1,1,2
Second	1,2,...,40	1,2	1
Third	1,4,5,...,40	2,3	1,2,3
Fourth	1,1,2,4,5,...,40	1,2	3
Fifth	4,5,...,40	1,1,2	1,1,2,3
Sixth	2,3,...,40	2,3	1,1
Seventh	2,3,...,38	39,40	1,1,39,40
Eighth	1,1,2,...,38	1,1	39,40

Note that apart from thieves 39 and 40, everybody is back on the near shore. Using the procedure, we can get thieves 37 and 38 across in another 8 crossings, and so on. Finally, one more crossing will accomplish the task. The total number of crossings is indeed $8 \times 19 + 1 = 153$. We now show that 153 crossings are necessary. Consider the crossings in pairs, one to the far shore and the very next one back to the near shore. A gain for this pair is defined as an increase of the number of thieves on the far shore after this pair of crossing is completed. When two thieves cross over, the gain is obviously 0 as two thieves must come back. The gain is 1 when three thieves cross over together. The very last crossing, which is to the far shore, generates a gain of 3. To get $n + 1$ thieves across, we must of course gain $n + 1$. So three thieves must cross over together $n - 1$ times. Obviously, this cannot happen in every crossing from the near shore, because in that case, thieves 1, 1 and 2 must all come back. Now two of them can come back right away, and one of them can cross over next time to fetch the third. So this can happen every other crossing from the near shore. So there must be at least 3 other crossings between two crossings from the near shore with three thieves. This means that we have to add $n - 2$ sets of 3 crossing to the total. So the minimum number of one-way crossings is $(n - 1) + 3(n - 2) = 4n - 7$. For $n = 40$, we have $4n - 7 = 153$.

Solution to Problem 2B.
For any pair of thieves who have 41 gold coins between them, at least one of them must either give up his gold coins or receive the gold coins from another thief. Since there were twenty such pairs, gold coins must change hands at least ten times. Ali Baba could have the twenty thieves with odd numbers of gold coins enter the prison in pairs. One thief in each pair gave all his gold coins to his cellmate and was released. Each of the thirty thieves in priosn had an even number of gold coins in his possession, and no group of them could have exactly 41 gold coins among them.

Solution to Problem 3A.
Number the veiled ladies 1 to 15 in clockwise order. Prince Boris kisses number 8. There are three cases.
Case 1. Number 8 is Princess Anna.
Then the mission is accomplished.
Case 2. Number 8 is a poisonous frog.
Then number 15 must be a poisonous toad and Princess Anna is not between these two. Prince Boris takes the first life-saving pill and kisses number 4. There are three subcases.
Subcase 2(a). Number 4 is Princess Anna.
Then the mission is accomplished.
Subcase 2(b). Number 4 is another poisonous frog.
Then number 5 to number 7 are all poisonous frogs. Prince Boris takes the second life-saving pill and kisses number 2. If it is Princess Anna, the mission is accomplished. If it is another poisonous frog, then number 1 is Princess Anna. If number 2 is a poisonous toad, then number 3 is Princess Anna. In either case, Prince Boris takes the last life-saving pill and accomplishes his mission.
Subcase 2(c). Number 4 is a poisonous toad.
Then number 1 to number 3 are all poisonous toads. Prince Boris takes the second life-saving pill and kisses number 6. The analysis is analogous to Subcase 2(b).
Case 3. Number 8 is a poisonous toad.
Then number 1 must be a poisonous frog and Princess Anna is not between these two. Prince Boris takes the first life-saving pill and kisses number 12. The analysis is analogous to Case 2.

Solution to Problem 3B.
Princess Anna partitions the 5×7 table into six 2×3 or 3×2 rectangles, two of which overlaps at the central square. She knows that the sum of all 35 numbers is 6×336 minus the number on the card in the central square. Confidently, she turns over that card and accomplishes her mission by annoucing the correct sum.



Sun Bin's Legacy

Dana Mackenzie
scribe@danamackenzie.com

Introduction

Sun Bin was a legendary Chinese military strategist who lived more than 2000 years ago. Among other exploits, he is credited with helping his patron, general Tian Ji, defeat the King of Qi in a match consisting of three horse races. If Tian Ji had simply raced his top horse against the King's top horse, his second against the King's second, and his third against the King's third, he would have lost all three races.

But Sun Bin had an idea. He told Tian Ji to race his *worst* horse against the King's best, his best against the King's second-best, and his second-best against the King's worst. In this way, Tian Ji won two out of three races.

Of course, Tian Ji was a little bit lucky. If we denote his horses by A, B, and C and the king's horses by a, b, and c, and if we rank the horses by speed, they happened to fall in the order $a > A > b > B > c > C$. It's easy to see with 20/20 hindsight that Sun Bin's strategy works here, because $A > b$ and $B > c$. Had the horses been in a different order, say $a > b > c > A > B > C$, then Sun Bin's strategy would not have worked (but neither would any other strategy!).

A key point to realize, though, is that Sun Bin's strategy does not depend on knowing the relative speeds of *all six* horses in advance. We only need to know the rankings of each side's horses: that is, we only need to know that $A > B > C$ and $a > b > c$. Given only this information, Sun Bin's strategy of racing C against a, A against b, and B against c is the optimal strategy in two different ways: It gives him the best odds of winning the match, and gives him the largest number of expected races won. I will justify these claims in this paper.

Now, let's bring Sun Bin's problem into the twenty-first century. What happens if we have a match of N horses against N horses? Can we find an optimal strategy for winning the match? How about for winning the largest expected number of match races? The answer to both of these questions is yes, and the goal of this paper is to answer both questions. The second one is much harder, and (in my opinion) much more interesting mathematically.

To define the problem a little better, I prefer to phrase it in terms of a card game. The deck has $2N$ cards, with face values from 1 to $2N$ ($2N$ being high). Player 1 receives N cards, face down, and player 2 likewise. Neither is allowed to see the cards in his hand or the other player's hand. However, the cards are placed in rank order in front of each player, so that each player knows the relative ranking of his own cards and the opponent's cards. On the first trick, player 1 plays one card (still face down) and player 2 plays one against it. Play continues in this fashion (player 1 always going first!) until all the cards have been played, and then both players reveal their cards. The winner is the one who takes the most tricks. What is player 2's optimal strategy?

It is easy to see that this is equivalent to an "idealized" version of the horse problem, in which the faster horse always wins the race. I prefer the card version because (as is well known) horses do not always race according to form, and faster horses sometimes lose to slower horses. In the card version, there are no such ambiguities: card 7 always beats card 6, and that is that.

To me, the *N*-card (or *N*-horse) version of Sun Bin's problem is extremely natural, and it is a bit of a mystery why it seems to be nearly absent from the mathematical literature. The only reference I have been able to find is [AGY], written in 1979, and even that reference is cursory. The authors simply noted that the problem reduces to a linear assignment problem, and therefore there exist fast computer algorithms (the Hungarian Algorithm) to solve it.

A college friend of mine, Howard Stern, independently posed this problem in his first year of graduate school, in 1980. He made, in my opinion, some extremely impressive progress towards a solution, and arrived at a correct conjecture for the general strategy, but he was unable to prove it. In the three decades since then, he has showed the problem to a number of mathematicians and computer scientists, always thinking that somebody would know a solution or a general theorem that would solve the problem. However, none of them did. Finally, he asked me in 2012, and very soon I was hooked on Stern's problem! The proof of his conjecture turned out to be a fascinating mix of group theory, probability, and combinatorics.

What's more, I believe that the solution has some relevance to the general theory of linear assignment problems. Stern's original, unpublished work from 1980 appears to be a novel observation about what I call mixed-Monge optimization problems. And the first step of my proof of Howard's conjecture also involves a general result about symmetric mixed-Monge optimization problems.

Gary Antonick wrote a post about Stern's problem in his "Numberplay" blog for the *New York Times* on January 13, 2014. It became his most-commented-on blog post in more than a year. I am indebted to one of his readers, a reader known to me only as Lee from London, who pointed out the ancient Chinese legend of Sun Bin. This is surely the first appearance of the problem in recorded history, so I think it is only appropriate to call it "Sun Bin's Legacy."

At this point I would strongly encourage readers to play the card game and see if they can figure out the strategy themselves, before reading on. Antonick's post [A] includes a wonderful applet by Gary Hewitt that will enable you to play against the computer online with 3 to 7 cards.

The main theorem of this paper is as follows:

Main Theorem. For sufficiently large *N*, the optimal strategy for player 2 is to play his cards in the order (1, 2, …, *k*, *N*, *N*-1, …, *k*+1) for some *k*. In other words, he plays his lowest card (1) against the opponent's highest, his second-worst card against player 1's second-highest, etc. Note that *k* represents the number of tricks that player 2 should (try to) "throw" or lose on purpose. This strategy is optimal in the sense of maximizing the *expected number of tricks won*. The optimal number $k = k^*(N)$ is given by the following formula:

$$k^*(N) = \sup\left\{k : \sum_{j=0}^{k-1} \binom{N}{j}^2 + \sum_{j=0}^{N-k}\binom{2N}{j} \geq \binom{2N}{N}\right\}.$$

Comments on the Main Theorem:

(1) There is also a somewhat simpler "approximate" formula $k \approx k(N)$:

$$k(N) = \sup\left\{k : \sum_{j=0}^{N-k} \binom{2N}{j} \geq \binom{2N}{N}\right\}.$$

It is approximate in the sense that $k^*(N) - k(N) = 0$ or 1. In fact, I do not know a single value of N for which $k(N) \neq k^*(N)$.

(2) Stern conjectured the general form of the optimal strategy in 1980, but did not make a conjecture for the optimal number $k^*(N)$ of tricks to throw. At that point there was not enough data to make a conjecture, and it is highly unlikely that anyone would have come up with the formula above anyway. It was a complete shock to me that I was able to derive an exact formula.

Here is a table of the optimal number of tricks to throw for small values of N:

N	k	N	k
2	n/a	9	2
3	1	10	2
4	1	11	2
5	1	12	2
6	1	13	3
7	2	14	3
8	2	15	3

And here are two "worked examples," showing the first two jumps in $k^*(N)$.

Example 1: $N = 7$. Here the approximate formula tells us to look up the 14-th row of Pascal's triangle and add the terms until we get a sum that is greater than the central element. We find that

$$1 + 14 + 91 + 364 + 1001 + 2002 = 3473 > 3432.$$

Then the approximate number of tricks to throw is (N+1) minus the number of terms added: in this case $(7 + 1) - 6 = 2$.

For the exact computation, we add a couple of squared terms from the 7-th row:

$$1 + 14 + 91 +364 + 1001 + 2002 + 1^2 + 7^2 = 3523 > 3432.$$

Thus the exact number of tricks to throw is at least 2. On the other hand, if we try throwing one more, we get

$$1 + 14 + 91 + 364 + 1001 + 1^2 + 7^2 + 21^2 = 1962 < 3432.$$

Thus the exact number of tricks to throw is at most 2, and hence the exact formula agrees with the approximate formula.

Example 2: $N = 13$. Now we look up the 26th row of Pascal's triangle and add up the terms until we get a sum that is greater than the center element. We find that

$$1 + 26 + 325 + 2600 + 14950 + 65780 + 230230 + 657800 + 1562275 + 3124550 + 5311735 = 10970722 > 10400600.$$

Therefore the approximate number of tricks to throw is (N + 1) minus the number of terms added, i.e. $(13 + 1) - 11 = 3$.

The exact computation involves adding some squared terms from the 13th row. Because

$$1 + 26 + 325 + 2600 + 14950 + 65780 + 230230 + 657800 + 1562275 + 3124550 + 5311735 + 1^2 + 13^2 + 78^2 > 10400600$$

we can be certain that the exact number of tricks to throw is at least 3. And because

$$1 + 26 + 325 + 2600 + 14950 + 65780 + 230230 + 657800 + 1562275 + 3124550 + 1^2 + 13^2 + 78^2 + 286^2 < 10400600$$

the exact number of tricks to throw is less than 4. Hence the exact number of tricks to throw is 3, which agrees with the approximate computation.

These two examples are completely typical. The additional "nuisance terms" from the N-th row of Pascal's triangle, even though they are squared, are dwarfed by the largest terms from the ($2N$)-th row. This is why the "approximate" formula agrees with the exact formula in every case I know of.

(3) It is also of interest to derive upper and lower bounds for the exact number of tricks to throw. After all, if you are playing the game with $N = 200$ cards, it may not be so easy to look up the 400^{th} row of Pascal's triangle! I prove the following estimate in this paper:

Theorem: If $N > 400$, then the optimal number of tricks to throw satisfies the inequalities

$$\sqrt{N \ln N / 4} < k^*(N) < \sqrt{N \ln N / 2}.$$

Computer calculations by Stern show that these inequalities hold for $400 \geq N \geq 91$ as well. The left-hand inequality is false for $N = 90$.

As $N \to \infty$, $k^*(N) \sim \sqrt{N \ln N / 2}$. It is interesting that this asymptotic limit is approached extremely slowly. Stern's computer calculations show that all the way up to $N = 500$, the ratio $k^*(N)/\sqrt{N \ln N}$ is closer to 0.5 than it is to 0.7071…, its eventual limit. (For $N = 500$, the ratio is 0.559…)

Finally, as mentioned briefly above, there is a second version of Sun Bin's Legacy, which is to find the strategy that guarantees the highest probability of winning a majority of tricks, regardless of the number of tricks won. Curiously, neither Stern nor I worked seriously on this question. In my case, this was because I expected the majority-of-tricks problem to be harder, because the objective function is nonlinear.

Imagine my astonishment when, within one day of Gary Antonick's post going up on the "Numberplay" blog, one of his readers found the optimal strategy for the majority-of-tricks problem! Here I assume $N = 2n+1$ is odd. Reader Bill Courtney showed that the optimal strategy is to throw n tricks. Thus player 2 pairs his top (n+1) cards against player 1's bottom (n+1) cards, in order. It is easy to see that if there is *any way at all* to win (n+1) tricks, then this strategy will do so. The proof is left to the reader (or see Courtney's comment to [A]).

While Courtney's strategy maximizes the probability of winning a simple majority, it is extravagantly wasteful on the level of tricks. It will on average lose nearly half the tricks. By contrast, the "Pascal's triangle" strategy described above will on average lose only about $\sqrt{N \ln N / 2}$ tricks. By playing with a large enough deck, you can win as close as you want to 100 percent of the tricks!

The outline of the rest of the paper is as follows:

I. Basic Results, Mixed Monge Matrices and the Shape Theorem.

This section sets up the problem as a linear assignment problem, shows that the objective function is given by a "mixed Monge matrix," and derives a weak form of the optimal strategy.

In particular, I show that the optimal strategy always involves throwing some tricks in reverse order, and playing the rest of the tricks in normal order. However, there may be "gaps" in the thrown tricks. Most of the work in this section is due to Stern (unpublished).

II. The Symmetry Lemma.

The objective function in section I leads to a mixed Monge matrix that is symmetric about the "anti-main diagonal," and skew-symmetric (after subtracting a constant from each entry) about the main diagonal. I exploit this symmetry to prove that *if* you have decided which tricks to throw (say tricks 1, 3, and 7) then the optimal strategy for these tricks is to play your *i*-th worst card against your opponent's *i*-th best card. Still, there may be gaps in the thrown tricks.

III. The No-Gaps Theorem.

In this section, which is the most technical one, I show that if N is large enough (at least 10 million) then the optimal strategy has no gaps. That is, you should throw tricks 1, 2, …, k for some k. Although I do not derive the best strategy in this section, the proof depends on knowing that a *very good* strategy is to sacrifice the first $\sqrt{N \ln N/2}$ tricks. Roughly speaking, this strategy beats any strategy with gaps in it.

IV. The Number of Tricks to Throw.

In the last section, I explain the wonderful and totally unexpected connection between the expected number of tricks won and the $2N$-th row of Pascal's triangle. I derive the exact number of tricks to throw, $k^*(N)$, the approximate number $k(N)$, and the asymptotic limit for both of them.

Bibliography.

[A] G. Antonick, "Stern-Mackenzie One-Round War," *New York Times* ("Numberplay" blog), Jan. 13, 2014, at http://wordplay.blogs.nytimes.com/2014/01/13/war/.
[AGY] A. Assad, Golden, B., and Yee, J. Scheduling players in team competitions. *Proceedings of the Second International Conference on Mathematical Modelling (St. Louis, MO, 1979),* Vol I, pp. 369-379. Rolla, MO: Univ. of Missouri-Rolla, 1980.

Addendum.

A much improved version of the proof (valid for all N, not just for $N > 10^7$) was published in R. Chatwin and D. Mackenzie, "How to Win at (One-Round) War," *College Math. Jour.* Vol. 46, No. 4, Sept. 2015, 242-253.

The story behind this paper may be of interest to Gathering for Gardner readers. One of the people who attended my G4G11 talk was Brian Hopkins, the editor of the *College Mathematics Journal* (published by the Mathematical Association of America). He invited me to submit a paper on the Sun Bin problem to the *College Mathematics Journal*.

In the meantime, I had started collaborating with Richard Chatwin, who had read about the Sun Bin problem in Gary Antonick's *New York Times* article referenced above. Chatwin, unlike me, is an expert on linear assignment problems. (He wrote his Ph.D. dissertation on airline over-booking, which involves this type of problem.) He asked the natural question, "What if we use the Hungarian algorithm?"

In the end, it turned out that the Hungarian algorithm *per se* does not solve Sun Bin's problem for all N, although it certainly can solve it for any individual value of N. The fundamental reason is that the algorithm repeatedly involves the step of finding the smallest element in a given row or column of a matrix and "pivoting" about that element. Identifying a particular element as the smallest amounts to proving a whole set of inequalities. The algorithm tells you *what inequalities you need to prove*, but not how to prove them! In fact, Chatwin found repeatedly that the inequalities he needed were precisely the ones already proved in my paper.

Nevertheless, Chatwin did make major improvements to certain parts of my proof, especially to Part III described above. The final (and definitive) result reduced the "at least 10 million cards" requirement to a much more manageable "at least 41 cards." That is, we can prove analytically that the Pascal's triangle strategy is optimal, provided that $N \geq 41$. For $3 \leq N \leq 40$, the analytic estimates are too inexact and we have to resort to a case-by-case analysis on the computer. Chatwin did the computer calculations necessary to prove that the Pascal's triangle strategy remains optimal. (Stern had already checked this for $N \leq 60$, but it was nice to have an independent verification.)

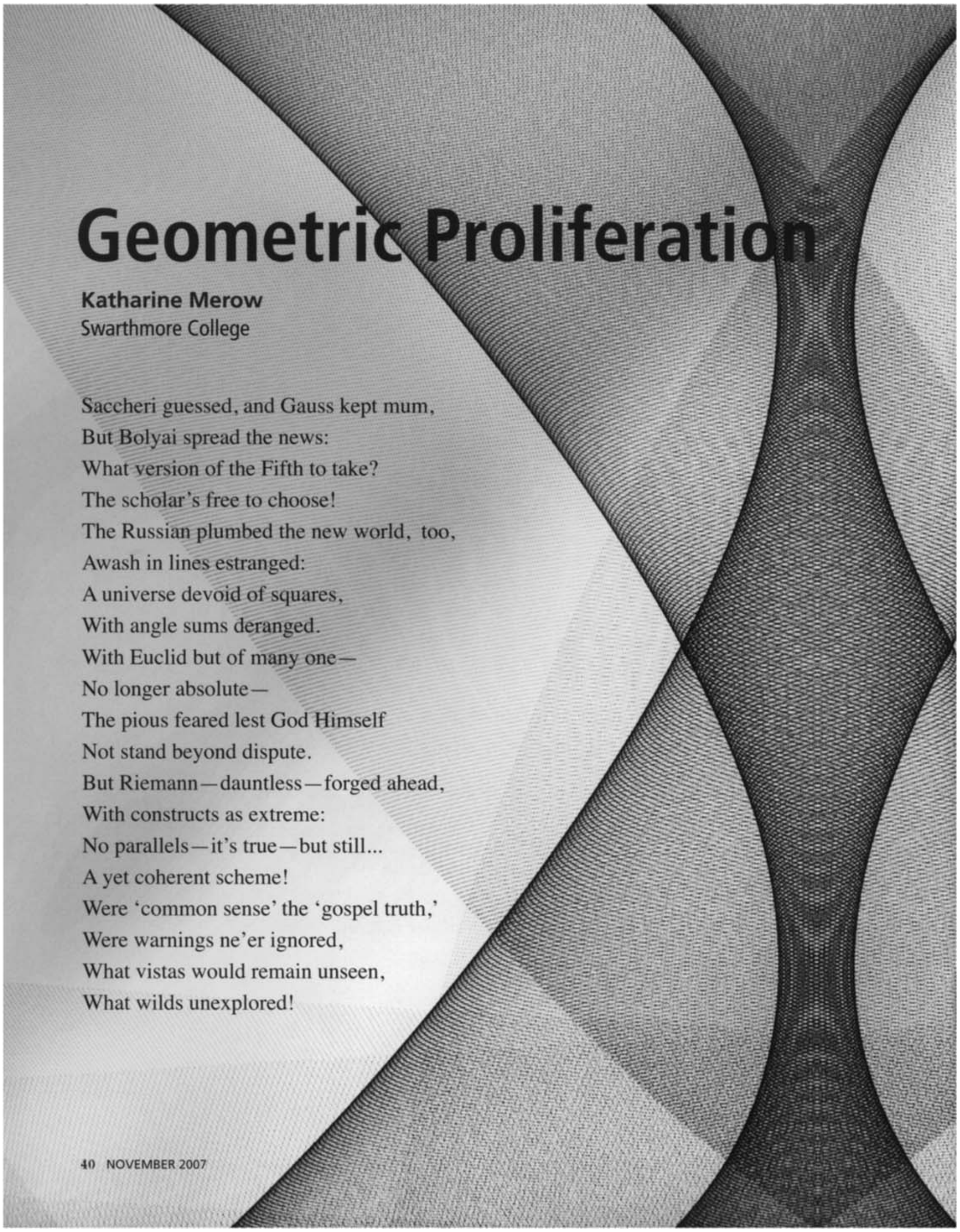
My original contribution to the G4G11 gift exchange was a 56-page paper with the proof that the Pascal's triangle strategy is optimal for $N > 10^7$. Because we now have a much better proof of a more complete result, it no longer seems necessary to me to have the entire 56-page paper reproduced in this volume. However, it does appear in the online version for any readers who might be interested in seeing the not quite fully-baked version of the proof.

I think that Martin Gardner would have approved of the way that a column written for the public (Gary Antonick's blog) put two mathematicians together who never would have been able to find each other otherwise; and the way that a conference in his honor put us in contact with the editor who eventually published our manuscript. So my main gift to the G4G11 exchange is to tell you that Martin Gardner's legacy (as well as Sun Bin's legacy) is alive and well!

Geometric Proliferation

by Katharine Merow | Mathematical Association of America

This poem, published in the MAA's student magazine Math Horizons in 2007, treats a topic fitting for this parallels-themed meeting: non-Euclidean geometry.



Geometric Proliferation

Katharine Merow
Swarthmore College

Saccheri guessed, and Gauss kept mum,
But Bolyai spread the news:
What version of the Fifth to take?
The scholar's free to choose!
The Russian plumbed the new world, too,
Awash in lines estranged:
A universe devoid of squares,
With angle sums deranged.
With Euclid but of many one—
No longer absolute—
The pious feared lest God Himself
Not stand beyond dispute.
But Riemann—dauntless—forged ahead,
With constructs as extreme:
No parallels—it's true—but still...
A yet coherent scheme!
Were 'common sense' the 'gospel truth,'
Were warnings ne'er ignored,
What vistas would remain unseen,
What wilds unexplored!

Stone Piles - A Variation of Nim

by Ryan Morrill | University Of Alberta

We will be analyzing a game that is similar to the usual game of Nim. Players Peter and Betty are alternating turns in a game with m piles of n stones each. Peter is allowed to take 1, 2, or 3 stones, all from any single pile. Betty is allowed to take 1, 2 or 3 stones, all from different piles. Determine for each n and m who has the winning strategy, according to who moves first.

Problem.
Peter and Betty take turns in a game with n stones in each of m piles. In his turn, Peter must take 1, 2 or 3 stones from any one pile. In her turn, Betty must take one stone from 1, 2 or 3 piles. Whoever takes the last stone overall is the winner. Determine for each n and m who has a winning strategy, according to who moves first.

First, let's consider the case when $n = m$.

Observation 1.
For $n = 1, 2, 3$, it is easy to prove that whoever moves first wins.

Observation 2.
For $n = 4$, whoever moves first loses.

(1) Suppose Peter moves first.

- (a) After Peter's first move, he leaves behind (4442), (4442) or (4441).
- (b) Betty responds by leaving behind (3333), (3332) or (3331).
- (c) After Peter's second move, he leaves behind (3332), (3331), (333), (3322), (3321), (332) or (3311).
- (d) Betty responds by leaving behind (2222), (2221), (222), (2211) or (221).
- (e) After Peter's third move, he leaves behind (2221), (222), (2211), (221), (22), (2111), (211) or (21).
- (f) Betty responds by leaving behind (1111), (111) or (11).
- (g) After Peter's fourth move, he leaves behind (111), (11) or (1).
- (h) Betty responds by leaving behind (0) and wins.

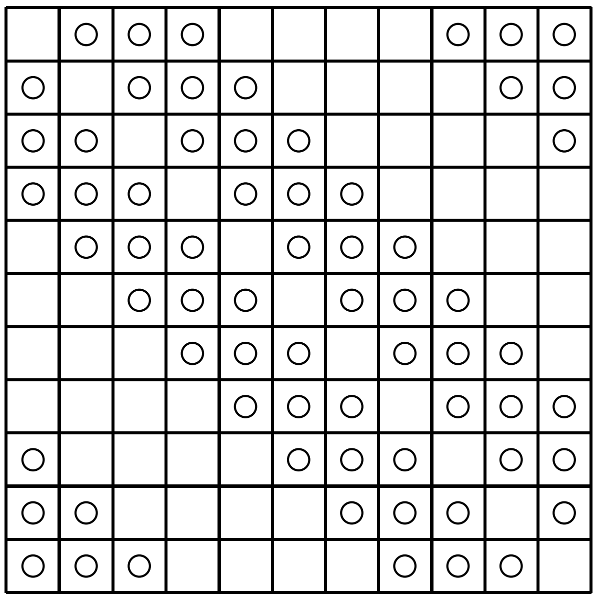
(2) Suppose Betty moves first.

- (a) After Betty's first move, she leave behind (4443), (4433) or (4333).
- (b) Peter responds by leaving behind (444), (443) or (433).
- (c) After Betty's second move, she leaves behind (443), (442), (433), (432), (422), (333), (332) or (322).
- (d) Peter responds by leaving behind (44), (43), (42), (33) or (32).
- (e) After Betty's third move, she leaves behind (43), (42), (41), (33), (32), (31), (22) or (21).
- (f) Peter responds by leaving behind (4), (3) or (2).
- (g) After Betty's fourth move, she leaves behind (3), (2) or (1).
- (h) Peter responds by leaving behind (0) and wins.

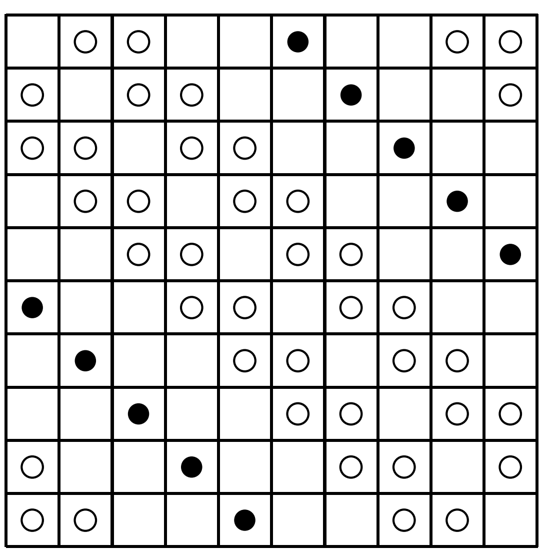
Conjecture.
For $n \geq 5$, Betty wins no matter who moves first.

Generalizing as $m > n$:

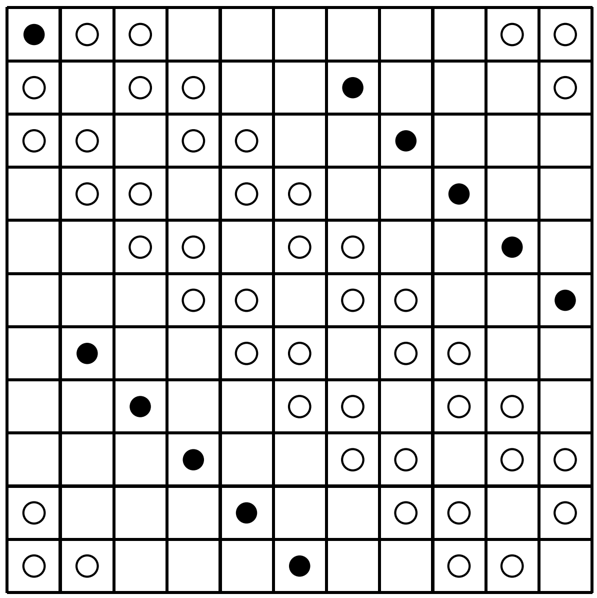
Solution:
We first assume that Peter goes first. We consider three cases.
Case 1. n is even.
Draw an $m \times m$ grid, with each row representing a pile. Place stones on (i, j) if and only if $|i-j| \leq \frac{n}{2}$ or $|i-j| \geq m - \frac{n}{2}$. Then there are n stones in each of the m rows, none lying on the main diagonal where $i = j$. The entire configuration is symmetric about this main diagonal. Betty's strategy is to take (j, i) whenever Peter takes (i, j). By symmetry, Betty gets the last stone and wins. The diagram below illustrates the case $m = 11$ and $n = 6$.



Case 2. n is odd but m is even.
Here $n - 1$ is even. So we temporarily replace n by $n - 1$, and draw an $m \times m$ grid as in Case 1. We now add a stone to each row at (i, j) where $|i - j| = \frac{m}{2}$. The argument is then the same as in Case 1. The diagram below illustrates the case $m = 10$ and $n = 5$.



Case 3. n is odd and m is also odd.
Here $n-1$ is even. So we temporarily replace n by $n-1$, and draw an $m \times m$ grid as in Case 1. We now add a stone to each row except the first at (i,j) where $|i-j| = \frac{m-1}{2}$, and another stone at (1,1). As in Case 1, the overall configuration is symmetric about the main diagonal, even though there is now a stone at (1,1). Since each pile has the same number of stones at the start, Betty may assume that Peter takes stones from the first pile, and that they include the one at (1,1). If Peter takes at least two stones, Betty can take the stones in matching positions except for (1,1). If Peter takes (1,1), Betty can take (1,2) and (2,1). After this, she can use the symmetry strategy and wins. The diagram below illustrates the case $m = 11$ and $n = 5$.



Suppose Betty goes first. In Case 3, she just removes (1,1) the lone stone on the main diagonal. In Cases 1 or 2, she removes any pair of stones symmetric about the main diagonal. Thereafter, she uses the symmetry strategy as before.

All that is left is to consider when $m < n$

Computer Search Results by Brian Chen

Black circle means Peter wins. White circle means Betty wins. Rows are piles. Columns are stones.

Peter moves first:

P	1	2	3	4	5	6	7	8	9	10	11	12
1	●	○	○	○	○	○	○	○	○	○	○	○
2	●	●	○	○	○	○	○	○	○	○	○	○
3	●	●	●	○	○	○	○	○	○	○	○	○
4	●	●	●	○	○	○	○	○	○	○	○	○
5	●	●	●	●	○	○	○	○	○	○	○	○
6	●	●	●	●	●	○	○	○	○	○	○	○
7	●	●	●	●	●	○	○	○	○	○	○	○
8	●	●	●	●	●	●	○	○	○	○	○	○
9	●	●	●	●	●	●	●	○	○	○	○	○
10	●	●	●	●	●	●	●	○	○	○	○	○
11	●	●	●	●	●	●	●	●	○	○	○	○
12	●	●	●	●	●	●	●	●	●	○	○	○

Betty moves first:

B	1	2	3	4	5	6	7	8	9	10	11	12
1	○	○	○	○	○	○	○	○	○	○	○	○
2	●	○	○	○	○	○	○	○	○	○	○	○
3	●	●	○	○	○	○	○	○	○	○	○	○
4	●	●	●	⊙	○	○	○	○	○	○	○	○
5	●	●	●	●	○	○	○	○	○	○	○	○
6	●	●	●	●	○	○	○	○	○	○	○	○
7	●	●	●	●	○	○	○	○	○	○	○	○
8	●	●	●	●	●	○	○	○	○	○	○	○
9	●	●	●	●	●	○	○	○	○	○	○	○
10	●	●	●	●	●	●	○	○	○	○	○	○
11	●	●	●	●	●	●	●	○	○	○	○	○
12	●	●	●	●	●	●	●	○	○	○	○	○

Closing Remarks

The motivation for this problem is from a problem presented in the mathematics contest Tournament of the Towns. It's origin is from the Fall 2013, A-level contest. It is a unique variation from the well known game of Nim.

Upon first inspection, it seems as though there is some symmetry between the players Peter and Betty, however it turns out not to be so simple. By the computer search results by Brian Chen, it seems as though Betty actually has an advantage over Peter, but the unsolved problem is to what extent does Betty have an advantage. In other words, what happens when $n > m$?

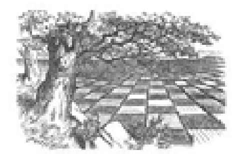
Another point of curiosity is when Betty goes first, and $n = m = 4$. This point is certainly an anomaly in the usual pattern. It is labelled with a dark dot and a circle.

We can come up with a solution for $n = m, n < m$, so all that is left is when $n > m$. This remaining case is much more difficult than the first two, however we can make some guesses about the trend, based off of the computer search results by Brian Chen.

Morning Math with Red Queen

by Marnie Muller | Universe Journey

MORNING MATH
not After-math
from Alice/Lacie
in One-Deux Land



Morning Math
(…not After-Math)
with
the Red Queen
&
Alice/Lacie
in
One-Deux
Land

Y
sum x's
5 - f
believed >
6 Impossible Things
B 4
Breakfast

from Marnie Muller, MLA in Asheville, North Carolina
excellentwordworks@gmail.com

Gift for G4G11: Gabriel's Paper Horn

David Richeson
Dickinson College
richesod@dickinson.edu

Gabriel's horn is the surface obtained by revolving the curve $y = 1/x$ ($x \geq 1/2$) about the x-axis (see figure 1). Mathematics professors wow introductory calculus students by sharing its paradoxical properties: it has finite volume, but infinite surface area. As they say, "you can fill it with paint, but you can't paint it."

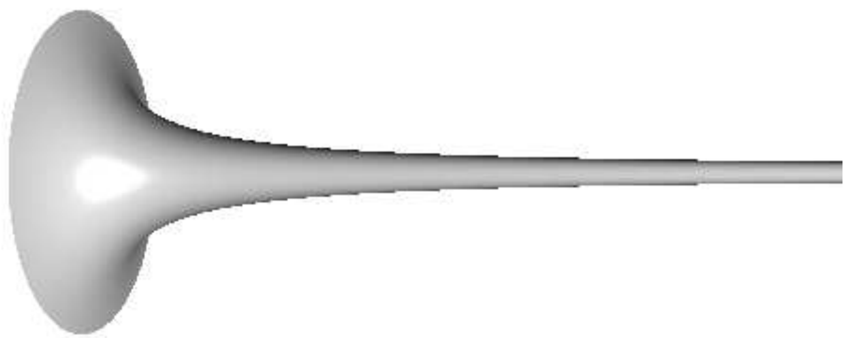


Figure 1. Gabriel's horn

At the end of this article we provide two templates (one color, one white) for making a model of Gabriel's horn out of paper cones, such as the one in figure 2. The instructions are simple: Cut out the sectors, tape them together to form cones, and stack the cones in numerical order. The approximation is good only for the first 1/4" of the last cone, so the stack of cones should be cut off at that height.

The rest of this article describes the mathematics used to make the sectors. In short: Take evenly spaced points along the curve, find the segments of the tangent lines between these points and the x-axis, and use them to generate the cones (see figure 3). We use $y = 1/x$ as the generating curve, but this procedure works for any curve that is positive, decreasing, and concave up.

First we use calculus to find the equation of the tangent line to the graph at the point $(a, 1/a)$:

$$y - \frac{1}{a} = -\frac{1}{a^2}(x - a).$$

Then we set $y = 0$ to find the x-intercept of the tangent line: $x = 2a$. From this we conclude that the radius of the base of the corresponding cone is $1/a$ and the slant height is $\sqrt{a^2 - 1/a^2}$.

Now, imagine cutting the cone and unrolling it into a sector with central angle θ. The arc of the sector is the circumference of the base of the cone,

FIGURE 2. Gabriel's horn made from paper cones

$2\pi/a$. The radius of the sector is the slant height of the cone, so the circumference of a full paper disk with this radius is $2\pi\sqrt{a^2+(1/a)^2}$. We use ratios to find θ:

$$\frac{\theta}{360^\circ} = \frac{\text{circumference(cone)}}{\text{circumference(paper)}} = \frac{2\pi/a}{2\pi\sqrt{a^2-1/a^2}}.$$

So, $\theta = \left(360/\sqrt{a^4+1}\right)^\circ$.

We want the cones to be evenly spaced along the surface. That is, we want the visible bands to have the same widths. To accomplish this we must find points that are evenly spaced along the curve (see figure 4). Again we

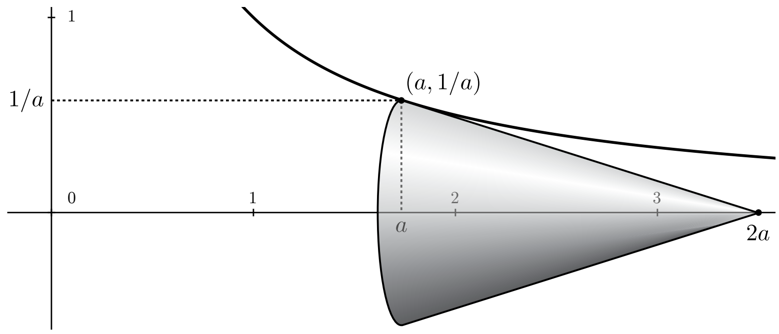


FIGURE 3. A tangent line segment revolved to obtain a cone

turn to calculus. The first cone corresponds to the point $(0.5, 2)$. The length of the curve from $(0.5, 2)$ to $(a, 1/a)$ is

$$\int_{0.5}^{a} \sqrt{1 + \frac{1}{x^4}}\, dx.$$

We used a computer algebra system to find the seventeen a-values so that this integral equals $0.25, 0.5, 0.75, \ldots, 4.25$.

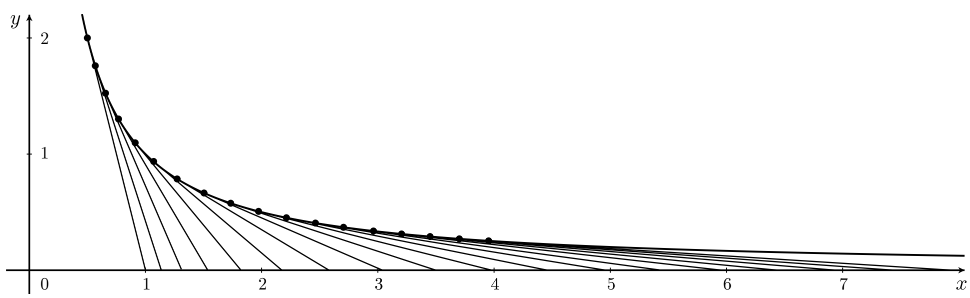


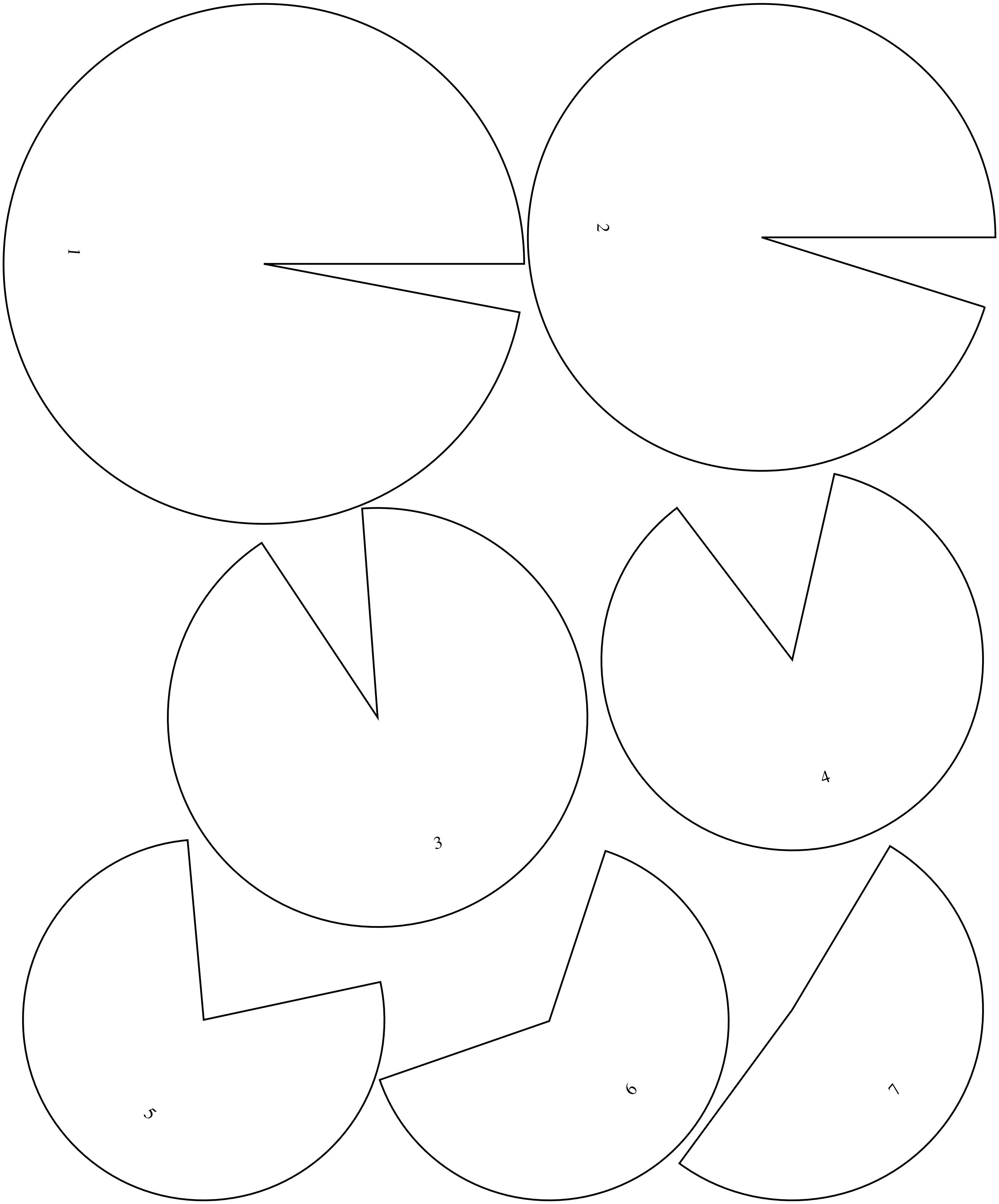
FIGURE 4. Evenly spaced line segments tangent to $y = 1/x$

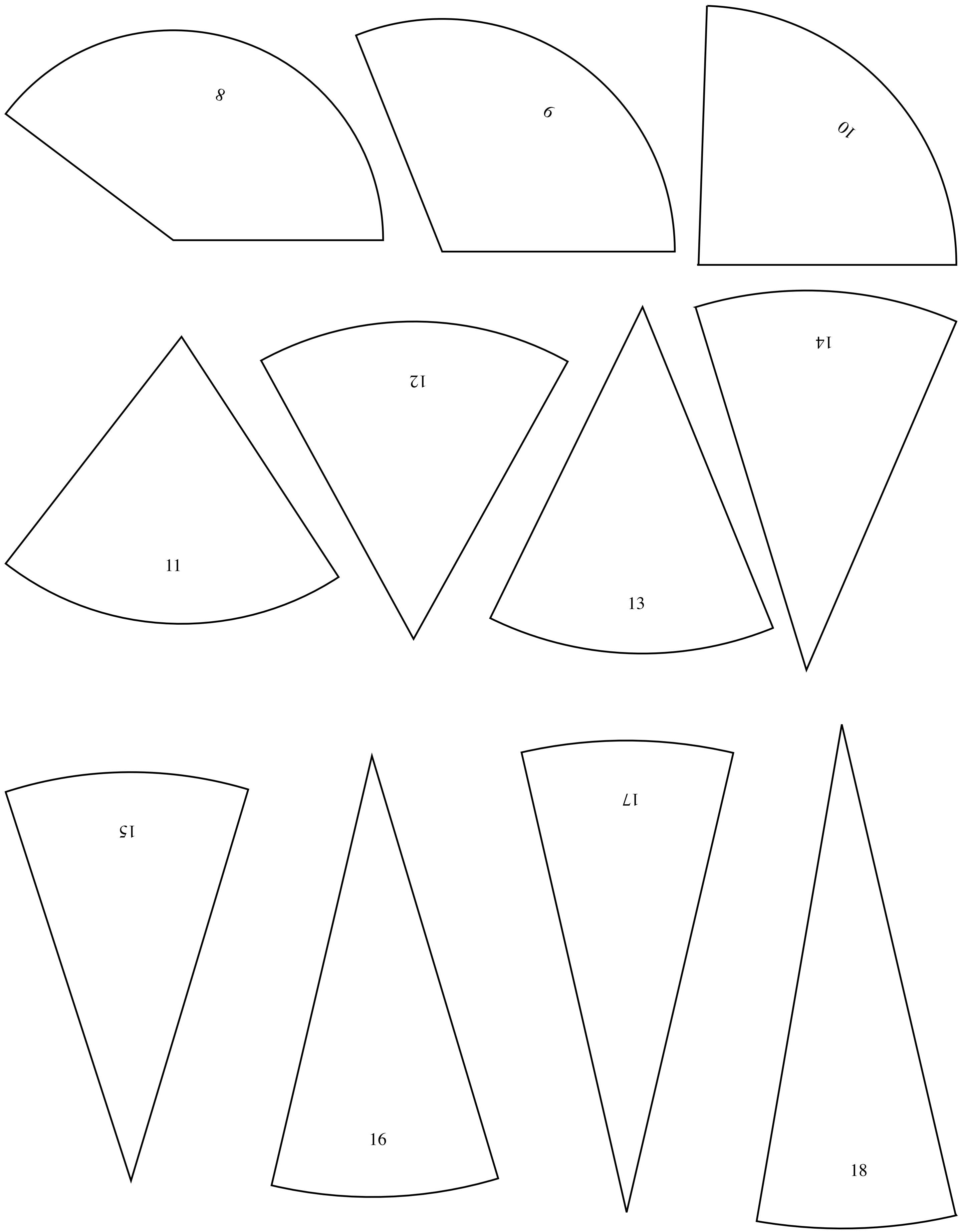
Finally, for each of these a-values, we construct a paper sector with radius $\sqrt{a^2 - 1/a^2}$ and central angle $\left(360/\sqrt{a^4+1}\right)^\circ$.

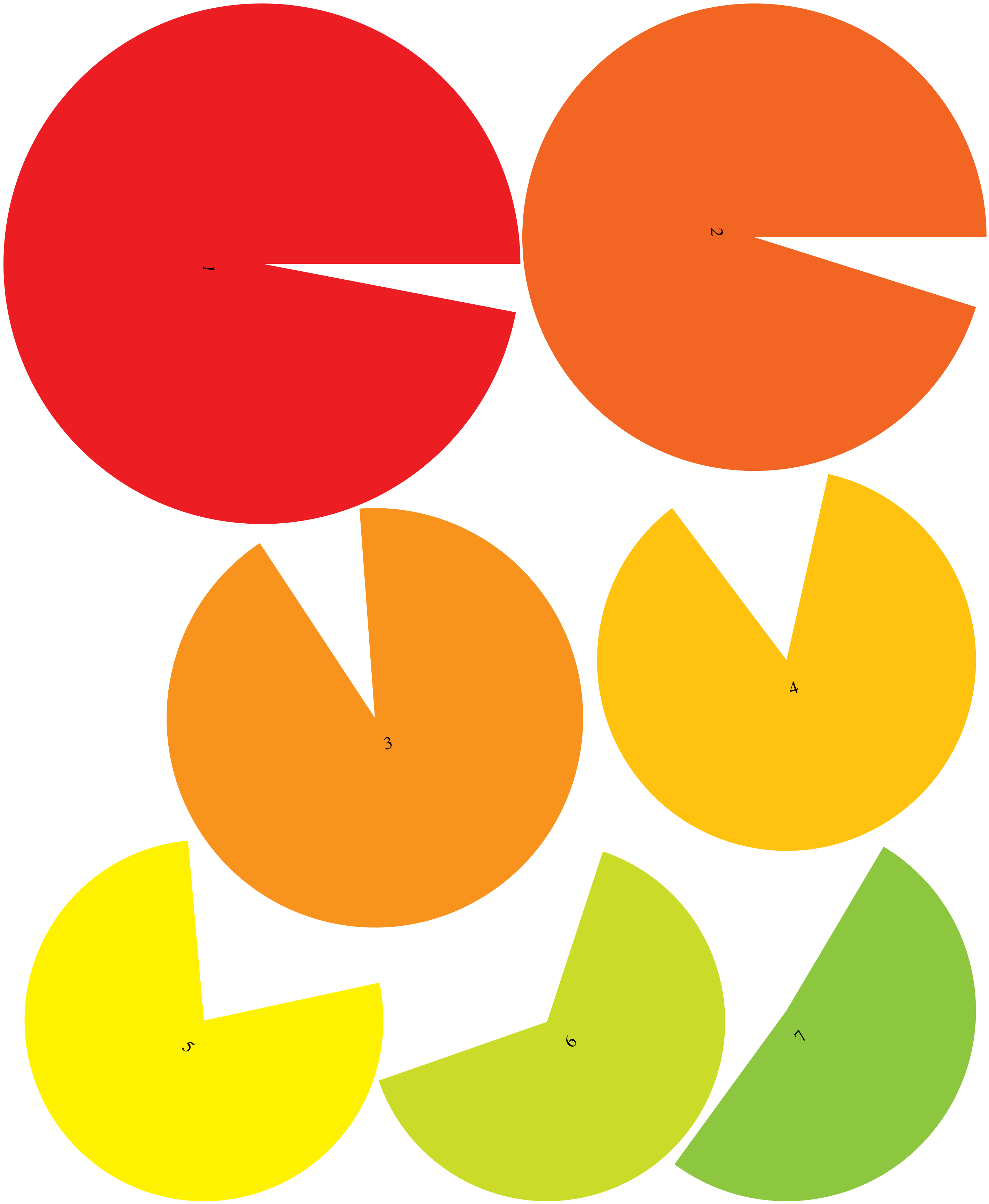
This project was inspired by Daniel Walsh [2] and Burkard Polster [1] who made pseudospheres out of paper cones. The pseudosphere is generated by revolving the tractrix about the x-axis. That model has the special property that the radii of the paper disks are all the same.

References

[1] Polster, Burkard "TracTricks," *Math Horizons*, April 2014, pp. 18–19.
[2] Walsh, Daniel, "Sudo make me a pseudosphere," December 11, 2012, http://danielwalsh.tumblr.com/post/2173134224/sudo-make-me-a-pseudosphere


1
2
3
4
5
6
7


8
9
10
11
12
13
14
15
16
17
18


1
2
3
4
5
6
7


8
9
10
11
12
13
14
15
16
17
18

What Happened a Few Minutes Before Day 0?

by Lisa Rougetet

Abstract:

I am currently a PhD student in history of mathematics and I am working on the history of the developments of combinatorial game theory. Quite naturally, I started my study with Charles Leonard Bouton's first article on the Nim game, and the several variations that appeared afterwards (Wythoff's Nim, Moore's Nim_k, Der Letzte gewinnt ! of Ahrens). Then, I went through the works between these articles and Sprague-Grundy's results to understand the construction and the completion of the impartial games theory. I realized that the German mathematician and chess player Emanuel Lasker brought many interesting ideas which definitely inspired Roland Sprague in his analysis of a global game seen as a sum of simpler ones. After that, Dawson's Chess played an important role for Richard Guy and Cedric Smith in the classification of octal games and their almost complete resolution. The work was much harder for Patrick Grundy and Cedric Smith when they tried to do the same with games in version misère, and even nowadays, the theory is far from being complete. Finally, I read interviews of Elwyn Berlekamp, Richard Guy and John Conway about the elaboration of Winning Ways for Your Mathematical Plays and interview of the latter about the writing of On Numbers and Games. I stopped my study at this point because I think On Numbers and Games reaches the highest level of mathematical abstraction applied to really concrete games.

Basically I know what John Conway said, but what I would like to know, is what he has not said… yet! How did he come to the creation of surreal numbers? What readings of the works behind Bouton's and Sprague-Grundy's results influenced him in his understanding of games? In his opinion, what mathematical resources helped him to construct his entire theory, and were missing to his predecessors? Who are Arthur, Bertha and Anne-Louise? And so many other questions!

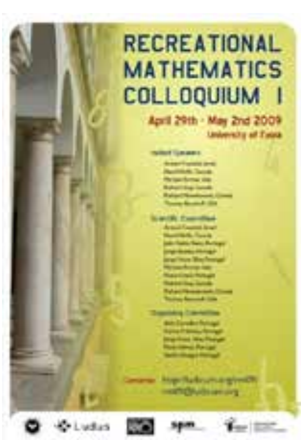
Gathering for Gardner

Carlos P. Santos

www.sites.google.com/site/cpshomepage

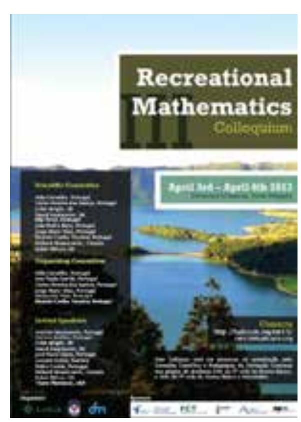
rmm.ludus-opuscula.org

Recreational Mathematics Magazine is published by the Ludus Association with the support of the Interuniversity Centre for the History of Science and Technology. The journal is electronic and semiannual, and focuses on results that provide amusing, witty but nonetheless original and scientifically profound mathematical nuggets. The issues are published in the exact moments of the equinox. It's a magazine for mathematicians, and other mathematics lovers, who enjoy imaginative ideas, non-standard approaches, and surprising procedures. Occasionally, some papers that do not require any mathematical background will be published.



ludicum.org/ev/rm/09

ex.ludicum.org/rm11

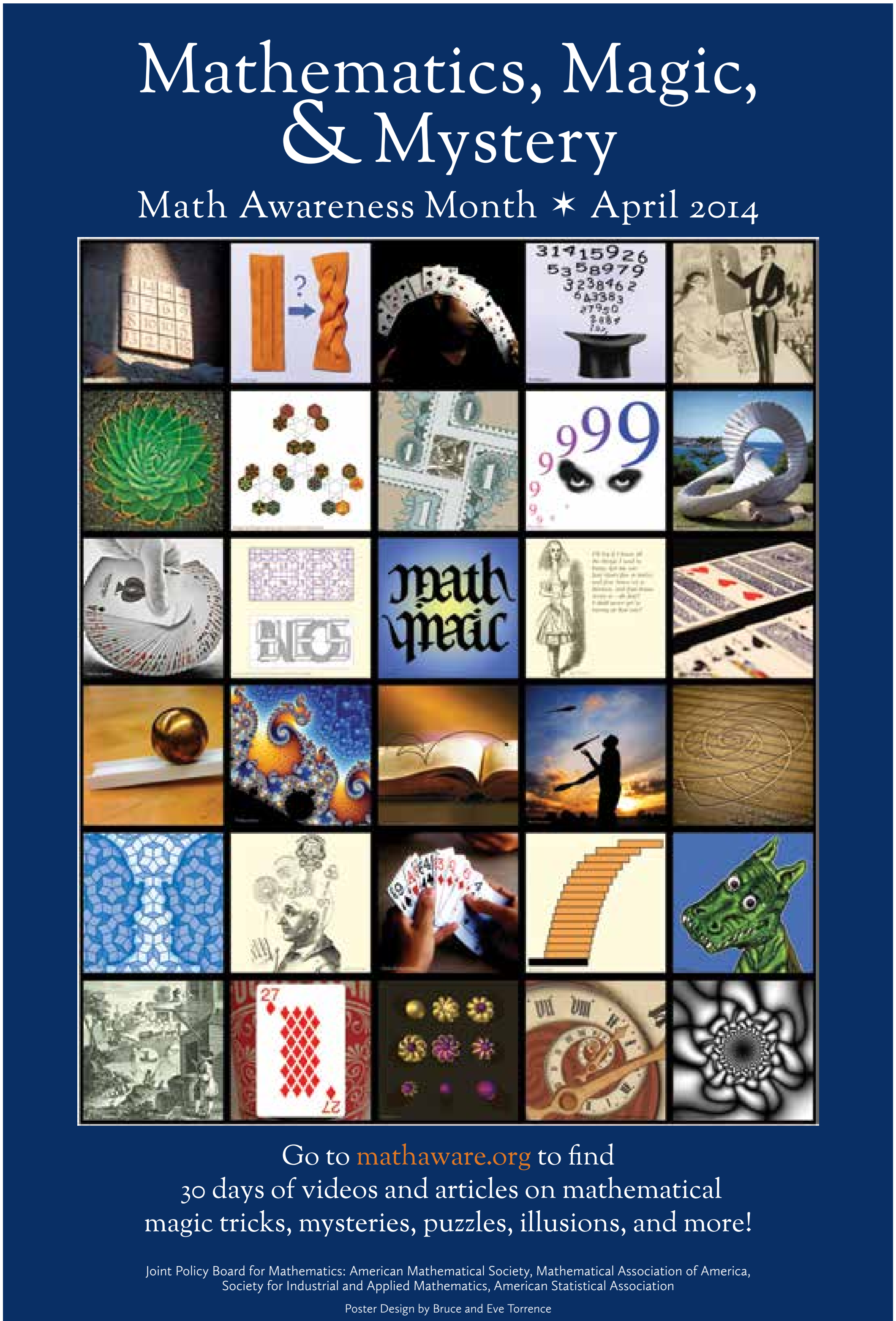


ludicum.org/ev/rm/13

Mathematics Awareness Month 2014 Poster

by Bruce Torrence & Eve Torrence | Randolph-Macon College

The Theme of MAM 2014 is Mathematics, Magic, and Mystery, after the 1956 book by MG. The poster shows teaser icons for the 30 topics that will be revealed, one each day, in April on the MAM website.



G4G11 Exchange Book

VOLUMES 1 & 2

ATLANTA, GEORGIA

March 19 - 23, 2014

Within these pages you will discover a world of magic, puzzles, mathematics and art. Individuals contributed to this book as part of the 11th Gathering 4 Gardner gift exchange in which attendees swapped puzzles, games, magic tricks, papers, artwork and novelty items. The biennial conference brings together a diverse group of highly enthusiastic mathematicians, magicians, puzzlers, artists, and other engaged and engaging attendees. With this book, each of them hopes to share with you insights into what sparks their interest and passion.

Published by:
Gathering 4 Gardner