The UMAP Journal

Vol. 16, No. 3

Vol. 16, No 3   1995

Table of Contents

Publisher's Editorial

Once More into the Breach

Solomon A. Garfunkel
Executive Director
COMAP, Inc.
57 Bedford St., Suite 210
Lexington, MA 02173
S.Garfunkel@mail.comap.com

This seems an extremely auspicious time to present an update on COMAP activities. The MCM issue of *The UMAP Journal* is always a cause for celebration at COMAP. It seems hard to believe that this coming February will mark the twelfth running of the contest. It also marks the publication of *MCM: The First Ten Years*, a review of the first ten years of the Mathematical Contest in Modeling. We hope that all of you who read the *Journal* and have followed this competition will receive a copy and appreciate as much as I do, the work of Paul Campbell in putting this volume together.

What else is doing at COMAP? Well, as many of you are aware, we are in the middle of a major secondary-school curriculum development project—ARISE. This project is producing an NCTM-Standards–based comprehensive grades 9–11 mathematics curriculum. We have already produced and are field-testing materials for grades 9 and 10 and hope to have all of the materials available for commercial publication in 1998. Needless to say, this project has consumed a great deal of our energies (and not an insubstantial amount of NSF funds).

But I want to focus on our major new undergraduate project, which is just coming to fruition. For the past four years, COMAP has been working on a new entry-level course for mathematics and science majors, which we have entitled *Principles and Practice of Mathematics*. This course, funded by the Division of Undergraduate Education of NSF, has resulted in a one-year mathematics text to be published by Springer-Verlag in early 1996. Our expressed purpose is to show students both the breadth and the depth of mathematics. Our goals are to attract and retain more students in serious mathematics courses of study and to provide students with a much broader early view of what our subject is about.

Perhaps the best way to describe this course is by listing the table of contents. This will also serve to recognize the contributions of a superb author team.

1. Change (Frank Giordano and Chris Arney, U.S. Military Academy, and Sheldon Gordon, Suffolk Community College, SUNY)
2. Position (Robert Bumcrot, Hofstra University)
3. Linear Algebra (Alan Tucker, SUNY at Stony Brook)
4. Combinatorics (Rochelle Wilson Meyer, Nassau Community College)
5. Graph Theory and Algorithms (Paul J. Campbell, Beloit College)
6. Analysis of Algorithms (Rochelle Wilson Meyer, Nassau Community College)
7. Logic and the Design of Intelligent Machines (Rochelle Wilson Meyer, Nassau Community College)
8. Chance (Michael Olinick, Middlebury College)
9. Modern Algebra (Joseph Gallian, University of Minnesota–Duluth)

The authors were encouraged and guided by editor Walter Meyer (Adelphi University), and Zaven Karian (Denison University) served as technology advisor. Moreover, we were fortunate from the beginning of this project to have the advice of a powerful and prestigious advisory committee, whose members included:

- Saul Gass, University of Maryland
- Andrew Gleason, Harvard University
- Zaven Karian, Denison University
- Joseph Malkevitch, York College, CUNY
- David Moore, Purdue University
- Henry Pollak, Bellcore (retired)
- Paul Sally, University of Chicago
- J. Laurie Snell, Dartmouth College
- Marcia Sward, Mathematical Association of America
- Alan Tucker, SUNY at Stony Brook
- Carol Wood, Wesleyan University
- Gail Young, Columbia Teachers College

The text has been field-tested in draft form at some 40 colleges across the country, and we are grateful to the brave faculty at these institutions for their efforts in making this a teachable and user-friendly book.

The publication of any new textbook always feels like the completion of a great deal of work, and *Principles and Practice* is no exception. But we are more than aware that in this case publication of the text is the beginning of a great deal of additional work. This is, after all, a revolutionary text. There are no courses with this title in college catalogs. Adoption of this text (in its entirety) implies a reconfiguration of the undergraduate math major. This sort of change doesn't happen over night. However, we are convinced that on reading this work, faculty will see many opportunities for enhancing the first undergraduate mathematics experience of majors. We hope that the charm of the ideas and the modernity of the applications will foster a will to experiment. We believe that through this experimentation we will take a new look at our curriculum and that real change can occur. What better way to describe what COMAP has always stood for?

About the Author

Sol Garfunkel received his Ph.D. in mathematical logic from the University of Wisconsin in 1967. He was at Cornell University and the University of Connecticut at Storrs for eleven years and has dedicated the last 20 years to research and development efforts in mathematics education. He has been the Executive Director of COMAP since its inception in 1980.

He has directed a wide variety of projects, including UMAP (Undergraduate Mathematics and Its Applications Project), which led to the founding of this *Journal*, and HiMAP (High School Mathematics and Its Applications Project), both funded by the NSF. For Annenberg/CPB, he has directed three telecourse projects: *For All Practical Purposes*, *Against All Odds: Inside Statistics*, and *In Simplest Terms: College Algebra*. He is currently co-director of the Applications Reform in Secondary Education (ARISE) project, a comprehensive curriculum development project for secondary-school mathematics.

Modeling Forum

Results of the 1995 Mathematical Contest in Modeling

Frank Giordano, MCM Director
Department of Mathematical Sciences
United States Military Academy
West Point, NY 10996–1786

Introduction

A total of 320 teams of undergraduates, from 194 schools, spent the third weekend in February working on applied mathematics problems. They were part of the eleventh Mathematical Contest in Modeling (MCM). On Friday morning, the MCM faculty advisor opened a packet and presented each team of three students with a choice of one of two problems. After a weekend of hard work, typed solution papers were mailed to COMAP on Monday. Seven of the top papers appear in this issue of *The UMAP Journal*.

Results and winning papers from the first ten contests were published in special issues of *Mathematical Modeling* (1985–1987) and *The UMAP Journal* (1985–1994). The 1994 volume of *Tools for Teaching*, commemorating the tenth anniversary of the contest, contains all of the 20 problems used in the first ten years of the contest and a winning paper for each. Limited quantities of that volume and of the special MCM issues of the *Journal* for the last few years are available from COMAP.

Problem A: The Single Helix

The problem consists of assisting a small biotechnological company in designing, proving, programming, and testing a mathematical algorithm to locate ”in real time” all the intersections of a helix and a plane in general positions in space (see **Figure 1**).

Similar programs for Computer Aided Geometric Design (CAGD) enable engineers to view a plane section of the object that they design, for example, an aircraft jet engine, an automobile suspension, or a medical device. Moreover, engineers may also display on the plane section such quantities as air flow, stress, or temperature, coded by colors or level curves. Furthermore,

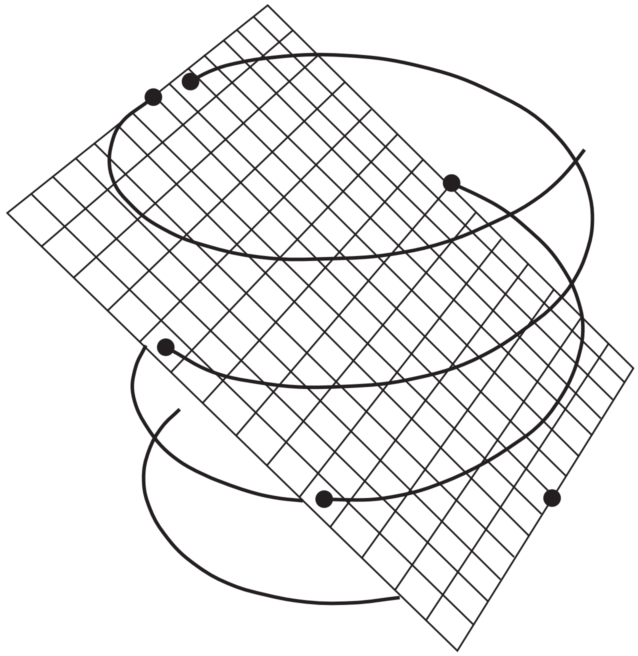

Figure 1. Some intersections of a helix with a plane.

engineers may rapidly sweep such plane sections through the entire object to gain a three-dimensional visualization of the object and its reactions to motion, forces, or heat. To achieve such results, the computer programs must locate all the intersections of the viewed plane and every part of the designed object with sufficient speed and accuracy. General "equation solvers" may in principle compute such intersections; but for specific problems, special methods may prove faster and more accurate than general methods. In particular, general software for Computer Aided Geometric Design may prove too slow to complete computations in real time, or too large to fit in the finished medical devices being developed by the company. The considerations just explained have led the company to the following problem.

> **Problem:** Design, justify, program, and test a method to compute all the intersections of a plane and a helix in general positions (at any locations and with any orientations) in space.

A segment of the helix may represent, for example, a helicoidal suspension spring or a piece of tubing in a chemical or medical apparatus.

The need for some theoretical justification of the proposed algorithm arises from the necessity of verifying the solution from several points of view. This can be done through mathematical proofs of parts of the algorithm, and through tests of the final program with known examples. Such documentation and tests will be required by government agencies for medical use.

Problem B: Aluacha Balaclava College

Aluacha Balaclava College has just hired a new Provost. Problems with faculty compensation at the college forced the former Provost to resign, so the new Provost needs to make the institution of a fair and reasonable compensation system her first priority. As a first step in this process, she has hired your team as consultants to design a compensation system that reflects the following circumstances and principles.

Circumstances

There are four faculty ranks: Instructor, Assistant Professor, Associate Professor and Professor, in ascending order. Faculty with Ph.D. degrees are hired at the rank of Assistant Professor. Faculty who are working on a Ph.D. are hired at the rank of Instructor and promoted automatically to Assistant Professor upon completion of their degrees. Faculty may apply for promotion from Associate Professor to Professor after serving at the rank of Associate for seven or more years. The promotion decisions are made by the Provost with recommendations from a faculty committee and are not your concern.

Faculty salaries are for the ten-month period September through June. Raises are always effective beginning in September. The total amount of money available for raises varies from year to year and generally is not known until March for the following year.

The starting salary this year for an Instructor with no prior teaching experience was $27,000 and for an Assistant Professor was $32,000. Faculty can receive credit, upon hire, for as much as seven years of teaching experience at other institutions.

Principles

- All faculty should get a raise any year that money is available.
- Faculty should get a substantial benefit from promotion. If one is promoted in the minimum possible time, the benefit should be roughly equal to seven years of normal (non-promotion) raises.
- Faculty who get promoted on time (after seven or eight years in rank) and have careers of 25 or more years should make roughly twice as much at retirement as a new Ph.D. starting off.
- Faculty in the same rank with more experience should be paid more than others with less experience. But the effect of an additional year of experience should diminish over time. In other words, if two faculty stay in the same rank, their salaries should tend to get closer over time.

The Project

First, design a new pay system without cost-of-living increases. Then incorporate cost-of-living increases. The final piece of this project is to design a transition process for existing faculty that will move all salaries towards your system without cutting anyone's salary. The existing faculty salaries, ranks and years of service, are in **Table 1**. Discuss any refinements that you think would improve your system.

The Provost has asked for a detailed pay system plan that she can use for implementation, as well as a short executive summary in clear language, which she can present to the Board and to the faculty. The summary should outline the model, its assumptions, its strengths and weaknesses, and the expected results.

The Results

The solution papers were coded at COMAP headquarters so that names and affiliations of the authors would be unknown to the judges. Each paper was then read preliminarily by two "triage" judges at Salisbury State University, Maryland. At the triage stage, the summary and overall organization are the basis for judging a paper. If the judges' scores diverged for a paper, the judges conferred; if they still did not agree on a score, a third judge evaluated the paper.

Final judging took place at Harvey Mudd College, Claremont, California. The judges classified the papers as follows:

	Outstanding	Meritorious	Honorable Mention	Successful Participation	Total
Helix Intersection	3	18	43	82	146
College Salaries	4	26	41	103	174
	7	44	84	185	320

The judges designated seven papers as Outstanding. They appear in this special issue of *The UMAP Journal*. We list those teams and the Meritorious teams (and advisors) below; the list of all participating schools, advisors, and results is in the **Appendix**.

Table 1.

Salary data for Problem B.

Case	Years	Rank	Salary	Case	Years	Rank	Salary	Case	Years	Rank	Salary
1	4	ASSO	54,000	2	19	ASST	43,508	3	20	ASST	39,072
4	11	PROF	53,900	5	15	PROF	44,206	6	17	ASST	37,538
7	23	PROF	48,844	8	10	ASST	32,841	9	7	ASSO	49,981
10	20	ASSO	42,549	11	18	ASSO	42,649	12	19	PROF	60,087
13	15	ASSO	38,002	14	4	ASST	30,000	15	34	PROF	60,576
16	28	ASST	44,562	17	9	ASST	30,893	18	22	ASSO	46,351
19	21	ASSO	50,979	20	20	ASST	48,000	21	4	ASST	32,500
22	14	ASSO	38,462	23	23	PROF	53,500	24	21	ASSO	42,488
25	20	ASSO	43,892	26	5	ASST	35,330	27	19	ASSO	41,147
28	15	ASST	34,040	29	18	PROF	48,944	30	7	ASST	30,128
31	5	ASST	35,330	32	6	ASSO	35,942	33	8	PROF	57,295
34	10	ASST	36,991	35	23	PROF	60,576	36	20	ASSO	48,926
37	9	PROF	57,956	38	32	ASSO	52,214	39	15	ASST	39,259
40	22	ASSO	43,672	41	6	INST	45,500	42	5	ASSO	52,262
43	5	ASSO	57,170	44	16	ASST	36,958	45	23	ASST	37,538
46	9	PROF	58,974	47	8	PROF	49,971	48	23	PROF	62,742
49	39	ASSO	52,058	50	4	INST	26,500	51	5	ASST	33,130
52	46	PROF	59,749	53	4	ASSO	37,954	54	19	PROF	45,833
55	6	ASSO	35,270	56	6	ASSO	43,037	57	20	PROF	59,755
58	21	PROF	57,797	59	4	ASSO	53,500	60	6	ASST	32,319
61	17	ASST	35,668	62	20	PROF	59,333	63	4	ASST	30,500
64	16	ASSO	41,352	65	15	PROF	43,264	66	20	PROF	50,935
67	6	ASST	45,365	68	6	ASSO	35,941	69	6	ASST	49,134
70	4	ASST	29,500	71	4	ASST	30,186	72	7	ASST	32,400
73	12	ASSO	44,501	74	2	ASST	31,900	75	1	ASSO	62,500
76	1	ASST	34,500	77	16	ASSO	40,637	78	4	ASSO	35,500
79	21	PROF	50,521	80	12	ASST	35,158	81	4	INST	28,500
82	16	PROF	46,930	83	24	PROF	55,811	84	6	ASST	30,128
85	16	PROF	46,090	86	5	ASST	28,570	87	19	PROF	44,612
88	17	ASST	36,313	89	6	ASST	33,479	90	14	ASSO	38,624
91	5	ASST	32,210	92	9	ASSO	48,500	93	4	ASST	35,150
94	25	PROF	50,583	95	23	PROF	60,800	96	17	ASST	38,464
97	4	ASST	39,500	98	3	ASST	52,000	99	24	PROF	56,922
100	2	PROF	78,500	101	20	PROF	52,345	102	9	ASST	35,798
103	24	ASST	43,925	104	6	ASSO	35,270	105	14	PROF	49,472
106	19	ASSO	42,215	107	12	ASST	40,427	108	10	ASST	37,021
109	18	ASSO	44,166	110	21	ASSO	46,157	111	8	ASST	32,500
112	19	ASSO	40,785	113	10	ASSO	38,698	114	5	ASST	31,170
115	1	INST	26,161	116	22	PROF	47,974	117	10	ASSO	37,793
118	7	ASST	38,117	119	26	PROF	62,370	120	20	ASSO	51,991
121	1	ASST	31,500	122	8	ASSO	35,941	123	14	ASSO	39,294
124	23	ASSO	51,991	125	1	ASST	30,000	126	15	ASST	34,638
127	20	ASSO	56,836	128	6	INST	35,451	129	10	ASST	32,756
130	14	ASST	32,922	131	12	ASSO	36,451	132	1	ASST	30,000
133	17	PROF	48,134	134	6	ASST	40,436	135	2	ASSO	54,500
136	4	ASSO	55,000	137	5	ASST	32,210	138	21	ASSO	43,160
139	2	ASST	32,000	140	7	ASST	36,300	141	9	ASSO	38,624
142	21	PROF	49,687	143	22	PROF	49,972	144	7	ASSO	46,155
145	12	ASST	37,159	146	9	ASST	32,500	147	3	ASST	31,500
148	13	INST	31,276	149	6	ASST	33,378	150	19	PROF	45,780
151	4	PROF	70,500	152	27	PROF	59,327	153	9	ASSO	37,954
154	5	ASSO	36,612	155	2	ASST	29,500	156	3	PROF	66,500
157	17	ASST	36,378	158	5	ASSO	46,770	159	22	ASST	42,772
160	6	ASST	31,160	161	17	ASST	39,072	162	20	ASST	42,970
163	2	PROF	85,500	164	20	ASST	49,302	165	21	ASSO	43,054
166	21	PROF	49,948	167	5	PROF	50,810	168	19	ASSO	51,378
169	18	ASSO	41,267	170	18	ASST	42,176	171	23	PROF	51,571
172	12	PROF	46,500	173	6	ASST	35,798	174	7	ASST	42,256
175	23	ASSO	46,351	176	22	PROF	48,280	177	3	ASST	55,500
178	15	ASSO	39,265	179	4	ASST	29,500	180	21	ASSO	48,359
181	23	PROF	48,844	182	1	ASST	31,000	183	6	ASST	32,923
184	2	INST	27,700	185	16	PROF	40,748	186	24	ASSO	44,715
187	9	ASSO	37,389	188	28	PROF	51,064	189	19	INST	34,265
190	22	PROF	49,756	191	19	ASST	36,958	192	16	ASST	34,550
193	22	PROF	50,576	194	5	ASST	32,210	195	2	ASST	28,500
196	12	ASSO	41,178	197	22	PROF	53,836	198	19	ASSO	43,519
199	4	ASST	32,000	200	18	ASSO	40,089	201	23	PROF	52,403
202	21	PROF	59,234	203	22	PROF	51,898	204	26	ASSO	[illegible],047

Outstanding Teams

Institution and Advisor	Team Members

Helix Intersections Papers

"A Specialized Root-Finding Method for Rapidly Determining the Intersections of a Plane and a Helix"	
Harvey Mudd College	Matthew Evans
Claremont, CA	Andrew Flint
David L. Bosley	Noah Kubow
"The Single Helix"	
Iowa State University	R. Robert Hentzel
Ames, IA	Scott Williams
Stephen J. Willson	
"Planes and Helices"	
Macalester College	Samar Lotia
St. Paul, MN	Eric Musser
A. Wayne Roberts	Simeon Simeonov

Faculty Salaries Papers

"Paying Professors What They're Worth"	
Harvey Mudd College	Jay Rosenberger
Claremont, CA	Andrew M. Ross
David L. Bosley	Dan Snyder
"The World's Most Complicated Payroll"	
North Carolina School of Science & Mathematics	Frank Thorne
Durham, NC	W. Garrett Mitchener
Dot Doyle	Marci Gambrell
"Long-Term and Transient Pay Scale for College Faculty"	
Southeast Missouri State University	Christena Byerley
Cape Girardeau, MO	Christina Phillips
Robert W. Sheets	Cliff Sodergren
"How to Keep Your Job as Provost"	
University of Alaska Fairbanks	Liam Forbes
Fairbanks, AK	Marcus Martin
John P. Lambert	Michael Schmahl

Meritorious Teams

Helix Intersections Papers (13 teams)
Baylor University, Waco, TX (Frank H. Mathis)
Beijing Institute of Tech., Beijing, China (Yan-ping Zhao)
California Polytechnic State Univ., San Luis Obispo, CA (Thomas O'Neil) (two teams)
Duke University, Durham, NC (David P. Kraines)
Harvard University, Cambridge, MA (Harry R. Lewis)
Lewis & Clark College, Portland, OR (Robert W. Owens)
Natl. Univ. of Defense Tech., Changsha, Hunan, China (Wang XiaoXing)
New Mexico Inst. of Mining and Tech., Socorro, NM (Brian T. Borchers)
South China University of Tech., Guangzhou, Canton, China (Lejun Xie)
Southeast University, Nanjing, Jiangsu, China (Huangjun Sunzhizhong)
Trinity College Dublin, Dublin, Ireland (Timothy G. Murphy)
University College Galway, Galway, Ireland (M. Tuite)
University of Alaska Anchorage, Anchorage, AK (Ted L. Gifford)
University of Utah, Salt Lake City, UT (Don H. Tucker)
Worcester Polytechnic Institute, Worcester, MA (Arthur C. Heinricher)
Xidian University, Xian, Shaanxi, China (Wang Yu Ping)
Xidian University, Xian, Shaanxi, China (Ma Yu Xiang)

Faculty Salaries Papers (26 teams)
College of William & Mary, Williamsburg, VA (Hugo J. Woerdeman)
Fudan University, Shanghai, China (Cao Yuan)
Fudan University, Shanghai, China (Tan Yongji)
Harbin Institute of Technology, Harbin, China (Shi Peilin)
Hiram College, Hiram, OH (James R. Case)
JiLin University, Changchun, Jilin, China (Lu Xian Yui)
Kenyon College, Gambier, OH (Dana N. MacKenzie)
Luther College, Decorah, IA (Reginald D. Laursen)
Mt. St. Mary's College, Emmitsburg, MD (Fred J. Portier)
Muhlenberg College, Allentown, PA (David A. Nelson)
Natl. Univ. of Defense Tech., Changsha, Hunan, China (Wu MengDa)
Shanghai Jiatong University, Shanghai, China (Longwan Xiang)
Southwestern University, Georgetown, TX (Therese Shelton)
Texas A & M Univ., College Station, TX (Denise E. Kirschner)
Trinity College Dublin, Dublin, Ireland (James C. Sexton)
U.S. Air Force Academy, USAF Academy, CO (Jeffrey S. Stonebraker)
Univ. of Alaska Fairbanks, Fairbanks, AK (Patricia A. Andresen)
Univ. of Colorado at Denver, Denver, CO (David C. Fisher)
Univ. of Missouri–Rolla, Rolla, MO (Roger H. Hering)
University of Dallas, Irving, TX (Charles A. Coppin)
University of Dayton, Dayton, OH (Ralph C. Steinlage)
Vilnius University, Vilnius, Lithuania (Ricardas Kudzma)
Wake Forest University, Winston-Salem, NC (Stephen B. Robinson)
Washington University, St. Louis, MO (Hiro Mukai)
Wheaton College, Wheaton, IL (Paul Isihara)
Xidian University, Xian, Shaanxi, China (Mao Yong Cai)

Awards and Contributions

Each participating MCM advisor and team member received a certificate signed by the Contest Director and the appropriate Head Judge.

The Institute for Operations Research and the Management Sciences (INFORMS) awarded to each member of two Outstanding teams a cash award and a three-year membership. The teams were from Macalester College (Helix Intersections Problem) and Harvey Mudd College (Faculty Salaries Problem). The teams made presentations at a special MCM session and were given cash awards. Moreover, INFORMS gave free one-year memberships to all members of Meritorious and Honorable Mention teams.

The Society for Industrial and Applied Mathematics (SIAM) designated one Outstanding team from each problem as a SIAM Winner. Each team member received a cash prize, and each team received a subsidized trip to the July 1995 SIAM Annual Meeting in San Diego, CA. The teams were from Iowa State University (Helix Intersections Problem) and from University of Alaska Fairbanks (Faculty Salaries Problem). These teams made presentations at a special modeling minisymposium.

Judging

Director
Frank R. Giordano, Dept. of Mathematical Sciences, U.S. Military Academy, West Point, NY

Associate Directors
Chris Arney, Dept. of Mathematical Sciences, U.S. Military Academy, West Point, NY
Robert L. Borrelli, Mathematics Dept., Harvey Mudd College, Claremont, CA

Helix Intersections Problem
Head Judge
Marvin S. Keener, Mathematics Dept., Oklahoma State University, Stillwater, OK
Associate Judges
Ben A. Fusaro (Triage), Dept. of Mathematical Sciences, Salisbury State University, Salisbury, MD
Patrick Driscoll, Virginia Polytechnic Institute and State University, Blacksburg, VA
Mario Juncosa, RAND Corporation, Santa Monica, CA
Veena Mendiratta, AT&T Bell Labs, Naperville, IL
Keith Miller, National Security Agency, Fort Meade, MD
Mike Moody, Harvey Mudd College, Claremont, CA

Lee Seitelman, Glastonbury, CT
Matthew Witten, University of Texas, Austin, TX
Daniel Zwillinger, Zwillinger & Associates, Arlington, MA

Faculty Salaries Problem
Head Judge
Maynard Thompson, Mathematics Dept., University of Indiana, Bloomington, IN
Associate Judges
Robert M. Tardiff (Triage), Dept. of Mathematical Sciences, Salisbury State University, Salisbury, MD
Karen Bolinger, Mathematics Dept., Arkansas State University, State University, AR
James Case, Baltimore, Maryland
William Fox, Dept. of Mathematical Sciences, U.S. Military Academy, West Point, NY
Jerry Griggs, University of South Carolina, Columbia, SC
Don Miller, Dept. of Mathematics, St. Mary's College, Notre Dame, IN
Peter Olsen, National Security Agency, Fort George G. Meade, MD
Judith Pastor, Haverly Systems, Inc., Houston, TX
Catherine Roberts, Mathematics Dept., University of Rhode Island, Kingston, RI
Theresa Sandifer, Mathematics Dept., Southern Connecticut State Univ., New Haven, CT
Michael Tortorella, Middletown, NJ

Triage Session
Director
Ben A. Fusaro
Head Judge, Helix Intersections Problem
Ben A. Fusaro
Head Judge, Faculty Salaries Problem
Robert M. Tardiff
Associate Judges
Homer W. Austin
Alfred S. Beebe, University of Maryland, Eastern Shore, Princess Anne, MD
E. Boyd, University of Maryland, Eastern Shore, Princess Anne, MD
Donald C. Cathcart
S.M. Hetzler
T.O. Horseman
Peter Olsen, National Security Agency, Fort George G. Meade, MD
Fatollah Salimian
Kathleen M. Shannon
Barbara A. Wainwright
M.E. Williams
W.J. Yurek, Worcester-Wicomico Community College, Salisbury, MD

Except as noted, the triage judges were from Salisbury State University, Salisbury, MD.

Sources of the Problems

The Helix Intersections Problem was contributed by Yves Nievergelt (Eastern Washington University, Cheney, WA), who describes its origin in his Author's Commentary in this issue. The Faculty Salaries Problem was contributed by Kathleen M. Shannon (Salisbury State University, Salisbury, MD); the data are public information from Salisbury State University.

Acknowledgments

MCM was funded this year by the National Security Agency, whose support we deeply appreciate. We thank Dr. Gene Berg of NSA for his coordinating efforts. The MCM is also indebted to INFORMS and SIAM, which provided judges, prizes, and forums for presentations of student papers.

I thank the MCM judges and MCM Board members for their valuable and unflagging efforts. Harvey Mudd College, its Mathematics Dept. staff, and Prof. Borrelli were gracious hosts to the judges.

Cautions

To the reader of research journals:

Usually, a published paper has been presented to an audience, shown to colleagues, rewritten, checked by referees, revised, and edited by a journal editor. Each of the student papers here is the result of undergraduates working on a problem over a weekend; allowing substantial revision by the authors could give a false impression of accomplishment. So these papers are essentially au naturel. Light editing has taken place: minor errors have been corrected, wording has been altered for clarity or economy, and style has been adjusted to that of *The UMAP Journal*. Please peruse these student efforts in that context.

To the potential MCM Advisor:

It might be overpowering to encounter such output from a weekend of work by a small team of undergraduates, but these solution papers are highly atypical. A team that prepares and participates will have an enriching learning experience, independent of what any other team does.

Appendix: Successful Participants

KEY:

P = Successful Participation
H = Honorable Mention
M = Meritorious
O = Outstanding (published in this special issue)
A = Helix Intersections Problem
B = Faculty Salaries Problem

INSTITUTION	CITY	ADVISOR	A	B
ALABAMA				
Univ. of Alabama	Huntsville	Claudio H. Morales	P	
ALASKA				
University of Alaska	Anchorage	Ted L. Gifford	M	
	Fairbanks	John P. Lambert	H	O
		Patricia A. Andresen		M
ARIZONA				
Northern Arizona U.	Flagstaff	Terence R. Blows	P	P
ARKANSAS				
Hendrix College	Conway	Ze'ev Barel		P
Williams Baptist Coll.	Walnut Ridge	Lana S. Rhoads		P
		Joy Holloway		P
CALIFORNIA				
Calif. Inst. of Tech.	Pasadena	Alexander S. Kechris	H	
Calif. Poly. State Univ.	S. Luis Obispo	Thomas O'Neil	M,M	
		Ernest Blattner		P
Calif. State Poly. Univ.	Pomona	James R. McKinney		P
Calif. State University	Northridge	Gholam Ali Zakeri		P
Harvey Mudd College	Claremont	David L. Bosley	O	O
Humboldt State Univ.	Arcata	Jeffrey B. Haag	P	
		Kathleen M. Crowe	H	
Loyola Marymount U.	Los Angeles	Thomas M. Zachariah	P	
Pomona College	Claremont	Amy Radunskaya	H	
Sonoma State Univ.	Rohnert Park	Clement E. Falbo	H	
Univ. of California	Berkeley	Allen M. Chen	P	H
Univ. of Redlands	Redlands	Alexander E. Koonce		P
COLORADO				
Metro. State College	Denver	Thomas E. Kelley	P	
Regis University	Denver	Diane M. Wagner		P

INSTITUTION	CITY	ADVISOR	A	B
U.S. Air Force Acad.	USAF Acad.	Jeffrey S. Stonebraker	H,H	M
University of Colorado	Denver	David C. Fisher		M,P
U. of Northern Colorado	Greeley	William W. Bosch		P
U. of Southern Colorado	Pueblo	Paul R. Chacon		P
CONNECTICUT				
Southern Conn. St. Univ.	New Haven	Edward F. Aboufadel		P
University of Bridgeport	Bridgeport	Ninygi Wang		H
		Natalia B. Romalis	P	
University of Hartford	W. Hartford	Diego M. Benardete	P	
Western Conn. St. Univ.	Danbury	Edward Sandifer		H
		Judith A. Grandahl	H	
DISTRICT OF COLUMBIA				
Georgetown University	Washington	Andrew Vogt	P	P
FLORIDA				
Florida Inst. of Tech.	Melbourne	Laurene V. Fausett	P	
Jacksonville University	Jacksonville	Robert A. Hollister	P	
Stetson University	Deland	Lisa O. Coulter	P	
University of S. Florida	Fort Myers	Charles E. Lindsey		P,P
GEORGIA				
Wesleyan College	Macon	Joseph A. Iskra		P
IDAHO				
Lewis-Clark State Coll.	Lewiston	Brent Bradberry		P
ILLINOIS				
Illinois College	Jacksonville	Darrell E. Allgaier	H	
Illinois Wesleyan Univ.	Bloomington	Lawrence N. Stout	P	
Wheaton College	Wheaton	Paul Isihara		M
INDIANA				
Rose-Hulman Inst. of Tech.	Terre Haute	Aaron D. Klebanoff	P	
Saint Mary's College	Notre Dame	Peter D. Smith	P	P
Valparaiso University	Valparaiso	Rick Gillman	P	
IOWA				
Clarke College	Dubuque	Carol A. Spiegel		P,P
Grinnell College	Grinnell	Anita E. Solow		P,P
Iowa State University	Ames	Stephen J. Willson	O	
Luther College	Decorah	Reginald D. Laursen		M
Maharishi Int'l Univ.	Fairfield	Cathy Gorini	P	P
Teikyo Marycrest Univ.	Davenport	Susan T. Youngberg		P
Univ. of Northern Iowa	Cedar Falls	Timothy L. Hardy	P	
		Gregory M. Dotseth		H

INSTITUTION	CITY	ADVISOR	A	B
KANSAS				
Benedictine College	Atchison	Jo Ann Fellin, O.S.B.	P	
KENTUCKY				
Asbury College	Wilmore	Kenneth P. Rietz	H	H
Bellarmine College	Louisville	John A. Oppelt	H	
Western Kentucky U.	Bowling Green	Douglas D. Mooney		P
LOUISIANA				
McNeese State Univ.	Lake Charles	Sid L. Bradley	P	
		George F. Mead	P	
MAINE				
Bowdoin College	Brunswick	Adam B. Levy		P
Colby College	Waterville	Amy H. Boyd	P	
University of Maine	Orono	Grattan P. Murphy	P	
MARYLAND				
Hood College	Frederick	John Boon		P
Loyola College	Baltimore	Dipa Choudhury		H,H
Mt. St. Mary's Coll.	Emmitsburg	Fred J. Portier		M
		Theresa A. Francis	P	
Salisbury State U.	Salisbury	Kathleen M. Shannon		P
		Steve M. Hetzler	P	
Univ. of Maryland	College Park	Michael C. Fu		P
MASSACHUSETTS				
Harvard University	Cambridge	Harry R. Lewis	M,P	
		Roger W. Brockett	H	
Smith College	Northampton	Ruth Haas		P
U. of Massachusetts	Amherst	Edward A. Connors	H	H
Worcester Poly. Inst.	Worcester	Arthur C. Heinricher	M	
		Bogdan Vernescu	H	
MICHIGAN				
Calvin College	Grand Rapids	Steven P. Dirkse	P	
Eastern Michigan U.	Ypsilanti	Christopher E. Hee		H
Lawrence Tech. Univ.	Southfield	Ruth G. Favro	P	
		Howard Whitston	P	
Southwest. Mich. C.	Dowagiac	Ronald Sawatzky		P
MINNESOTA				
Macalester College	St. Paul	Wayne A. Roberts	O	H
Moorhead State Univ.	Moorhead	Ronald M. Jeppson		P
MISSISSIPPI				
Jackson State Univ.	Jackson	David C. Bramlett	H	
		Carl Drake		P

INSTITUTION	CITY	ADVISOR	A	B
MISSOURI				
Crowder College	Neosho	Cheryl Ingram		P
Missouri Southern St. Coll.	Joplin	Patrick Cassens		P,P
Northeast Missouri St. U.	Kirksville	Steven J. Smith		H
Northwest Missouri St. U.	Maryville	Russell N. Euler		P
Southeast Missouri St. U.	Cape Girardeau	Robert W. Sheets		O
University of Missouri	Rolla	Roger H. Hering		M
Washington University	St. Louis	Hiro Mukai	H	M
Wentworth Mil. Academy	Lexington	J.O. Maxwell		P
NEBRASKA				
Hastings College	Hastings	David B. Cooke		H
Nebraska Weslyan Univ.	Lincoln	P. Gavin LaRose	P	
NEVADA				
Sierra Nevada College	Incline Village	Sue Welsch		P
NEW JERSEY				
New Jersey Inst. of Tech.	Newark	Bruce G. Bukiet		P
NEW MEXICO				
New Mexico State Univ.	Las Cruces	Marcus S. Cohen	H	
New Mexico Tech.	Socorro	Brian T. Borchers	M	
NEW YORK				
Hofstra University	Hempstead	R.N. Greenwell	P	P
Ithaca College	Ithaca	John C. Maceli		H
		James E. Conklin		H
Le Moyne College	Syracuse	William C. Rinaman		P
Nazareth C. of Rochester	Rochester	Ronald W. Jorgensen	H	
Pace University	Pleasantville	Robert A. Cicenia		P
Rensselaer Poly. Institute	Troy	Mark Levi	P	
Siena College	Loudonville	T.H. Rousseau	P	
Westchester Comm. Coll.	Valhalla	Rowan Lindley	P	
NORTH CAROLINA				
Appalachian State Univ.	Boone	Jaimie Hebert	P	
		Holly Hirst	H	
Duke University	Durham	David P. Kraines	M	
		Richard A. Scoville	H	
N.C. Schl. of Sci. & Math	Durham	Dot Doyle	H	O
Univ. of North Carolina	Wilmington	Russell L. Herman	H	
	Chapel Hill	Jon W. Tolle		H
Wake Forest University	Winston-Salem	Stephen B. Robinson		M
Western Carolina Univ.	Cullowhee	Jeff A. Graham		P
		Joseph S. Sportsman	H	

INSTITUTION	CITY	ADVISOR	A	B
NORTH DAKOTA				
Univ. of North Dakota	Williston	Wanda M. Meyer		P
	Grand Forks	David J. Uherka		P
OHIO				
College of Wooster	Wooster	Matthew Brahm		P
Hiram College	Hiram	Michael A. Grajek		P
		James R. Case		M
Kenyon College	Gambier	Dana N. MacKenzie	P	M
University of Dayton	Dayton	Ralph C. Steinlage		M
Xavier University	Cincinnati	Richard J. Pulskamp	H	
OKLAHOMA				
Oklahoma State Univ.	Stillwater	John E. Wolfe	H	
Southeastern Okla. St. U.	Durant	Brett M. Elliott		P
OREGON				
Lewis & Clark College	Portland	Robert W. Owens	M	
Southern Oregon St. C.	Ashland	Kemble R. Yates	P	
PENNSYLVANIA				
Bloomsburg University	Bloomsburg	Scott E. Inch	P	
Gannon University	Erie	Rafal F. Ablamowicz	P,P	
Gettysburg College	Gettysburg	James P. Fink		P
Messiah College	Grantham	D.C. Phillippy	H	
Muhlenberg College	Allentown	David A. Nelson		M,P
Univ. of Pittsburgh	Johnstown	Stephen J. Curran		P
Westminster College	N. Wilmington	Carolyn Cuff		P
RHODE ISLAND				
Rhode Island College	Providence	D.L. Abrahamson		P
SOUTH CAROLINA				
Central Carolina Tech. C.	Sumter	Karen G. McLaurin		P,P
The Citadel	Charleston	Kanat Durgun	P	
Coastal Carolina Univ.	Conway	Prashant S. Sansgiry	P	
Columbia College	Columbia	Scott A. Smith		P
Francis Marion Univ.	Florence	Catherine A. Abbott		P
Midlands Tech. College	Columbia	John R. Long	P	
		Rick Bailey		P
SOUTH DAKOTA				
Northern State Univ.	Aberdeen	A. S. Elkhader		H
TENNESSEE				
David Lipscomb Univ.	Nashville	Mark A. Miller	P	P

INSTITUTION	CITY	ADVISOR	A	B
TEXAS				
Baylor University	Waco	Frank H. Mathis	M	H
Southwestern University	Georgetown	Therese Shelton		M
Texas A & M University	Coll. Station	Denise E. Kirschner		M
University of Dallas	Irving	Charles A. Coppin		M
U. Texas–Pan American	Edinburg	Roger A. Knobel	P	P
UTAH				
University of Utah	Salt Lake City	Don H. Tucker	M,P	
Utah State University	Logan	Michael C. Minnotte		P
Weber State University	Ogden	Afshin Ghoreishi	P	
VERMONT				
Johnson State College	Johnson	Glenn D. Sproul		P,P
VIRGINIA				
Coll. of William & Mary	Williamsburg	Lawrence M. Leemis	P	
		Hugo J. Woerdeman		M
James Madison Univ.	Harrisonburg	James S. Sochacki	P	
Roanoke College	Salem	Jeffrey L. Spielman		P
Thomas Jefferson H. S. for Sci. & Tech.	Alexandria	Geraldine F. Oliveto	P	
		Peter J. Braxton	H	
University of Richmond	Richmond	Kathy W. Hoke	H	
Virginia Pol. Inst. & St. U.	Blacksburg	Joel A. Nachlas		H
		Sheldon H. Jacobson	P	
WASHINGTON				
Eastern Wash. Univ.	Cheney	David Jabon	P	
Univ. of Puget Sound	Tacoma	Robert A. Beezer		P
		Martin Jackson		H
Western Wash. Univ.	Bellingham	Tjalling J. Ypma	H	P
WISCONSIN				
Alverno College	Milwaukee	Susan F. Pustejovsky		P,P
Beloit College	Beloit	Philip D. Straffin	P	P
Northcentral Tech. Coll.	Wausau	Frank J. Fernandes	P	
		Robert J. Henning	P	
		Carmen Olson		P
Ripon College	Ripon	Robert J. Fraga		H,P
St. Norbert College	De Pere	John A. Frohliger		P
University of Wisconsin	Madison	Howard E. Conner	H	
	Oshkosh	Andrew E. Long	P	
		K. L. D. Gunawardena		P

INSTITUTION	CITY	ADVISOR	A	B
University of Wisconsin	Platteville	Sherrie Nicol		H
		Clement T. Jeske		H
	Stevens Point	Norman Curet		H
BULGARIA				
Bulgarian Acad. of Sci.	Sofia	Jordan B. Tabov	P	
		Petar S. Kenderov		H
CANADA				
Scarborough College, University of Toronto	Toronto, Ont.	Paul S. Selick	P,P	
University of Calgary	Calgary, Alb.	David R. Westbrook		P
University of Toronto	Toronto, Ont.	Luis A. Seco	H,H	
CHINA				
Anhui University	Hefei	Teng Yaoqing	P	
		Wang Huimin	P	
Automation Eng'ng Coll. of Beijing Union Univ.	Beijing	Ren Kai-long		P
		Wang Xin-feng	H	
Beijing Institute of Tech.	Beijing	Zhao Yan-ping	M	
		Qin Hongxun	H	
Beijing Normal University	Beijing	Liu Laifu	H	
		Zeng Wenyi		P
		Di Zhengru		P
Beijing U. of Post & Tel.	Beijing	Ding Jinkou	P	
		Luo Shou Shan	P	
Beijing U. of Sci. & Tech.	Beijing	Wang Bingtuan		P
		Chen Mingwen		P
		Yang Xiaoming	P	P
China Pharmaceutical U.	Nanjing	Yang Jing Hua	P	
		Qiu Jia Xue		H
Chongqing University	Chongqing	Chu Gong		H
		Ren Shanqiang		H
		Qu Gong	P	
		Liu Qiongxun		P
Dalian University of Tech.	Dalian	Yu Hong Quan		H
		He Ming Feng	P	
E. China U. of Sci. & Tech.	Shanghai	Xu Sanbao		H,H
		Yuanhong Lu	P,P	
Fudan University	Shanghai	Cao Yuan		M
		Tan Yongji		M
		Ye Yao-hua		H
		He Ye		P

INSTITUTION	CITY	ADVISOR	A	B
Harbin Engineering Univ.	Harbin	Zhang Xiaowei		P
		Shen Jihong		H
		Gao Zhenbin		P
Harbin Institute of Tech.	Harbin	Shi Peilin	P	M,P,P
Hefei University of Tech.	Hefei	Du Xueqiao		H
Huazhong U. Sci. & Tech.	Wuhan	Qi Huan		P
JiLin University	Changchun	Lin Zhenghua	H	
		Lu Xian Yui		M
		Zhao Shuven		P
Jinan University	Guangzhou	Zeng Yijun		P
		Fan Suohai		P
Nanjing U. Aero. & Astro.	Nanjingu	Zhang Yi	H	
		Gu Yudi		H
		Zhou Yiqian		H
Nanjing U. of Sci. & Tech.	Nanjing	Qian Xiong Ping		P
		Zhao Chong Gao	P	
Natl. U. of Defense Tech.	Changsha	Wu MengDa		M
		Wang XiaoXing	M	
		Wu Yi	H	
Northwestern Poly. Univ.	Xian	Su Chaowei	P	
		Xu Wei	P	
		Yang Jian	P	
		Rong Haiwu	P	
Peking University	Beijing	Lei Gongyan	H	
		Wu Wei Min		P,P
Shanghai Jiatong Univ.	Shanghai	Xiang Longwan	H	M
		Chen Zhihe	H	H
Shanghai Teachers' Univ.	Shanghai	Dan Sha		H
South China U. of Tech.	Guangzhou	Xiao Renyue		H
		Xie Lejun	M	
		Zhu Fengfeng		P
Southeast University	Nanjing	Deng Jian-ming		H
		Huangjun Sunzhizhong	M	
Tsinghua University	Beijing	Gao Celi		P
		Song Binheng		P
U. of Sci. & Tech. of China	Hefei	Wang Shuhe	H	
		Xu Junming		H
Xian Jiaotong University	Xian	He Xiaoliang		P
		Dong Tianxin	P	
		Zhou Yicang	P	
		Dai Yonghong	P	

INSTITUTION	CITY	ADVISOR	A	B
Xidian University	Xian	Cai Mao Yong		M
		Wang Yu Ping	M	
		Ma Yu Xiang	M	
Zhongshan University	Guangzhou	Zhang Lei	P	P
		He Yuanjiang		P
HONG KONG				
Hong Kong Baptist University	Kowloon	Wai Chee Shiu		P
		Li Zhi Liao		H
IRELAND				
Trinity College Dublin	Dublin	James C. Sexton		M
		Timothy G. Murphy	M	
University College Galway	Galway	Patrick M. O'Leary	H	
		M. Tuite	M	
LATVIA				
University of Latvia	Riga	Janis P. Vucans	P,P	
LITHUANIA				
Vilnius University	Vilnius	Algirdas Zabulionis	P	
		Ricardas Kudzma		M
ZIMBABWE				
University of Zimbabwe	Harare	James Preen		P

A Specialized Root-Finding Method for Rapidly Determining the Intersections of a Plane and a Helix

Matthew Evans
Andrew Flint
Noah Kubow
Harvey Mudd College
Claremont, CA 91711
{mevans, aflint, nkubow}@hmc.edu

Advisor: David L. Bosley

Introduction

Our problem is to locate all of the intersections between a helix and a plane that are in general position.

The problem statement leaves several potentially significant parameters unspecified. In the most general case, solutions may be entirely intractable. Certainly, such cases would be computationally difficult and therefore inappropriate for real-time simulations. At the onset of our investigation, we made the following assumptions:

- **Software application.** The problem is motivated by a desire to predict intersections using computer software, and therefore any solution should be proposed as if it were the computational engine for a larger package. We further assume that all relevant information about the plane and helix is passed into this engine from the user interface.

- **Nonzero tolerance.** The engine should be expected to locate approximate points of intersection within a certain numerical tolerance. Exact solutions are not required for graphical applications. Particularly when a rapid sequence of solutions is desired, the tolerance should increase, to minimize computation time.

- **Frame-by-frame animation.** We approach the problem of real-time simulation as a finite sequence of discrete static instances of the general problem. For example, a 90° rotation of the plane is simulated by a handful of fixed relative orientations, which the engine solves sequentially. Each solution set is then used to construct a single frame in the sequence, which is animated for the user in real time.

- **Nondegenerate, regular finite helix.** We assume that the helix is circular, of finite height, and with a uniform pitch and nonzero radius at any fixed time.

- **Effectively infinite plane.** We consider only an infinite plane, since the possibility of a helix sneaking around the edge of a plane segment adds considerable difficulty to the problem, and we do not believe that is in the spirit of the problem as stated.

- **Relative coordinate system.** The positions of any intersections are to be generated by the engine in its own coordinate system and are then passed out to the calling function. If necessary, the calling function then rescales and translates these points to a coordinate system appropriate for the user interface, including projecting these points in three-space onto a two-dimensional screen.

Constructing the Model

On-screen, a helix and plane may be oriented in virtually any manner. In constructing a model, however, we are concerned only with the relative orientations and positions of the two objects. In light of this, we are free to choose an arrangement that is easiest to treat mathematically.

Assigning a Coordinate System

We begin by assigning a coordinate system to the model. First, we choose one end of the helix to be the initial point and the other end to be the terminal point. The z-axis is the axis of the helix, and the x-axis contains the initial point. The origin of the coordinate system is thereby fixed at the bottom of the helical axis. The orientation of the helix in this coordinate system is shown in **Figure 1**. For the purposes of discussion, we consider only right-handed helices, but the model can easily be extended to left-handed ones.

To locate the plane in this coordinate system, we take $(0, 0, z_0)$ to be the point of intersection between the plane and the z-axis. This intersection is guaranteed to occur provided that the vector normal to the cutting plane is not perpendicular to the z-axis (a special case, which we address later).

Since we have located the point $(0, 0, z_0)$, the plane is completely specified by the angles θ_0 and ϕ_0. In our representation, θ_0 is the angle formed by the line of intersection of the cutting plane with the plane $z = z_0$, as measured counterclockwise from the x-axis. The angle ϕ_0 is the angle of declination of the cutting plane from the z-axis. One could think of this as creating a plane specified by z_0, θ_0, and ϕ_0 by starting with a vertical

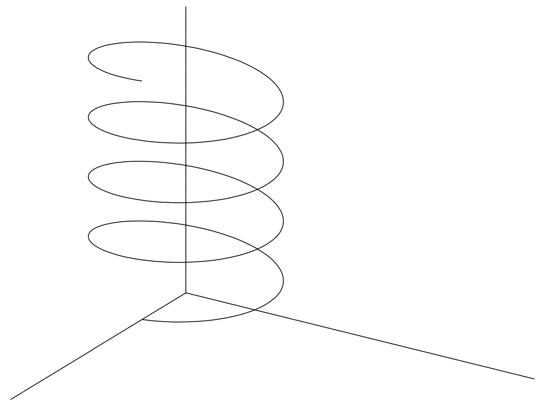

Figure 1. The helix-defined coordinate system.

plane along the xz-axis, then rotating the plane about the z-axis counterclockwise through θ_0 radians, rotating back from the z-axis by ϕ_0 radians, and finally translating directly upward to z_0. The orientation of the plane in the coordinate system is shown in **Figure 2**.

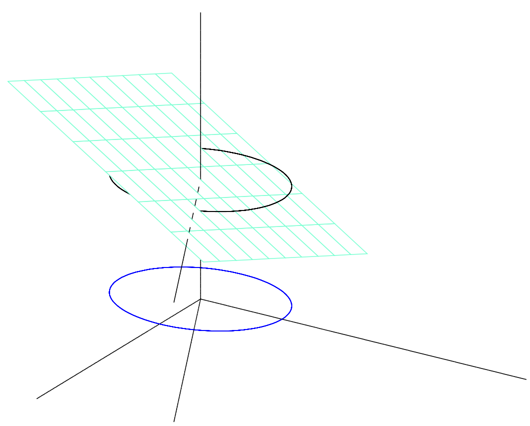

Figure 2. Fixing the plane in the coordinate system.

Parametrizing the Problem

By defining our coordinate system in this way, we can preserve any relative orientation of a helix and a cutting plane while fixing the helix in a vertical position. In these coordinates, we can easily describe these two objects with explicit equations. If the helix has radius R, pitch p, and length L, then it is given in rectangular coordinates by the parametric equations

$$x = R\cos 2\pi t, \qquad y = R\sin 2\pi t, \qquad z = pt, \qquad 0 \le t \le L/p. \tag{1}$$

For $R > 0$, this generates a right-handed helix.

Similarly, the plane can be described directly in rectangular coordinates by

$$z = -\sin\theta_0 \tan\left(\tfrac{\pi}{2} - \phi_0\right) x + \cos\theta_0 \tan\left(\tfrac{\pi}{2} - \phi_0\right) y + z_0. \qquad \textbf{(2)}$$

For a derivation of this result, see **Appendix A**.

Now consider a right circular cylinder of radius R and height L, centered at the origin and resting on the xy-plane. This cylinder is given by

$$x = R\cos 2\pi t, \qquad y = R\sin 2\pi t, \qquad z = s, \qquad 0 \le s \le L. \qquad \textbf{(3)}$$

This cylinder contains the helix, and the intersection of the cylinder and the cutting plane must contain all of the intersections between the helix and the plane. The intersection of the cylinder and the plane is just the projection of the cylinder onto the cutting plane, which is the curve

$$z = R\tan\left(\frac{\pi}{2} - \phi_0\right)\sin(2\pi t - \theta_0) + z_0,$$

as determined by the simultaneous solution of the equations for z in **(2)** and **(3)**. See **Appendix A** for details.

A Visual Interpretation

Further insight into the problem of finding intersections between the helix and the cutting plane can be gained from the following illustration. Suppose that we place a bead on the top of the helix and allow it to slide downward along the helix at a constant rate. If there is a bright light directly above the bead, we will see the bead's shadow repeatedly tracing out an elliptical curve on the cutting plane. We further endow the bead with the magical ability to slide directly through the plane without stopping. Each time the bead comes into contact with its shadow, the bead must be passing through the plane.

In mathematical terms, the motion of the bead is described by **(1)**—if you invert time, or reverse the influence of gravity, anyway—and the motion of its shadow is given by **(3)**, simply allowing t to run on as in **(1)**. Clearly, then, since the x- and y-components of the bead and its shadow are equivalent for all t, we really search only for those instances when they share the same z-component.

Thus, we would like to find all solutions to the equation

$$pt = R\tan\left(\tfrac{\pi}{2} - \phi_0\right)\sin(2\pi t - \theta_0) + z_0, \qquad 0 \le t \le L/p. \qquad \textbf{(4)}$$

Unfortunately, this is a transcendental equation whose solutions cannot be found analytically for $\phi_0 \neq 0$. When $\phi_0 = 0$, this equation is not a valid description of the helical intersections with the cutting plane; we have found analytical solutions to the problem in this case (see **Appendix B**).

Despite the fact that **(4)** has no analytical solutions when $\phi_0 \neq 0$, we have developed a solution algorithm that takes advantage of the unusual characteristics of the equation. A presentation and discussion of our algorithm is given in the following section. Four views of the intersection of a sample helix and plane are shown in **Figure 3**.

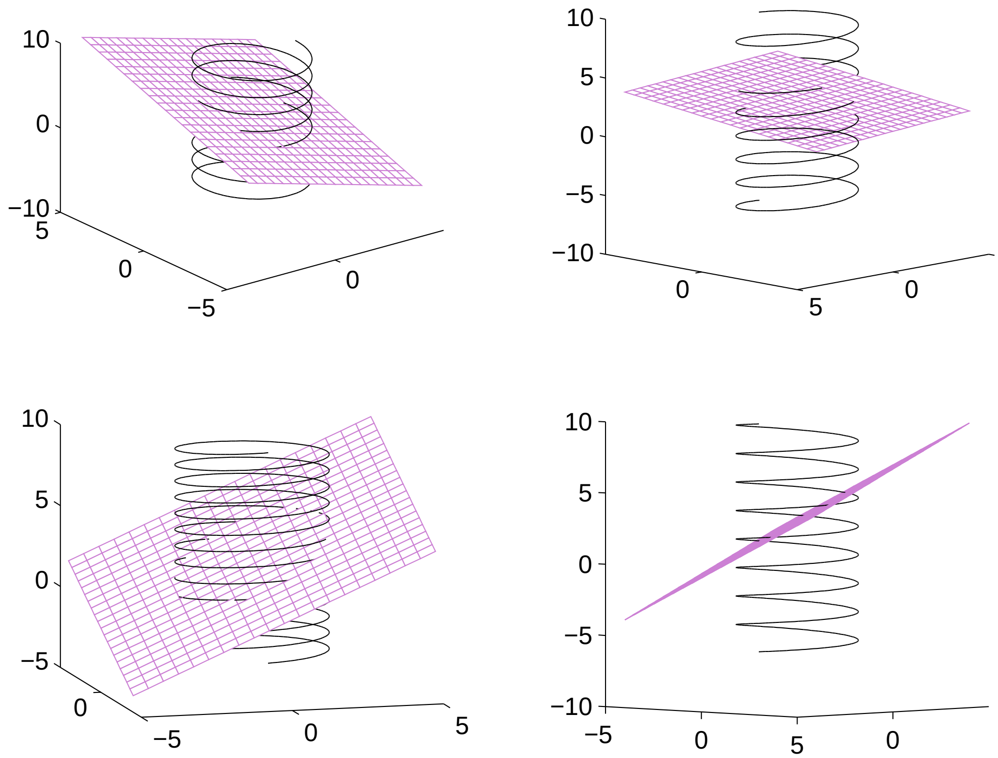

Figure 3. Four views of a typical plane-helix intersection.

Nondimensionalizing the Problem

Eq. **(4)** can be somewhat difficult to deal with. To allow for easier analysis, we nondimensionalize the equation using the definitions

$$\tau = 2\pi t, \qquad \beta = \frac{2\pi_0}{p}, \qquad \sigma = \frac{2\pi R \tan\left(\frac{\pi}{2} - \phi_0\right)}{p}.$$

This nondimensionalization allows us to consider the equation

$$\tau = \sigma \sin(\tau - \theta_0) + \beta.$$

We then define a function

$$f(\tau) = \sigma \sin(\tau - \theta_0) + \beta - \tau \qquad \textbf{(5)}$$

and attempt to solve $f(\tau) = 0$. By finding the roots of f, we find the points of intersection between the helix and the plane. Having found a root $\tau^*$of f, we simply divide this value by 2π and substitute into the parametric equations in **(1)** to locate the points of intersection in Cartesian coordinates.

Analysis of the Model

In proceeding to analyze this model, we first produced a plot of f, as shown in **Figure 4**. The curve represents the vertical separation between the bead and its shadow for any τ.

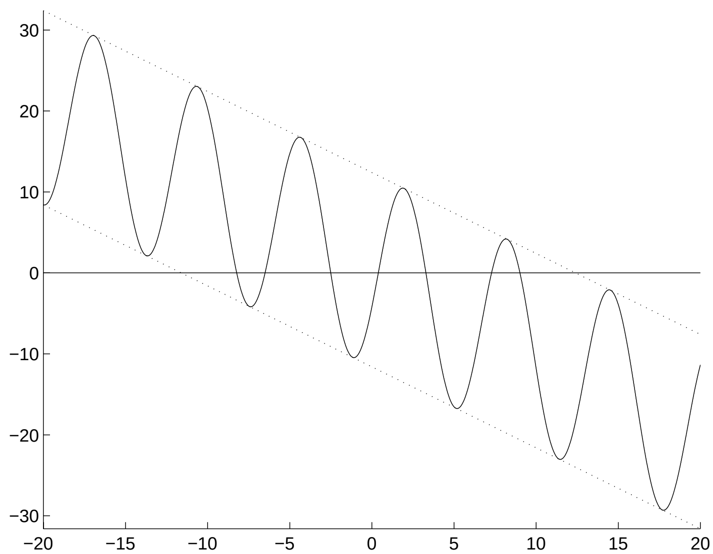

Figure 4. A plot of the function f, of vertical separation between the bead and its shadow, as a function of τ.

First note that the function f is bounded above by $\sigma \cdot 1 + \beta - \tau$ and below by $\sigma \cdot (-1) + \beta - \tau$, the dashed lines in **Figure 4**. These lines cross the τ-axis at $\beta \pm \sigma$, so we can limit our search for roots of f to τ in the interval $(\beta - \sigma, \beta + \sigma)$. Of course, for real-time applications, even limiting the search for intersections to this interval may not let us achieve sufficiently fast solutions. Furthermore, the size of this interval depends on variables controlled by the user and cannot be guaranteed to be small. Moreover, because of the rapid oscillations of the function, most standard root-finding algorithms, such as Newton's method or the bisection method, will not perform adequately on this function.

A Fast Approximation

As a first approximation, we devised a root-finding technique that constructs a linear sketch of f. The local maxima and minima of f can be found analytically by locating the roots of its first derivative. The solutions of

$$\frac{df}{dt} = \sigma \cos(\tau - \theta_0) - 1 = 0$$

are given by

$$\tau = \theta_0 \pm \arccos \frac{1}{\sigma} + 2\pi n, \qquad \text{for } \sigma > 1.$$

(The special case when $|\sigma| \leq 1$ is addressed in **Appendix B**.) We use this family of points to create a jagged linear approximation g of f, by connecting each maximum to its immediately surrounding minimum, and vice versa. **Figure 5** illustrates g. From this we can easily (and quickly) estimate the roots of f by finding the roots of g. It is especially important to note that because g is hinged on the local maxima and minima of f, there will be precisely the same number of roots of g as of f. In other words, *this method of root approximation can never miss a real intersection, or introduce an artificial intersection, of the helix and the plane*.

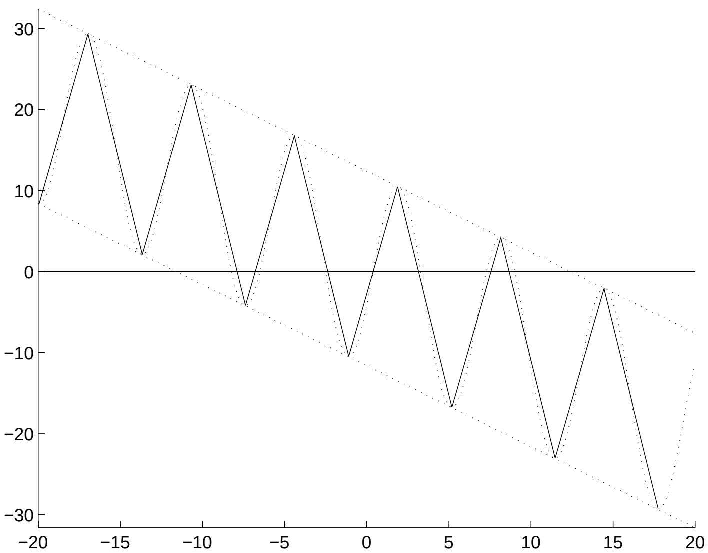

Figure 5. The function g, which approximates f.

From **Figure 4**, we see that the roots of f occur, in increasing τ, either between a maximum and the following minimum, or between a minimum and the following maximum. We refer to the former as a "descending" root and the latter as an "ascending" root. According to the derivation in **Appendix C**, all of the ascending roots are separated by a constant interval in τ. The constant interval is $w\pi(1 + 1/s)$, where s is the slope of the line connecting (in the case of ascending roots) a maximum to the following minimum. Descending roots are separated by a different constant interval in τ, with the same expression for τ but with s being the slope between a maximum and the following minimum.

For example, having located the first ascending root of g at τ_1, we know that the $(n+1)^{\text{st}}$ ascending root occurs at $\tau_1 + 2\pi n(1 + 1/s)$. In this way, we can generate the complete collection of g's roots almost immediately. Only those n roots corresponding to roots on the finite helix need to be considered, thus limiting n. The limits are derived in **Appendix C**.

This calculation of the roots of g constitutes a rough but very fast estimate of the roots of f and in some cases may actually suffice for real-time graphical applications. If the calling program allows a generous enough tolerance in the coordinates of the helical intersections, this initial collection of approximate roots will be adequate, and the problem can be considered solved for the current frame in the animation sequence.

It should further be noted that as more intersections occur, the linear approximation should generate increasingly accurate estimates. In terms of the model, the user may increase the number of intersections by either increasing R (the radius of the helix), decreasing p (the pitch of the helix), or decreasing ϕ_0 (the inclination of the cutting plane with the helical axis). Note that all of these changes cause an increase in the nondimensional parameter σ. Therefore, the vertical distance between adjacent maxima and minima increases, and the linear approximation g becomes more and more accurate as an indicator of the roots of f. Thus, as the apparent complexity of the problem increases, our algorithm experiences only a nominal increase in runtime and achieves more accurate first approximations.

A Rapid Root Search

In many cases, however, the roots of the linear approximation to f may not satisfy the accuracy requirements for the current frame. In this event, the algorithm engages in a more precise method for finding the roots of f.

Taking as seed values the roots of g, we use a modified Newton's method to zero in on the roots of f. The method is modified by taking an approximation of the derivative of f at the seed point. Essentially, we use two approximations to the derivative of f. The constant approximation to an upward slope is given by $\pi/2\sigma$, and $-\pi/2\sigma$ is used to approximate a negative slope. These constants are used for improving the location of descending and ascending roots, respectively. The determination of these approximations is given in **Appendix C**.

We use this alternative to a true Newton's method for two reasons.

- Each computation of the derivative entails computing a cosine function, which is orders of magnitude more time-consuming than a simple variable lookup. By using two constant values, we significantly reduce the computation time.

- Newton's method often has trouble locating the roots of functions with periodic derivatives, such as f.

As in Newton's method, we evaluate the function f at a seed point τ_s to determine the direction and magnitude of the error there. The next approximation is $f(\tau_s) \times \pi/2\sigma$ plus a small perturbation. This process is repeated until the approximation is sufficiently close to the root, yielding an estimated intersection within the desired tolerance. The perturbation has period three, so our linear method is unlikely to fall into an indefinite oscillation about a root. Additionally, the multiplicative factor $\pi/2\sigma$ is well below π/σ, which is the limit on jump size that prevents the algorithm from skipping too far from the approximate root and missing the actual root all together. A more detailed discussion of this technique and the parameters chosen in presented in **Appendix C**.

Testing the Model

We coded our algorithm in C++ and ran several test cases to confirm its root-finding capabilities for this particular problem. Our trials suggest not only that the engine is very rapid in its approximations of the roots of f, but also that it can attain a great level of accuracy with a nominal time penalty. The code itself is presented in **Appendix D**. [EDITOR'S NOTE: Omitted.]

Runtime

We programmed a series of ten frames, with each frame representing a $10°$ rotation of the cutting plane about the z-axis. The angle of declination ϕ_0 was fixed at $45°$, and all parameters describing the helix were held constant. We feel this might represent a typical course of duties demanded by the user. On average, the algorithm calculated about ten points of intersection between the plane and helix in each frame and was able to generate all ten frames in 0.4 seconds. This indicates an average speed of about 25 frames per second.

To put the algorithm to a more demanding test, we then programmed a series of 100 frames. This time, all parameters were permitted to vary randomly (within appropriate bounds) between frames. We found basically the same runtime estimate: 20–25 frames per second.

Accuracy

In each of our trial runs, we specified a tolerance level that corresponds to an allowable error in the distance between an estimated intersection and an actual intersection in three-space. Within the algorithm, this tolerance is transformed into an upper bound on the allowable error in the root approximation. In every case tested, our algorithm was able to satisfy this accuracy test. This is because the algorithm will always succeed in finding an actual root within the precision of the machine or that requested by the calling program: Because we use a period-three perturbation in the root-finding iterations, the algorithm cannot bounce indefinitely around the root.

Furthermore, our root-finding method outperformed the root-finding routines of Mathematica (which uses Newton's method), which gave approximations always farther from the true root than our algorithm's approximations.

A Graphical Test

The graphical results of a simulation are presented in **Figure 6**.

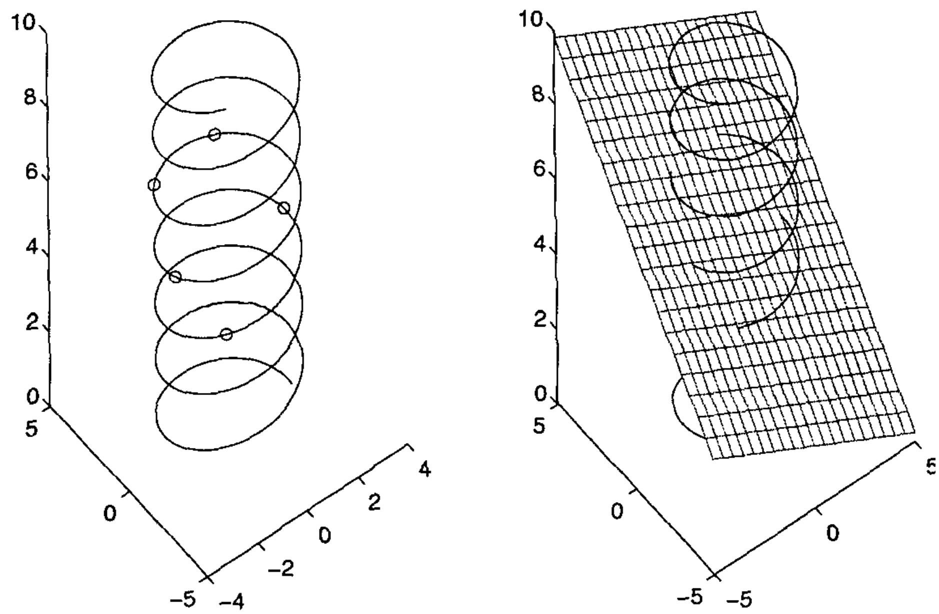

Figure 6. Intersection points as generated by our algorithm.

Critique of the Model

Strengths of the Model and Algorithm

By fixing the coordinate system about the helix, we can easily construct parametric equations to describe the helix in Cartesian coordinates. This representation in turn allows us to construct a relatively simple equation, the roots of which correspond to actual intersections between the helix and the cutting plane. These roots can easily be translated to give the locations of the intersections in Cartesian coordinates.

Our algorithm for finding the roots of **(5)** also has several notable features. First, by using a linearized approximation to f, we are guaranteed never to miss a root. That is, the number of approximate roots found from the linearization g will always equal the true number of roots of f. Furthermore, our root-searching technique can always improve on these estimates of the roots of f to within the desired accuracy. That is, for computational purposes, the algorithm always finds the correct roots.

In addition, since we compute all the parameters of **(5)** for each new frame, any parameter describing the system can vary arbitrarily between consecutive frames. For example, although the radius of the helix is assumed to be uniform along its entire length in one frame, the radius may increase or decrease in the successive frame with no performance penalty. Similarly, both the length and the pitch of the helix and the relative orientation of the plane to the helix can vary between frames without disturbing the root-finding capabilities of the engine.

Our algorithm can determine roots very quickly. In fact, as the number of intersections—and thus the number of roots—increases, the speed with which these roots are determined increases, because the first approximation of the roots is more likely to satisfy the accuracy criterion.

Finally, the algorithm exhibits very little sensitivity to the input parameters. That is, there are no pathological cases for which the algorithm fails. The model represented by **(5)** fails to describe the desired situation only when the cutting plane is parallel to the helical axis, for which case we give analytical results.

Weaknesses of the Model

Any weaknesses of the model are present in the assumptions of the model. The model assumes a helix of only finite length, with a uniform radius and pitch along its entire length. Provided the plane is not parallel to the helical axis, intersections can occur over only a finite stretch of the helix. More important, the model does not succeed in representing a helix that varies in radius or pitch, though such a case may not even qualify as a true helix to some.

Another drawback is that the model considers only a static relationship between the helix and the plane. No attempt has been made to incorporate, for example, a rate of change of declination of the plane from the helical axis. Instead, our model assumes that relative motion between the two objects occurs in consecutive discrete steps.

The algorithm to find the roots of equation **(5)** also has some drawbacks. As with any computational routine, runtime is bound to increase when greater accuracy is desired. For primarily graphical applications, however, extreme accuracy is seldom required. Furthermore, the runtime limitations are dramatic only in the case of σ approaching 1 from above.

In addition, computational techniques inherently introduce numerical errors in solutions; in some cases this may cause misleading results. For example, consider a situation where the cutting plane doesn't cut directly through the helix but instead contains only a single point where the helix is tangent to the plane. This would correspond to a point in **(5)** where $f = 0$ and $df/d\tau = 0$. In this case, our algorithm would find two roots, one corresponding to an ascending root and the other to a descending root. However, the roots would be identical within machine error (not the user-requested error) and thus would yield two identical points of intersection in three-space.

Appendix A: Derivations of Equations

Derivation of the Plane Equation

A plane is determined by a normal vector (a, b, c) and a point (x_0, y_0, z_0) in the plane. The z-coordinate of the plane can be written as a function of x and y as

$$z = \frac{a}{c}(x - x_0) + \frac{b}{c}(y_0 - y) + z_0.$$

The cutting plane is defined so that it contains $(0, 0, z_0)$. Furthermore, a vector normal to the cutting plane is given by the cross product

$$\vec{n} = (\cos\theta, \sin\theta, 0) \times \left(-\cos\left(\tfrac{\pi}{2} - \phi_0\right)\sin\theta, \cos\left(\tfrac{\pi}{2} - \phi_0\right)\cos\theta, \sin\left(\tfrac{\pi}{2} - \phi_0\right)\right).$$

This product reduces to

$$\vec{n} = \left(\sin\theta \sin\left(\tfrac{\pi}{2} - \phi_0\right), -\cos\theta\sin\left(\tfrac{\pi}{2} - \phi_0\right), \cos\left(\tfrac{\pi}{2} - \phi_0\right)\right).$$

Thus, the z-coordinate of the cutting plane as a function of x and y is

$$z = -\sin\theta_0 \tan\left(\tfrac{\pi}{2} - \phi_0\right) x + \cos\theta_0 \tan\left(\tfrac{\pi}{2} - \phi_0\right) y + z_0.$$

The Equation of the Elliptical Intersection

Combining the equations for x and y given in **(1)** and the equation for z in **(2)** yields the intersection of the cutting plane and the cylinder enclosing the helix:

$$\begin{aligned} z_e &= -\sin\theta_0 \tan\left(\tfrac{\pi}{2} - \phi_0\right) R\cos 2\pi t + \cos\theta_0 \tan\left(\tfrac{\pi}{2} - \phi_0\right) R\sin 2\pi t + z_0 \\ &= R\tan\left(\tfrac{\pi}{2} - \phi_0\right)\left(\cos\theta_0 \sin 2\pi t - \sin\theta_0\cos 2\pi t\right) + z_0 \\ &= R\tan\left(\tfrac{\pi}{2} - \phi_0\right)\sin(2\pi t - \theta_0) + z_0. \end{aligned}$$

Appendix B: Special Cases

The Case of a Vertical Cutting Plane

When $\phi_0 = 0$, the cutting plane is parallel to the z-axis, and the parameter z_0 has no interpretation. To describe the plane, let the perpendicular distance from the plane to the z-axis be r_0. The equation of the plane becomes

$$r = \frac{r_0}{\sin(2\pi t - \theta_0)}.$$

Note that if $r_0 > R$, there are no intersections between the plane and the helix. Otherwise, finding the intersections of a vertical cutting plane with the helix is equivalent to finding all t that satisfy either of the following equations

$$R = \begin{cases} \dfrac{r_0}{\sin(2\pi t - \theta_0 - 2\pi n)} \\ \dfrac{r_0}{\sin(\pi - 2\pi t + \theta_0 - 2\pi m)}, \end{cases}$$

where $m, n \in \mathbb{Z}$. Solving for t yields infinite families of solutions

$$t = \begin{cases} \dfrac{\theta_0 + \arcsin \frac{r_0}{R}}{2\pi} + n \\ \dfrac{\theta_0 - \arcsin \frac{r_0}{R}}{2\pi} + m + \dfrac{1}{2}. \end{cases}$$

Having found these values of t, we need only substitute them back into the equations that define the helix to determine all of the intersection points exactly. Because the helix is finite, we expect that this will hold for only a finite selection of ms and ns, the bounds on which are given in **Appendix C**.

The Case $\sigma \leq 1$

For $\sigma \leq 1$, the derivative $df/d\tau = \sigma \cos(\tau - \theta_0) - 1$ is nonpositive for all τ. Thus, f is monotonically nonincreasing and crosses the axis exactly once. Since this root must occur between $\beta + \sigma$ and $\beta - \sigma$, a simple bisection search can be used. Our algorithm starts the bisection search at β and has an initial step size of $\sigma/2$.

Appendix C

Constructing a Linear Approximation of f

The Slope of g

Finding the slopes of the line segments from which g is constructed can be broken down into two parts: finding the slope of the ascending segments and finding the slope of the descending segments.

First we find the slope of the ascending segments. Since all of these segments are parallel, we may choose any one of them as the basis for our problem. Let the endpoints of the segment be denoted by τ_{M} (a maximum) and τ_{m} (the preceding minimum). From the roots of $df/d\tau$ we find

$$\tau_{\mathrm{M}} = \theta_0 - \arccos\frac{1}{\sigma}, \qquad \tau_{\mathrm{m}} = \theta_0 + \arccos\frac{1}{\sigma}.$$

The slope of the line connecting these two points is given by

$$s_d = \frac{f(\tau_{\mathrm{M}}) - f(\tau_{\mathrm{m}})}{\tau_{\mathrm{M}} - \tau_{\mathrm{m}}},$$

which, upon application of the definition of f (and a little algebra), reduces to

$$s_d = \frac{\sqrt{\sigma^2 - 1}}{\arccos\frac{1}{\sigma}} - 1.$$

A similar approach finds the slope of the descending segments, using the same τ_{M} but with τ_{m} being the minimum that follows it rather than precedes it. Again, from the roots of $df/d\tau$, we have

$$\tau_{\mathrm{M}} = \theta_0 - \arccos\frac{1}{\sigma} + 2\pi, \qquad \tau_{\mathrm{m}} = \theta_0 + \arccos\frac{1}{\sigma},$$

which, using the definitions of slope and of the function f, yields

$$s_d = \frac{-\sqrt{\sigma^2 - 1}}{\arccos\frac{-1}{\sigma}} - 1.$$

The Roots of g

Once the slopes for g have been found, finding the equations of the lines from which g is made is a simple matter of using a point that is known to be on one of the lines in combination with the slope. Evaluating f at $\tau = \theta_0$ yields the point $f(\theta_0) = \beta - \theta_0$ on the descending line segment used in finding the slope. Thus, this line has a root at $\tau_d = \theta_0 - (\beta - \theta_0)/s_d$. In the same vein, we have $f(\theta_0 + \pi) = \beta - \theta_0 - \pi$, this time on the ascending line segment used in finding the slope. Thus, this line has a root at $\tau_d = \theta_0 + \pi - (\beta - \theta_0 - \pi)/s_d$. Finally, the equations of the lines are

$$g_{d_0}(\tau) = s_d(\tau - \tau_d), \qquad g_{a_0}(\tau) = s_a(\tau - \tau_a).$$

Since $f(\tau + 2\pi) = f(\tau) - 2\pi$, all of the lines used in the construction of g are given by

$$g_{d_n}(\tau) + 2\pi n = s_d(\tau - \tau_d - 2\pi n), \qquad g_{a_m}(\tau) + 2\pi m = s_d(\tau - \tau_a - 2\pi m),$$

where $m, n \in \mathbb{Z}$, which can be reduced to

$$\begin{aligned} g_{d_n}(\tau) &= s_d \left[\tau - \tau_d - 2\pi n \left(1 + \frac{1}{s_d} \right) \right], \\ g_{a_m}(\tau) &= s_d \left[\tau - \tau_a - 2\pi m \left(1 + \frac{1}{s_a} \right) \right]. \end{aligned}$$

This formulation shows that the roots of the lines from which g is constructed simply repeat in τ every $2\pi(1 + 1/s)$. Thus, the set of all of these roots is given by

$$\begin{aligned} \tau_{d_n} &= \theta_0 - \frac{\beta}{s_d} + 2\pi n \left(1 + \frac{1}{s_d} \right), \\ \tau_{a_m} &= \theta_0 - \frac{\beta}{s_a} + 2\pi m \left(1 + \frac{1}{s_a} \right) + \pi. \end{aligned} \qquad \textbf{(6)}$$

Since f can have roots only between $\beta \pm \sigma$, we need concern ourselves only with the roots of g that lie between these bounds. Further limits can be placed on the range of possible roots by recalling that the helix is of finite length and thus intersections can occur only in $0 \leq \tau \leq 2\pi L/p$. Calling the upper bound τ_h (the lesser of $2\pi L/p$ and $\beta + \sigma$) and the lower bound τ_ℓ (the greater of 0 and $\beta - \sigma$), the limits on the integers n and m are given by

$$\begin{aligned} \tau_\ell \leq \quad & \theta_0 - \tfrac{\beta}{s_d} + 2\pi n \left(1 + \tfrac{1}{s_d} \right) && \leq \tau_k \\ \tau_\ell \leq \quad & \theta_0 - \tfrac{\beta}{s_d} + 2\pi m \left(1 + \tfrac{1}{s_a} \right) + \pi && \leq \tau_k. \end{aligned}$$

Solving these equations for n and m yields

$$\begin{aligned} \frac{\frac{\beta}{s_d} - \theta_0 - \tau_\ell}{2\pi\left(1 + \frac{1}{s_d}\right)} \quad & \leq n \quad \leq \frac{\frac{\beta}{s_d} - \theta_0 + \tau_k}{2\pi\left(1 + \frac{1}{s_d}\right)} \\ \frac{\frac{\beta}{s_a} - \theta_0 - \tau_\ell - \pi}{2\pi\left(1 + \frac{1}{s_a}\right)} \quad & \leq n \quad \leq \frac{\frac{\beta}{s_a} - \theta_0 + \tau_k - \pi}{2\pi\left(1 + \frac{1}{s_a}\right)}. \end{aligned}$$

This range in n and m, when substituted back into **(6)**, produces all of the roots of g, and thus we have a set of approximations to all of the roots of f.

Quick and Safe Approximation of the Slope of f

Since the period of $df/d\tau$ is 2π, a change in τ that is greater than π could move us from the search for one ascending (or descending) root to another. To prevent this, we use $\pi/2$ as the limit for possible changes in our approximation of τ (for any single step). Since f is bounded by $g \pm \sigma$, the maximum value for a given root of g is $|\sigma|$. Thus, a search routine that produces successive approximations by $\tau_n = \pm(\pi/\sigma) f(\tau_{n-1})$ (positive when searching on an ascending segment of g, and negative when on a descending segment) would never move away from its intended target root. However, machine errors could cause this search method to converge to the wrong root, so we use one-half of that quantity.

Furthermore, we add a small (less than 10%) period-three oscillation to the constant to prevent the search algorithm from oscillating about a root with a period other than an integer multiple of three.

Reference

Edwards, C.H., and David E. Penney. 1990. *Calculus and Analytic Geometry.* Englewood Cliffs, NJ: Prentice-Hall.

The Single Helix

R. Robert Hentzel
Scott Williams
Iowa State University
Ames, IA 50011
topquark@iastate.edu
williams@ttd.teradyne.com

Advisor: Stephen J. Willson

Introduction

We present an iterative algorithm that finds all points of intersection between a finite, infinitely thin helix and an infinite plane, ordered along the helix. The algorithm has constant space complexity and has time complexity proportional to the number of intersections and the logarithm of the desired precision.

Assumptions

- The helix is of finite length.
- The helix is infinitely thin.
- The plane is infinite in extent.
- The helix has constant radius.
- No more than one helix needs to be considered simultaneously; if multiple helices must be modeled, the routine may be run sequentially on each.

Input and Output

Input to the program consists of:

- the endpoints of the central axis of the helix,
- one point which is on the helix,
- the winding number of the helix,
- the handedness of the helix,

- the specification of a plane, and
- the desired precision of the solutions.

The *winding number* of a helix is the number of times that a point traveling on the helix makes a complete circle while its projection on the helix axis advances by one unit. In other words, it is the number of coils of the helix contained in one distance unit parallel to its axis. **Figure 1** illustrates a helix with a winding number of 2.0.

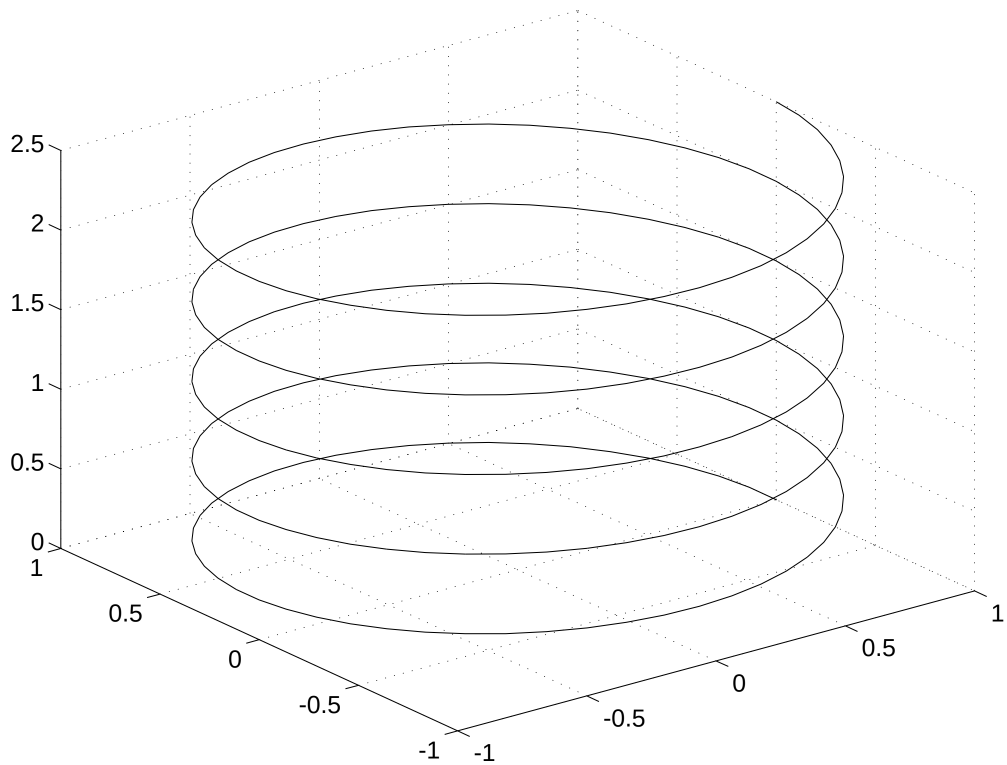

Figure 1. A helix with winding number of 2.0 (note the axes!).

Output consists of an ordered list of intersections, giving for each the distance along the axis of the helix, the coordinates of the intersection point, and the distance from the helix point to the plane (which will be less than the prescribed precision.

Helicoidal Normal Form

We will do all of our computations with geometric figures (helices and planes) which have been put into *helicoidal normal form* (HNF), described below. Doing so takes advantage of the symmetries of helices and of the infiniteness of planes to simplify the resulting calculations and increase the speed of calculating intersection points. The transformation is invertible, so the coordinates of intersection points obtained may be transformed back into the original system.

Definition of Helicoidal Normal Form

In HNF, helices and planes have the following properties:

- The axis extends from $(0, 0, 0)$ to $(0, 0, 2\pi\eta L)$, where L is the length of the original helix and η is the original winding number of the helix.
- The point $(1, 0, 0)$ is on the helix; that is, the coordinate system is oriented so that the lowest point on the helix is on the x-axis and the radius is unity.
- The winding number of the helix is $1/2\pi$.
- The plane's normal vector is normalized to unit length.
- The helix is right-handed.
- The z-component of the plane's normal vector is nonnegative.

Thus, a normalized helix may be parameterized as

$$r(t) = 1, \qquad \theta(t) = t, \qquad z(t) = t$$

in radial coordinates or as

$$x(t) = \cos t, \qquad y(t) = \sin t, \qquad z(t) = t$$

in Cartesian coordinates. The equation of a plane is given by

$$ax + by + cz + d = 0,$$

with the triple (a, b, c) representing a unit normal vector, so that

$$a^2 + b^2 + c^2 = 1 \qquad \text{and} \qquad c > 0,$$

as per the definition of the helicoidal normal form.

Transformation to Helicoidal Normal Form

The helix and plane are specified by the user with the following data:

- the endpoints of the symmetry axis of the helix:
  $$\vec{x}_0 = (x_0, y_0, z_0), \qquad \vec{x}_1 = (x_1, y_1, z_1);$$
- any point on the helix, $\vec{p} = (x_p, y_p, z_p)$;

- the winding number of the helix, $\eta > 0$;
- any (a, b, c, d) quadruplet representing the plane, so that the normal vector $\vec{n}$ is (a, b, c); and
- the handedness of the helix.

The transformation to HNF consists of seven steps:

1. translation of one axis endpoint to the origin,
2. rotation of coordinates to bring the other endpoint to the z-axis,
3. rotation of coordinates to eliminate any initial phase of the helix,
4. space inversion to ensure right-handedness of helix,
5. scaling of coordinates to normalize the helix radius,
6. scaling of coordinates to normalize the helix winding number, and
7. normalizing the normal vector of the plane.

The details of each step follow.

Translation of one axis endpoint to the origin

We translate the coordinate system to bring x_0 (one end of the helix's symmetry axis) to $(0, 0, 0)$:

$$\begin{aligned} \vec{x}_0 &\leftarrow \vec{x}_0 - \vec{x}_0 = \vec{0} \\ \vec{x}_1 &\leftarrow \vec{x}_1 - \vec{x}_0 \\ \vec{p} &\leftarrow \vec{p} - \vec{x}_0. \end{aligned}$$

We can find the effect on the representation of the plane as follows. If $\vec{x}$ is originally on the plane, then

$$\vec{n} \cdot \vec{x} + d = 0,$$

so the translated point $\vec{x} - \vec{x}_0$ satisfies the condition

$$\vec{n} \cdot (\vec{x} - \vec{x}_0) + (d + \vec{n} \cdot \vec{x}_0) = 0.$$

Accordingly, we transform

$$d \leftarrow d + \vec{n} \cdot \vec{x}_0.$$

Rotation of coordinates to bring the other endpoint to the z-axis

Taking

$$\theta = \arctan(x_1/y_1), \qquad \phi = \arctan\left(\frac{\sqrt{x_1^2+y_1^2}}{z_1}\right),$$

we can construct the rotation matrix

$$R = R_{yz}R_{xy} = \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which can be applied to bring x_1 to coincide with $(0,0,L)$, where L is the original length of the helix, given by

$$L = \sqrt{\vec{x}_1 \cdot \vec{x}_1}.$$

That is, we transform

$$\vec{x}_0 \leftarrow R\vec{x}_0 = \vec{0}, \qquad \vec{x}_1 \leftarrow R\vec{x}_1 = (0,0,L), \qquad \vec{p} \leftarrow R\vec{p}.$$

We can find the effect on the representation of the plane as follows. If $\vec{x}$ were originally on the plane, then

$$\vec{n} \cdot \vec{x} + d = 0.$$

Now the rotated point $R\vec{x}$ satisfies the condition

$$R\vec{n} \cdot R\vec{x} + d = 0,$$

so we make the transformation

$$\vec{n} \leftarrow R\vec{n}.$$

Rotation of coordinates to eliminate any initial phase

We now look at the point p on the helix and note its parametric representation as

$$r_p = \sqrt{x_p^2 + y_p^2}, \qquad \theta_p = \arctan(y_p/x_p), \qquad z_p = z_p.$$

Thus, in going downward along the helix's axis from $z = z_p$ to $z = 0$ (the bottom end), we will go through ηz_p rotations, bringing the initial angular position of the helix (θ_0) to

$$\theta_0 = \theta_p - 2\pi\eta z_p.$$

We wish this to coincide with the x-axis, so we apply an additional counterclockwise rotation of $-\theta_0$ about the z-axis to $\vec{x}_1, \vec{x}_2, \vec{p}$ and the representation of the plane, using the same techniques as in the previous section.

Scaling of coordinates to normalize helix radius

We can now parameterize our helix as

$$x(t) = r\cos t, \qquad y(t) = r\sin t, \qquad z(t) = t,$$

with $r = r_p = \sqrt{x_p^2 + y_p^2}$. We note that if any point $\vec{x}(t)$ coincides with the plane, then

$$a \cdot r\cos t + b \cdot r\sin t + ct + d = 0;$$

so the radius-normalized point $(\cos t, \sin t, t)$ satisfies

$$(ar)\cos t + (br)\sin t + ct + d = 0,$$

and we can make the transformation

$$a \leftarrow ar, \qquad b \leftarrow br, \qquad r \leftarrow 1.$$

Space inversion to ensure right-handedness of helix

If the helix is originally left-handed, it can be made right-handed by effecting a spatial inversion of the coordinate system about the xz-plane. This is accomplished by negating the y-coordinate of the plane's normal vector:

$$b \leftarrow -b, \qquad \text{handedness} \ \leftarrow \ \text{right}.$$

Scaling of coordinates to normalize helix winding number

The final parameter to normalize is the helix winding number. We do this by forcing the helix to advance one rotation per 2π advance along its axis. To compensate, we scale our helix.

If the helix originally advanced η turns per unit axis advance and had a length of L, then the same number of turns will be made in a length of $2\pi\eta L$ with a winding number of $1/2\pi$. Thus:

$$L \leftarrow 2\pi\eta L, \qquad \eta \leftarrow \frac{1}{2\pi}.$$

We can find the effect on the representation of the plane as follows. If a point $\vec{x} = (x, y, z)$ is originally on the plane, then

$$ax + by + cz + d = 0.$$

Then, since the transformed point $(x, y, \eta z)$ fulfills the condition

$$ax + by + \frac{c}{2\pi\eta}(2\pi\eta z) + d = 0,$$

we can make the transformation:

$$c \leftarrow \frac{c}{2\pi\eta}.$$

Normalizing the normal vector of the plane

We make the transformations

$$
\begin{aligned}
a &\leftarrow \frac{a}{\sqrt{a^2+b^2+c^2}}, & b &\leftarrow \frac{b}{\sqrt{a^2+b^2+c^2}}, \\
c &\leftarrow \frac{c}{\sqrt{a^2+b^2+c^2}}, & d &\leftarrow \frac{d}{\sqrt{a^2+b^2+c^2}}.
\end{aligned}
$$

If $c < 0$, we negate each component of the normal vector (but not d).

Locating Intersections

Transcendental Lemma

We need solutions of the transcendental equation

$$f(t) = a\cos t + b\sin t + c = 0$$

subject to the restriction that $a^2 + b^2 + c^2 = 1$. We observe that

$$
\begin{aligned}
&\sqrt{a^2+b^2}\sin\left(t + \arctan\frac{b}{a}\right) \\
&\qquad = \sqrt{a^2+b^2}\left[(\sin t)\left(\cos\arctan\frac{b}{a}\right) + (\cos t)\left(\sin\arctan\frac{b}{a}\right)\right] \\
&\qquad = \sqrt{a^2+b^2}\,(\sin t)\left(\frac{a}{\sqrt{a^2+b^2}}\right) + \sqrt{a^2+b^2}\,(\cos t)\left(\frac{b}{\sqrt{a^2+b^2}}\right) \\
&\qquad = a\sin t + b\sin t.
\end{aligned}
$$

So, we attack the original problem as:

$$
\begin{aligned}
\sqrt{a^2+b^2}\sin\left(t + \arctan\frac{b}{a}\right) &= -c, \\
\arcsin\left(\sin\left(t + \arctan\frac{b}{a}\right)\right) &= \arcsin\left(\frac{-c}{\sqrt{a^2+b^2}}\right), \\
t &= \arcsin\left(\frac{-c}{\sqrt{1-c^2}}\right) - \arctan\left(\frac{b}{a}\right).
\end{aligned}
$$

We note:

- If $c > \sqrt{2}/2$, then the initial $\arcsin$ will have an argument larger than unity and will be nonexistent.
- The $\arcsin$ function will yield two distinct values in $[0, 2\pi)$, both of which must be checked.
- If t is a solution of $f(t) = 0$, then $t + n2\pi$ must also be a solution for integers n.

Difference Function

The phrase "the (helix) point t" will be taken to mean the helix point parameterized by t, namely $(\cos t, \sin t, t)$ for $t \in [0, L]$. Since all points $\vec{x}$ on the plane satisfy

$$ax + by + cz + d = 0,$$

we see that finding the points of intersection between the helix and the plane amounts to finding t such that

$$f(t) = a\cos(t) + b\sin(t) + ct + d = 0$$

for $t \in [0, L]$.

We name f our *difference function* since it provides a measure of the distance between the helix and the plane for a given t. In fact, $f(t)$ is the perpendicular distance from the helix point t to the plane. We will eventually seek intersections by attempting to minimize the absolute value of $f(t)$.

Slope of the Difference Function

The slope of the difference function at the helix point t is given by:

$$f'(t) = \left.\frac{df(x)}{dx}\right|_{x \leftarrow t} = -a\sin t + b\cos t + c$$

We note:

- For $c > \sqrt{a^2 + b^2}$ (which is equivalent to $c > \sqrt{2}/2$), the slope is everywhere positive, so the difference function is monotonically increasing.
- The form of the slope function is that studied in the **Transcendental Lemma** above, so we know that we can find zeros of the slope function, and hence local extrema of the difference function, analytically. Here the two values for the $\arcsin$ correspond to a local maximum and a local minimum of the distance function.

Upper and Lower Bounds

It is easy to find bounds on t that contain all of the possible intersection points. At an intersection point, we have

$$a\cos t + b\sin t + ct + d = 0,$$

so

$$t = \frac{-a\cos t - b\sin t - d}{c}.$$

Simple calculus, plus using $c = \sqrt{a^2+b^2}$, $|\cos t| \le 1$, and $|\sin t| \le 1$, shows that

$$-\sqrt{a^2+b^2} \le -a\cos t - b\sin t \le \sqrt{a^2+b^2},$$

which provides bounds on t:

$$\ell_b \equiv \frac{-\sqrt{a^2+b^2}-d}{c} \le t \le \frac{\sqrt{a^2+b^2}-d}{c} \equiv u_b.$$

All intersection points t must lie within these bounds.

If $c = 0$ (the discontinuity in the above formulae), the plane runs exactly parallel to the axis of the helix and would intersect an infinite helix either infinitely often (inside the radius) or never (outside the radius). In this case, ℓ_b may be taken to be $-\infty$ and u_b to be $+\infty$, since both values will be truncated by the finiteness of the helix, as mentioned in the next subsection.

Intervals and Subintervals

We define the *search interval lower bound* ℓ_B to be the larger of 0 (one end of the helix) and ℓ_b. We define the *search interval upper bound* u_B to be the smaller of L (the other end) and u_b. All intersections must lie in the *search interval* $[\ell_B, u_B]$.

We divide the interval $[\ell_B, u_B]$ into subintervals, broken by the local extrema, both maxima and minima. Each subinterval, except the leftmost and rightmost, is bounded by adjacent local extrema. The leftmost and rightmost are bounded on their "inner" side by a local extremum and on their "outer" side by either ℓ_B or u_B. We can do this because we can apply the **Transcendental Lemma** to the slope function to calculate two local extrema, one minimum and one maximum. Since the spacings of minima and maxima are both 2π, we can now locate all local extrema within the search interval by repeated addition/subtraction of 2π from the original maximum and minimum (see **Figure 2**).

It may be the case, if $c > \sqrt{2}/2$, that there are no local extrema. Then there is only one subinterval, $[\ell_B, u_B]$, which contains the only intersection. (There can be only one intersection, since f is monotonically increasing in this case.)

Searching the Subintervals

We now consider each subinterval separately. We let ℓ_t and u_t represent the initial left and right endpoints of the subinterval. We evaluate $f(\ell_t)$ and $f(u_t)$ and compare the signs. If the signs agree, then there *cannot* be an intersection in the subinterval; if there were, there would have to be a local extremum *within* the subinterval, but all local extrema fall on subinterval boundaries by construction. However, if the signs are different, there *must*

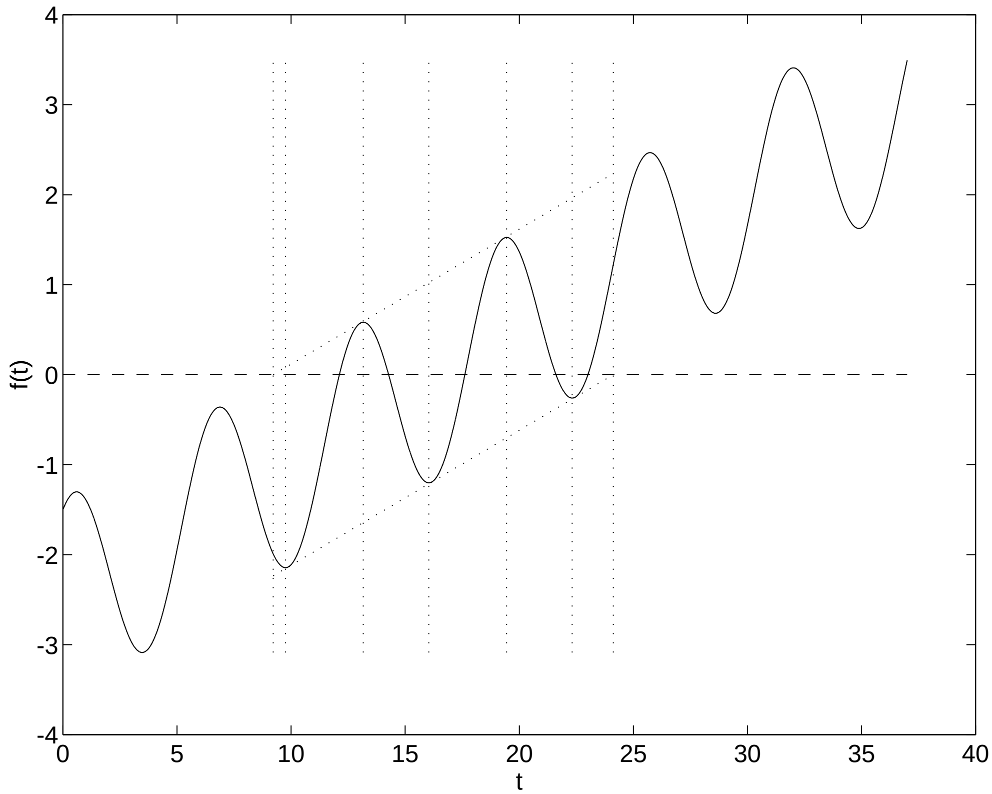

Figure 2. Difference function $f(t)$ with search interval limits shown as the outermost dotted vertical lines, zero shown as a dashed horizontal line, and subinterval boundaries shown as dotted vertical lines. Note the small(er) subintervals at the ends.

be an intersection point, by application of the Intermediate Value Theorem to the continuous function f, which is positive at one endpoint and negative at the other.

If endpoints show that no intersection can exist (i.e., both have the same sign), then the subinterval is immediately discarded (this can occur only with the leftmost and rightmost subintervals). Otherwise, we begin a binary search for the intersection point, provided that the endpoint itself is not a point of intersection within the desired precision.

We evaluate f at the midpoint of the subinterval $(\ell_t + u_t)/2$ and compare its sign with the endpoints. By the Intermediate Value Theorem, the intersection must lie between the midpoint and whichever endpoint differs in sign from it (they cannot *both* differ, since they have different signs). We can therefore define a new subinterval equal to either the left or right half of $[\ell_t, u_t]$ as appropriate and repeat this process. Eventually, evaluation of f at the midpoint must yield a number within the prescribed accuracy from zero. At this point, the binary search terminates, having located the intersection with sufficient precision.

We note that we apply the binary search *only* in subintervals that are guaranteed to contain an intersection by virtue of the Intermediate Value

Theorem.

When intersections are found, the proper inverse transformations are performed to return the point to its original coordinate system, and it is displayed.

After completing a subinterval, processing proceeds with the next and terminates following the rightmost. Because of the left-to-right processing of subintervals, the intersections are generated *in order* of increasing t along the helix axis.

Space and Time Complexity

Space Complexity

In our implementation, each subinterval is calculated and searched subsequently, obviating the need to store large arrays of points found or of edges of subintervals. Thus, the storage space needed is independent of the input. For our implementation, it is approximately 300 bytes.

Time Complexity

The time to put the helix and plane into HNF is independent of the input and, in the current implementation, involves approximately 118 floating-point multiplies/divides and 36 floating-point trigonometric functions and square roots. Floating-point adds, subtracts, and comparisons take negligible time relative to multiplication, division, square roots, and trigonometric function calls.

The time per intersection is 40 floating-point multiplications/divisions and 8 complicated functions, plus 4 multiplications and 3 complicated functions per iteration required. The largest that a subinterval may be is 2π, so to divide this by halving into a slice as small as the desired accuracy ϵ requires $\log_2(2\pi/\epsilon)$ steps. Thus, for each intersection, $40 + 4\log_2(2\pi/\epsilon)$ multiplications and $8 + 3\log_2(2\pi/\epsilon)$ complicated functions are required. For an accuracy of one part in 10^6, this amounts to 131 multiplications and 76 complicated functions per intersection. These are absolute bounds.

There are sufficiently few operations that on a DECstation 5000/240, the processing can be carried out in a tiny fraction of a second for as many as 200 intersection points (we did not test larger numbers).

We believe that this is sufficiently fast to be used in modeling large numbers of helices or as part of a real-time rendering engine.

Algorithm Analysis

Strengths

- The algorithm requires little memory.
- The algorithm is guaranteed to find all intersection points within the desired precision, unless this is finer than the machine's floating-point precision.
- Intersections are generated quickly enough for real-time applications.
- The algorithm requires time linear in the number of intersection points and logarithmic in the desired precision. Hence, an order-of-magnitude increase in accuracy costs only 12 multiplications and 9 complicated functions per intersection.

Weaknesses

- The algorithm does not take previous intersections into account when searching a subinterval; there is evidence to think that doing so could provide a moderate increase in speed, perhaps 15%.
- The number of multiplications necessary per intersection can be cut by 18, by combining the three separate rotation matrices into a single one.
- The algorithm currently weights each endpoint equally when choosing a new subinterval during the binary search. Using a linear interpolation that weighted each endpoint by the value of f evaluated at that endpoint could increase speed significantly, because of the smoothness and near-linearity of f.

Planes and Helices

Samar Lotia
Eric Musser
Simeon Simeonov
Macalester College
St. Paul, MN 55105
{slotia, ssimeonov}@macalstr.edu
emusser@prodea.com

Advisor: A. Wayne Roberts

Introduction

We are asked to design, implement, and test a mathematical algorithm that locates in real time all of the intersections of a helix and a plane located in general positions in space. In addition, we must prove that our algorithm is mathematically and computationally correct.

Assumptions

- **"Real time."** We assume that "real time" means that the time to solve a reasonably "difficult" problem must be very small. For example, an algorithm that is used for the back end of a Computer Aided Geometric Design program must not impose unacceptable delays on the user. When the user repositions the helix or plane, the user expects an immediate refresh of the screen—usually in only a fraction of a second. For other applications, our algorithm should take little time compared to the calling program. To meet these requirements, the algorithm must have a time complexity linear in the number of intersections; our algorithm is indeed linear.
- **Helix/Plane.** We assume the strict mathematical definition of a single helix: that it is infinite and nonelliptical, and that it "wraps around" a cylinder. We believe that our algorithm will work for other "helices," such as those having elliptical bases; but we have not tested these cases to any extent. We also assume that the plane extends infinitely.
- **Correctness.** We assume that twelve digits of precision is acceptable for all calculations. This is sufficient for most biotechnological applications. Twelve digits means that we have error less than the radius of some atoms [Chang 1994]. For some applications, engineers often expect fewer

digits of accuracy; so our program allows the user to change the desired precision.

Summary of Approach

Our approach to the problem involves

- definition of design requirements,
- development of a mathematical model for the problem,
- design and implementation of an algorithm,
- debugging and testing, and
- evaluation.

Definition of Design Requirements

The algorithm will be used as a support routine for mission critical processes, where its failure to produce correct results, or failure to produce results on time, could have dire consequences. This fact leads us to the following design requirements, in decreasing order of importance:

- **Correctness.** The algorithm must produce correct results.
- **Robustness.** The algorithm must handle exceptions well and not terminate abnormally.
- **Performance.** The algorithm must execute in real time.
- **Efficiency.** The algorithm must spare system resources, as long as the above three requirements are not adversely affected.
- **Flexibility.** The algorithm must allow users to formulate the problem in different ways, e.g., it must allow more than one way of defining a plane or a helix. Also, users must be able to fine-tune the algorithm to improve performance.
- **Portability.** The algorithm must be machine-independent, written in a common programming language, and easy to transfer to a different programming language of choice (e.g., to include it in an embedded system).

Because algorithms execute on physical entities (computers) that have some finite working precision, it is possible that a computationally correct algorithm may produce incorrect results due to roundoff and compound errors. In our case, two types of errors are possible: skipping an existing root and reporting a root that does not exist. It is difficult to claim that one type of error should be preferred to another; we choose, if possible, to minimize the second kind of error given the possibility of some error of the first type.

Development of a Mathematical Model

Initial Development

The General Case

Having examined the problem in general Cartesian coordinates, we found that a graphical or vector analysis–related approach would fail us, in the sense that it would be incredibly hard to program. Hence, we attempted to reduce the problem to an algebraic problem, since algebraic techniques are generally much more suited to programming. In effect, we "projected" the helix onto the plane in the following manner.

Consider the general parametric equations of a helix in space (for derivation, see the **Appendix**):

$$\begin{aligned} x &= a_{11}\cos\alpha t + a_{12}\sin\alpha t + a_{13}t + a_{14} \\ y &= a_{21}\cos\alpha t + a_{22}\sin\alpha t + a_{23}t + a_{24} \\ z &= a_{31}\cos\alpha t + a_{32}\sin\alpha t + a_{33}t + a_{34}. \end{aligned}$$

The equation of a plane in space may be written as

$$ax + by + cz - d = 0.$$

Our transformation replaces x, y, and z on the left-hand side in this equation with the corresponding parametric forms of the helix, thus returning an expression in t, which we call $f(t)$:

$$f(t) = A\cos t + B\sin t + Ct - D,$$

where A, B, and C are appropriately transformed coefficients. The α in the $\cos$ and $\sin$ terms has been incorporated into C via a change of parameter from t to t/α.

The task now is to solve the equation $f(t) = 0$. After perusing relevant literature, we concluded that this equation must be solved numerically [Plybon 1992]. Hence, we developed a numerical technique that, given the parameters A, B, C, and D, attempts to locate all the roots of the equation. Many well-documented algorithms guarantee convergence to roots—given certain bounds on the problem—and give an easily computable bound on the error in the result. The numerical technique that we employ is heavily influenced by information that we have about the equation $f(t) = 0$. For instance, we know that the extrema of the function (if any) occur periodically, and we use this fact at several stages of our method, hence ensuring an efficient algorithm. We essentially built a robust, *strong* equation solver (one that uses problem-specific information to maximize the efficiency of the algorithm).

The first step in our approach is to examine the general form of the function, as in **Figure 1**. We note that all roots must be located between two extrema that are on opposite sides of the t-axis (assuming a continuous function). The only other case, which we handle separately, is when a root and an extremum or inflection point occur simultaneously.

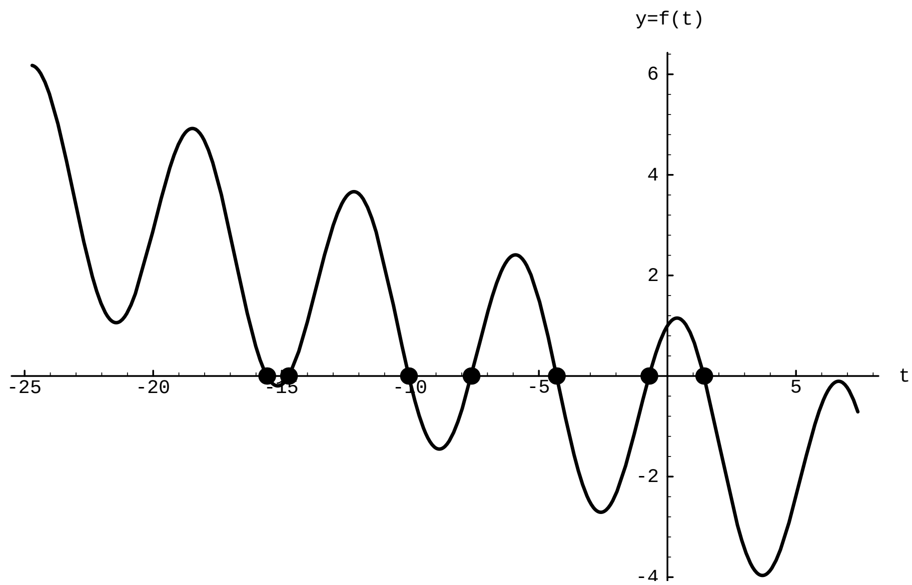

Figure 1. General functional form.

We begin by locating the minima and maxima of the function $f(t)$, which are periodic with period 2π. These are found by solving the equation

$$f'(t) = -A\sin t + B\cos t + C = 0.$$

From

$$\cos t = \frac{A\sin t - C}{B},$$

by using $\cos^2 t + \sin^2 t = 1$ we find

$$\sin t = \frac{ac \pm b\sqrt{a^2 + b^2 - c^2}}{a^2 + b^2}.$$

However, this method returns some extraneous roots because of the squaring (just as does squaring $t = 1$ and solving $t^2 = 1$). These are discarded via a simple test, namely, substituting the values back into $f'(t)$ and checking whether or not the derivative is indeed zero.

We then interpolate the root by connecting the two extrema via a line segment (as in **Figure 2**) and pass the interpolated value to our root-finding algorithm.

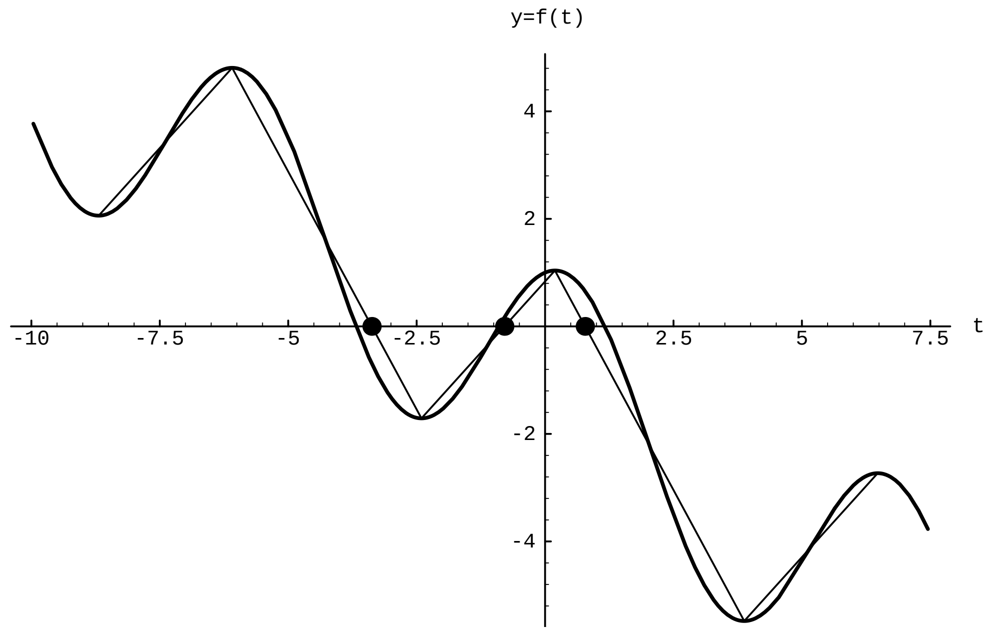

Figure 2. The interpolation method.

We must now judiciously choose a value of t that guarantees roots in its immediate neighborhood. We choose the value $t_0 = D/C$, as we are ensured that—if roots exist—there is one within 2π of this value of t (see the **Appendix** for a detailed argument).

Certain Special Cases

If the coefficient C is 0, then the function $f(t)$ is periodic and oscillates to within $\sqrt{A^2 + B^2}$ (the maximum possible value of $A \cos t + B \sin t$) about the line $g(t) = D$. Thus, if $|D| > \sqrt{A^2 + B^2}$, then the function never intersects the t-axis and we have no roots; the plane is parallel to the helix and outside the "reach" of its radius. In such a case, our program returns the message "No Roots." If, on the other hand, $|D| \leq \sqrt{A^2 + B^2}$, then we have infinitely many roots; this case is handled appropriately.

Another important case is that of only one root (see **Figure 3**). This occurs when $a^2 + b^2 - c^2 \leq 0$. Then the equation has either one real solution or none. If a solution is found, it is at an inflection point. This condition is recognized by our algorithm and appropriately dealt with by calling the bisection method instead of the usual Newton-Raphson (which exhibits very slow convergence when dealing with simultaneous occurrence of roots and extrema/inflection points). The bisection method guarantees convergence as long as we bracket the root properly. We are confident that we do so, as we give bisection an interval of radius 2π about the point $t_0 = D/C$. The bisection method achieves our prescribed goal of 12-digit accuracy in no more than 44 iterations.

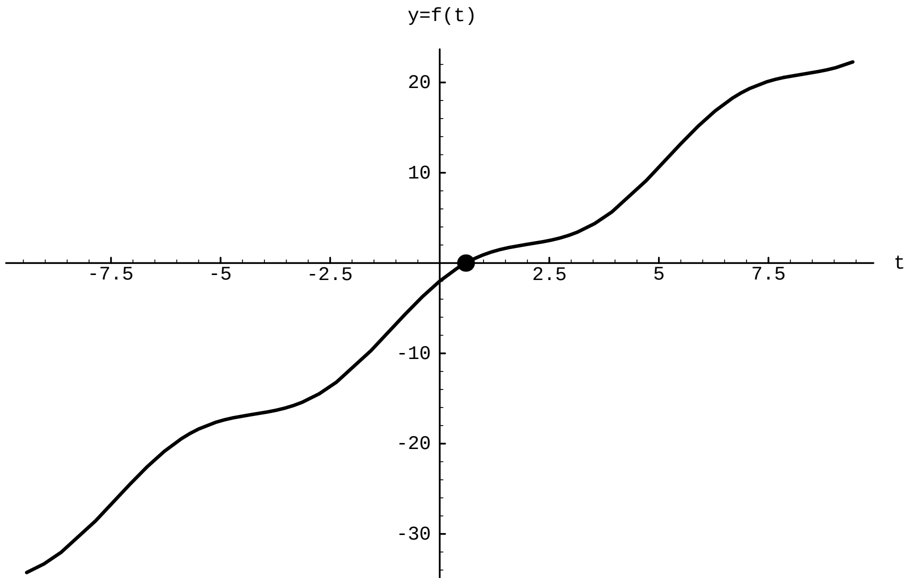

Figure 3. The single-root case.

Algorithm Description

To facilitate understanding, we provide four levels of abstraction in the description of our algorithm. At the top level, we use a linearized flowchart that shows the main subproblems that need to be solved. [EDITOR'S NOTE: Because of space considerations, we do not reproduce the flowchart.] Parallel to the flowchart, at the second level of abstraction, we offer comments that provide more detail of the workings of the algorithm. They refer to the third level of abstraction, mathematical proofs and detailed explanations. Finally, at the lowest level of abstraction is the C++ code of our program, which includes many comments with Mathematica code and references to relevant literature.

Before we go into the details of the algorithm implementation, we mention some general conventions.

- **Input**
  - Planes can be defined by the user in three ways: by general Cartesian equation of the form $ax + by + cz = d$, by two vectors and a point, or by three points.
  - Helices can be defined by the user in two ways: by general parametric equations, or by the three Eulerian angles and the translation vector

that map the z-axis to the central axis of the helix.

- **Output**
  - If the helix does not intersect the plane, no roots are returned.
  - If there are infinitely many solutions (a case in which the plane is parallel to the helix axis), sufficient information is provided so that the user can generate all of the intersection points.
  - Otherwise, a structure containing the x, y, and z coordinates of the points of intersection is produced.
- **Accuracy of Estimation.**
  The default working precision of calculation in our C++ program is twelve digits, but it can be changed by modifying a single variable in the code. The maximum working precision is limited by the floating point precision of the computer.
- **Portability.**
  - Our algorithm is implemented in ANSI C++, ensuring portability across most computing platforms. The code does not use machine-dependent features, and it could be translated to any procedural language.

Testing and Quality Control

We devoted more than half of our algorithm design and implementation efforts to testing. We checked the correctness of ideas and implementations at four different levels:

- **Math Model.** All transformations and function forms that we used were generated symbolically using Mathematica's standard features and the `Vector Analysis` and `Rotations` packages. All symbolic solutions to equations were checked using Mathematica. Whenever possible, we simplified expressions, sometimes by applying trigonometric substitution rules manually.
- **Algorithm Design.** Our root-finding procedure was carefully chosen always to find a bracketed root (if necessary, by invoking bisection).
- **Implementation.** We applied the function `CForm` to convert Mathematica expressions to C code when transferring expressions to our implementation. That minimized the chances of erroneous expression entry. The root-finding procedure that we use was taken from Plybon [1992]. The procedure was independently tested against the built-in Mathematica

routine `FindRoot`, which implements a combination of Newton-Raphson and the secant methods [Wolfram 1991]. Our procedure never failed to find a root and never reported a root where there was none. In several cases where `FindRoot` failed, our procedure correctly managed to find a root.

- **Runtime.** We performed three different types of checks on the output of our program:
  - We generated more than 50 functions of the form $f(t) = A\cos t + B\sin t + Ct - D$ and checked whether our program correctly found their roots. We implemented a set of Mathematica routines that finds the roots of the equation $f(t) = 0$ by the same algorithm as our program but with higher accuracy. In all test cases, the output of our program agreed with Mathematica's output, suggesting that round-off error is not a major problem in our implementation. In most cases, we visually inspected the graph of $f(t)$ to ensure than no roots were missed and that no false roots were introduced. The tests included a mix of test cases including none, one, many, and infinitely many roots. We considered potentially problematic cases, such as roots at tangency points and roots at inflection points. We explored and tested all control paths of the algorithm. Debugging output was generated and investigated carefully.
  - We inputted more than fifty helices and planes in various input formats and used our algorithm to find the coordinates of the intersection points between them. Then for each test run we used Mathematica to do a 3-D plot of the plane, the helix, and the intersection points (see **Figure 4**). We inspected the 3-D plots from various viewpoints to ensure that no intersection points were missed and that no extraneous points were plotted. Our program passed all tests.
  - In the testing, often we were uncertain about the actual location of the intersection points. So we designed an additional battery of tests to check the results of our program against a known (sub)set of the intersection points. We obtained the known intersection points by starting with an arbitrary helix and defining an intersecting plane either by choosing three points on the helix or by choosing a point on the helix and an arbitrary vector. In the first case, we knew the coordinates of at least three intersection points. In the second case, we could position the plane so as to experiment with different patterns of intersection. We then checked the results of our program against subset of known roots. In all of the more than 50 test cases, our algorithm performed correctly.

In *all* test cases, our program performed successfully. This result—added to the fact that our algorithm is closely based on a rigorously proven mathematical model—gives us a high degree of confidence that our program is indeed computationally correct. Also, we have seen no evidence that roundoff or compound errors are a significant source of error.

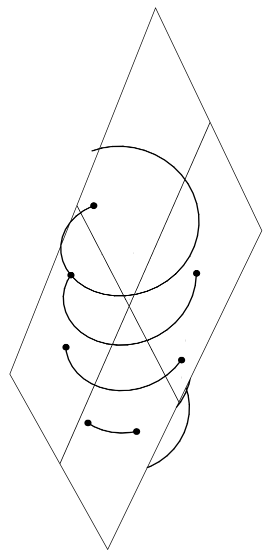

Figure 4. Our solutions plotted on the given helix and plane.

Evaluation

Correctness

Our tests have shown that mathematically and computationally our algorithm is correct. Numerically, however, problems can arise. The possibility of compounding error and loss of precision is inevitable in some cases (though rare).

Compounding error is introduced in routines that map our original helix onto a helix located about the z-axis (i.e., computing the coefficients A, B, C, and D, for $f(t)$). We are uncertain of the error bounds on these calculations, but comparisons with high-precision Mathematica runs show that the error is small (less than 10^{-12}).

Fortunately, given correct values of A, B, C, and D, we can guarantee that the roots found are accurate to the desired working precision. This is because we use Newton-Raphson and the bisection method as our root-finding techniques. With Newton's method, it is possible to get some idea of absolute error of a root x_0 simply by looking at the value of $f(x_0)$ (it should be 0 if we have a root). Also, error is not compounded with Newton's method; each iteration actually decreases the error. Newton-Raphson converges quadratically [Plybon 1992], and near a root the number of significant digits approximately doubles with each step.

Detrimental error can also enter the problem when calculating the brackets for each root. If error in numerical computation places a bracket on the wrong side of a root, that root may be lost. We have never encountered such a case but expect that cooked-up data could produce such an error. A possible remedy (not guaranteed to work in all cases) is first to approximate the bracket using our exact formulas and then to run a root-finding method on them to reduce the error. We did not implement this strategy, because we feel that the chance of this error occurring does not warrant the loss in speed that will occur.

Further, in searching for solutions of $f(t) = 0$, if two roots are extremely close, they can unfortunately be mistaken as one. However, if two extremely close roots are found, it may not even make sense to consider them as distinct, since they may be the product of numerical error.

Robustness

Our algorithm checks extensively for exception cases and will not terminate abnormally due to a computation error or lack of system resources. We implemented checks for special cases that in effect trap errors. One example: We check for an infinite number of roots before we search for roots. We handle "special" cases such as tangencies, inflection points, double roots, etc.

Our algorithm ensures that all roots are bracketed. This prevents Newton's method from accidentally going off and finding another root. Our implementation of Newton's method incorporates the bisection method in cases where Newton's does not converge fast enough or Newton's method departs the bracketed interval. The bisection method is guaranteed to find a root within a bracketed interval [Plybon 1992].

In our algorithm, there is an inherent limit on the number of possible roots; but this limit can be easily increased or even eliminated (allowing memory of the computer to be the only limitation).

Performance/Efficiency

The performance of our algorithm is linear in the number of intersections. This is because we do a single root-find for each intersection (including bracketing). The complexity of a root-finding method is dependent on the how the function is shaped and the number of digits desired. Typically, when Newton's method is used, 5 or 6 iterations are required to find each root to 12 digits of precision, while the bisection method can be shown to take fewer than 44 iterations.

For mapping the helix, finding first-order roots, etc., the time required is relatively constant. Therefore, the time complexity of the algorithm is dominated by the number of intersections.

In addition, our code is efficient in use of space, since the space complexity is also linear in the number of intersections.

Suggestions for Improvement

As the general organization of our algorithm is closely based on a mathematical model, we believe that it is not possible to improve it without a thorough revision of the underlying methodology. However, the *implementation* of the algorithm can be improved in several ways:

- One can attempt to modify the evaluation of expressions so as to reduce the compound and roundoff errors. This often necessitates understanding and use of architecture-specific features of the processor on which the algorithm is executing, thus limiting portability.
- One can attempt to find bounds for the error introduced during the calculation of the A, B, C, and D coefficients for $f(t)$. An estimation of the relationship between this error and the error generated by the root-finding algorithm will also be helpful.
- A useful yet hard-to-implement feature would be the inclusion of internal validation routines that improve correctness and robustness by monitoring for unacceptable computational errors while the algorithm is executing. Such routines would adversely affect performance.
- The input procedures could potentially be extended to include alternative definitions of a helix. However, our explorations of the relevant literature yielded no definitional forms different from the ones that we use.
- The Mathematica routines used in the testing process could be extended to infinite-precision calculation, running in batch mode to check at random the correctness of solutions provided by the real-time algorithm.
- The algorithm could be extended to handle other types of helices: alpha helices, double helices, etc. The underlying methodology of the algorithm should remain unchanged.

Appendix

General Parametric Equations for a Helix

Consider a general 3×4 rotation-translation matrix:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} & a_{34} \end{bmatrix}$$

and the "unit" helix (in vector form)

$$(\cos(\alpha t - t_0), \sin(\alpha t - t_0), t).$$

Applying the matrix produces any general helix:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} & a_{34}. \end{bmatrix} \begin{bmatrix} \cos(\alpha t - t_0) \\ \sin(\alpha t - t_0) \\ t \\ 1 \end{bmatrix},$$

giving the general parametric equations of a helix in space:

$$\begin{aligned} x &= a_{11}\cos(\alpha t - t_0) + a_{12}\sin(\alpha t - t_0) + a_{13}t + a_{14} \\ y &= a_{21}\cos(\alpha t - t_0) + a_{22}\sin(\alpha t - t_0) + a_{23}t + a_{24} \\ z &= a_{31}\cos(\alpha t - t_0) + a_{32}\sin(\alpha t - t_0) + a_{33}t + a_{34}. \end{aligned}$$

Upon expanding the $\sin$ and $\cos$ terms, we obtain the same format as presented in the text of the paper.

Why $t_0 = D/C$?

The equation

$$f(t) = a\cos t + B\sin t + Ct - D$$

can be thought of as the intersection of a helix with parametric equations

$$x = \cos t, \qquad y = \sin t, \qquad z = t$$

with a plane with equation

$$Ax + By + Cz = D.$$

As far as the root-finder is concerned, we have a "vertical" helix with radius 1 and the plane as above. Thus, we say that the point at which the helix's central axis (also the z-axis for the root-finder) meets the plane $Ax+By+Cz = D$ is the point around which the roots are distributed almost symmetrically.

The justification for this claim is as follows. The intersection of a plane and a cylinder in space is an ellipse; as the helix lies on a cylinder, its intersections with a plane must lie on an ellipse. The center of the ellipse is the point of intersection of the plane and the helix's central axis. In **Figure 5**, we show the cylinder that the helix sits on, the ellipse of intersection with the plane, and the helix itself. The curve represents the ellipse of intersection and the lines are the helix. The whole picture shows the cylinder "unfolded."

It can be shown that if any roots exist, they must do so within one complete rotation of the center of the ellipse. Hence, we use the center as the starting point for our root-finder.

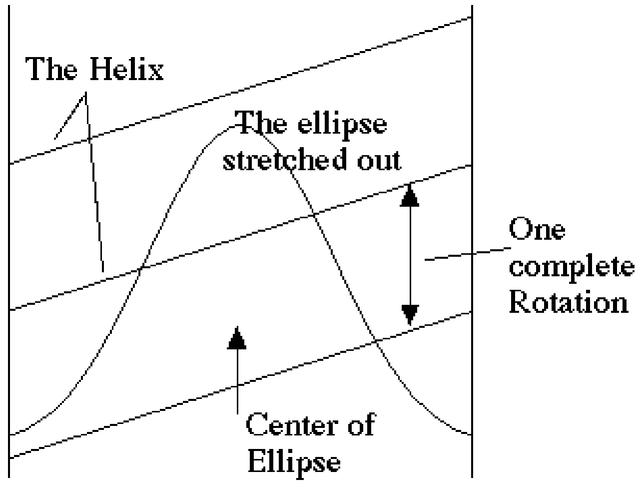

Figure 5. Representation of the intersection of the helix and the plane.

References

Canale, Raymond P., and Steven C. Chapra. 1985. *Numerical Methods for Engineers.* Reading, MA: Addison-Wesley.

Chang, Raymond. 1994. *Chemistry.* 5th ed. New York: McGraw-Hill.

Englefield, M.J. 1987. *Mathematical Methods for Engineering and Science Students.* Edward Arnold.

Press, William F., et al. 1990. *Numerical Recipes in Pascal: The Art of Scientific Computing.* New York: Cambridge University Press.

Plybon, Richard F. 1992. *Applied Numerical Analysis.* Boston: PWS-Kent.

Wolfram, Stephen. 1991. *Mathematica: A System for Doing Mathematics by Computer.* 2nd ed. Reading, MA: Addison-Wesley.

Judge's Commentary: The Outstanding Helix Intersections Papers

Daniel Zwillinger
Zwillinger & Associates
Newton, MA 02165
`zwilling@world.std.com`

Introduction

Typical industrial tasks for applied mathematicians are varied, and many require a computational approach to solve a relatively simple problem. The Helix Intersections Problem was representative: The problem statement, solution techniques to be used, interpretation of the result, and techniques for checking the answer were all straightforward. Most of the submissions did, in fact, perform nearly all of the above steps. The judging criteria focused on how well each step was carried out as well as on the overall organization and clarity.

Thorough Analysis of All Cases

A plane and a helix can have no intersections, any finite number of intersections, or an infinite number of intersections (in the case of an infinite helix or a degenerate helix with zero pitch). A computer program asked to find all the intersection points must respond appropriately to each of these cases.

Simplicity of the Resulting Numerical Problem

Given a helix and a plane, it is straightforward to write down the parametric equations for the helix, depending on one variable, and the equation of the plane. Substituting the parametric equations into the equation for the plane results in an equation with a single variable. Finding the roots of this single variable equation is far simpler than finding the roots of a multiple-variable equation (as some teams proposed).

Numerical Solution of the Problem

A bisection method is guaranteed to find a zero if appropriate endpoints are given, but the method is slow. Newton's method is the method of choice for nonlinear problems, since it is so much faster. However, several teams apparently did not know that Newton's method does not always converge (for example, near multiple roots).

Since appropriate numerical bounds can be found, and bisection can be made to work, several teams used it. The judges preferred a bisection technique, with provable bounds on the results, to a Newton iteration with no mention of possible convergence problems. The team from Macalester College used both a Newton iteration and a bisection method (when the Newton method failed).

Testing the Results

There are many ways in which the results of the computer program created to solve this task could be tested. Some teams used graphical methods, while others used the result of a more general-purpose equation solver (such as Mathematica). There often seemed to be confusion about the reliability of the results of programs such as Mathematica.

Responsiveness to the Question

The original problem statement was concerned about the computational speed of locating the intersection points. The judges looked for statements about the computational requirements of the algorithms presented. There are many ways in which this issue could be addressed: the computer time per intersection point, the computer time saved when compared to a more general mathematical solver (such as Mathematica), or the computational complexity of the algorithm.

Conclusion

Of course, really outstanding papers not only solve the problem but also consider possible extensions and possible limitations. Do these make the problem easier or harder? Do they make the problem applicable to another field? Restricting the problem to a finite-length helix (which is more physically reasonable) was considered by several teams, including the teams from Harvey Mudd College and Iowa State University. Using a finite-area sweeping plane was considered by the team from Harvey Mudd

College. Additionally, one team considered more general helices, such as a spiral drawn on a cone.

About the Author

Daniel Zwillinger received an undergraduate degree in mathematics from MIT and a Ph.D. in applied mathematics from Caltech. His Ph.D. research dealt with the focusing of waves as they travel through random media. He taught at Rensselaer Polytechnic Institute for four years, was in industry for several years, and has been managing a consulting group for the last few years. His work areas have included many industrial mathematics needs: radar, sonar, communications, visualization, statistics, and computer-aided design. He is the author of several mathematical reference books.

Practitioner's Commentary: The Outstanding Helix Intersections Papers

Pierre J. Malraison
Manager, Design Constraints
Autodesk, Inc.
111 McInnis Parkway
San Rafael, CA 94903
pierre.malraison@autodesk.com

The problem statement is straightforward and clear, except for the description of the helix. All of the Outstanding papers assumed a cylindrical helix, which is probably the intent of the problem, but an elliptical helix could be used.

The mathematical reduction used by the Macalester College team is the usual way to do surface-curve intersections in general: Implicitize one surface (in this case the plane), and parametrize the curve (the helix). The intersection points can then be expressed by solving the equation generated when the parametrized form satisfies the implicit equation.

An alternative to the numerical solution of the resulting equation is to use a rational quadratic parametrization (instead of trigonometric) for the helix and end up with a polynomial function to solve.

The team from Iowa State University used a similar strategy but with a different root-finding approach.

The team from Harvey Mudd College used an approach that gets closer to (but didn't quite find) a different alternative: Intersect the cylinder the helix lies on with the plane, and then intersect the helix and that ellipse. The approach that the team took is correct but is limited to finite pieces of helix.

In summary, all three teams provide correct solutions, with slightly different limiting assumptions. The main differences in the solutions are in root-finding strategies.

About the Author

Dr. Malraison started out in (professional) life as a category theorist but saw the error of his ways and has been working in geometric modeling and CAD for the last 18 years. His current interests are generative languages and geometric constraints.

Author's Commentary: The Outstanding Helix Intersections Papers

Yves Nievergelt
Dept. of Mathematics
Eastern Washington University
Cheney, Washington 99004
`ynievergelt@ewu.edu`

The problem of computing all the intersections of a plane and a helix in general positions in space arose at a small company in the western U.S. that designs medical technology. The problem came in the design of a helicoidal part of a device that doctors and technicians together will have to manufacture to fit the particular measurements of each patient. With x-ray data from the patient loaded in a computer with numerical and three-dimensional graphics capabilities, and with a program to compute the requested intersections, doctors and technicians can quickly vary the parameters of the helix, view the helicoidal part superimposed in space with a model of the patient, and examine critical locations by sweeping a plane section through them.

The mathematically accurate yet medically vague description given in the problem statement typifies a common situation of real applications of mathematics: The small start-up company does not want anyone else to know the object of its current research and development. Even the company's name must remain secret, lest anyone else conduct a computed search of the publications of the company's staff and thence piece together a good guess of the objective. Such a situation explains, in part, the dearth of real applications of mathematics in textbooks.

Nevertheless, because the mathematical problem fits in most undergraduate curricula in the mathematical sciences, one solution is scheduled to appear in 1995 in *SIAM Review*, published by the Society for Industrial and Applied Mathematics. The solution was developed in part with support from the National Science Foundation's grant DUE–9255539.

Instructors interested in designing similar material for their own classes are encouraged to contact the author to participate in either of two workshops: 17–21 June 1996 in Spokane, WA, or 26–30 August 1996 in Seattle, WA. Through grant DUE–9455061, the National Science Foundation will pay for participants' room, board, and academic credit, and some summer stipends will be available for participants who would like to submit their material for publication.

About the Author

Yves Nievergelt graduated in mathematics from the École Polytechnique Fédérale de Lausanne (Switzerland) in 1976, with concentrations in functional and numerical analysis of PDEs. He obtained a Ph.D. from the University of Washington in 1984, with a dissertation in several complex variables under the guidance of James R. King. He now teaches complex and numerical analysis at Eastern Washington University.

Prof. Nievergelt is an associate editor of *The UMAP Journal*. He is the author of several UMAP Modules, a bibliography of case studies of applications of lower-division mathematics (*The UMAP Journal* 6 (2) (1985): 37–56) (in which the Brain-Drug Problem was discussed explicitly), and *Mathematics in Business Administration* (Irwin, 1989).

Prof. Nievergelt was also the author of previous MCM problems: the Water Tank Problem (1989), the Brain Drug Problem (1990), and the Optimal Composting Problem (1993).

Paying Professors What They're Worth

Jay Rosenberger
Andrew M. Ross
Dan Snyder
Harvey Mudd College
Claremont, CA 91711
{Jay_Rosenberger, Andrew_Ross, Daniel_Snyder}@hmc.edu

Advisor: David L. Bosley

Introduction

We develop a model with two variations: One encourages early retirement, the other does not.

Our model is generous to those who are promoted or retire later than is typical. It does not allow the college to change gracefully the real starting salaries, although nominal starting salaries are adjusted each year for inflation. The model also takes into account those that it considers to be overpaid. Instead of not giving them any raises, it gives them cost-of-living raises.

When the model is applied to the existing faculty at Aluacha Balaclava College with simulated hirings, promotions, and retirements, the faculty is separated into clearly different salary bands by the year 2010 (with the exception of two overpaid full professors), replacing the present muddle.

Constraints

For convenience in reference, we note here the constraints:

1. If there is enough money for raises, then everyone gets a raise.

2. Instructors who are promoted according to the usual schedule of seven years as an assistant professor and seven years as an associate professor and who work 25 years or more should receive at retirement twice as much as a new assistant professor's salary.

3. Although there should be a reward for years of experience, the salaries of two faculty members with equal rank should approach each other as they gain experience.

4. The salary for a newly promoted faculty member should be about what it would have been in seven years, without the promotion.

Assumptions

Although payments are actually made throughout the school year, and salary decisions are made in March, we assume that the decisions are made between discrete yearly salary payments.

We also assume that when the decisions are made, the Provost has the budget for next year and an estimate of the cost-of-living increase. However, no information is available for years beyond the one for which salaries are being decided.

Since we are prohibited from decreasing anyone's salary when moving the current faculty to the new scheme, we assume that there is always enough money to pay everyone's salaries from the previous year. That is, there might be no money for raises, but we can at least pay the faculty at last year's nominal level.

We must give everyone a raise if anyone gets a raise, but we assume that we can give unequal raises. Otherwise, we would just split the money evenly among the faculty, and the current salary system would not change very much at all.

Although Constraint 2 mentions 25 years and retirement, instructors are not forced to retire either at 25 years of experience or at 65 years of age. However, we assume an upper limit of 60 on the number of years of experience.

Constraints 2 and 4 refer to real dollar values. Otherwise, it would be extremely difficult to guarantee promising new Ph.D.s that they will retire at twice their current salary. Besides, doubling the nominal salary in 25 years won't even keep up with 3% inflation.

Constraints 1 and 3 refer to nominal dollar values. The college always deals in nominal amounts when calculating budgets. There might not be enough to give the full cost-of-living increase; this would amount to a decrease in real salary.

Faculty members who take longer to be promoted than usual may receive more than a seven-year raise if and when they finally do get promoted. That is, there is no penalty for being promoted late other than the salary lost while waiting.

New faculty are hired only if there is enough other money to pay their starting salaries, which take into account any previous experience and are not taken from the amount available for raises.

The salaries of retiring faculty do not get thrown into the pool for raises but can be applied to hiring new faculty.

Since no information was given about the transition from assistant professor to associate professor, we will assume that one must have seven years of experience total (not necessarily with this college, or as an assistant professor) in order to be promoted.

Analysis of the Problem

It is currently late in the winter of 1995, too late for next year's salaries to be decided by the model proposed in this paper. The first year that will be on the new salary model is 1996–97. Even then, salaries will gradually move toward the target curves, since some faculty members are overpaid and cannot have their salaries reduced.

One solution is to set target salaries so high that everyone needs a raise to attain the target. This is clearly not a solution that the college would favor, although it would be rather popular with the faculty.

It is difficult to conceive of a salary scheme that had no view of the future yet managed to reliably satisfy the constraints regarding promotions and retirement. Therefore, the model should contain an overall view of how much a faculty member at a certain rank with a certain number of years of experience should be paid.

This feature is certain to cause some friction between the college and the faculty; how much is a full professor with ten years of experience worth? Moreover, at what rate should salaries increase, within the framework of the constraints? Should the model be set up to encourage or discourage retirement? Can someone be hired as a full professor with no experience? These are political questions for the Provost and faculty to negotiate; the model must handle whatever answers that negotiation produces. Faculty members must become accustomed to the the Provost placing a certain value on them.

Design of the Model

Our model presents an overall goal for a salary system and then adapts it to the real world. We offer a pair of core systems (logarithmic and linear) for the administration and faculty to decide between. These cores are what the college will pay the faculty if it has enough money to do so. These cores are in real dollars, not adjusted to inflation. They are then adjusted for inflation (both historical and predicted) and adjusted to meet a finite budget. There are two options to take care of faculty members who, according to the new system, are being overpaid. Finally, we deal with the unlikely event of a budget excess.

Variables

Let t be the current year; $t+1$ is the year for which salaries are being computed. Let t_i be the year that faculty member i started at the college, adjusted by the number of years of experience credited upon entrance into the plan. Thus, if faculty member i joined the college in 1994 with four years of experience, then $t_i = 1990$.

Let $T(i,t)$ denote the amount in real dollars that faculty member i should get paid in year t, the *T*arget for i for that year. This target will depend on the rank and number of years of experience of i.

In the cores that follow, a_0, b_0, c_0, d_0 denote the initial salaries of a full professor, associate professor, assistant professor, and instructor, respectively. This is the amount paid to a faculty member with no experience on entering the plan. According to the problem statement, no one can be hired at a rank higher than assistant professor; however, the model could easily handle such an event. Indeed, one needs to estimate the "initial" salaries of associate professors and full professors to start the salary system.

The initial salaries for no experience, $d_0 =$ \$27,000 for an instructor and $c_0 =$ \$32,000 for an assistant professor, are of some concern. We must have $a_0 < 2c_0$, or else full professors would have a decreasing salary in order to hit $2c_0$ at 25 years. Convention forces $d_0 < c_0 < b_0 < a_0$. It will turn out (see the **Appendix**) that $a_0 =$ \$40,000 and $b_0 =$ \$36,000 are good estimates.

The Two Cores

According to Constraint 3, faculty members of equal rank but different experience should have their salaries approach each other as time goes on. This leaves two possibilities: Either the absolute difference goes to zero (logarithmic core), or the ratio goes to one (linear core).

One might expect a model to have a core that has a horizontal asymptote, so that a faculty member's salary has a clear upper bound. However, as careers are limited to sixty years (Prof. Methuselah does not work at Aluacha Balaclava College), salaries are bounded. As long as the proper rate constants are chosen, no faculty member's salary will get too large for the college to handle.

The first core, the logarithmic, increases more rapidly at the beginning of someone's career than at the end. Larger raises occur early, but by the time one has gained twenty-five years of experience, the salary curve has really flattened out. Here is the logarithmic core:

$$T(i,t) = \begin{cases} d_0 \log_{10}(d(t-t_i)+10), & \text{for } i \text{ an instructor;} \\ c_0 \log_{10}(c(t-t_i)+10), & \text{for } i \text{ an assistant professor;} \\ b_0 \log_{10}(b(t-t_i)+10), & \text{for } i \text{ an associate professor;} \\ a_0 \log_{10}(a(t-t_i)+10), & \text{for } i \text{ a full professor.} \end{cases}$$

The $+10$ term inside the logarithm allows the instructor's starting salary to be the coefficient of the logarithm expression. Indeed, the equation becomes (taking c as an example) $c_0 \log_{10}(0+10) = c_0 \cdot 1$, so the factors out front are the initial salaries, with no scaling necessary.

For the linear core, we have:

$$T(i,t) = \begin{cases} c_0 + c(t - t_i - 7), & \text{for } i \text{ an instructor;} \\ c_0 + c(t - t_i), & \text{for } i \text{ an assistant professor;} \\ b_0 + b(t - t_i), & \text{for } i \text{ an associate professor;} \\ a_0 + a(t - t_i), & \text{for } i \text{ a full professor.} \end{cases}$$

The variables a, b, c, and d are determined by Constraints 2 and 4. This core guarantees that a full professor retiring at twenty-five years after the usual promotions will make twice as much (in real dollars) as a new Ph.D. entering as an assistant professor. It also guarantees the equivalent of a seven-year raise for someone who gets promoted. The calculation of these coefficients is a matter of applying Constraints 2 and 4; they depend only on the initial salaries. (See the **Appendix** for a brief derivation.)

Note that in the linear core, the instructor salary is a seven-year time shift of the of the assistant professor salary. Since there is uncertainty about when instructors will receive their Ph.D.s and become assistant professors, we shift the assistant professor salary curve to use it as an instructor salary curve. A seven-year shift makes the initial instructor salary \$27,000, as it should be.

In **Figures 1** and **2**, we graph the ideal salaries for all ranks of faculty. Note how the logarithmic core tapers off after twenty-five years, giving more experienced instructors less and less of a raise each year, while the linear core keeps giving them the same raise. It is in this way that the logarithmic core encourages retirement.

The Real World

Inflation

Let γ be the cost-of-living function from one year to the next, so that $\gamma(t+1)$ is cost-of-living factor from year t to year $t+1$; a typical value would be 1.03 for 3% inflation. Each person's real target salary for year t is multiplied by the accumulated cost-of-living increases to produce the nominal target salary $N(i, t+1)$. If the new plan started in year t^0, then the nominal target salary is

$$N(i,t+1) = T(i,t+1)\hat{\gamma}(t+1) \prod_{j=t^0}^{t} \gamma(j),$$

where $\hat{\gamma}(t+1)$ is an estimate of the cost-of-living factor from the current year to the next.

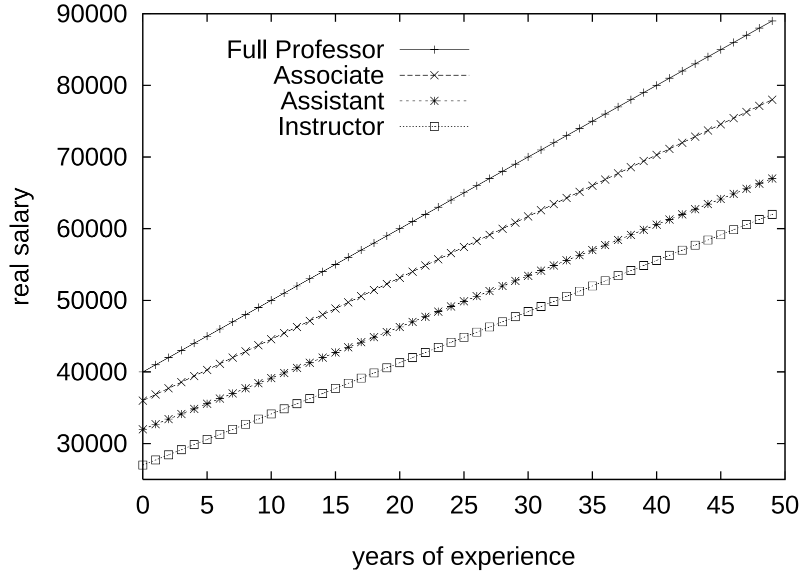

Figure 1. The ideal salaries using the linear core.

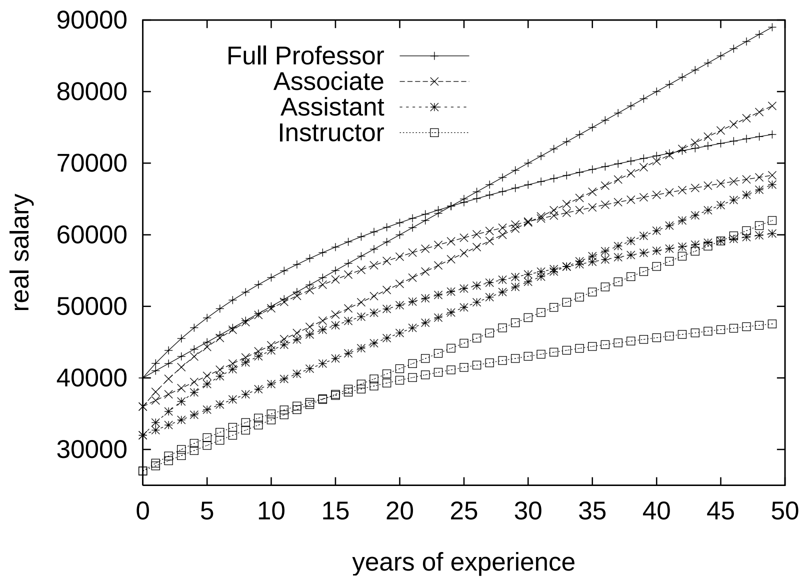

Figure 2. The ideal salaries using the logarithmic core, with linear core salaries shown for comparison.

Finite Budgets

Of course, the college will not always have enough money to give each faculty member the full raise each year. This calls for some way to portion out the raises in accordance with who deserves them the most. If two faculty members each make $40,000 in 1996, and (according to one of the salary plans) the first has a target $41,000 in 1997 while the second's target is only $40,100, then the first should get a larger raise.

Let $n(i,t)$ be the amount in nominal dollars that faculty member i gets paid in year t. Since budgets are limited, this will be at most $N(i,t)$ if nobody is overpaid. We have only some small amount of money, $M_r(t+1)$, for raises next year. This number is given to us by the outside world (the college's treasury). Let $M_n(t+1)$ be the amount needed next year for raises if all targets are to be met. That is,

$$M_n(t+1) = \sum_i \left[N(i,t+1) - n(i,t)\right],$$

where all Ms are measured in nominal dollars. Usually, $M_n > M_r$: There isn't enough to give the faculty the raises that they deserve.

A fair way to give raises is to give each person a raise in proportion to how much one is needed, where the proportion is the amount available for raises over the amount needed for raises. Thus, next year's salary for instructor i is

$$n(i,t+1) = n(i,t) + \text{raise}(i,t),$$

where i's raise from year t to year $t+1$ is

$$\text{raise}(i,t) = \frac{M_r(t+1)}{M_n(t+1)} \left[N(i,t+1) - n(i,t)\right].$$

In this manner, instructors who are getting paid far less than their target will get a larger portion of the raises, bringing them closer to their target.

Bruised Egos

The former salary plan has given some faculty members more money than they deserve under the new system. Thus, some salaries need to be reduced. We can't actually cut salaries; indeed, if there is money available, everyone has to get a raise. However, raises can be unequal. To make things simple, we could give an overpaid faculty member an ϵ-dollar raise each year until the target salary catches up with the actual salary. However, this is likely to bruise a few (overpaid) egos.

A good way to placate the overpaid would be to give them a new nominal target $O(i,t)$ that corresponds only to the projected cost of living increase:

$$O(i,t+1) = \hat{\gamma}(t+1)n(i,t).$$

We then treat the overpaid who are underneath their new target O just like those who are underneath their original target N. If there is no positive inflation that year, just give the overpaid instructors some small amount each, to make sure they get a raise. Now, recompute the amount of money needed for raises of both types, and portion out the money that we have according to who needs it the most:

$$M_n(t+1) = \sum_i \begin{cases} N(i,t+1) - n(i,t), & \text{if } i \text{ would be underpaid;} \\ O(i,t+1) - n(i,t), & \text{if } i \text{ would be overpaid.} \end{cases}$$

The new salaries are then computed as

$$n(i,t+1) = n(i,t) + \text{raise}(i,t),$$

where the raise from year t to year $t+1$ is

$$\text{raise}(i,t) = \frac{M_r(t+1)}{M_n(t+1)} \times \begin{cases} N(i,t+1) - n(i,t), & \text{if } i \text{ would be underpaid;} \\ O(i,t+1) - n(i,t), & \text{if } i \text{ would be overpaid.} \end{cases}$$

Excess Funds

Perhaps this belongs under an "Unreal World" section, but on the off-chance that there is more money available than is needed to put everyone on target, there are several options:

- *Raise everyone's salary.* This has a negative consequence: it could put people over their targets for the next year, if the excess is very large. It could put the college into dire financial straits in the future, when the faculty members' salaries cannot be cut. However, if the excess is not very large, instructors will still be below target for the year after the excess, and not much harm is done.

- *Give everyone bonuses.* This would take care of the excess without raising the faculty's expectations for years to come. This is a better option from the college's point of view than raising salaries, and it is a common practice in industrial settings.

- *Give it to the General Fund,* perhaps to caffeine grants for sleep-starved students.

Model Verification

We have projected the performance of the proposed models over the next fifty years. We analyzed the model both with and without such influences as limited budgets, inflation, hiring of new faculty, promotion of faculty members, and retirement. We have also analyzed the effect on this performance of changes in the chosen constants a_0 and b_0.

Figures 3 and **4** show the long-term effects on existing faculty of our model, using the linear and logarithmic cores. These figures assume no hiring, promotion, or retirement, and they ignore cost-of-living increases and possible monetary constraints.

From these two figures, we can see how our model will move the faculty toward a uniform salary system over time. Faculty members with current salaries below the model's target are given raises to bring them up to target. Faculty members with current salaries above target are held to a constant salary (since there is no inflation) until the target catches up to them.

Figures 5 and **6** show how our model, using the linear and logarithmic cores, behaves in the presence of 3% annual inflation. In these graphs, the faculty retire according to the schedule described later, which explains why the graphs become more sparse at the left side. The graphs assume that the college has unlimited funds. Note that these two graphs are essentially identical for large t, when the inflation terms dominates the model.

Figures 7 and **8** show the behavior of the model, with the linear and logarithmic cores, when faculty are promoted, eventually retire, and are replaced by new hires who also are promoted and eventually retire. As before, the model brings faculty into a coherent salary structure over time.

Figures 9 and **10** show the effect of budgetary constraints on the salary of an individual faculty member over time, for each of the cores. These graphs do not include promotions. In **Figure 9**, under the linear core, an early difference between the actual and target salaries gets magnified, because the yearly raise never decreases under the linear model but there is never enough money to give a full raise. In **Figure 10**, under the logarithmic core, an early difference between the actual and target salaries is eliminated as the yearly raises get smaller. These graphs demonstrate the model's ability to cope with limited-money situations.

Sensitivity Analysis

To analyze the sensitivity of our model to changes in a_0 and b_0, we varied these constants and examined the effect on the salaries of faculty members at each rank with fifty years of service. We held c_0 and d_0 constant because they were provided in the problem statement. **Tables 1** and **2** show our results.

Table 1 shows that the fluctuation in salary is higher with the linear core, as is to be expected. It is still only 8%, though. **Table 2** shows that in spite of 10% variation in a_0 and b_0, the fifty-year salaries of all four levels of faculty fluctuate by at most 2% under the logarithmic core. We conclude that our model, especially with the logarithmic core, is relatively insensitive to its initial parameters.

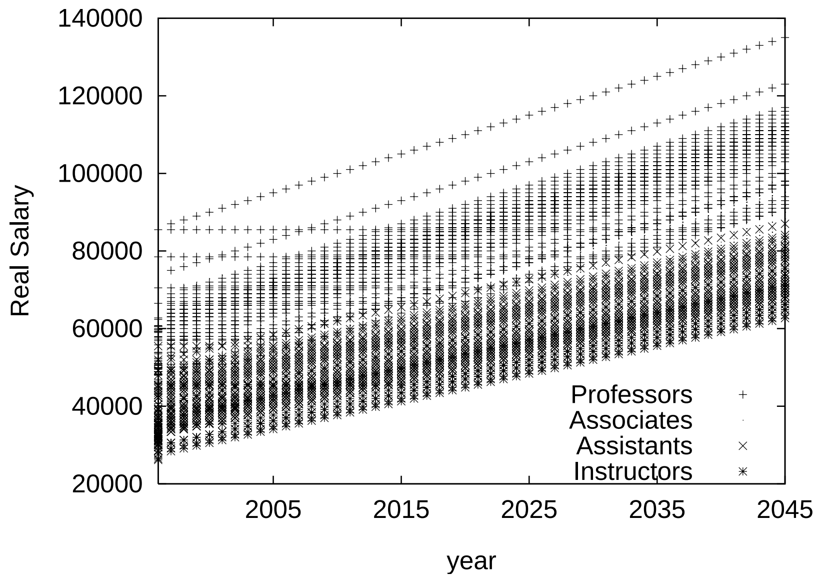

Figure 3. Long-term transition with linear core.

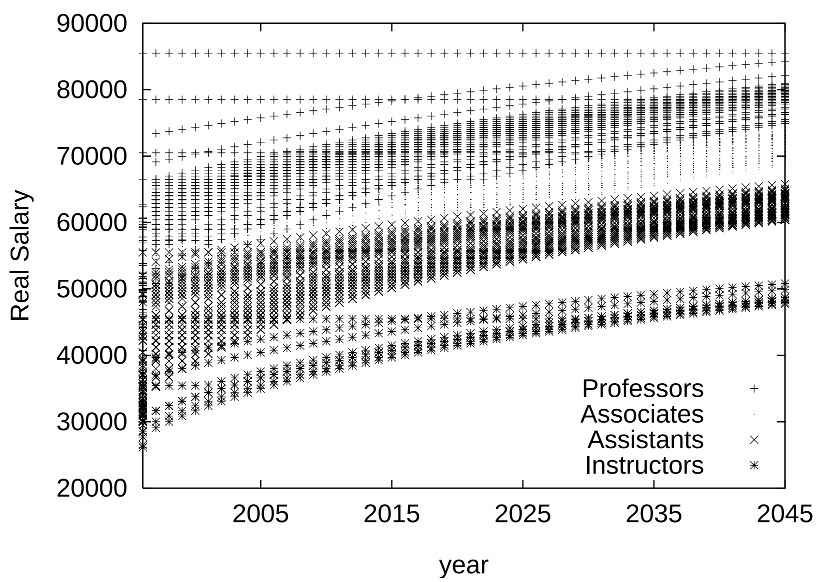

Figure 4. Long-term transition with logarithmic core.

Table 1.

Effect of variations in initial conditions on salary after fifty years of service, using the linear core.

a_0	b_0	Professor	Associate Prof.	Assistant Prof.	Instructor
40,000	36,000	89,000	78,000	67,000	62,000
42,000	36,000	86,917	79,944	67,972	62,833
40,000	38,000	89,000	75,333	71,667	66,000
42,000	38,000	86,917	77,278	72,639	66,833
38,000	34,000	91,083	78,722	61,361	57,166

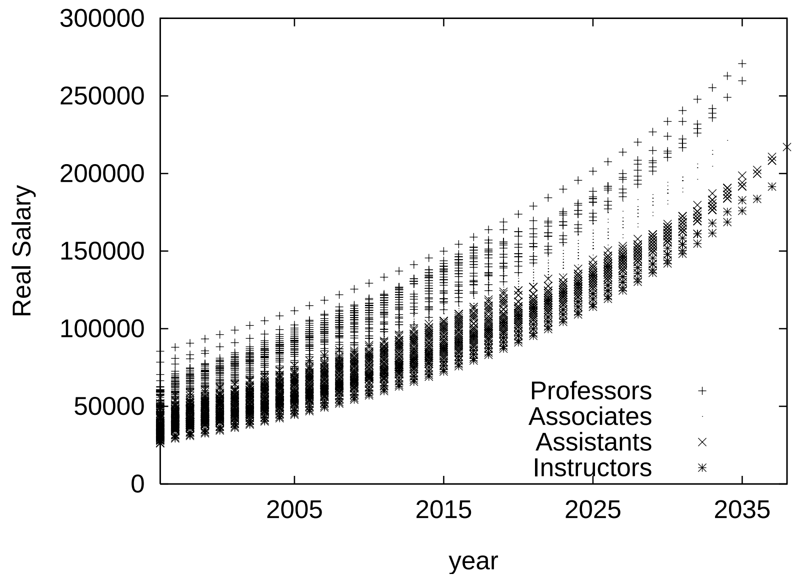

Figure 5. Linear core with retirement and inflation.

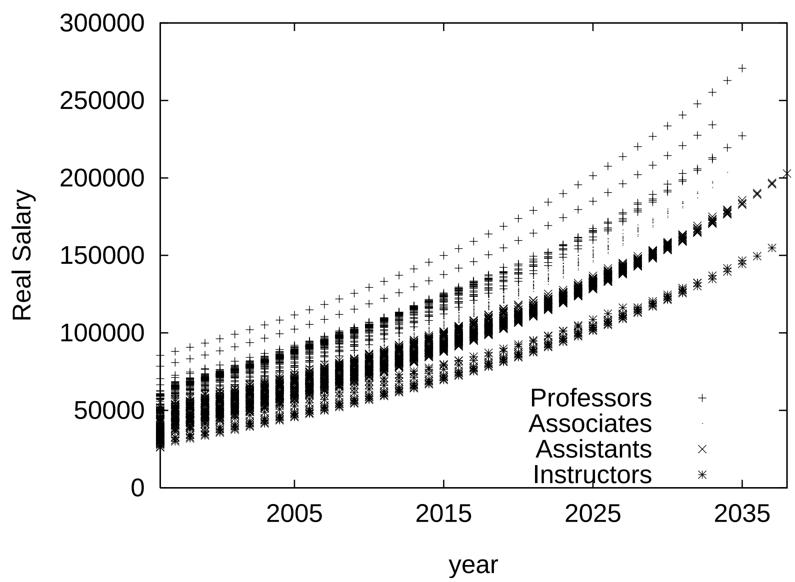

Figure 6. Logarithmic core with retirement and inflation.

Table 2.
Effect of variations in initial conditions on salary after fifty years of service, using the logarithmic core.

a_0	b_0	Professor	Associate Prof.	Assistant Prof.	Instructor
40,000	36,000	74,107	68,312	60,180	47,556
42,000	36,000	73,996	68,545	60,365	47,601
40,000	38,000	74,017	68,306	60,898	47,744
42,000	38,000	73,996	68,548	61,061	47,788
38,000	34,000	73,938	68,048	59,530	47,391

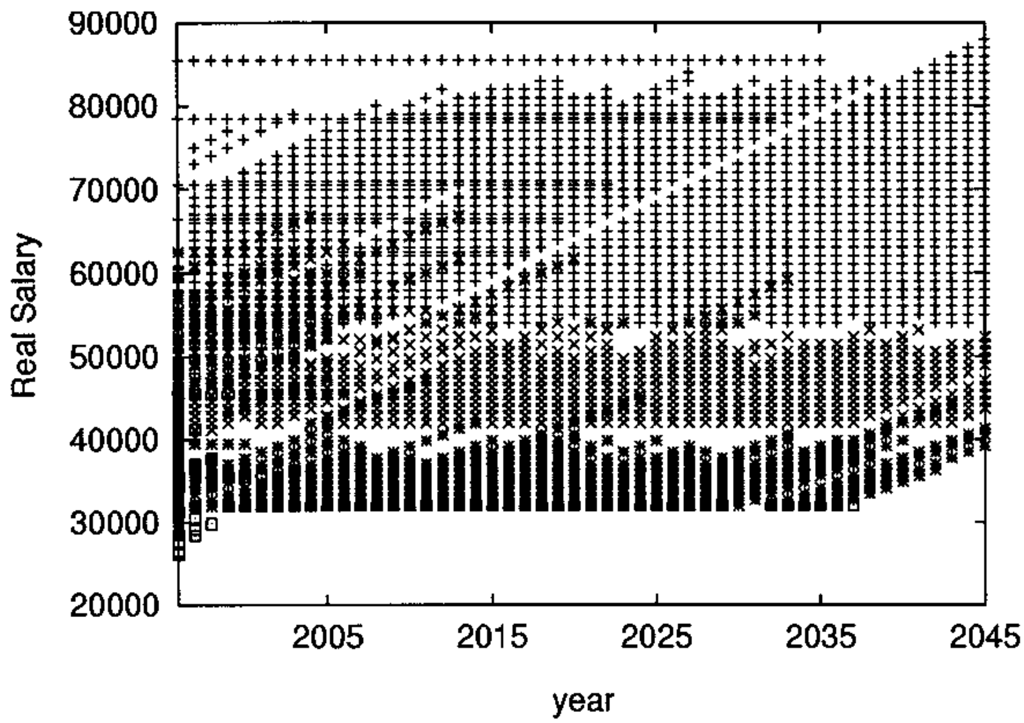

Figure 7. Hiring, promotion, and retirement, under the linear core.

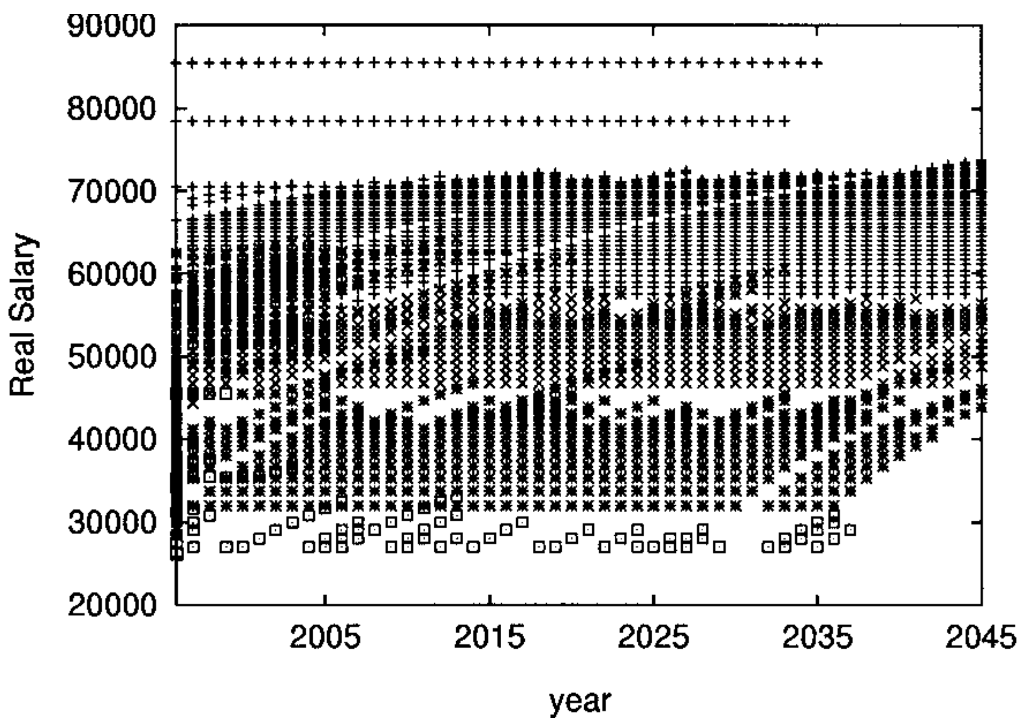

Figure 8. Hiring, promotion, and retirement, under the logarithmic core.

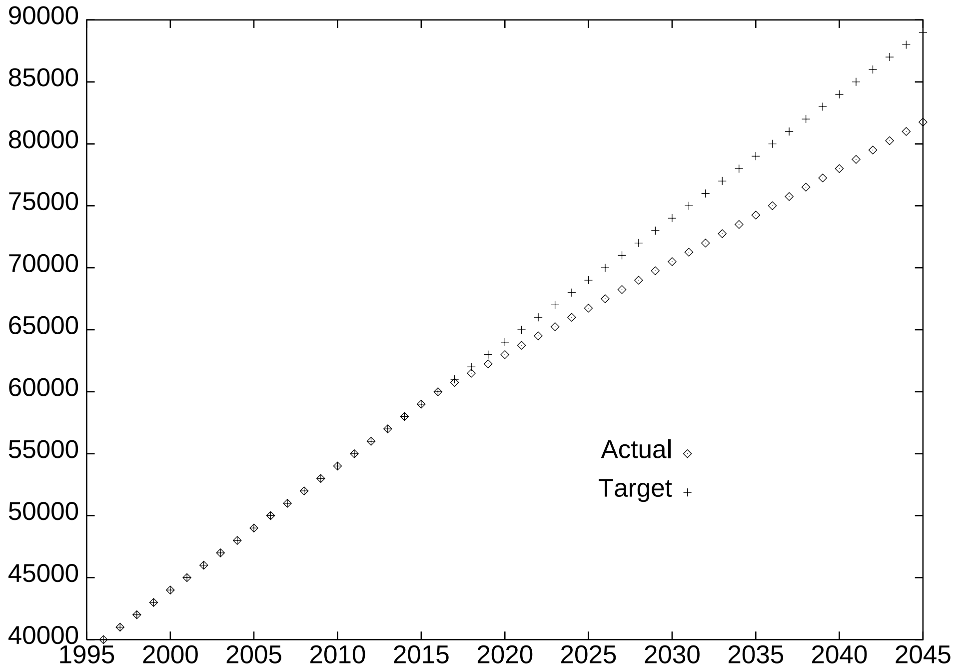

Figure 9. Effect of monetary pressure on an individual faculty member, under the linear core.

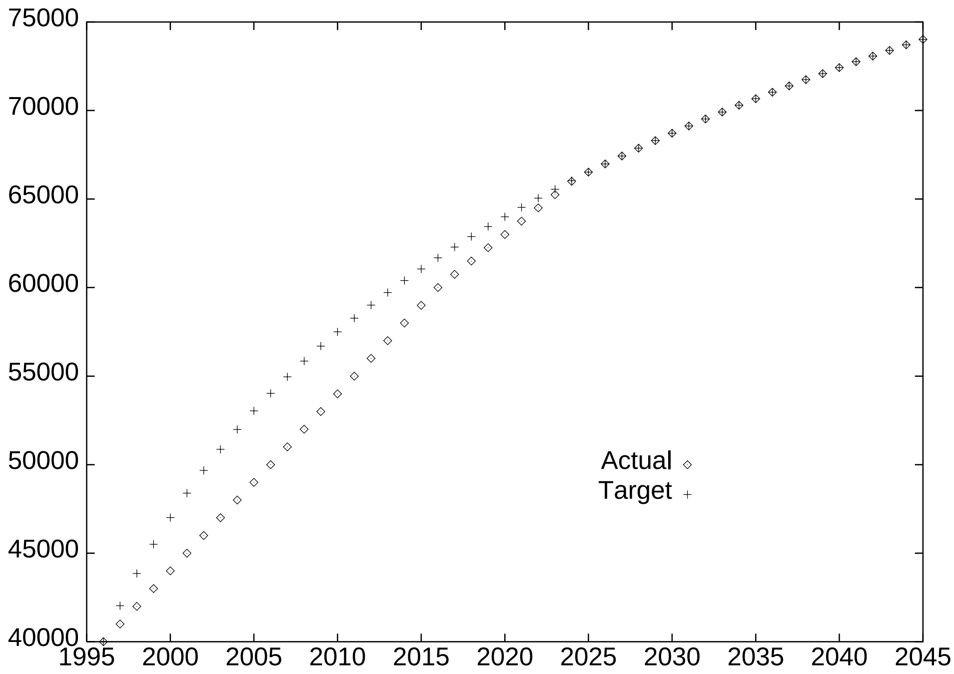

Figure 10. Effect of monetary pressure on an individual faculty member, under the logarithmic core.

Long-Term Performance

In making long-term predictions, many real-world influences on our model had to be simulated. Specifically, we had to decide how many faculty members would change status each year, including new hires, promotions, and retirements. Also, we had to simulate the money available for promotion and arrive at a reasonable cost of living factor for each year.

By examining the initial data, we concluded that that the college hires a mean of 9 faculty per year, with a standard deviation of 5 on a discretized normal distribution. We decided more or less arbitrarily that faculty will retire after working for 40 years, with a standard deviation of 2 years, again on a discretized normal distribution. For promotions, we decided that 50% of assistant professors would become associate professors in 7 years, 25% would become associate professors in 8 years, and so on with continued halving. We used the same probability distributions for associate professors being promoted to full professor.

We used a constant 3% inflation per year, or $\gamma_i = 1.03$ for all i. We also assumed that the college would always be able to predict this value accurately for the next year. Small fluctuations in any γ_i will have negligible effects on the model's performance, as will small inaccuracies in the college's yearly predictions. Finally, we analyzed the model in the presence of both limited and unlimited money supplies. For limited money, we made a constant amount of money available for raises and chose this constant such that it would be inadequate soon after the initial year of the simulation.

Strengths and Weaknesses

The logarithmic option encourages instructors to retire earlier than the linear model, since the logarithm curve flattens out at higher values. This means that the raise for a professor with forty years of experience is small relative to the raise for someone with twenty-six years of experience. This gives the Provost a way to encourage or discourage retirement, based on the needs of the college.

If the college wishes to adjust the real values of the starting salaries, everyone's salaries will change, not just those hired after the change is made.

The faculty must be willing to settle for a higher potential salary, instead of a guaranteed higher salary, for promotions and retirement. As long as the college cannot guarantee enough money for everyone's raises, this will remain.

Salary increases were calculated according to an faculty member on a track with the minimum years between promotions. This causes late promotions to receive the equivalent of an eight- or nine-year raise at their current rank. For example, an instructor promoted in eight years will receive a raise greater than the raise awarded for a promotion in seven years.

Furthermore, all future promotions (if any) will have larger raises. Likewise, a professor retiring at twenty-six years experience, instead of twenty-five years will receive a salary greater than twice the salary of a beginning assistant professor (see the **Appendix**).

Appendix

The two cores allow the initial salaries to be set by the outside world; however, the rate at which salary increases depends on these initial salaries. The constraints apply in this manner:

Constraint 2: Consider a new Ph.D. with no prior experience. If promotions occur on time, then years 0 through 6 are spent as assistant professor, years 7 through 13 as associate professor, and years 14 and on as full professor. Thus, the Ph.D.'s first year as associate professor is year 7; it should correspond in salary to the year 14 as assistant professor:

$$T(\text{associate}, 7) = T(\text{assistant}, 14).$$

Similarly, the Ph.D. becomes a full professor at year 14; that should correspond to the year 21 as associate professor:

$$T(\text{professor}, 14) = T(\text{associate}, 21).$$

We don't have to track inflation and budget constraints through the actual year because we are dealing in real dollars, and the salary curves don't change: Ideally, Associate Professor X with 15 years of experience in 1997 makes the same real amount as Associate Professor Y with 15 years of experience in 2010.

Constraint 4: Similar to the previous constraint:

$$2T(\text{assistant}, 0) = T(\text{professor}, 25).$$

Since one's salary increases with time, those who retire after year 25 receive more than twice $T(\text{assistant}, 0)$, which fits the constraint.

Earlier, we discussed the reasoning behind choosing the instructor salary curve in the linear core as a time shift: There is no explicit method of solving for a linear rate constant for the Instructor salary. We run into the same problem for the logarithmic model, except that a time shift would not produce a suitable result for the given starting salaries of an assistant professor and an instructor. Thus, we need to choose an average year in which to promote instructors to assistant professors. An instructor may be promoted after exactly one year; if we base the salary curve on a longer period, instructors who take longer than expected to earn their Ph.D.s would receive less than a seven-year raise upon promotion. Therefore, we promote instructors after one year.

Substitution and simplifying produce the following rate constants:

Logarithmic Core:

$$\begin{aligned} a &= \left[10^{2c_0/a_0} - 10\right]/24 \\ b &= \left[(14a+10)^{a_0/b_0} - 10\right]/21 \\ c &= \left[(7b+10)^{b_0/c_0} - 10\right]/14 \\ d &= \left[(c+10)^{c_0/d_0} - 10\right]/8 \end{aligned}$$

Linear Core:

$$\begin{aligned} a &= (2c_0 - a_0)/24 \\ b &= (14a + a_0 - b_0)/21 \\ c &= (7b + b_0 - c_0)/14 \end{aligned}$$

How do the starting salaries $a_0 = \$40{,}000$ and $b_0 = \$36{,}000$ work out so well? For the linear core, look at the initial salary for instructors:

$$c(0-7) + c_0 = -7c + c_0.$$

The quantity c is determined by b and b_0, while b is determined by a and a_0. So, a_0 and b_0 affect the instructor's initial salary. With a_0 and b_0 chosen as they were, the starting instructor's salary comes out to \$27,000, exactly as specified in the problem statement.

The World's Most Complicated Payroll

Frank Thorne
W. Garrett Mitchener
Marci Gambrell
North Carolina School of Science and Mathematics
Durham, NC 27705

Advisor: Dot Doyle

Introduction

We present a model for paying the faculty that we believe is fair and consistent with the requirements, as well as ways of dealing with cost-of-living increases, budget shortages, and the transition process.

We use a two-track salary model: All of the instructors are on one track while all of the Ph.D.s are on the other. To accommodate both seniority and rank elements into our model, we make salary a function of "quality points," an index that incorporates both seniority and rank for Ph.D.s but just seniority in instructors. We use part of a root function as our salary function for both tracks and make the slope of the instructor function half that of the Ph.D. function.

Furthermore, we designed a transition process for instructors promoted to assistant professor. To incorporate cost-of-living raises, our salary functions work in constant dollars. To deal with budget deficits, we let promotions happen as normally but also give out fractional quality points and make up the difference later. Our model does not provide for drastic deficits; these would require major changes, such as cuts or layoffs, which we leave as the business of the administration.

To create a transition process, we give all faculty making more money than they should (as determined by the model) a minimum $100 raise each year (in real, not constant, dollars) until they are on the track as they should be, and divert the remaining money into extra raises for the financially challenged. Estimating some data that were not provided, we determined that 95% of the overpaid faculty would be on the proper payroll track within 4.6 years, meaning that with the exception of a few grossly overpaid faculty, the problem of unfairness would be solved.

Our model is highly flexible and adaptable; many of the parameters can be changed to fit the wishes of the administration. Unfortunately, many of the estimates that we produced are based on simulations of data that were not given. We do, however, include the templates we used, so better estimates can be made using data which should be in the college's files.

What follows is our recommendation to the Provost for the new faculty compensation system, including the of cost-of-living increases, the policy when there is not enough money to fully support the model, and the transition from the current system to the new, improved system.

Major Assumptions

- *All other colleges from which faculty might be transferred have a faculty ranking system similar to ABC's.* Thus, when a faculty member transfers from another school, we have a good estimate both of her status in the system and of her experience.
- *An assistant professor or associate professor may not be promoted until she has spent seven years at her current rank.* The circumstances stipulate that an associate professor must work seven years before promotion but do not say anything about an assistant professor. Furthermore, we consider promotion after exactly seven years in a rank to be "on time."
- *We have enough money to implement our model, and there is no inflation or deflation.* ("The world is perfect.") Later we will consider the effect of limited funds and inflation.
- *Every year that a faculty member receives a promotion benefit, she also receives a normal year's raise* (the raise first and then the benefit). We assume that all promotions take place in between academic years.
- *With no inflation, the starting salaries are $27,000 for instructors and $32,000 for assistant professors.* In this model, a full professor who has worked at ABC for twenty-five years starting as assistant professor, with promotions after seven and fourteen years, will receive a salary that is twice the starting salary of an assistant professor (from Principle 3), or $64,000.
- *The number of faculty members in each rank does not change significantly from one year to the next.*

Problem Analysis and Development of the Model

We graph the current salary data for each of the faculty ranks in **Figure 1**.

The most obvious problem with the data is that the number of years on the chart is only the number of years at ABC. Many faculty members may

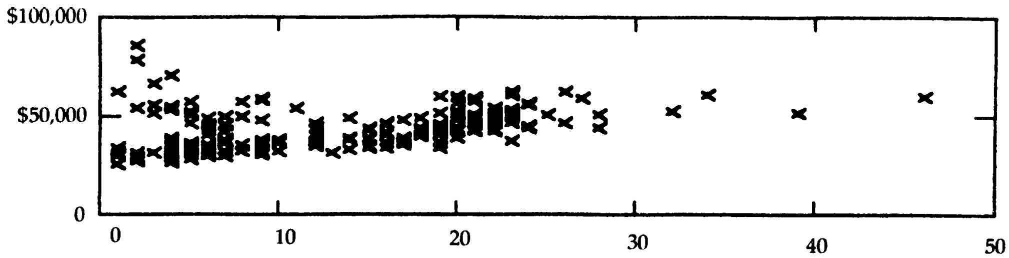

Figure 1. Scatterplot of salary vs. years at ABC.

have transferred in from some other institution with credit for up to seven years of experience, but we do not know how much credit they have been given. We also do not know the number of years spent in each rank for each faculty member moving through the system. We would expect that a faculty member with a 24-year career, for example, who spent 7 years as an assistant professor, 7 years as an associate professor, and 10 years as a full professor, would be making more than another faculty member who had spent 15, 8, and 1 years, respectively. In the data, however, both are listed just as 25-year full professors.

Since the data are incomplete, we designed a model independent of the data. The first model we tried involved using several salary tracks, one for each rank, each with the same slope. On promotion, an instructor would move an extra seven years worth of salary on her own track and then move up to the next one. This complicated system resulted in several problems; for example, faculty who were promoted earlier would earn less money after a while than those promoted later. We wanted faculty promoted earlier to earn more money. Furthermore, we wanted a model that would be easy to present, easy to implement, and reasonably flexible.

We opted for a different alternative. Since the salary system for assistant/associate/full professors is a single track with consistent rules, we deal with the instructor track separately, thus giving a two-track system. We tackle inflation by simply vertically stretching the curves by the next year's projected cost-of-living increase.

The next problem was what to do if there is not enough money for each faculty member's expected raise. We wanted to see the effects of this change on the actual faculty at ABC, but lack of the prior work history of each faculty member stops us from being able to calculate quality points and the corresponding salary. Instead, we generated a possible set of work histories for the faculty. We used a MathCAD template to generate at random 204 (the number of current faculty members) sets of four points (the numbers of years spent in each rank) and then ran these through our salary function to find the salaries. We calculated the amount needed to pay everyone, the expected raises, and amounts for each faculty member when there is not enough to give everyone a full raise. For application at ABC, actual histories could be entered into the template. We also created a special plan for the transition between the current system and the new one, involving giving a minimum raise to those making more than they should and dividing the extra among the others.

The Ph.D. Track

For the ranks of faculty with a Ph.D., it is much clearer how the salary increases than in the instructor phase. We are explicitly told that a promotion is "on time" for a Ph.D. if it happens after seven years and that the benefit should be the same as seven years of raises. For an instructor, there is no clear definition of "on time" or minimum time to promotion. Someone could stay instructor for her entire teaching career by never getting a doctorate. Furthermore, being an assistant professor is a prerequisite to reaching the higher levels, while being an instructor is not.

We solve the problem of the difference by having the two salary schemes lie on different lines. Since we put all Ph.D.s on one salary track, we must find a way of dealing with promotion benefits and years of experience as two parts of the same variable, since salary can be a function of only these two factors. We introduce "quality points," which take into account the following rules:

- Faculty receive 1 point for every year worked at ABC.
- When a faculty member is promoted, she receives 7 points and the correspondingly higher salary if she is promoted on time (after seven years). If she is promoted later than the minimum amount of time, the reward should be correspondingly less, to satisfy the constraints. We increase the number of points by $49/t$, where t is the number of years spent in the rank from which she is promoted.
- Employees hired from other colleges receive 14 points if they have reached the rank of associate professor and an additional 14 points if they have reached the rank of full professor when they transfer in. These are the minimum numbers of points for these ranks. Furthermore, in whatever rank the employee enters, we give her up to seven years (quality points) of credit in that rank. For example, if an employee comes in with 30 years of assistant professor experience and 2 years of associate professor experience, we make her an associate professor with $14+2 = 16$ points; if another employee comes in with 30 years as an assistant professor and 9 years as an associate professor, we make her an associate professor with $14 + 7 = 21$ points.

Given these conditions, we observe that a new Ph.D. who has worked 25 years and has made all of her promotions on time has earned 39 points and thereby should earn $64,000. We should therefore fit our function $f(x)$, where x is the number of quality points and $f(x)$ is dollars in thousands, to the following constraints:

- The function must always be increasing, to satisfy the constraint that more experience in a rank results in more money: $f'(x) > 0$.
- The function should always be concave down, so that its slope decreases. Therefore, the effect of a difference in experience will narrow over time, as prescribed: $f''(x) < 0$.
- $f(0) = 32$ and $f(39) = 64$.
- The slope should be reasonably large (but not too large) at the beginning, so new employees get reasonably large (but not huge) raises.

We use part of a power function to determine our function. In general form, the equation is part of a p^{th} root equation, where p is some (not necessarily integral) number greater than 1. Performing an affine transformation of sorts, we map $(0, 32)$ to $(1, 1)$, and $(39, 64)$ to $(2, 2^{1/p})$. This was accomplished via the function (see **Figure 2**)

$$f(x) = 32\left[1 + (2^p - 1)\,\frac{x}{39}\right]^{\frac{1}{p}}.$$

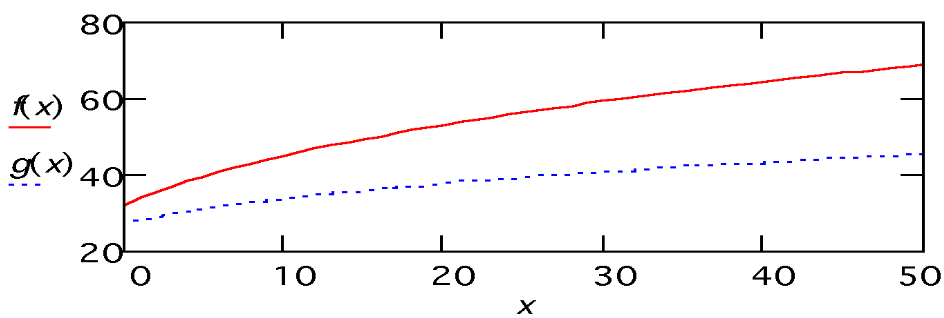

Figure 2. Salary vs. quality points: Ph.D. track ($f(x)$, solid line) and instructor track ($g(x)$, dotted line), with $p = 2$.

This family of functions has the advantage that we can adjust p to produce a salary function that is very steep initially and later very narrow, or one that is close to a line.

The Instructor Track

The instructor track cannot be defined as easily as the Ph.D. track, because we do not have any information about the expected value of the yearly raises. Arbitrarily, we set the raise for an instructor to be half of that of a Ph.D. with the same number of quality points. So the instructor salary $g(x)$ for x quality points is

$$g(x) = \tfrac{1}{2} f(x) + 11 = 16 \left[1 + (2^p - 1)\, \tfrac{x}{39}\right]^{\frac{1}{p}} + 11$$

(see **Figure 2**). We add the 11 to force the function to go through $(0, 27)$. The multiplier of one-half can be changed easily if necessary. The number of instructor quality points equals the number of years as an instructor, and an instructor moving into ABC from another institution can receive as many as 7 quality points. Note that between the initial salaries for instructor and for assistant professor, we initially have a \$5,000 gap, which grows with quality points. We assume that if an instructor becomes an assistant professor very quickly, then her salary should be very close to that of someone who has had the same length of career and has always been an assistant professor, whereas an instructor who takes very long to get her Ph.D. should not be very close in salary to an assistant professor who has worked for the same number of years. Therefore, we suggest that upon the completion of a Ph.D., we give the new assistant professor an immediate salary raise of \$5,000 (not enough to completely make up the difference) and then reevaluate her quality points on the Ph.D. scale. The function

$$q_{\text{Ph.D.}} = f^{-1}(g(q_{\text{Inst.}}) + 5)$$

calculates the position on the Ph.D. curve for an instructor receiving a promotion and receiving a \$5,000 benefit.

What Will It Cost?

For small p, the salary functions look too much like lines and salaries do not converge much; for high p, the cost of paying the faculty is high. We feel that $p = 2$ is a reasonable value that ABC might want to use.

Since we do not have the actual histories of the employees, we cannot calculate the actual sum of the salaries on the new system. Instead, we created randomized histories to get an idea of the monetary demands of the new system. This required a simulation satisfying a number of constraints, the details of which are in the **Appendix**.

Several iterations of the simulation on different random faculties estimate that a highly experienced faculty with approximately the same number of faculty members in each rank would cost about \$$10.0 \pm 0.1$ million to pay for a value of $p = 2$. This is not an average case estimate, but more of a "worst of the average" case estimate, because the simulated faculty have all been at ABC for their entire careers. Since most of the actual faculty have only been at ABC for a few years, they are making less money as a whole than the simulated faculties, so \$10.0 million is more than the college should expect to pay under our system.

The least the college could expect to pay out in a year would correspond to replacing the entire faculty with people from other schools with no experience in the rank at which they enter ABC. The faculty would come in with no credit, so each faculty member would make the baseline salary for her rank. The total minimum salary for the new system with $p = 2$ is \$8.7 million, which is roughly the same as the current payroll (\$8.8 million).

The 1993 national average salaries for faculty members, from instructor to full professor, were \$28,300, \$38,600, \$46,900, and \$68,700 for a private college (as ABC must be with a name like that). For $p = 2$, the baseline salaries at ABC for each category with promotions on time (in the same order) would be \$27,000, \$32,000, \$46,100, and \$56,800. These values are all less than the national average, but they are just the baseline values. Because of years in rank, many faculty members will be making more, creating an average closer to the national one.

The Cost of Living

We have determined the salary functions for faculty members with a Ph.D. and for instructors in terms of the number of quality points (we have not, however, determined a value for the parameter p). We need to incorporate annual increases in the cost of living into this model. The Consumer Price Index (CPI) measures well the year-to-year changes in the inflation rate [Parkin 1993, 628].

We propose to increase everybody's salary by a proportion equal to the increase in the cost of living every year. So, if after one year, the cost of living has increased 4%, then we just multiply $f(x)$ and $g(x)$ by 1.04.

Note that the \$5,000 salary increase that promoted instructors receive must also be adjusted for the cost of living.

When Funds Run Short

Up until this point, we have been assuming that ABC has enough money to pay everyone on the new salary schedule. Currently, ABC is paying out \$8.8 million; unfortunately, according to our model, ABC has been paying out too little. This raises the question of what to do if ABC cannot afford the new system: how to pay everyone if funding is low for a year.

We must make several assumptions about priorities.

- Give the usual raise and a bonus to faculty who are promoted.
- Give the usual raise (if there is enough money) to faculty members not getting a promotion.
- Spend any excess money (if there is any after giving raises) on replacing faculty members who leave (about 7% each year; see the **Appendix**).
- If there is not even enough to give everyone their raise, do not hire replacements and give each faculty member a fraction of a quality point instead of a whole one.
- If the budget cut is very serious, other measures, such as cutting salary for some and laying off others, may need to be taken. We cannot dictate firing people, but the method for reducing salary that would be best is to subtract the same number of quality points from each faculty member.

Since we do not have the actual salary histories of the employees at ABC, we created another random data set to experiment with. This data set, unlike the previous one, creates a population of faculty members in which some have spent their whole career at ABC and some have not (see the **Appendix**). Using our random data, we then created a MathCAD template to calculate the salary of each faculty member when there is not enough money. The data can easily be changed to calculate these amounts based on the actual numbers.

If there is not enough money to pay everyone the full expected raise (not taking into account cost-of-living increases), we first pay everyone who is being promoted and then divide the remaining money among the remaining people by assigning each faculty member the same fractional quality point of a yearly increase. If the administration decides that it is not proper to hire someone during the money shortage, that faculty member's history should just be left out of the data file.

We ran a simulation with a budget of $9.45 million for 1994–95. The amount needed to give everyone their raise and fulfill promotion raises was $9.50 million. Since there is not enough money, the template calculates the appropriate number of quality points to give out after raises. The result was to give each faculty member 0.735 points.

The Transition

We created the following procedure for moving current low-paid faculty members to the new scale without cutting anyone's salary. We must assume that there is enough money for the transition. The major points are the following:

- Anyone making more than they should receives a minimum raise (because the basic structure requires that each faculty member must receive a raise any year that money is available). We make this minimum raise \$100. We do not, however, include an increase for cost of living. In adding the minimum raise, we never allow an employee's actual salary to drop below her deserved salary.

- First, we pay everyone their previous year's salary, and people making more than they should are paid the minimum raise.

- We distribute the remaining money as raises to the people making less than they should, proportional to the amount they are below the salary that they deserve, including a cost-of-living increase.

- We assume that the college budgets \$10 million in 1994–95 dollars each year until the current salaries catch up to the new payment system.

In order to create and test a template, we again need actual values. We use the current salaries and generate possible histories to go along with them. We calculate how long it would take for each overpaid faculty member to receive only as much money as she deserves. (We wanted to calculate how long it would take each underpaid faculty member to catch up, but we could not fix an error in our program.) Five runs of the program were combined. The number of people with a certain "catch-up" time appears to be an exponential distribution with a mean of 3.55. A quick computation shows that it should take 4.6 years for 95% of the overpaid people to catch up to their salaries, although some extremely overpaid people take much longer.

Strengths

- Our model is simple. Using a single salary curve for all doctorates is much simpler than using a separate curve for each rank. The concept of quality points is a very convenient and simplifying assumption and is one that anybody could use. Furthermore, the cost-of-living adjustment that we make is simple.

- The model rewards those instructors who advance in rank quickly, as we believe is appropriate, without making the penalty for late promotions too drastic—late promotees still get a raise, just not as much.

- The template that we used with simulated data can easily be used with actual data.

- Our model is flexible. The parameters to which we assigned values (such as $p = 2$) can easily be changed,

- Our model conforms to all of the principles and circumstances of the problem statement.
- The model seems stable and consistent. Different random sets of data do not produce large changes in the results, and our results agree with values for average faculty salary in the U.S.
- If so desired, the concept of quality points could be extended to situations beyond the scale of this model. For example, if the administration wanted to give a faculty a special raise/cut for extraordinary performance, or to make especially good offers to desirable faculty from other schools, it could adjust a faculty member's number of quality points.

Weaknesses

- The parameters used in our model are arbitrary, since we are missing data that would be useful in identifying them.
- We have no basis for setting the instructor slope at half of the Ph.D. slope.
- It is very difficult to perform sensitivity analysis on our model. We have no way of estimating how much our simulated data, upon which we should not rely too heavily, differs from the real data.
- We assumed that the number of instructors in each rank remains reasonably constant. It might not, which would alter our results.
- Implementing our model assumes adequate funding. If money is a problem, we could decrease p, but this may not totally remedy the situation.

Appendix

Creating Random Faculty

Since it does not seem to take very long to be promoted from instructor to assistant professor, the number of years as instructor is generated by taking a random number from an exponential distribution with a mean of 3. (This means that more than half of all faculty members spend 3 years or less as instructors, but a few spend much longer.) The number of years as an assistant professor is similarly taken to be 7 plus a random number from an exponential distribution with a mean of 3. The 7 must be added because it takes a minimum of 7 years to be promoted to associate professor. Similarly, the number of years as an associate professor is generated from the same distribution. The number of years as a full professor is generated by 1 plus a random number from an exponential distribution with a mean of 5.

Not very much information is given in the problem to support this model. However, it does have some basis. For assistant professors, the average number of years at ABC is roughly exponentially distributed, with a mean of 9.4 (see **Figure 3**). This suggests that exponentially distributed random numbers are decent estimates of this number. Most of the simulated times are between 7 and 10 years, which roughly agrees with the minimum and average time that faculty members spend as assistant professors. The time spent as an associate professor is roughly the same, so we use the same random variable for years as an associate professor.

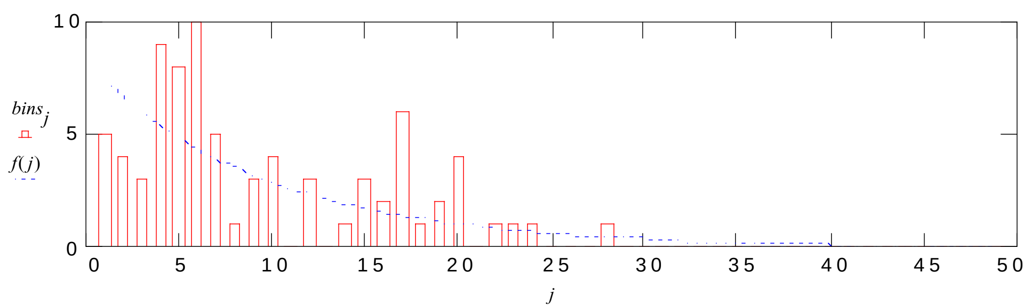

Figure 3. Number of assistant professors vs. years at ABC. The dotted curve is the function used in the simulations.

The mean of 5 for years-in-rank as a full professor comes from a total career of about twenty-five years: $7 + 3\text{-ish} + 7 + 3\text{-ish} + 5\text{-ish} = 25\text{-ish}$.

The simulation creates the same distribution of positions as the given salary list; we assume that the makeup of the faculty does not change.

The second simulation creates a random number of years that each instructor has been at ABC, based on the given data and using an exponential distribution with a mean of 13.3 (the mean of the data) (see **Figure 4**).

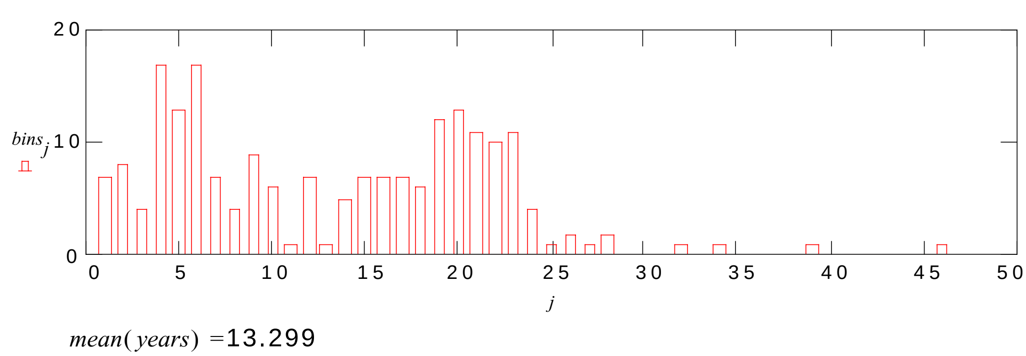

Figure 4. Number of faculty members vs. years at ABC.

The number of faculty members who leave after t years is the difference between the numbers there t years ago and $t-1$ years ago, which is approximately equal to the derivative of the density function at t. By explicit calculation, or from properties of the exponential distribution, the duration of stay is exponential with the same mean.

How Many People Are Replaced Each Year?

We assume that the same fraction of the faculty is replaced each year and that a faculty member stays at ABC an average of 13.3 years.

For a continuous model, we use an exponential distribution with probability density function

$$f(t) = \frac{e^{-t/\lambda}}{\lambda}, \qquad t > 0,$$

which has mean λ. The area lying to the right of λ is e^{-1}. We set $\lambda = 13.3$, so that after 13.3 years, the fraction of faculty remaining is $e^{-1} \approx .368$. For the exponential "decay" in the number of faculty, let μ be the half-life and r the decay constant; r is the fraction of faculty who remain from one year to the next. Let N_0 be the initial number of faculty in a cohort, and let N_μ be the number of those faculty remaining after one "half-life." We have

$$N_\mu = N_0 r^\mu = \frac{1}{e} N_0, \qquad r = e^{-\mu}.$$

For $\mu = 13.3$, we have $r = 0.928$, so roughly 93% of the faculty remain from one year to the next, and on average $.072 \times 204 = 15$ faculty members are replaced every year. Depending on the ranks involved, 15 faculty members cost anywhere from $405,000 (inexperienced instructors) to $900,000 (experienced full professors). If the college runs short of money, some money could be saved by not replacing these faculty members.

For a discrete model, we use a geometric distribution with probability mass function

$$f(n) = p(1-p)^{n-1}, \qquad n = 0, 1, 2, \ldots/,$$

where p is the fraction of faculty who leave each year. The mean of this distribution is $1/p$. We set $1/p = 13.3$, getting $p = .075$, in good agreement with the continuous model.

References

U.S. Department of Labor, Bureau of the Census. 1994. *Statistical Abstract of the United States: 1994.* 114th ed. Washington, DC: U.S. Government Printing Office.

Parkin, Michael. 1993. *Economics.* New York: Addison-Wesley.

Long-Term and Transient Pay Scale for College Faculty

Christena Byerley
Christina Phillips
Cliff Sodergren
Southeast Missouri State University
Cape Girardeau, MO 63701

Advisor: Robert W. Sheets

Introduction

Our assignment was to design a salary system for the college's faculty without cost of living increases, then incorporate cost-of-living increases, and finally design a transition model to move all current faculty towards our model. We decided that one of the most important issues to consider was the quality of education at the college. Hence, we established a model to push faculty to work toward promotion or else their raise possibilities are minimized. This should raise faculty incentive and increase publications, research, and instruction, which will raise the overall quality of education at Aluacha Balaclava.

The equation for our model without regard to rank is

$$P_x = P_1 + m\left(1 - e^{-k(x-1)}\right).$$

This model satisfies the given criteria and also leaves room for variations.

To incorporate cost-of-living increases into our model, we use the Consumer Price Index, a national measure of inflation.

For the transition, we take a faculty member's current salary and years in current rank and determine a salary curve that starts at that point and follows our model in the best possible way. We tested the transition model as well as possible with insufficient data by modeling salaries for fictional faculty. Results show that the transition model moves all faculty from their current salaries toward our desired system.

Assumptions

- In the data, number of years of service means years of teaching at that institution, not overall career and not teaching at that rank. However,

Table 1.
List of variables.

Variable	Meaning
CPI_{x-1}	cost of living factor from the previous year
F_{required}	funds needed to give all faculty their full raise
F_x	available funds for raises in the xth year
k	decay constant
m	multiplicative factor
P_1	entry-level salary at a given rank
P_x	salary in the xth year
P'	salary the year before the transition plan begins
x	years in rank
x'	years at a given rank when the transition plan begins

many equations throughout our model make use of the variable x', which is the number of years that a faculty member has been at his or her current rank. This information was not given to us, so it will be the responsibility of the Provost to obtain it.

- Each rank should have a minimum base salary and a maximum base salary for newly hired faculty. We also presumed that some faculty with Ph.D.s could be hired above the assistant professor level, based on experience.
- The minimum time for an instructor to complete the Ph.D. degree and be promoted is approximately two years, and there is a four-year minimum of service at the assistant professor level before promotion to associate professor.
- A "normal raise," in reference to our requirement of promotion benefit being equivalent to seven years of raises, could be determined approximately by the average raise for the first twenty years, $(P_{21} - P_1)/20$.
- A decaying exponential curve gives a good basis for the model.

Motivation for the Model

We choose a decaying exponential model for a number of reasons.

- It allows for considerable raises at the beginning of a rank, but as time passes, the raises decrease. We think that this is important because it gives faculty the incentive to work toward promotion by contributing to the college through research, publications, or excellence in teaching. Without being promoted after a certain amount of time, their salary will top out at a value that reflects their rank. Not only would promotion offer a raise, but it would also result in higher future raises.

- By decreasing the multiplicative factor while at the same time increasing the starting salary, we establish a model that causes salaries in the same rank to grow closer over time, as required.

The motivation behind choosing the decay constant is the number of years of experience that we want the faculty to have before the exponential term begins to level off and salary raises decrease dramatically from one year to the next. For each rank, we decide on the the number of years for which a faculty member at that rank would receive 50% of his or her total raises at that rank and solve for the decay constant k:

$$0.5 = 1 - e^{-k(x-1)}.$$

We want an instructor to receive 50% of his or her raises after seven years in an attempt to motivate the instructor to work for promotion. If a promotion does not occur, raises decrease more rapidly.

For assistant and associate professors, we want this time to be longer, because they are less likely to be promoted. It takes a couple of years longer than an instructor to receive about 50% of their raises—by their tenth year of service at that rank.

For a full professor, because there is no possibility for promotion, this amount of time should be even longer. A full professor will receive about 50% of his or her raises by his or her twelfth year in rank.

In choosing the high and low salaries for the first year, we took into consideration a number of things. We averaged the current salaries of faculty, getting \$31,919, \$35,908, \$44,286, and \$54,228 for the ranks in ascending order. We set our base salaries taking these averages into consideration so that new faculty would be paid reasonably. We also took into consideration the statistics from the May 1994 *Occupational Outlook Handbook* [U.S. Department of Labor, Bureau of Statistics, 1994], which gives national average salaries as \$27,700, \$36,800, \$44,100, and \$59,500.

Another source that we took into account was the the *Statistical Abstract of the United States: 1994* [U.S. Dept. of Labor, Bureau of the Census, 1994], which lists average beginning salaries offered to candidates according to their degree level and their field of concentration. By taking all of these statistics into account, we attempt to establish a fair window for entry-level salaries for all ranks and a top salary at each rank below full professor (see **Table 2**).

The reason that we have a different model for full professors is that we think that there should not be a maximum salary for them because they have no possibility for promotion; if we kept the same type of model, there would be a ceiling. So, although the salaries of full professors at the minimum and the maximum bases for full professors do not tend toward each other together as quickly as with the other ranks, they do get slightly closer over time.

Table 2.
Entry-level base salaries and top salaries.

Rank	Minimum base	Maximum base	Top salary
Instructor	27,000	37,000	42,000
Assistant Professor	32,000	47,000	52,000
Associate Professor	37,000	52,000	62,000
Full Professor	47,000	62,000	—

To determine the raise for promotion from one rank to another, we looked at the time to promotion and the maximum and minimum entry levels at the next rank. If a faculty member is promoted on time, he or she should get a raise equivalent to seven years' worth of normal raises (according to the problem statement). We chose $3,500 as a raise for a promotion achieved on time. This is a compromise between the calculations we made for normal raises over seven years, which were between $2,000 and $5,000, depending on the entry-level salary for the next rank. We chose $,3500 because that amount keeps faculty receiving promotions within the entry-level salary range of the next rank and is in between the high and low normal raises over seven years.

Another issue is how to allocate available funds. The first thing to consider is how much would be required to give everyone the raise that coincides with his or her salary-scale curve. If the required amount is available, then each faculty member is given his or her expected raise. If the required amount is not available, then each faculty member is given a proportion of his or her raise. If an excess is available, and if all faculty members are where they should be on the curve according to entry salary and the number of years at that level, then the excess is held over until the next year.

The Model

Once the starting salary has been established, our model gives a salary curve for the faculty member to follow, as shown in **Table 3**. See **Figure 1** for maximum and minimum salary curves in each rank and how they converge. We also present a salary schedule for minimum and maximum salary according to year and rank. [EDITOR'S NOTE: We omit the detailed salary schedule.]

What If There Isn't Enough Money?

When determining the amount of the raise, it is necessary to take into consideration the available funds for the year vs. the amount that it would

Table 3.
Salary curves.

Rank	Curve without inflation
Instructor:	$P_1 + (42{,}000 - P_1)\left(1 - e^{-0.10(x-1)}\right)$
Assistant Professor:	$P_1 + (52{,}000 - P_1)\left(1 - e^{-0.08(x-1)}\right)$
Associate Professor:	$P_1 + (62{,}000 - P_1)\left(1 - e^{-0.08(x-1)}\right)$
Full Professor:	$P_1 + \left[10{,}000 \times \frac{62{,}000 - P_1}{15{,}000}\right]\left(1 - e^{-0.07(x-1)}\right)$

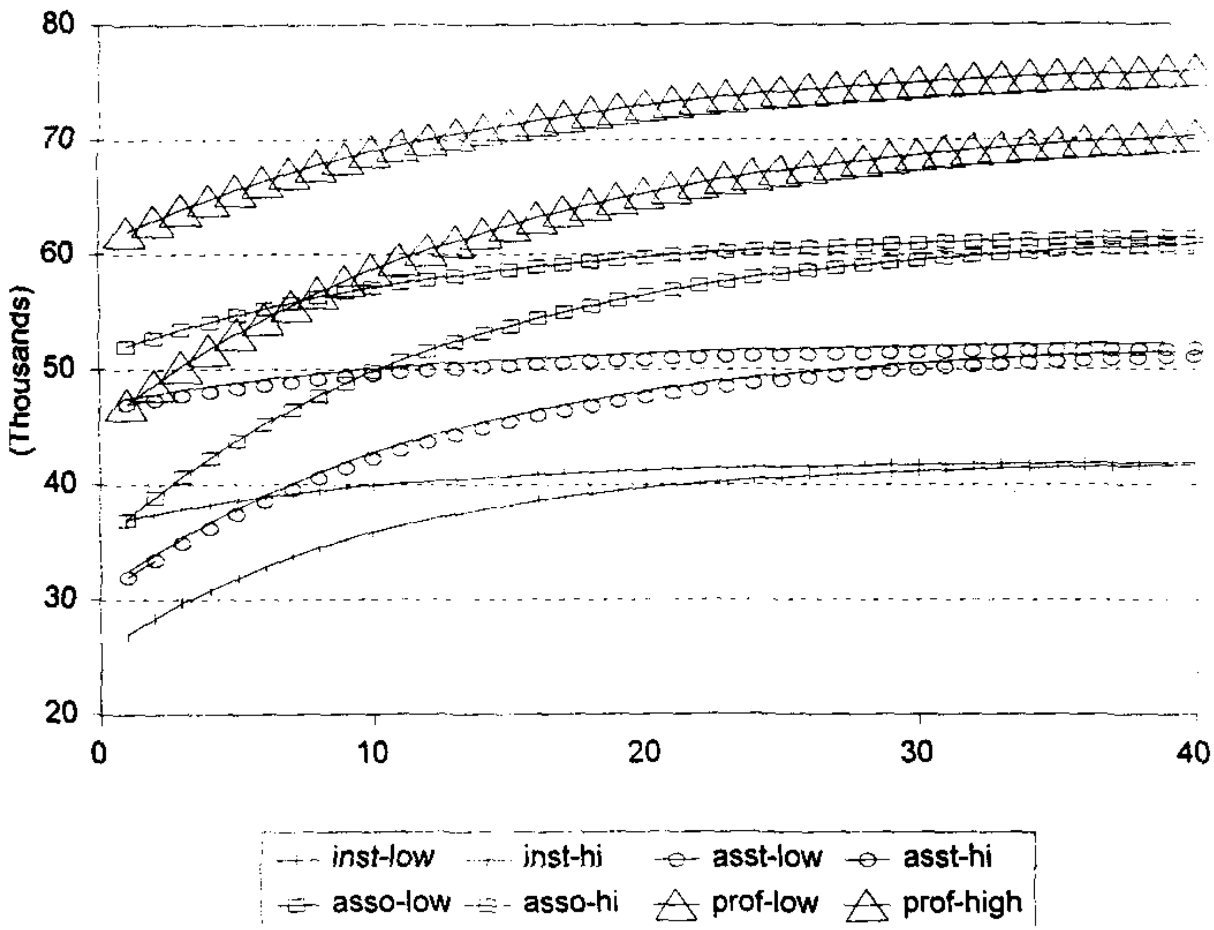

Figure 1. Maximum and minimum salary curves for each rank.

cost to give everyone the raise according to his or her curve. If the amount required exceeds funds available, then the raise for each faculty member will be determined in the following manner.

For each faculty member, divide the expected raise by the total of expected raises and multiply this fraction by the available funds. An exception: If the available amount is less than 10% of the required amount, then no raises are given and the funds are held over until the next year. (This ensures that only substantial raises will be given, i.e., no one will receive a $0.40 raise.) We multiply each of the formulas in **Table 3** by

$$\frac{(P_x - P_{x-1})F_x}{F_{\text{required}}}.$$

If adequate funds are not available in a certain year, then everyone will be below their salary curve. The following year, the way to determine who gets what percentage of the available funds is again to use the formulas to proportionate.

A faculty member who receives a promotion will jump to the equation for the next rank. The entry-level salary at the new rank will be previous salary plus a raise of:

- $3,500, if promoted in the least amount of time (i.e., two years for instructor to assistant professor, four years for assistant professor to associate professor, and seven years for associate professor to full professor);
- $2,500, if promoted within five years of the minimum number of years;
- $1,000, if promoted any time later.

Cost of Living

The model is the same as the previous model with salaries multiplied by a factor that takes cost of living into consideration. We chose the Consumer Price Index (CPI) because it is gives an all-encompassing measure of the percentage increase of goods and services in the U.S. We use the CPI from the previous year to determine the rise in cost of living for the current year. Many other indices can be inaccurate because they are based on projections of what is expected to happen in the future. We multiply each of the formulas in **Table 3** by $(1 + \text{CPI}_{x-1})$.

Transition

The transition model takes a current faculty member and finds a salary curve that fits our model while considering current salary and years in their rank. For each rank, current salary and years in rank will fall

- above our maximum salary curve for that rank,
- below the minimum curve, or
- between the curves.

We consider each possibility for each rank.

Above the Maximum Salary Curve

We cannot fit faculty who currently are above the maximum salary curve for their rank into our salary range, because we cannot cut their salary. Because we have to allow them to receive raises but do not want them to receive very large ones, we let them increase at only the same rate as the maximum salary curve for their rank. To find their salary curve, we need to project backwards to determine a corresponding P_1. We use the formulas in **Table 3**, substituting P' (salary the year before the transition begins) for P_x and x' (number of years at a given rank when the transition begins) for x.

Below the Minimum Salary Curve

For those who fall below the minimum salary curve for their rank, we increase their salaries so that over a five-year period they move into the salary range for their years at that rank. Our model for this transition calculates what the faculty member's salary should be in five years to fit the minimum salary curve, then divides the difference into five equal increments, for an equal raise each year.

We have

$$P_x = P' + (x - x')\,\frac{P_1 + m\left(1 - e^{-k(x'+4)}\right) - P'}{5}.$$

This equation is used for only five years after the implementation of our model. At the end of this five-year period, all faculty originally falling below the minimum salary curves will be caught up to these minimums, and their salaries will follow the minimum curve from then on.

Between the Two Curves

If a faculty member's current year in rank and salary fall between our maximum and minimum salary curves, we implement our original model. For each rank, we can find the curve that fits our model and passes through the faculty member's current point. To do this, we just project backwards to find the value of P_1 for such a curve, then substitute into the appropriate formula.

Testing the Model

As a preliminary test of our model, we developed tables of salaries for the minimum and maximum entry salaries for each rank based on the number of years at that rank. These tables were the basis for **Figure 1**.

Since we do not know years in rank for actual faculty members in the given data, we could not examine how their salaries would change under the transition model (which would be ideal). Instead, we generated several random contrived faculty members in each rank who are above the maximum salary curve, below the minimum salary curve, or between the two curves. We modeled the curves for these fictitious faculty and observed good transition from current salaries to our model.

For further analysis, we could also test if the model financially coincides with the capabilities of the college. If not, then the model could be scaled down by lowering the entry salaries or by increasing the decay constant.

Strengths and Weaknesses

The most successful way to improve an institution's prestige and quality of their degrees is to improve the quality of the faculty at the institution. This is one of the most important strengths in our model. We have attempted to improve the quality of the faculty at Aluacha Balaclava College by creating a large window for entry salaries, pushing faculty toward promotion, and setting salaries comparable to the national averages.

Our window of entry has established a very wide range of salaries for prospective faculty. This gives Aluacha Balaclava College the opportunity to hire the best available faculty. In turn, these new faculty will boost the quality of teaching and raise the overall rating of the college. At the same time, our minimum entry salary helps keep the faculty salary budget down by not overpaying existing faculty who meet only average criteria.

This window of entry salary may also be a weakness, because it will not allow the college to bring in more-prestigious instructors who expect a higher salary than the scale allows. We compensate with an "extraordinary circumstances" clause. When a faculty member does enter at a salary above what our maximum curve allows for, he or she will follow a salary scale equivalent to that of an existing faculty member who is making more than our curve allows during transition.

The strength of the transition phase is that it brings current faculty up to a fair salary in a relatively short amount of time (five years). The shortness of the transition period should help alleviate any animosity among current faculty that has been caused by their salaries and resulted in the departure of the previous Provost. The shortness, however, may also produce a financial crunch on the college, which is a weakness. A longer transition period would lengthen the transition period and lessen the financial burdens in the first five years.

Our model does not take economic deflation into consideration. Although deflation rarely occurs in the U.S., if it does, we do not want to give the faculty a salary cut. Instead, we use an inflation multiplier of 1 instead of $(1 + \mathrm{CPI}_{x-1})$. The CPI factor also poses another weakness because it is a national and not local average of cost-of-living increase. If there is an index available that estimates the cost-of-living increase in the surrounding area of Aluacha Balaclava College, then perhaps it would be a more accurate estimation of inflation. Another weakness is that by using the previous year for reference, there is a lag in cost-of-living adjustments.

References

U.S. Department of Labor, Bureau of the Census. 1994. *Statistical Abstract of the United States: 1994.* 114th ed. Washington, DC: U.S. Government Printing Office.

U.S. Department of Labor, Bureau of Statistics. 1994. *Occupational Outlook Handbook*. May, 1994. Washington, DC: U.S. Government Printing Office.

How to Keep Your Job as Provost

Liam Forbes
Marcus Martin
Michael Schmahl
University of Alaska Fairbanks
Fairbanks, AK 99775
lforbes@arsc.edu
{fsmgm, fxmgs}@aurora.alaska.edu

Advisor: John P. Lambert

Introduction

Our salary system proposal quickly brings faculty salaries into line without cutting any salaries. In accordance with the principles given us, we have created a function that assigns each faculty member an ideal salary that they will be paid if the budget is large enough.

Bringing salaries up to scale is easy if there is enough money. For the quite likely case of not enough money, we present an algorithm that brings as many faculty as possible up to the curve as rapidly as possible, within the limits of the available budget. We call this algorithm for figuring annual raises the *scaling method*.

After giving everyone a minimum nominal raise, the scaling method finds a scaled-down version of the ideal salary function and brings all faculty below this curve up to the curve, using the rest of the money available for raises in the process.

Cost of living is applied each year before the annual raises. The first method applies the cost-of-living increase to the ideal salary curve and allows the scaling method to naturally bring faculty salaries up to this curve. The second method applies the cost-of-living increase simultaneously to the ideal curve and the current faculty salaries. We recommend the first method.

The transition phase will be very easy. There is no difference in our algorithm over the long term and for the short-term transition period. Our method instantly brings those faculty who are drastically behind scale up to a salary similar to their peers and continues to collapse the discrepancies between faculty at the same rank and experience.

Table 1.
Glossary of symbols.

Symbol	Meaning
a, b	parameters in the logistic equation
c	scaling factor for the ideal salary curve
e	base of the natural logarithm, approximately 2.718.
COL	cost-of-living increase, expressed as a percentage
D_i	salary deficit of faculty member i
i	a given faculty member
$f(x)$	the general logistic function
$I(x)$	the ideal salary function
$I'(x)$	a scaled-down version of $I(x)$, with parameter k' instead of k
k	a parameter in the logistic equation; the absolute max value for the curve
k'	the scaled-down constant for the scaled-down ideal salary curve
$k_{\text{hi}}, k_{\text{low}}$	the high and low endpoints in a binary search for an optimal k'
M	the amount of money available for raises, after nominal raises
R	the minimum annual raise
S_i, S_i^+	the current and next year's salaries for faculty member i
T_i^+	temporary guess for next year's salary for faculty member i
x, x_i	the faculty indexing function; the index of a given faculty member i

Assumptions and Justifications

- *No individual faculty member's salary may ever be decreased.* This is a clarification and expansion of the principle given that existing faculty salaries cannot decrease during the transition. If this is true of the transition period, then by natural extension it should be true in the future.

- *Not all the money available for raises each year necessarily needs to be distributed, if all faculty are at their ideal salary.* It is unlikely that an administration would want to pay faculty members more than the salary system suggests that they deserve.

- *Salary rates for faculty are not based on merit but solely on longevity and rank.* There was no mention of merit in the instructions from the Provost. In our salary system, all faculty with the same longevity and the same rank have the same ideal salary.

Building the Salary System

The Ideal Salary Function

The first step in designing a salary system is deciding what salary each faculty member would receive if there were always as much money in the budget as needed. We develop a curve that satisfies some basic criteria.

Curve Evaluation Criteria

1. A promotion is worth seven years of salary increases (Principle 3).
2. New instructors are paid \$27,000 (Principle 1).
3. New assistant professors are paid \$32,000 (Principle 1).
4. Full professors with 25 or more years in service should make approximately \$64,000 (Principle 4).
5. Full professors with more than 25 years of service should also make approximately \$64,000 but more than a full professor with only 25 years of service (Principles 4, 5).
6. Salary increases should diminish over time (Principle 4).

We created an indexing system whereby faculty members are given a number equal to their years in service plus seven times their rank. The rank values are 0 for instructors, 1 for assistant professors, 2 for associate professors, and 3 for full professors. Thus the index x equals seven times the current rank number plus the number of years in service. A single salary system is used for all of the ranks, and a promotion is equivalent to seven years of service. This index is the input for the salary functions that will be developed in this paper.

Equations of the Curve Evaluation Criteria

- $x = \text{years } + 7 \times \text{ rank}.$
- Ideal salary is a function $I(x)$, with

  $I(0) = \$27{,}000.$

  $I(7) = \$32{,}000.$

  $I(46) = \$64{,}000 \pm 5\%.$

  $I(71) = \$64{,}000 \pm 10\%; \quad I(71) > I(46).$ (The number 71 is arbitrary; it represents a professor with fifty years in service.)

  $d^2I/dx^2 < 0$ for large x.

We examined polynomial, square-root, cube-root, exponential, power, and logistic curves. We choose the logistic curve as the ideal salary curve because it meets all of the criteria.

$$\textbf{Logistic function:} \qquad f(x) = \frac{k}{1 + ae^{bx}}.$$

The logistic function meets Criterion 6 for $a, k > 0$ and $b < 0$; it is bounded above by k as $x \to \infty$, so it serves as a cap on the ideal salary function. If k is known, then a and b can be solved for analytically using Criteria 2 and 3.

We explored a range of possible k values, calculating $I(46)$ and $I(71)$ for each k. These values are used to extrapolate the appropriate error bounds at both indices. The two ranges did not quite overlap, so a value for k was chosen between the two ranges. The parameters a and b were solved for at this k, yielding

$$k = 83{,}000, \qquad a = 56/27 = 2.07, \qquad b = -0.0376.$$

This procedure generated the ideal salary function:

$$\textbf{Ideal Salary Function:} \qquad I(x) = 83{,}000\left(1 + \frac{56e^{-0.0376x}}{27}\right) - 1.$$

Figure 1 shows the ideal salary curve and the faculty salaries as a function of index. Notice that a few points are drastically above the curve, while most are a few thousand dollars under the curve. Those faculty who are above the curve will henceforth be called *Red-Circle* faculty [Henrici 1980].

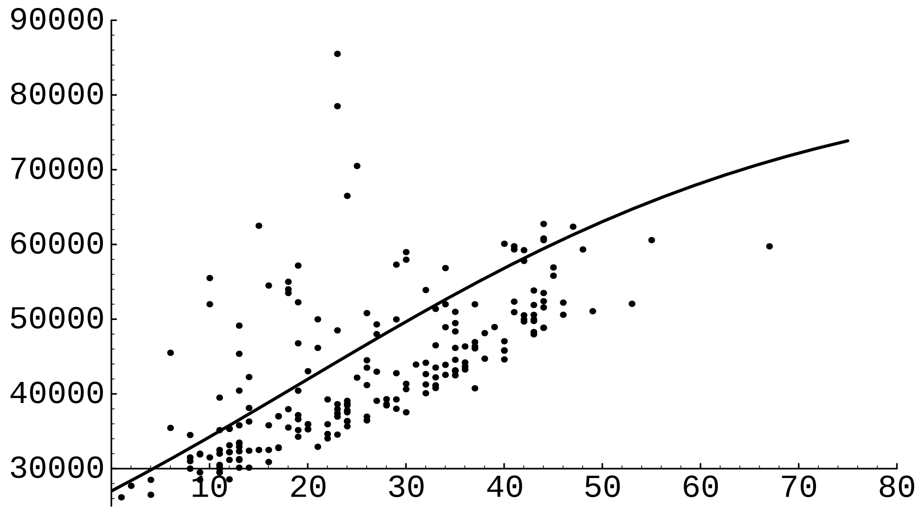

Figure 1. Ideal salary curve and current faculty salaries.

The Annual Raise

Apportioning the money available for raises is a matter of deciding how to move people toward the ideal salary curve. All faculty must get a raise in a year where money is available. The principles do not state how much each faculty member must get, so we leave this up to the Provost. Whatever minimum raise is decided upon we will call R.

After minimum raises, the difference between where the salaries are and where the ideal salary curve would have them is called the *salary deficit*:

$$D_i = I(x_i) - (S_i + R).$$

The money left over after minimum raises will be called M. In both methods, if there is not enough money to give everybody the minimum raise, then the money available for raises is divided evenly among the faculty.

The Deficit-Proportion Method

This method gives everyone a mandatory raise (R) and allocates the remaining money (M) among the faculty in proportion to their deficit. For each faculty member i, next year's salary S_i^+ will increase from the current year's salary S_i according to

$$S_i^+ = S_i + R + D_i \, \frac{M}{\sum D_i},$$

where $\sum D_i$ is the sum of the deficits of the entire faculty. If a faculty member has a surplus (a negative deficit), it is treated as zero for this sum.

The Scaling Method

A second scheme is to attempt to bring as many people as possible up to the same fraction of the ideal salary curve. This is done by scaling the ideal salary curve down so that the amount needed to raise everyone whose salary is below the curve up to the scaled-down curve equals the amount available for raises. The model again starts with the mandatory raise and then uses a binary search to find this scaled curve. The scaled function is called $I'(x) = cI(x)$. Under this method, we have

$$S_i^+ = \max\{S_i + R, cI(x)\} \qquad \text{for some } c.$$

The value of c, which is a constant between 0 and 1, must be found heuristically. Because of this fact and roundoff error, we cannot necessarily find a scalar c that will allow the total ideal salaries to precisely match the actual salaries. The largest possible roundoff error for N faculty, using whole dollars, is the case where all faculty would otherwise be given 50 cents over a whole dollar amount. The amount of roundoff error in that extreme case gives a tolerance value of $\$N/2$.

Notice that $cI(x) = ck/(1+ae^{bx})$. Instead of looking for a value for c, our algorithm looks for a value for $k' = ck$. The upper bound of k' is the actual value of k. The lower bound is $I(0)$, because we can be sure that all salaries are above that value. Solving gives a new function $I'(x) = k'/(1+ae^{bx})$.

Procedure for the Scaling Method:

1. If there is no money available for raises, then finished.

2. If there is money, but not enough to give the full mandatory raise, then divide the money equally among all faculty, and we are finished.

3. If there is excess after the mandatory raise, then find a k' that will put the total salaries within $\$N/2$ of the given salary budget, using the following binary search:

   (a) $khi \leftarrow k$.

   (b) $k_{\text{low}} \leftarrow I(0)$.

   (c) $k' \leftarrow (k_{\text{hi}} + k_{\text{low}})/2$.

   (d) For each faculty member, assign a temporary salary $T_i^+ = \max\{S_i + R, I'(x)\}$.

   (e) If the sum of the T_i^+s is within $\$N/2$ of the salary budget, then end procedure.

   (f) Otherwise, if $\sum T_i^+ > \text{Budget}$, then $k_{\text{hi}} \leftarrow k$ and return to Step d.

   (g) $k_{\text{low}} \leftarrow k$; return to Step d.

4. $S_i^+ \leftarrow T_i^+$.

Comparison

The deficit-proportion method gives faculty whose salaries are below the ideal salary curve a raise proportional to the amount below the curve. This is fair in that it gives a larger raise to people who are further away from their deserved salary. It is unfair in that when two faculty at the same index start at different salaries, the person who starts with the higher salary will continue to have the higher salary until they both reach the ideal salary curve. **Figure 2** shows this method applied to the original data set. Note that although all of faculty salaries have risen, those that were the furthest below the curve are still behind their peers.

Like the deficit-proportion method, the scaling method is fair in that it gives a larger raise to people further away from their deserved salary. But it does not have the drawback that faculty with lower salaries will be perpetually lower than their peers. The scaling method equalizes the salaries of faculty with the same index much more rapidly than the deficit-proportion method (see **Figure 3**). For these reasons, the scaling method is superior.

Cost-of-Living Increases

According to Henrici [1980, 162], "Some employers, when they adjust the salary curve upward, give everyone a 'general increase' at the same time. Others prefer to carry this adjustment into individual salaries throughout the year." Therefore, there are two ways of granting cost-of-living increases. Both involve moving the salary curve upward. One involves moving only

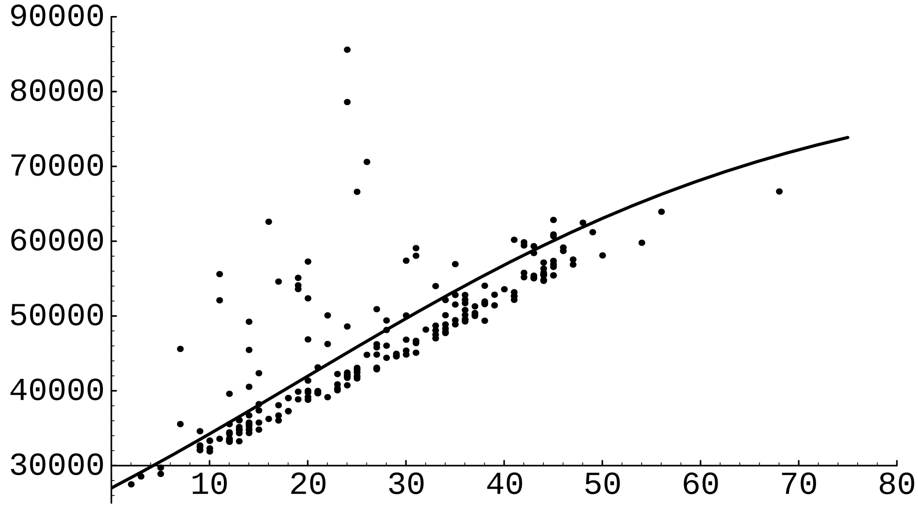

Figure 2. Deficit-proportional method applied to the faculty data. Budget: \$9.5 million.

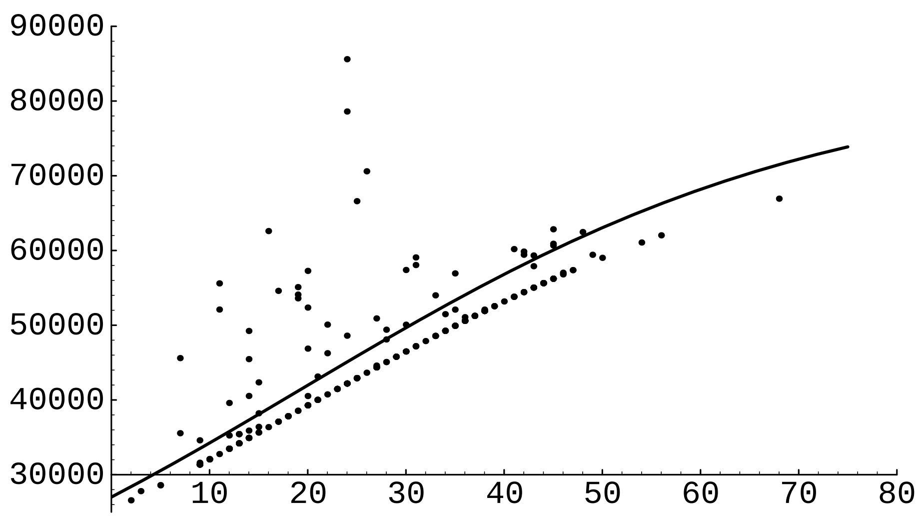

Figure 3. Scaling method applied to the faculty data. Budget: \$9.5 million.

the salary curve upward (changing k), and the other involves moving the ideal salary curve upward and simultaneously raising current salaries by the same factor. We present both options and discuss their advantages.

COLA Method I: Raising the Curve

For a given cost-of-living increase COL, set $k \leftarrow k(1+\mathrm{COL})$. This has the effect of raising the ideal salary curve by a factor of COL (see **Figure 4**). All other parameters are unchanged. The corresponding raises occur naturally at the end-of-year salary computations, which are done with the scaling algorithm.

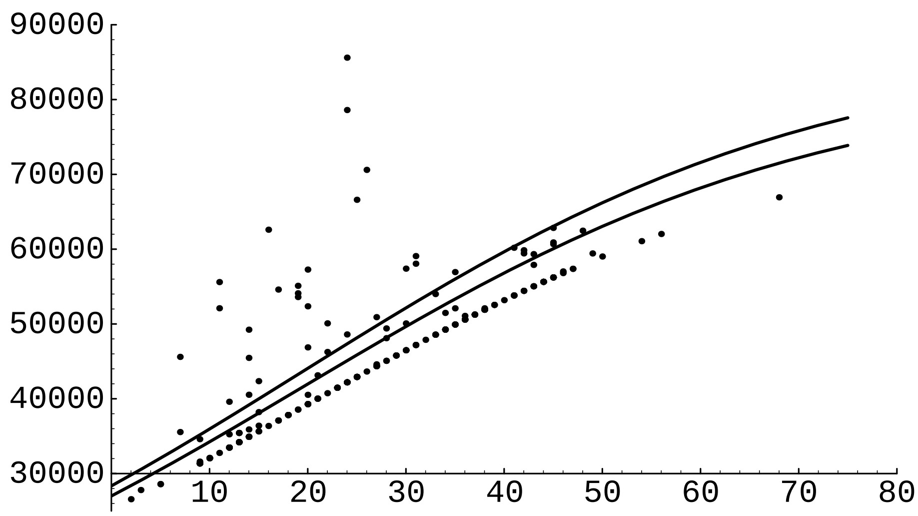

Figure 4. Scaling method with cost-of-living applied to the salary curve. Cost of living: +5%. New budget: $9.5 million.

COLA Method II: Raising Salaries, Too

For a given cost-of-living increase COL, set $k \leftarrow k(1 + \mathrm{COL})$. Then, if there is enough money, raise all salaries by multiplying them by $(1 + \mathrm{COL})$ and then applying the scaling algorithm. Otherwise, raise all salaries by the maximum percentage possible:

$$S_i^+ = S_i \, \frac{\sum S + M}{\sum S}.$$

Comparison

By raising the curve but not immediately raising salaries, Method I brings the Red Circles (those faculty being paid above their ideal) into line more quickly. This method does not immediately increase the cost of salary, which is good because we are not sure there is money available when cost-of-living is factored in. If there is money, Method I will bring more people up more quickly than Method II will, since Method II depletes the discretionary budget. Method II increases the salary of all faculty, including the Red-Circle faculty. Method I does not, and giving a "cost-of-living" increase selectively could upset these Red Circles, even though they are already making an above-scale salary.

Once the Red Circles have left or retired, all faculty will be at or below the ideal curve. This neutralizes some of the advantages of Method I. In the ideal case, where all faculty are being paid on the ideal curve and there is enough money to bring them to the new ideal curve, the two methods are equivalent, since the result will always be to bring them up to the curve.

Transition

The transition period for the deficit-proportion and the COLA–II models is poor unless there is a large influx of money. The scaling and COLA–I models are better adapted to lower salary budgets. The length of the transition period is dependent upon how long the Red-Circle employees remain at the college. This is because the people below the curve can make the transition quickly if there is enough money in the budget, but there is no quick way to bring those people who are far above the curve back down.

The procedure during the transition period is the same as the procedure after the transition period. It is not even necessary to make such a distinction. This is one of the major strengths of our methods.

Testing and Analysis

Creating a Population Model

To test our methods, we created a program for each salary scheme. These programs model the movement of faculty through the college and include changing faculty indices based on accumulated years teaching, earning promotions when eligible, retiring eligible faculty, and hiring new faculty (to eliminate any gaps in numbers of faculty at a given rank).

Program Assumptions

- *The number of faculty in each rank should be constant.* We have a decidedly optimistic view of the development of the college. The computer programs are based on the college remaining at least at the current level of faculty employment. Losing faculty would have an adverse affect on the college's income potential and the number of classes offered to students. Expansion of the college can easily be incorporated into the model by allowing more faculty to be hired at each rank. Also, by requiring that the number of faculty at each rank remain constant, we can study the population and salary shifts as individuals move through the system.

- *After 25 years, any faculty are eligible to retire, regardless of rank.* There is no reason to prevent instructors, assistant professors, and associate professors from retiring once they have achieved longevity. Anyone who reaches twenty-five years of teaching is eligible for the college's retirement plan (whatever that may be). This assumption comes from observing the common practices of universities and businesses across the nation. Retirement is a separate issue from an individual deciding to move on to other job possibilities, which also must be dealt with in the computer program. Retirement is handled as part of promotion considerations.

- *There is no forced retirement.* In order to allow maximum flexibility in the model, we do not impose forced retirement. This allows the college to add and modify its own policies to the model with a minimum of effort. Also, forcing retirement adds nothing of benefit to our model.

- *Promotions are generated on a probabilistic basis.* The problem specifically states that promotions are decided by the Provost herself. In order to model how promotions will affect an individual's salary, we needed to generate a probability that an individual will be promoted if eligible. This way we could model a random determination of promotions or tightly control how often promotions are earned. The college could also use this program as a tool to review a wide range of promotion schemes by altering certain parameters in the program.

- *An individual eligible for promotion who is passed over will leave the college.* This is a weakness of the program, but we used it to simplify our model of human behavior. It is hard to say how people will react to being passed over for promotion, so we decided for simplicity that someone passed over for promotion will be more likely to accept a better position somewhere else. This way, the program can simulate individuals leaving for reasons other than retirement.

- *Promotion from assistant professor to associate professor is independent of the number of years served.* The problem states specific rules for all promotions except this one; therefore, in order to allow the most flexibility in the models, we do not impose any requirements other than one must be an assistant professor in order to be promoted to associate professor. This way, the college can add requirements or expand them as necessary. This same expansion can easily be applied to existing promotion requirements, due to the modular structure of the final programs.

- *The college may hire faculty from outside the college into any rank.* Sometimes gaps in the numbers of personnel at a particular rank will appear. In order to keep a constant number of faculty at each rank, whenever a gap occurs, a new faculty member is automatically hired from the outside community. The starting salary for this new person is determined by placing them directly on the ideal salary curve. In this way, the hiring of new faculty members helps develop the transition from the old salaries to the new compensation scheme.

The Flux Model

The program model resulting from the above assumptions allows a great deal of built-in flexibility. We feel this is crucial, to show that the model and the testing programs can be applied to current and future situations of the college. The programs themselves can be used to examine the models and changes in data supplied to the models.

The fluctuation, after which this program is modeled, is the combined effect of retirement, individuals leaving, promotion, and new hires. **Figure 5** shows the origin and direction of each of these conditions.

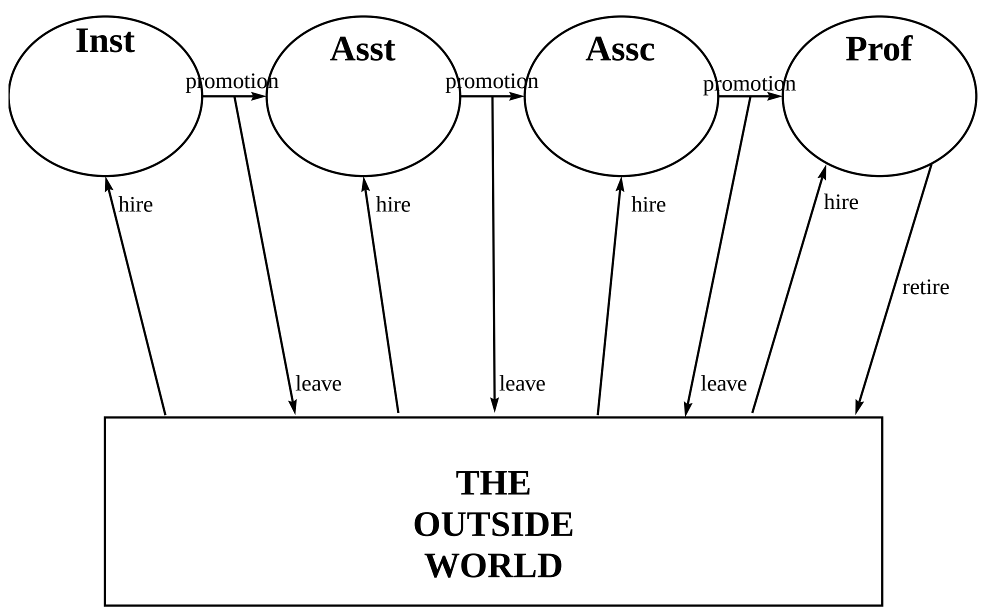

Figure 5. The fluctuation model—a basis for programming the four models.

Faculty Fluctuations and Program Results

An initial analysis of the output and the performance of the programs was done using Mathematica and its animation functions. The data were plotted, one frame per year. Then the frames were animated using a built-in animation function. In this way, we could see how the faculty population moved towards the ideal salary curve as their indices rose because of time and promotion. Often, unexplainable data movement led to a re-analysis of the program and sometimes even of the model.

The conclusion obtained from the animation of series of data sets over 10, 20, and 100 years is that the programs and models do create a feasible salary scheme that the Provost can use. However, determining which method was the most fair and reasonable required a more rigorous examination of the output data and the models. Once again, Mathematica was used to do statistical analysis of the models.

Analyzing Results of the Models

A fair salary system gives equal rewards for equal services. Using the index system, all people at a given index are considered to have the same value. Thus, a fair system minimizes the differences in salary between people of the same index. To determine which model is the most fair, and therefore the best, our programs were used to simulate the effects of the different salary schemes on the relative standard deviation of salaries at any given salary level.

Each salary scheme was run for ten years at budget increases of 2% and 5% per year. The relative standard deviation of salaries in the same index was calculated for each year. Then the relative standard deviations over the entire range of indices was averaged. Those indices that had 0 or 1 people were not included in this average, since a standard deviation could not be calculated. The results are shown graphically in **Figures 6–9** for a 2% increase in budget per year (results are similar for 5% per year). Ten replicates of this experiment were performed to show the consistency of this calculation.

The scaling and the COLA–I models were consistent over multiple runs and decreased relatively quickly and smoothly to a relative standard deviation of 2%. The COLA–I model was superior to the scaling model when there was a larger yearly increase in the budget; because the COLA–I model scales up over time, it can use all of the money that is available. The deficit-proportional model decreased to 2% relative standard deviation almost as well as the scaling and COLA–I models, but it was more prone to temporary jumps up to a higher relative standard deviation than the other two models. The COLA–II model was the least consistent and took ten years to decrease to a relative standard deviation of 7.5%.

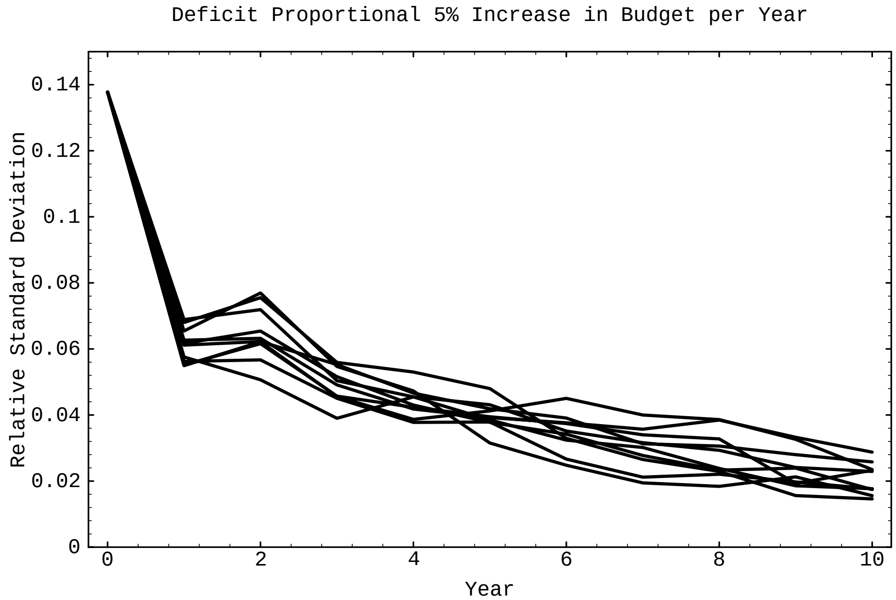

Figure 6. Average relative standard deviations of the **deficit-proportional model**, run at a 2% increase in budget per year. Ten different runs are displayed to show the uncertainty.

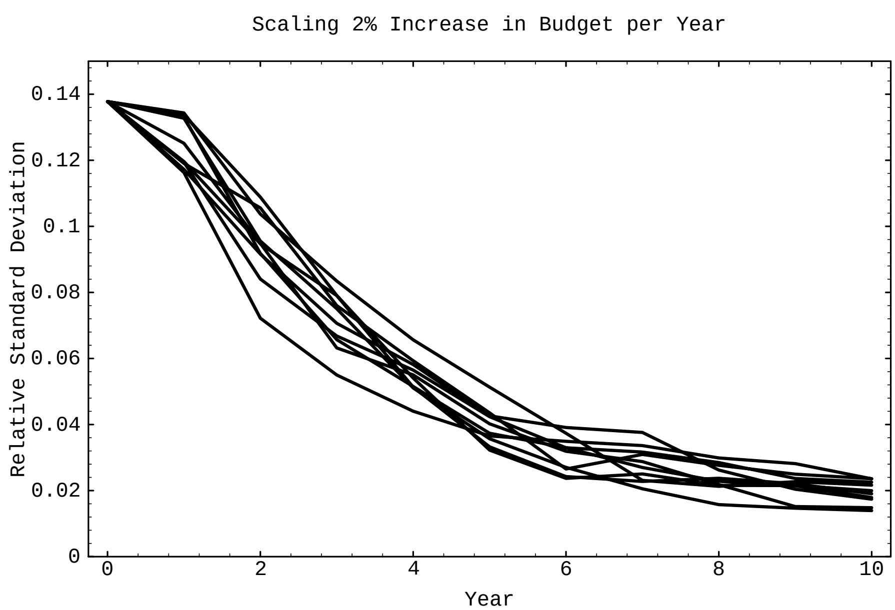

Figure 7. Average relative standard deviations of the **scaling model**, run at a 2% increase in budget per year. Ten different runs are displayed to show the uncertainty.

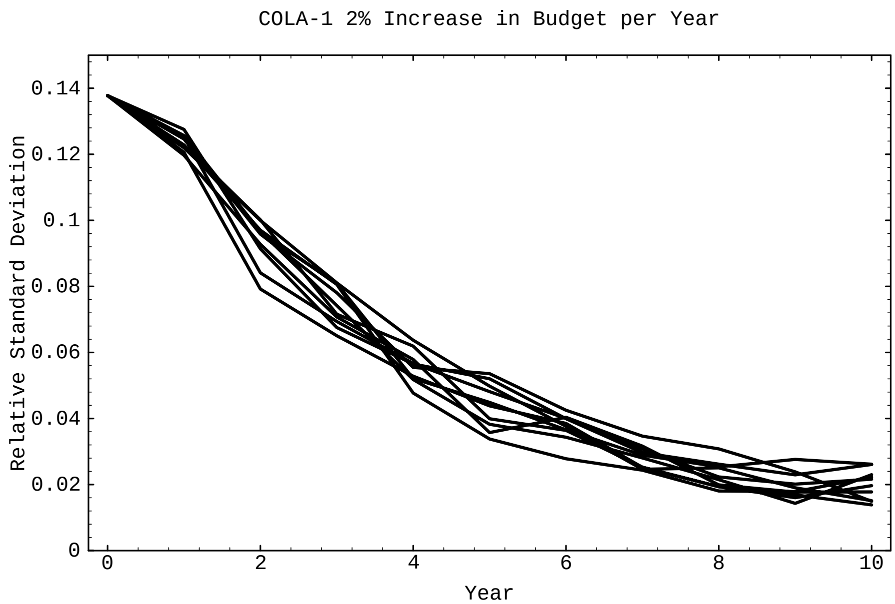

Figure 8. Average relative standard deviations of the **cost-of-living model (COLA–I)**, run at a 2% increase in budget per year and a 3% increase in the COLA value per year. Ten different runs are displayed to show the uncertainty.

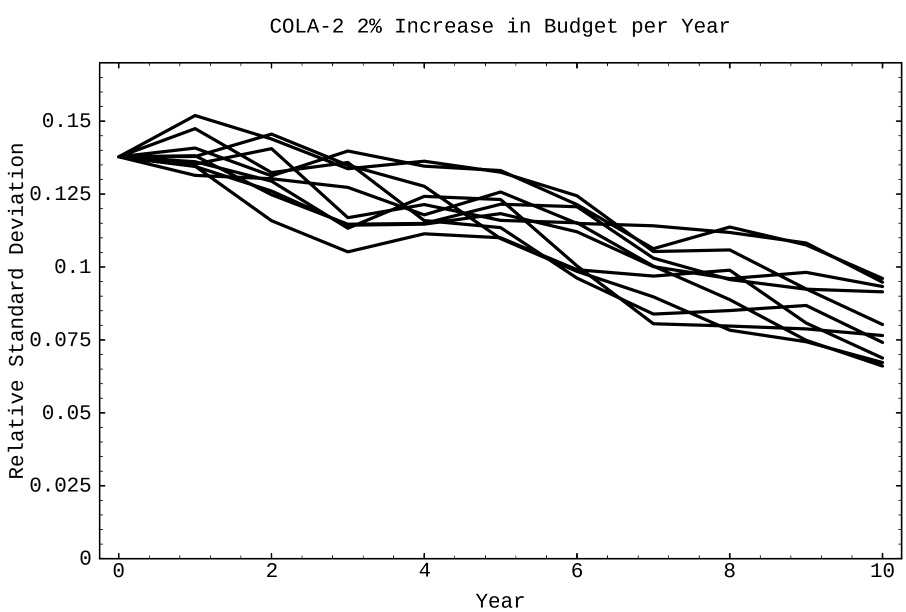

Figure 9. Average relative standard deviations of the **cost-of-living model COLA–II**, run at a 2% increase in budget per year and a 3% increase in the COLA value per year. Ten different runs are displayed to show the uncertainty.

These results show that the scaling and COLA–I models are superior to the deficit-proportional and COLA–II models, because they have less variance in the salaries of equivalent faculty members.

References

Henrici, Stanley B. 1980. *Salary Management for the Nonspecialist.* New York: Amacom.

Marshall, Don R. 1978. *Successful Techniques for Solving Employee Compensation Problems.* New York: Wiley.

Judge's Commentary: The Outstanding Faculty Salaries Papers

Donald E. Miller
Dept. of Mathematics
Saint Mary's College
Notre Dame, IN 46556
dmiller@saintmarys.edu

The substance of this problem was recognizable by the judges from academic institutions, causing many of them to comment on how similar it is to the situation at their own institutions. The basic problem of designing a new pay system satisfying the specified criteria was simple enough that most teams recognized it as a curve-fitting problem. They also, with varying degrees of success, were able to find a model that satisfied most of the criteria.

Implementing the model, however, with limited resources, on salaries of current faculty at Aluacha Balaclava turned out to be deceptively more difficult. Most teams recognized that the salary curve would need to have a negative second derivative in order for salary increments to decrease with experience. Models included variations of logistic, power, exponential, root and polynomial functions, with the exponential function appearing most often. In their Outstanding paper, the team from University Alaska Fairbanks indicated that they had experimented with six different functions before settling on the logistic model because it met all the criteria.

Some teams used an appropriate model for the first twenty-five years of experience, then froze everyone's salary, thus violating the principle that all faculty should get a raise any year that money is available. Most teams used the same model for all four ranks and entire tenure of the faculty member. However, others recognized the instructor rank as different from the other ranks and developed a separate model for it. Still others developed a separate model for the full professor rank, arguing that it was different because a full professor has no chance for promotion. The Outstanding paper from Southeast Missouri State University even developed a separate model for each rank.

Modelers can frequently gain creditability by demonstrating that they understand the problem in its context. With our salary data, this could have been accomplished by recognizing some of the salaries as outliers that would need to be dealt with individually. The datum point representing a professor with two years of experience and a salary of $85,500 is clearly an outlier and will be virtually impossible to bring into line with any reasonable model. This person may be some sort of a "superstar" who is not expected to fit into the model salary structure. One of the judges suggested that this salary belonged to the football coach. Whatever the case, the modeler should be willing to point out or question such unusual situations, but few of the teams did. One exception was the Outstanding paper from the North Carolina School of Science and Mathematics. This team showed that, "...with the exception of a few grossly overpaid faculty, the problem of unfairness would be solved" in five years.

Better papers were distinguished by a complete and mature treatment of the assumptions as well as a precise plan of implementation. The team from University of Alaska at Fairbanks even offered two plans, showed graphically how they differed, and then recommended one over the other. All the Outstanding papers also did some sensitivity analyses and included cost of living in their implementation plan. Another feather of these papers was a well-written executive summary.

About the Author

Donald Miller is Associate Professor and of Chair of Mathematics at Saint Mary's College. He has served as an associate judge of the MCM for three years and prior to that mentored two Meritorious teams. He has done considerable consulting and research in the areas of modeling and applied statistics. He is currently a member of SIAM's Education Committee and president of the Indiana Section of the Mathematical Association of America.