# Industrial Mathematics Modeling Workshop for Graduate Students

**July 27 - August 4, 1998**

*Edited by:*
**Pierre A. Gremaud**
**Zhilin Li**
**Ralph C. Smith**
**Hien T. Tran**

**Center for Research in Scientific Computation**
**North Carolina State University**
**Raleigh, NC**

# Contents

# Participants

## A. Graduate Students

1. Karen Bliss, University of Missouri-Rolla, kbliss@umr.edu
2. Philippe Bremond, University of New Orleans, pbremond@math.uno.edu
3. Jim Buchanan, University of California, Santa Barbara, buchanan@math.ucsb.edu
4. Danielle Bundy, University of Colorado, Danielle.Bundy@Colorado.EDU
5. Chongsheng Cao, University of California, Irvine, ccao@hillary.lanl.gov
6. Michelle Capozzoli, University of New Hampshire, mhopkins2christa@unh.edu
7. Jerawan Chudoung, Virginia Tech, jerawan@calvin.math.vt.edu
8. Timothy Chevalier, University of New Hampshire, twc@kepler.unh.edu
9. Linda Connolly, University of New Orleans, lconnoll@math.uno.edu
10. Steven J. Coorpender, University of New Orleans, scoorp@pop3.lsumc.edu
11. Daniel Finkel, North Carolina State University, definkel@unity.ncsu.edu
12. Simone Flory, Louisiana State University, flory@marais.math.lsu.edu
13. Jayadeep Gopalakrishnan, Texas A&M
14. Jeffrey Grant, North Carolina State University, jdgrant2@unity.ncsu.edu
15. Kevin Hogan, University of Texas, Austin, yang@ticam.utexas.edu
16. Henrik Kalisch, University of Texas, Austin, Kalisch@fireant.ma.utexas.edu
17. David Kan, University of California, dkan@math.ucla.edu
18. Jennifer Kyzar, University of Tennessee, kyzar@mathsun1.math.utk.edu
19. Brian Lewis, North Carolina State University, bmlewis@unity.ncsu.edu
20. Anastasios Liakos, Pittsburgh University, alist@math.pitt.edu
21. Renee Yong Ran Liu, University of California, Berkeley, reneeliu@math.berkeley.edu
22. Jordan E. Massad, North Carolina State University, jemassad@unity.ncsu.edu
23. John Matthews, North Carolina State University, jvmatthe@unity.ncsu.edu
24. Shailesh Naire, University of Delaware, naire@dipper.math.udel.edu
25. Lakshmi Puthanveetil, University of Massachusetts, puthanv@math.umass.edu
26. Sumithra Ramaswamy, Ohio State, sumithra@stat.ohio-state.edu
27. Bob Sims, University of Alabama at Birmingham, sims@vorteb.math.uab.edu
28. Alexandra B. Smirnova, Kansas State University
29. Mohamed Tawhid, University of Maryland, tawhid@sun4.math.umbc.edu
30. Troy Tokarchik, University of Tulsa, TTM24542@centum.utulsa.edu
31. Ilker Tunay, Washington University, St. Louis, ilker@rodin.wustl.edu
32. Shree Yvonne Whitaker, North Carolina State University, sywhitak@eos.ncsu.edu
33. Steve Wilmarth, Texas A & M, wilmarth@math.tamu.edu
34. Shannon Wynne, University of California, Irvine,
35. Igor Zhislin, Middle Tennessee State University, igor@frank.mtsu.edu

## B. Problem Presenters and Faculty

1. H.T. Banks, North Carolina State University, htbanks@eos.ncsu.edu
2. Negash Begashaw, Clark Atlanta University
3. Chad Bouton, Battelle Memorial Institute, Bouton@battelle.org
4. Bryan Cunningham, FACE International Corp., bryan@faceco.com
5. Zhong Ding, Ortho-Clinical Diag, Johnson & Johnson, Zding@cldx.com
6. Rich Fabiano, University of North Carolina, Greensboro, fabiano@uncg.edu
7. Pierre Gremaud, North Carolina State University, gremaud@math.ncsu.edu
8. Kazi Ito, North Carolina State University, kito@unity.ncsu.edu
9. Zhilin Li, North Carolina State University, zhilin@eos.ncsu.edu
10. Hung Ly, North Carolina State University, hly@eos.ncsu.edu
11. Negash Medhin, Clark Atlanta University, medhin@hubble.cau.edu
12. Karla Mossi, FACE International Corp., kmossi@faceco.com
13. Tien Nguyen, Aerospace Corp., tien.m.nguyen@aero.org
14. Norvell Rose, FACE International Corp., norvell@faceco.com
15. Tony Royal, Jenike & Johanson, Inc., royalta@jenike.com
16. Jeff Sachs, D.H. Wagner Assoc., jeff.sachs@alum.mit.edu
17. Ralph C. Smith, North Carolina State University, rsmith@eos.ncsu.edu
18. Hien T. Tran, North Carolina State University, tran@math.ncsu.edu

# Preface

This volume contains the proceedings of the Industrial Mathematics Modeling Workshop for Graduate Students that was held at the Center for Research in Scientific Computation at North Carolina State University (NCSU), Raleigh, North Carolina, July 27 to August 4, 1998. This workshop which was the fourth one held at NCSU, brought together 36 graduate students from various graduate programs across the country (see participants section for a list of students and their institutions). These graduate students represented a large number of graduate programs across the country including the University of Missouri-Rolla, University of New Orleans, University of Colorado, University of New Hampshire, Virginia Tech, University of California (Santa Barbara, Irvine, Berkeley, and Los Angeles), Texas A&M, Louisiana State University, University of Texas, University of Tennessee, Worcester Polytech, University of Delaware, University of Massachusetts, Kansas State University, University of Alabama at Birmingham, University of Tulsa, Middle Tennessee State University, Washington University, University of Maryland, North Carolina State University, Pittsburgh University and Ohio State University. The students were divided into six teams working on "industrial mathematics" problems presented by industrial scientists. These problems were not the neat, well-posed academic exercises typically found in coursework, but rather challenging real world problems from industry or applied science. The problems, which were explained to the students on the first day of the workshop, required fresh, new insight for their formulation and solution. Each group spent the first nine days of the workshop working on their problem and then reported on their findings in half-hour public seminars on the last day of the workshop. The following is a list of the presenters and the problems which they presented.

- **Chad Bouton** (Battelle Memorial Institute) *Maximizing Deflection While Minimizing Power Consumption in an Unloaded Circular Piezoelectric Actuator Design.*
- **Bryan Cunningham, Karla Mossi and Norvell Rose** (FACE International Corporation) *THUNDER Curvature Modeling.*
- **Tien M. Nguyen** (The Aerospace Corporation) *Mathematical Modeling of High Power Amplifier (HPA) and HPA Linearization Through a Predistorter.*
- **T. Antony Royal** (Jenike & Johanson, Inc.) *Time Dependent Consolidation of Fine Powders.*
- **Jeff Sachs** (D.H. Wagner Associates) *High Precision, High Accuracy Control of Fluid Microdispenser Under Variable Load.*
- **Zhong Ding** (Ortho-Clinical Diagnostics, Inc., Johnson & Johnson Co.) *Mass Transport and Surface Reaction of Immunoassay.*

These problems represent a very broad spectrum of mathematical topics and applications. The readers will observe that traditional topics such as number theory, algebra, probability and statistics, fluid dynamics, ODE and PDE, optimization and numerical analysis play a very important role in the solution of such industrial problems. Although nine days is a short time for a full investigation of some aspects of the problems, the reports illustrate remarkable progress on all problems.

We, the organizers, strongly believe that this workshop (and previous workshops) provided substantive non-academic research-related experiences for graduate students while contributing to the research efforts of industrial participants. In addition, these experiences facilitated the development of graduate students' ability to communicate and interact with scientists who are not traditional mathematicians but require and employ various mathematical tools in their research. Finally, the

experience gained in the workshop will help to improve graduate student's vision about the type of mathematics and science that should be pursued by students who aspire to both non-academic and academic postgraduate career.

The success of the workshop was greatly enhanced by active participation in a very friendly atmosphere and almost uninterrupted work during the nine days of attendance, and the organizers are most grateful to participants for their contributions. The organizers would like to thank the National Security Agency, the National Science Foundation, the Center for Research in Scientific Computation, and the Department of Mathematics at North Carolina State University for their generous financial support of this and previous workshops. The financial support from the National Science Foundation this year has made it possible for us to provide much needed travel support to some graduate students who, otherwise, would not be able to participate. Special thanks and appreciation are due to the faculty and staff of the Center for Research in Scientific Computation, the Department of Mathematics and North Carolina State University for the provision of excellent facilities and services. Finally, we would like to thank Michelle Hein, Rory Schnell, and Brenda Smith for their efforts and help in all administrative matters and Jerry Batzel and Scott Beeler for providing transportation for the participants.

Pierre Gremaud
Zhilin Li
Ralph C. Smith
Hien T. Tran
Raleigh, 1999

# PROBLEM 1: MAXIMIZING THE DEFLECTION TO VOLTAGE RATIO IN A CIRCULAR PIEZOELECTRIC ACTUATOR DESIGN

Steven Coorpender[1], Daniel Finkel[2], Jennifer Kyzar[3], Bob Sims[4], Alexandra Smirnova[5], Mohamed Tawhid[6]

Problem Presenter:
Chad Bouton
Battelle Memorial Institute

**Abstract**

Piezoelectric actuators are useful in many applications, particularly when small, accurate deflections are required. The amount of deflection that can be attained is dependent on many factors including the amount of applied voltage and the geometry of the design. Efficient actuators are especially important when power is limited. We developed an ODE model of a thin circular piezoelectric patch attached to a single layer circular mounting surface (a unimorph-type actuator). All material properties and boundary conditions were specified in the design. The model was used to optimize the design of the two-layer structure by varying the patch radius and mounting layer thickness.

## 1 Introduction

Battelle Memorial Institute, based in Columbus, Ohio, is a non-profit research and development organization with laboratories throughout the world. The institute develops innovative technology for government and industry that is integrated into a variety of products. Examples of these products range from hologram technology to drug developments for health needs. Chad Bouton, a research engineer at Battelle, was invited by the workshop organizers to present a current technical problem. The problem Mr. Bouton presented focussed on the modeling of a circular piezoelectric actuator with multiple layers and optimization of its performance by varying model parameters. Similar devices have been investigated by members of the CRSC including Ralph Smith, H.T. Banks and their students.

A piezoelectric actuator is an electromechanical device consisting of a patch bonded to one or both surfaces of a single- or multi-layered mounting. The patch is composed of a piezoceramic material, while the mounting is usually a stronger material, e.g., hardened brass in our experimental data. When a voltage is applied to the piezoceramic material, dipoles expand in the direction of the electric field causing a change in the length of the material. This strain, combined with the stiffness of the mounting layers, causes the entire actuator system to bend, or deflect. The geometry, materials, and number of mounting layers of the actuator can vary and are application dependent. In the present design, a single patch is attached to the mounting (unimorph), only one mounting

[1]University of New Orleans
[2]North Carolina State University
[3]University of Tennessee
[4]University of Alabama at Birmingham
[5]Kansas State University
[6]University of Maryland, Baltimore County

layer is considered, and the geometry is specified as circular, i.e., both the patch and mounting are cylindrical disks. The edge of the mounting was considered to be clamped.

Our goal for the workshop was to create a model of the actuator and then use this model to find the optimal geometry, i.e., the patch radius and mounting layer thickness that would give the maximum deflection. The original problem included a consideration of power consumption but because of time limitations, this was not incorporated into the model.

Much of the first few days was spent learning basic structural mechanics. To achieve this, we examined in detail the derivation of the equation governing vibrations in a cantilever beam [1]. The next step was to examine the circular geometry present in the actuator problem. Once this was understood, we were then able to apply many of the concepts learned from the beam to the actuator design. The primary difficulty lay in the derivation of appropriate boundary conditions. Once boundary conditions were determined, we were able to find the analytical solution which led to the computation of the optimal parameters for the design. In this report we present the derivation of the model, the solution, and the results of the optimization.

# 2 Model Development

## 2.1 The Basics

As a first step toward the development of a model which quantifies the deflection of two circular plates with different material properties, we first considered the deflection of one circular plate with uniform properties. By uniform, we mean that the plate is composed of only one material, i.e., the physical properties (such as flexural rigidity and conductivity) of this material remain constant throughout and that measured quantities are axially symmetric.

Briefly, we arrived at an expression for the deflection of this uniform circular plate by summing the moments about an infinitesimal sectoral wedge. In terms of the transverse deflection $w(r)$, the flexural rigidity $D$, and Poisson's ratio $\nu$, the radial and tangential moments for a circular plate are given by

$$M_r = D\left(\frac{d^2w}{dr^2} + \frac{\nu}{r}\frac{dw}{dr}\right)$$

$$M_\theta = D\left(\nu\frac{d^2w}{dr^2} + \frac{1}{r}\frac{dw}{dr}\right)$$

(see [2]). Furthermore, summing the moments and neglecting higher order terms yields the expression

$$Q_r = \frac{1}{r}M_r + \frac{dM_r}{dr} - \frac{1}{r}M_\theta,$$

for the shear force. A simple substitution now yields a third order differential equation for the deflection of a uniform circular plate

$$\frac{d^3w}{dr^3} + \frac{1}{r}\frac{d^2w}{dr^2} - \frac{1}{r^2}\frac{dw}{dr} = \frac{Q_r}{D}. \tag{1}$$

For our purposes, the shearing force was assumed to be due only to a uniformly distributed pressure, $q$, and so we have the relation

$$2\pi r Q_r = \pi r^2 q.$$

Using this, (1) can be expressed as

$$\frac{d}{dr}\left[\frac{1}{r}\frac{d}{dr}\left(r\frac{dw}{dr}\right)\right] = \frac{qr}{2D},$$

which can be easily integrated [3]. The result is

$$w(r) = \frac{qr^4}{64D} + \frac{C_1 r^2}{4} + C_2 \ln\left(\frac{r}{R_o}\right) + C_3$$

where $R_o$ is the outer radius of the plate.

### 2.2 The Problem

Our problem was to model the deflection of a circular piezoelectric patch bonded to a mounting material that is clamped at the outer edge. The relevant differences between this problem and the one discussed above lie in the fact that there are two plates with unequal radii and different material properties (i.e., different $D$'s and $\nu$'s), and there is an external moment caused by the application of voltage to the piezoceramic patch (this is in addition to the previously mentioned uniform pressure).

These differences represent the three issues that were resolved so that we could employ previous results. The issue of unequal radii was the most fundamental and was the first to be resolved. We treated the entire structure as two separate components which were to be brought together in the end through the matching of boundary conditions. Specifically, we modeled the deflection of the center piece (the composite comprised of piezoceramic patch bonded to the substrate) and the outer piece (an annulus consisting of only the mounting) as separate entities, both of which have a constant radius throughout their height. By doing this, the problem was reduced to two simpler problems: that of modeling the annulus, for which we could immediately use our previous results (the annulus is simply a uniform plate with restricted domain: $R_i \leq r \leq R_o$, inner radius to outer radius), and that of modeling the composite, which is composed of two materials, and has an external moment.

After we obtained the results on both pieces (the annulus and the composite), we needed to find a way to ensure compatibility. In mathematical terminology, we needed to find the proper boundary conditions to get a unique solution that was applicable for describing deflection at any point of our two-plate structure.

### 2.3 Modeling the Composite

In order to use our previous results for a uniform plate, we needed to find an expression for the flexural rigidity, $D$, of the composite plate. The flexural rigidity for a plate made of one material (such as the annulus) is given by $\frac{Eh^3}{12(1-\nu^2)}$, where $E$ denotes the Young's modulus for the material and $h$ is the thickness of the plate. For a composite material special considerations must be made. An expression for the composite flexural rigidity may be derived through the bending moment-curvature relationship, and involves the generalized isotropic Hooke's Law with two-dimensional stress and strain tensors. Fortunately, this was available to us from the problem presenter and is given by

$$D_c = \frac{1}{3}\left(\frac{E_p}{1-{\nu_p}^2}\right)\left[(h_m + h_p - z_{ns})^3 - (h_m - z_{ns})^3\right] + \frac{1}{3}\left(\frac{E_m}{1-{\nu_m}^2}\right)\left[(h_m - z_{ns})^3 - (-z_{ns})^3\right]$$

where $z_{ns}$ is the vertical coordinate of the neutral surface measured from the bottom of the mounting and the $h$'s are the thicknesses of the layers denoted by the subscripts $p$ (patch) and $m$ (mounting). For a uniform plate, the location of the neutral surface is simply the centroid of the cross-sectional area, but for a composite, it depends on the stiffness of the material as well as the thickness of each material. To this end, we derived an expression for the neutral surface from the basic principles of pure bending. Pure bending is the condition in which a structural member is in equilibrium under the action of a bending moment alone. In a horizontal beam subjected to pure bending, vertical

plane sections (perpendicular to the beam axis) will tilt slightly, while horizontal plane sections will curve. The neutral surface is defined as the horizontal plane section which does not change in length, i.e., $\varepsilon = 0$. This assumption is also valid for circular plates in equilibrium under pure bending. Since the strain varies linearly with its distance from the neutral surface, we can calculate the position of the neutral surface using equilibrium requirements.

Consider a differential element (wedge) of the composite circular plate in equilibrium subjected to a pure bending moment. Since the sum of all forces in the radial direction equals zero we have

$$\sum F_r = \int_A \sigma_r dA = 0.$$

Substituting for the stress using Hooke's Law gives

$$\int_A E\varepsilon_r dA = 0.$$

Since the plate is axisymmetric, we may assume that the radial strain, Young's modulus, and Poisson's ratio are constant in the tangential direction. Taking the bottom of the mounting layer as the origin, we have

$$\int_0^{h_m+h_p} E\varepsilon_r dz = 0.$$

Substituting for the strain gives

$$\int_0^{h_m+h_p} \frac{E}{\mathcal{R}}(z - z_{ns})dz = 0$$

where $\mathcal{R}$ is defined as the radius of curvature of the neutral axis. Separating the integral into the sum of two integrals over each material gives

$$\int_{h_m}^{h_m+h_p} \frac{E_p}{\mathcal{R}}(z - z_{ns})dz + \int_0^{h_m} \frac{E_m}{\mathcal{R}}(z - z_{ns})dz = 0.$$

Finally, integrating and then solving for the neutral surface location gives

$$z_{ns} = \frac{E_p h_p(\frac{h_p}{2} + h_m) + E_m h_m(\frac{h_m}{2})}{E_p h_p + E_m h_m}.$$

It is worth noting that this formula is easily generalized to any number of layers of different materials and would be given by

$$z_{ns} = \frac{\sum_{k=1}^{N} E_k h_k z_k}{\sum_{k=1}^{N} E_k h_k}$$

where N is the total number of layers, $E_k$ is Young's modulus of the $k^{th}$ layer, $h_k$ is the thickness of the $k^{th}$ layer, and $z_k$ is the distance from zero to the center of the kth layer.

The last issue to be resolved was the calculation of the external moment generated by the application of a voltage to the piezoceramic patch. Piezoelectric material elongates in the direction of an applied electric field, causing a subsequent contraction in the plane normal to the field. The relationship between this strain and the voltage is given by a material "strain" constant, $d_{31}$, in units of [(m/m)/(V/m)] [4]. Thus the "free" strain, i.e. the strain for an unmounted piezoelectric patch, is given by

$$\varepsilon_p = \frac{V}{h_p} d_{31}.$$

In order to derive a relationship for the external moment as a function of voltage, we simply substitute the free strain into the definition for the moment [5], to give

$$(M_r)_{pe} = \int_A \frac{E_p}{1-\nu_p} \varepsilon_p (z - z_{ns}) dA = \int_A \frac{E_p}{1-\nu_p} \left(\frac{V}{h_p} d_{31}\right)(z - z_{ns}) dA.$$

Now we must integrate over the cross-section of the patch normal to the radial direction. In order to be consistent with our expression for the moments from Section 2.1, which are actually moments per unit length, we must also divide both sides by the arc-length. Redefining $M_p$ to be the moment per unit length, we have

$$(M_r)_{pe} = \int_{h_m}^{h_m + h_p} \frac{E_p}{1-\nu_p} \left(\frac{V}{h_p} d_{31}\right)(z - z_{ns}) dz$$

which may be easily integrated to give

$$(M_r)_{pe} = \frac{E_p}{2(1-\nu_p)} V d_{31} (h_p - 2z_{ns} + 2h_m)$$

Incorporating the external moment into our model to get a total moment for the composite was simply a matter of subtracting it from the "reaction" moment term in our original equation: $\mathcal{M}_r = M_r - (M_r)_{pe}$. Once we had brought the external moment into our model and had an expression for the composite flexural rigidity, $D$, we were able to use our previous equation for the uniform plate to obtain the following results:

| **Region:** | **Equation:** | **Domain:** |
|---|---|---|
| Annulus | $w(r) = \frac{qr^4}{64D} + \frac{C_1 r^2}{4} + C_2 \ln\left(\frac{r}{R_o}\right) + C_3$ | $R_i < r \leq R_o$ |
| Composite | $\tilde{w}(r) = \frac{qr^4}{64D_c} + \frac{C_4 r^2}{4} + C_5 \ln\left(\frac{r}{R_o}\right) + C_6$ | $0 \leq r \leq R_i$ |

## 2.4 Boundary Conditions

Once we had both equations for the deflection: one in the annular region and one for the composite, we needed to link these together so that we would have one equation for the combined structure. To do so, we applied boundary conditions that enabled us to solve for the six unknown constants of integration. The first boundary condition is that of a fixed, or clamped, mounting, which yields: $w(R_o) = 0$ and $w'(R_o) = 0$. Next, the position and slope of the two components must be equal where they meet, i.e., $w(R_i) = \tilde{w}(R_i)$ and $w'(R_i) = \tilde{w}'(R_i)$. Then, from the assumption of axial symmetry, we have an implicit, geometrical boundary condition at the center, namely $\tilde{w}'(0) = 0$. The final boundary condition is due to equilibrium conditions at the boundary between the annulus and the composite. Specifically, we have $\mathcal{M}_r(R_i) = M_r(R_i)$; that is, the total moment of the composite (which incorporates the voltage term) should be equal and opposite to the moment of the annulus at the boundary, $R_i$. Summarizing these results yields the boundary conditions

| | |
|---|---|
| 1) $w(R_o) = 0$ | 4) $w(R_i) = \tilde{w}(R_i)$ |
| 2) $w'(R_o) = 0$ | 5) $w'(R_i) = \tilde{w}'(R_i)$ |
| 3) $\tilde{w}'(0) = 0$ | 6) $\mathcal{M}_r(R_i) = M_r(R_i)$ |

The above conditions allowed us to calculate the unknown constants in our equations

$$c_1 = \frac{2}{1+\nu+(1-\nu)\frac{R_0^2}{R_{pe}^2} - \frac{D_c}{D}(1+\nu_c)(1-\frac{R_0^2}{R_{pe}^2})} \left\{ \frac{qR_{pe}^2}{16D}\left[\nu_c - \nu + D_c(1+\nu_c)\left(\frac{1}{D}-\frac{1}{D_c}\right)\right] + \frac{qR_0^4}{16DR_{pe}^2}\left[\nu - 1 - \frac{D_c}{D}(1+\nu_c)\right] - \frac{M_{pe}}{D}\right\}$$

$$c_2 = \frac{-qR_0^4}{16D} - \frac{c_1 R_0^2}{2}$$

$$c_3 = \frac{-qR_0^4}{64D} - \frac{c_1 R_0^2}{4}$$

$$c_4 = \frac{qR_{pe}^2}{8}\left(\frac{1}{D}-\frac{1}{D_c}\right) - \frac{qR_0^4}{8R_{pe}^2 D} + c_1\left(1-\frac{R_0^2}{R_{pe}^2}\right)$$

$$c_5 = 0$$

$$c_6 = \frac{-qR_0^4}{64D}\left(\frac{1}{D}-\frac{1}{D_c}\right) + \frac{qR_0^4}{64D} - \frac{qR_0^4}{16D}\ln\left(\frac{R_0^2}{R_{pe}^2}\right) - \frac{c_1 R_0^2 \ln\left(R_0^2/R_{pe}^2\right)}{2}.$$

# 3 Optimization

Our goal was to maximize the deflection to voltage ratio of the piezoelectric actuator model by finding the optimal mounting layer thickness and piezoceramic patch radius. Since we were able to obtain an analytical solution to the circular plate equation for the annulus and composite structures, the optimization process was easily achieved using Matlab. We assumed a one volt input and a pressure of zero. We used hardened brass as the mounting layer.

## 3.1 Variation of Mounting Layer Thickness and Patch Radius

The radius of the mounting layer was held constant at 4.0 mm and the thickness of the piezoelectric patch was held constant at 200 microns. We varied the thickness of the mounting layer and the radius of the piezoceramic patch in order to find the maximum deflection to voltage ratio, which can be seen in Figures 1 and 2. We found the optimal thickness of the mounting layer to be 49.85 microns and the optimal radius of the piezo layer to be 3.67 mm.

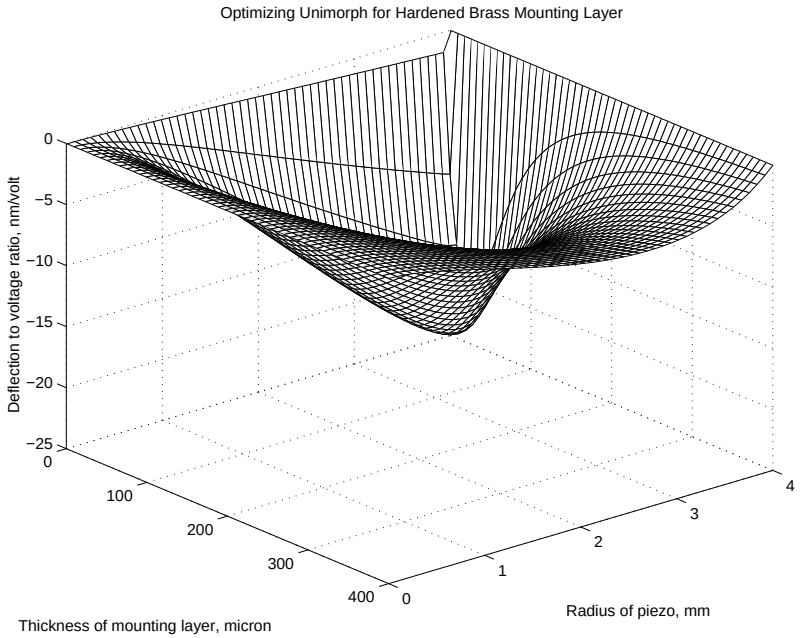



Figure 1. Relationship between mounting layer thickness and patch radius.

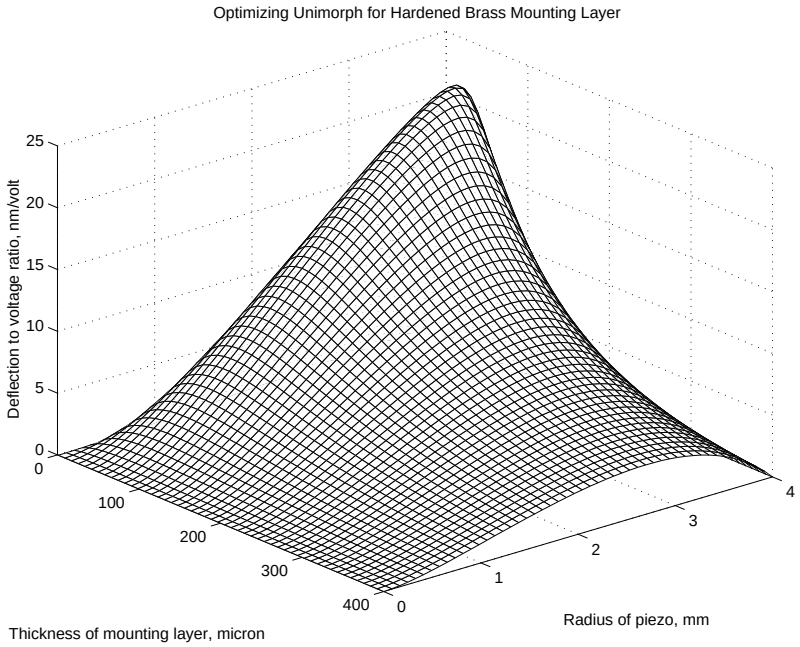



Figure 2. Relationship between mounting layer thickness and patch radius.

### 3.2 Constant Thickness of the Mounting Layer

Once the optimal thickness of the mounting layer was found, we wanted to investigate the effect of varying the radius of the piezoceramic patch. Figure 3 shows a cross section of deflection to voltage ratio versus the radius of the piezoceramic, with a constant mounting layer thickness set at the optimal value of 49.85 microns. Notice that the deflection drastically decreases when the radius extends beyond the optimal value of 3.67 mm.

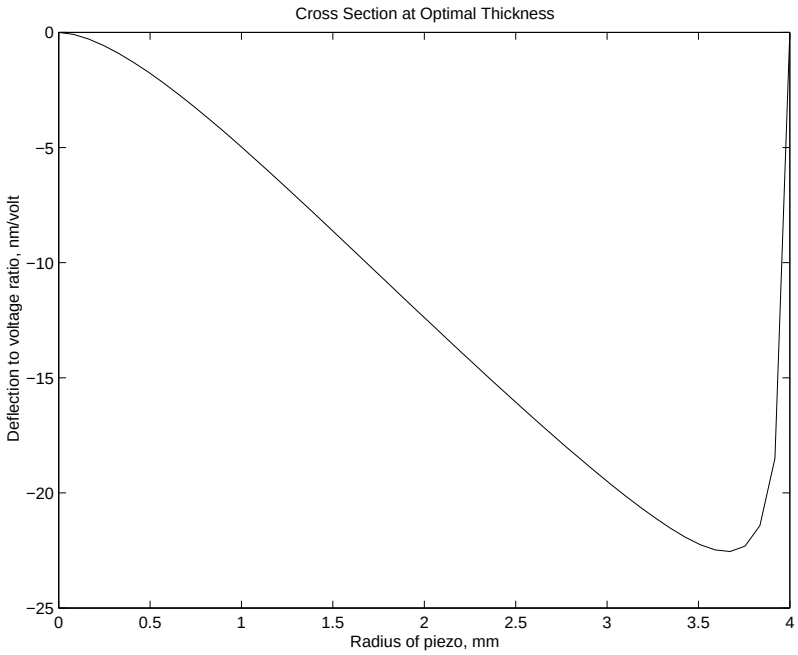



Figure 3. Deflection as a function of radius at the optimal patch thickness.

### 3.3 Constant Radius of Piezoceramic Patch

Similarly, Figure 4 shows a cross section of deflection to voltage ratio versus the thickness of the mounting layer, with the radius of the piezoceramic held constant at its optimal value of 3.67 mm. Notice the optimal thickness is around 49 microns. The amount of deflection will decrease if we allow the thickness to become too small.

We found the deflection for the optimal design to be -22.54 nm, which is illustrated in Figure 5. Based on Chad Bouton's past experience in testing piezoceramic actuators, our model results look excellent. He compared some of our results with data that he had collected in previous work; however, we are unable to publish this information due to the confidentiality between Battelle Memorial Institute and their clients.

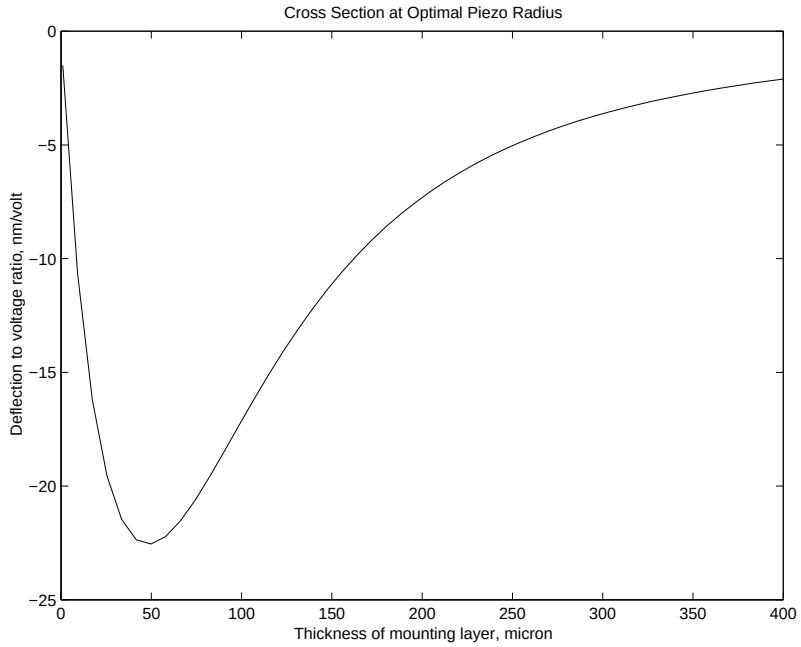



Figure 4. Deflection as a function of thickness at the optimal patch radius.

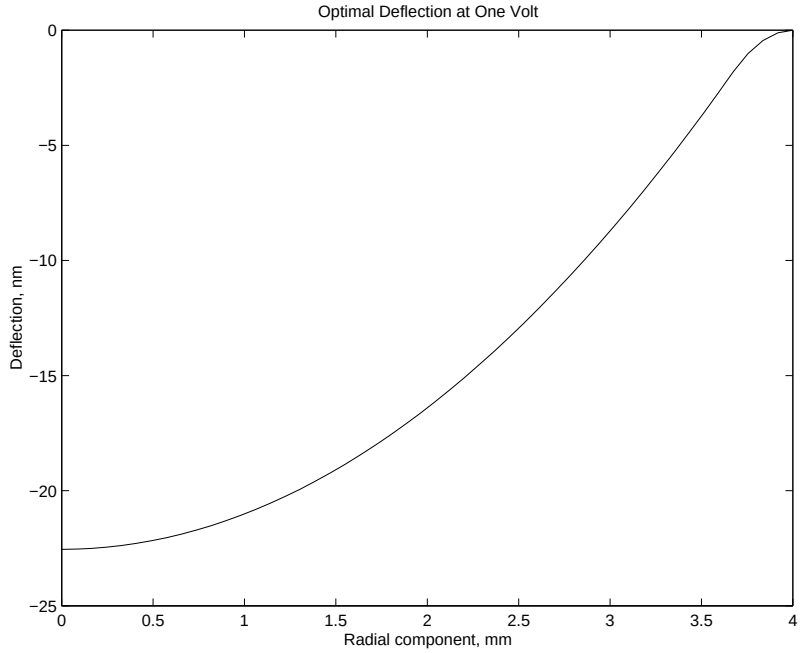



Figure 5. Optimal deflection at one volt

.

# 4 Conclusion

We have presented a model for the specified piezoelectric actuator system and optimized the design with regard to certain aspects of the geometry. In the problem presented to us the radius of the piezoceramic patch and the thickness of the mounting layer were the only variable parameters. With this model, it was also possible to investigate changes in deflection which would result from adjusting other parameters such as material stiffness, patch thickness, and mounting layer radius. In addition, the model could be easily generalized to include additional piezoceramic or mounting layers, and an improvement in accuracy could be achieved through a consideration of bonding materials. The original goal for this workshop was to maximize the deflection while minimizing the power consumption in a circular piezoelectric design. Unfortunately, we did not have time to incorporate power consumption into the model. This is a future consideration and a model similar to ours will most likely be used in this development.

### Acknowledgements

We would like to thank Chad Bouton, Rich Fabiano, and Ralph Smith for their invaluable assistance during the workshop. We would also like to thank everyone from North Carolina State University and the Center for Research in Scientific Computation for their generous hospitality and the organizers for giving us the opportunity to participate.

# References

[1] R.B. Guenther and J.W. Lee, *Partial Differential Equations of Mathematical Physics and Integral Equations,* Prentice Hall, Englewood Cliffs, NJ, 1988.

[2] H.T. Banks, R.C. Smith and Y. Wang, *Smart Material Structures: Modeling, Estimation and Control*, Masson/John Wiley, Paris/Chichester, 1996.

# PROBLEM 2: MODELING OF THE RADIUS OF CURVATURE FOR THUNDER ACTUATORS

Michelle Capozzoli[1], Jayadeep Gopalakrishnan[2], Kevin Hogan[3], Jordan Massad[4], Troy Tokarchik[5], Steve Wilmarth[6]

Problem Presenters:
Bryan Cunningham, Karla Mossi and Norvell Rose
FACE International Corporation
http://www.faceco.com

**Abstract**

THUNDER is a composite ferroelectric structure. This composite is formed of an aluminum layer, a PZT-5A ceramic, a stainless steel layer, and a hot-melt adhesive sheet between each layer. This assembly is heated to an elevated temperature under pressure and then cooled to room temperature. The adhesive solidifies at a high temperature. Further cooling imparts a compressive pre-stress to the surfaces of the ceramic. This process significantly enhances the resulting actuator's performance and ruggedness. Due to the resulting internal stress-state, the actuator possess a radius of curvature which to some extent determines its characteristic behavior. The radius of curvature is the result of the properties of the constituent materials, the geometry of the constituent materials, and the temperature at which the adhesive solidifies. An ideal modeling algorithm would predict a radius of curvature and its impact on the actuator's properties. With this tool such variables as switching from stainless steel to another material, or changing the thickness of the substrate could be modeled and evaluated. THUNDER elements could easily be custom designed for specific applications.

**Keywords and Phrases:** modeling, composite materials, surface effects

## 1 Introduction

Headquartered in Norfolk, Virginia, FACE International Corporation is a developer and manufacturer of actuators and sensors. In 1997, FACE corporation obtained the manufacturing license to THUNDER (THin-layer composite UNimorph ferroelectric Driver and sEnsoR) developed by NASA. This technology allows the production of compact, adaptable and extremely rugged piezoelectric devices. Since 1997, FACE has introduced a variety of high performance actuators and sensors in a range of shapes and sizes. Companies representing a wide range of industries have expressed strong interest in using these devices. Practical applications exist in the automotive to the aerospace industries as well as in smart structures to industrial process control,

During the manufacturing process, THUNDER is pre-stressed which is responsible for greatly enhancing the actuator's performance and ruggedness. Due to the internal stresses, the actuator

[1]University of New Hampshire
[2]Texas A & M
[3]University of Texas
[4]North Carolina State University
[5]University of Tulsa
[6]Texas A & M

possess a natural radius of curvature, which is associated with THUNDER's characteristic behavior. FACE is very interested in developing a mathematical model which predicts a radius of curvature based on the properties of the materials used. The company's long term goal is to use this model in automating the manufacturing process, ultimately allowing them to easily produce devices tailored to individual customers' requirements.

In this report, we develop a model that predicts the radius of curvature based on the different materials used and their thickness. Section 2 discusses the manufacturing process during which the THUNDER actuator is pre-stressed. In Section 3, we introduce our notation, discuss our assumptions and develop our model. We also characterize the geometry of the actuator and develop a feasible definition of radius of curvature. Section 4 contains the results from our model, which are compared to the initial experimental results found by FACE International Corporation. The last section contains our recommendations and a discussion of further research possibilities.

# 2 Manufacturing Process

The raw materials that comprise the composite material are a piezoelectric ceramic wafer, a metallic backing material, and a hot-melt adhesive sheet. Aluminum, stainless steel, titanium and brass are some of the metals that are used as the backing material. The adhesive used by FACE is LARC-SI (Langley Research Center-Soluble polyImide), which was also invented by NASA. LARC-SI is a wholly aromatic, high performance thermoplastic with a unique combination of physical, mechanical and adhesive properties [1]. The materials are assembled in the following order of layers: the metallic backing material, LARC-SI, piezoelectric ceramic wafer, LARC-SI, top metallic layer (optional). It is possible to vary the number of layers and/or the materials used depending upon the underlying application.

The pre-stressed condition that causes the radius of curvature is created when the composite material is heated in an autoclave. Before being placed in the autoclave, the assembled THUNDER is placed in a bag that will act as a vacuum (14.5 psi) during the heating process. The assembly is then heated from room temperature to $204^o$ C at a rate of $5.6^o$ C/min. At this point, 35 psi (nitrogen) is added to the chamber. When the oven reaches a temperature of $325^o$ C, the composite is "soaked" for 30 minutes. After the 30 minutes have elapsed, the heat is then reduced by a rate of $5.6^o$ C/min to $52^o$ C. The oven-pressure is then released(to zero in about 2 minutes). In addition, the vacuum is also released (from 14 psi to 0 psi in about 30 seconds. It should be noted that the pressure and the vacuum are released gradually to ensure that the composite material does not bend too fast and break.

It is assumed that the bending occurs when the LARC-SI solidifies at $270^o$ C. The piezoelectric ceramic wafer and the metallic backing attach to each other which, due to differing thermal coefficients, causes the composite material to bend as it is cooled. In addition, the coefficient of thermal expansion for LARC-SI is given for its solid form. Therefore, we decided to model the process when the bending occurs, from $270^0$ C to $25^o$ C.

# 3 Model

## 3.1 Assumptions and Notation

The following notation will be used throughout the paper. Figure 1 contains further explanation of the notation.

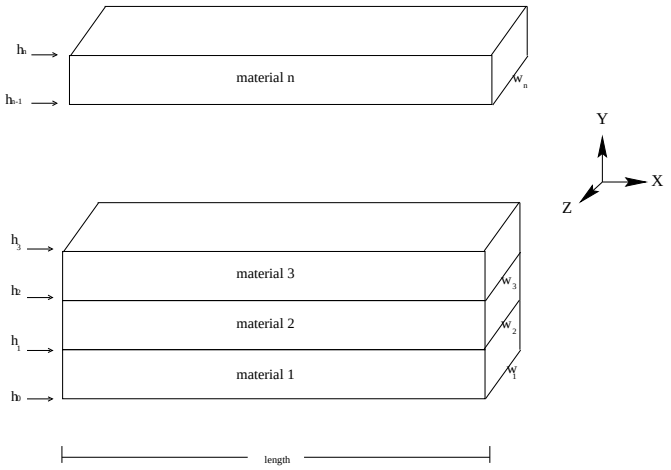



Figure 1. Construction of the Composite Material.

*Notation*

| | |
|---|---|
| $n$ | number of layers in THUNDER |
| $j$ | j$th$ layer of THUNDER $j = 1, \ldots, n$ |
| $\epsilon_o$ | the strain at the $y = 0$ position |
| $k(x)$ | curvature measured at the neutral axis prior to poling |
| $y$ | height assuming that the bottom of the bottom layer is $y = 0$ |
| $E_j$ | modulus of elasticity for a given layer $j$ |
| $E(y)$ | $E_j$ when y is in the j$th$ layer |
| $\alpha_j$ | coefficient of thermal expansion for a given layer $j$ |
| $\Delta T$ | change in temperature |
| $A_x$ | the cross section at any point $x$ |
| $\sigma$ | stress |
| $w_i$ | width of the j$th$ layer |
| $h_j - h_{j-1}$ | height of the j$th$ layer |

In the development of our model, we made the following assumptions.

*Assumptions*

1. The materials are homogeneous and isotropic.
2. The materials have the same length.
3. End effects are ignored.
4. The strains in the $y$ and $z$ direction are zero: $\epsilon_y = \epsilon_z = 0$
5. The cross section of the composite is assumed to be uniform along the length of the material.
6. The thermal strain and the mechanical strain are additive.
7. Temperature heating process is uniform.
8. The materials are free to expand. (i.e., There are no external forces applied.)
9. The increase in curvature due to poling is ignored.

10. The mechanical strain, $\epsilon_x$, is linearly relates to $y$.

11. The effects of the Curie temperature on the material is being ignored.

12. Hookes's Law is valid for each material.

13. Thermal expansion is proportional to temperature change.

14. All cross sections remain plane during the deformation.

### 3.2 The Model

We model the THUNDER device as a slender composite beam of different materials stacked in layers. The adhesive is considered as a separate layer and all layers are assumed to be in ideal contact. We ignore all variations in cross-section in the longitudinal direction. Hence, we may assume that the beam after bending has constant curvature, $k$. The model predicts the value of $k$.

Our analysis is based on the analysis of a bimetallic strip [2]. The method obtains two formulas for the longitudinal strain $\epsilon_x$ (after the temperature change) on the beam; one is in terms of the final curvature of the beam derived purely from geometric considerations, and one in terms of the internal stress and temperature changes that cause the bending. From these equations, we can solve for the internal stress, $\sigma_x$. Force and moment balance for the stresses in each cross-section provide constraints that determine the curvature.

The basic assumption in the model is that the longitudinal strain $\epsilon_x$ varies *linearly* in the transverse direction, i.e., if $y$ is the height of a point from the base of the beam,

$$\epsilon_x(y) = \epsilon_0 - my. \tag{1}$$

This is reasonable when transverse strains are negligible. We will now show that the geometry of deformation forces $m$ to be $k$. Consider a piece of the beam of length $\Delta s = AC$ as depicted in Figures 2 and 3. We assume that the plane containing points $A$ and $C$ becomes the plane containing $A'$ and $C'$ after bending, and similarly for $B$ and $D$. Let $\Delta\theta$ be the angle subtended by the plane containing $A'$ and $C'$ with that containing $B'$ and $D'$. Since $\epsilon_0$ is the strain at $y = 0$, and since $C'D'$ is at a height $\epsilon_0/m$,

$$(1 + \epsilon_0)\Delta s = \left(R + \frac{\epsilon_0}{m}\right)\Delta\theta.$$

Since $\Delta s/\Delta\theta = R$, we get that $(1 + \epsilon_0)R = R + \epsilon_0/m$, from which we conclude that $m = 1/R$ if $\epsilon_0$ is nonzero. In the event $\epsilon_0$ is zero, we may construct a similar argument based on another segment, say $A'B'$ to arrive at the same conclusion. Thus (1) follows with $m = k$. Note that here and henceforth by $k$ we mean the inverse of the length measured from the point $O$ to the line of zero strain.

We next relate the strain $\epsilon_x(y)$ to the stress $\sigma_x(y)$ and change in temperature $\Delta T$. Assuming the thermal expansion is linear for each layer, the total strain $\epsilon_x(y)$ after temperature change $\Delta T$ is given by

$$\epsilon_x(y) = \frac{\sigma_x(y)}{E(y)} + \alpha(y)\Delta T \tag{2}$$

where $\alpha(y)$ and $E(y)$ are piecewise constant functions such that $\alpha(y) = \alpha_j$, and $E(y) = E_j$ when $y$ is in layer $j$. Note that we assume here that the thermal and mechanical strains are additive.

The stress $\sigma_x(y)$ must obey force and moment balance on each cross-section. Combining equations (1) and (2) and solving for $\sigma_x(y)$ yields

$$\sigma_x(y) = (\epsilon_0 - k(x)y - \alpha(y)\Delta T)E(y).$$

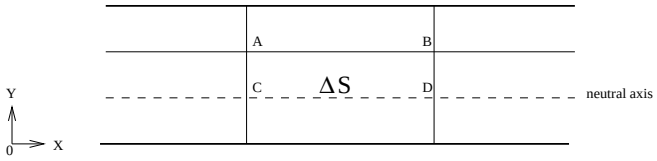



Figure 2. Composite Material Before Bending.

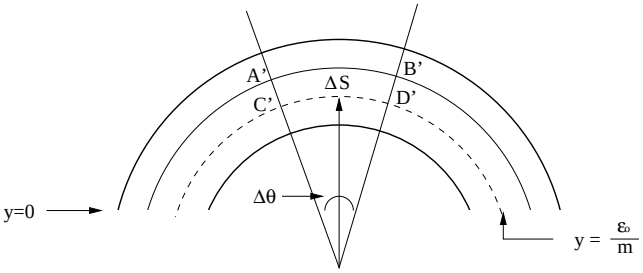



Figure 3. Composite Material After Bending.

Force balancing on cross-section $A_x$ requires that

$$\begin{aligned} 0 &= \int\int_{A_x} \sigma_x(x,y)\,dA \\ &= \sum_{j=1}^{n} \int_{y=h_{i-1}}^{h_i} \int_{z=0}^{w_i} [(\epsilon_0 - k(x)y - \alpha_j \Delta T) E_j]\,dydz \\ &= \sum_{j=1}^{n} E_j w_j \left[ (h_j - h_{j-1})(\epsilon_0 - \alpha_j \Delta T) - \frac{1}{2} k(x)(h_j^2 - h_{j-1}^2) \right]. \end{aligned}$$

Similarly, moment balancing yields

$$\begin{aligned} 0 &= \int\int_{A_x} y\,\sigma_x(x,y)\,dA \\ &= \sum_{j=1}^{n} \int_{y=h_{i-1}}^{h_i} \int_{z=0}^{w_i} E_j y(\epsilon_0 - k(x)y - \alpha_j \Delta T)\,dydz \\ &= \sum_{j=1}^{n} E_j w_j \left[ \frac{1}{2}(h_j^2 - h_{j-1}^2)(\epsilon_o - \alpha_j \Delta T) - \frac{1}{3} k(h_j^3 - h_{j-1}^3) \right]. \end{aligned}$$

These equations yield the following linear system which can be easily solved for $\epsilon_0$ and $k(x)$ using Maple, Matlab or C.

$$A \cdot \begin{bmatrix} \epsilon_0 \\ k(x) \end{bmatrix} = B$$

where the $2 \times 2$ matrix $A$ and the vector $B$ are given by

$$\begin{aligned} A &= \sum_{j=1}^{n} E_j w_j \begin{bmatrix} (h_j - h_{j-1}) & -\frac{1}{2}(h_j^2 - h_{j-1}^2) \\ \frac{1}{2}(h_j^2 - h_{j-1}^2) & -\frac{1}{3}(h_j^3 - h_{j-1}^3) \end{bmatrix} \\ B &= \sum_{j=1}^{n} E_j w_j \alpha_j \Delta T \begin{bmatrix} (h_j - h_{j-1}) \\ \frac{1}{2}(h_j^2 - h_{j-1}^2) \end{bmatrix}. \end{aligned}$$

### 3.3 The Relationship between the Theoretical and Experimental Radius of Curvature

The radius of curvature for a THUNDER actuator must be explicitly defined. In this section, we describe the geometry of typical THUNDER actuators, derive and define their radii of curvature, and describe the measurements that need to be taken to compute those radii. Typical actuators consist of several top layers of equal length, with a longer bottom layer as depicted in Figure 4. In our model, the top layers extend the length of actuator that is curved, and the remaining backing material (called tabs) on either side of the top layers remain unbent. In Section 3, we assume that the curved section of the actuator admits a circular arc. Therefore, we define the radius of curvature of a THUNDER actuator to be the radius of the circle corresponding to the circular arc maintained by the bottom boundary of the actuator's center layers (See Figure 5).

This radius of curvature, along with the tab-length, characterizes the total geometry of the actuator. Considering Figure 5, the horizontal length, $l$, and the vertical height, $h$, of the circular arc are all one needs to *measure* the curvature of a pre-stressed actuator. Using the constant radius $r$ of the circular arc, simple analysis yields the relation

$$r = \frac{h}{2} + \frac{l^2}{8h}\,. \tag{3}$$

FACE International company calculates the radius of curvature using (3).

Originally, radii of curvature were calculated by relation (3) using the *total* dimensions of the actuator with tabs, $l'$ and $h'$. With the radii calculated in this way, one is treating the entire actuator as one circular arc. Nevertheless, our model *attempts to* predict the radius of curvature we have defined here. It may not be feasible for the manufacturer to make direct measurements of $h$ and $l$ accurately from an actuator with tabs. To correct for this discrepancy, we derived an expression that relates the reported radius of curvature $r'$ to the model-predicted radius $r$, the center (top layers) length $s$, and the tab-lengths $t$. The manufacturer reported

$$r' = \frac{h'}{2} + \frac{l'^2}{8h'}\,. \tag{4}$$

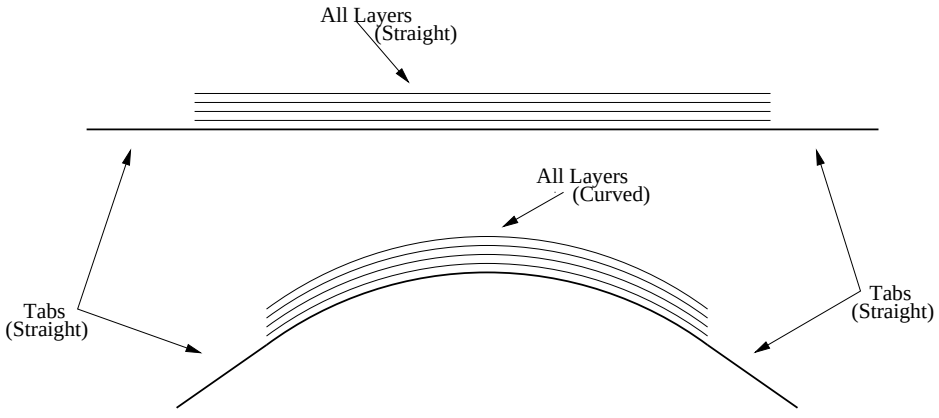



Figure 4. Non-bent actuator with thicknesses.

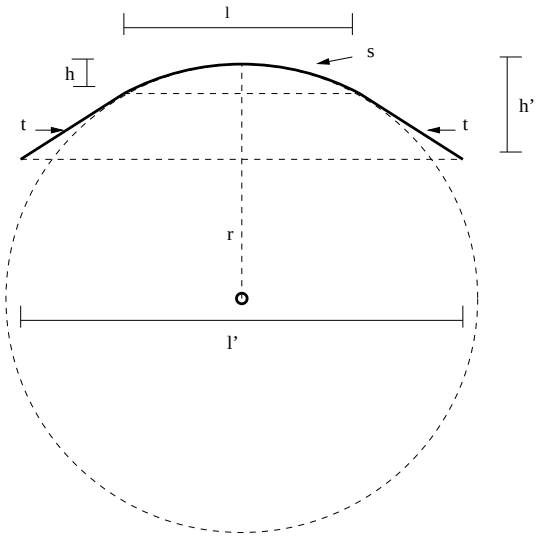



Figure 5. Pre-stressed actuator: radius of curvature.

Using expressions for $l'$ and $h'$ in terms of our given parameters $r$, $s$, and $t$ we get

$$r'(r,s,t) = \frac{t^2 + 2r^2 - 2r^2 cos(\frac{s}{2r}) + 2trsin(\frac{s}{2r}) - trsin(\frac{2}{r}) + tsin(\frac{2}{r})}{r + tsin(s/2r) - rcos(\frac{s}{2r})}\,. \tag{5}$$

Using (5), we can compare our predicted radii of curvature with sets of data consisting of $r'$. Figure 6 shows the relative difference between $r$ and $r'$ versus $r$in cm. With typical radii ranging from 5 cm-20 cm, the reported radii are 19.1%-19.5% (+-0.01%) larger than the defined radii of curvature.

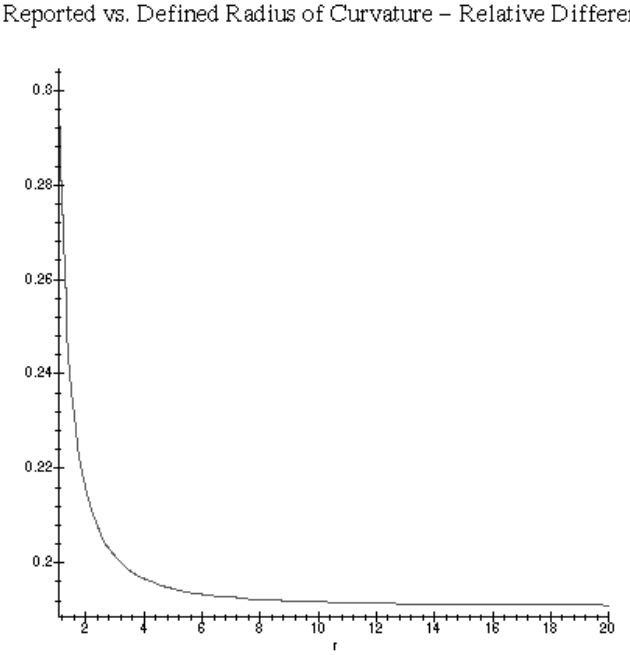



Figure 6. Relative difference between $r$ and $r'$.

Our defined radius characterizes the geometry in a way that is useful to the manufacturer in spite of the presence of tabs. We note that the tabs exist primarily to provide a means to install the actuator in the product for which it was designed. (For instance, one would clamp the actuator at the tabs.) Then, it is the curved center section that the manufacturer should design to fit the product and it is this curvature that our model attempts to predict.

# 4 Model Testing

In order to test the validity of our model, we obtained data from FACE Corporation. The $2 \times 2$ linear system was solved using a Maple program. Table 1 summarizes the values that were used for the coefficients of expansion and the modulus of elasticity [1]. Note that the coefficient of thermal expansion for LaRC-SI is given in two ranges. Although our model did not account for this possibility, only the thermal strain term $\alpha T$ appears in our formulas, so we simply replace this with the sum of the strains for the LaRC-SI over the two subranges. No specification was found for temperatures over $200^o\,C$ so we assume that the given coefficient simply extends to $270^o\,C$. With the values of the other constants as shown, the thermal expansion of the LaRC-SI layers is almost completely irrelevant to the output of the model.

The values that our model predicted are summarized in Table 2. As can be noted, the values vary only slightly from the measured values. As noted in Section 3, we originally did not take into account the side tabs. The adjusted values for the radius of curvature can be found in parentheses below the values obtained from our model. A comparison between the adjusted and experimental values illustrates that a slight difference still occurs. Further investigation indicated that the THUNDER actuators are subject to a repoling process to realign electric dipoles in the PZT ceramic wafer. FACE corporation reported an approximately 20% decrease in height after this process which will in turn cause an increase in the radius of curvature. From our initial results, there appears to be a 2% to 10% increase in the radius of curvature. Unfortunately, our model does not account for this discrepancy and further model development is necessary.

| Materials | Coefficient of Expansion ($10^{-6}/^oC$) | Modulus of Elasticity (GPa) |
|---|---|---|
| Aluminum 3003 | 24.00 | 68.95 |
| Stainless Steel 304 | 17.30 | 193.10 |
| PZT 3195HD | 3.00 | 67.00 |
| LARC-SI | 52.86 | 3.45 |

Table 1. Values Of constants for specific materials.

| Layer | Materials | Thickness (cm) | Radius of Curvature Model (cm) | Radius of Curvature Experimental (cm) |
|---|---|---|---|---|
| Layer 1 | Stainless Steel 304 | 7.62E-03 | 8.44 | 17.02 |
| Layer 2 | LaRC-SI | 7.62E-03 | (10.28) | |
| Layer 3 | PZT 3195HD | 2.032E-03 | | |
| Layer 4 | LaRC-SI | 7.62E-03 | | |
| Layer 1 | Stainless Steel 304 | 1.524E-02 | 14.17 | 17.78 |
| Layer 2 | LaRC-SI | 5.08E-03 | (17.12) | |
| Layer 3 | PZT 3195HD | 2.032E-03 | | |
| Layer 4 | LaRC-SI | 5.08E-03 | | |
| Layer 5 | Aluminum 3003 | 2.54E+00 | | |
| Layer 1 | Stainless Steel 304 | 2.54E-02 | 17.23 | 12.7 |
| Layer 2 | LaRC-SI | 3.81E-03 | (21.17) | |
| Layer 3 | PZT 3195HD | 3.81E-02 | | |
| Layer 4 | LaRC-SI | 3.81E-03 | | |
| Layer 5 | Aluminum 3003 | 2.54E-03 | | |
| Layer 1 | Stainless Steel 304 | 2.032E-02 | 13.46 | 17.78 |
| Layer 2 | LaRC-SI | 3.81E-03 | (16.16) | |
| Layer 3 | PZT 3195HD | 2.54E-02 | | |
| Layer 4 | LaRC-SI | 3.81E-03 | | |
| Layer 5 | Aluminum 3003 | 2.54E-04 | | |
| Layer 1 | Stainless Steel 304 | 1.524E-02 | 10.71 | 14.73 |
| Layer 2 | LaRC-SI | 3.81E-03 | (12.97) | |
| Layer 3 | PZT 3195HD | 2.032E-02 | | |
| Layer 4 | LaRC-SI | 3.81E-03 | | |
| Layer 5 | Aluminum 3003 | 2.54E-03 | | |

Table 2. Theoretical and experimental results.

# 5 Conclusions and Further Research

Although our model appears to fit the experimental data fairly well, there are still several issues that should be addressed before we can consider the model complete. First and foremost, our analysis ignores some physical effects which contribute to the final curvature of these devices. For example, at the end of the manufacturing process, THUNDER devices undergo a repoling step which realigns the electric dipoles of the PZT and this step has been observed to cause significant changes in the measured dimensions of these devices.

In addition to considering these physical effects, we also encountered some difficulties in comparing our computed results to the experimental data. Primarily, we felt that the number of usable data points provided was inadequate to properly guage the accuracy of our model. Furthermore, the presence of tabs on the ends of the measured devices caused some discrepancy between the reported and the actual curvatures (see Section 3.3).

In order to make the next step in analyzing the curvature of these devices, we would like to make the following recommendations to our colleagues at FACE:

- In order to adjust our model for the effects of poling, the following data should be collected: curvatures reported before and after poling, height and length data taken before and after poling, data on the deformation of PZT due to poling (stress, strain, and elasticity changes).

- In order to test our model properly, we need more data, preferably from a systematic set of experiments that investigate the dependence of curvature on the various material parameters (number of layers, thicknesses, elastic moduli, thermal expansion coefficients, etc.) This data would give us a better idea of how well our model matches physical reality.

Finally, in order to make this model more useful in the custom design process, we need to design an algorithm to the solve the inverse problem. That is, given a specific curvature, we need an algorithm that will predict the range of material properties that will produce the curvature. Possible algorithms for this problem include direct search or various iterative algorithms such as Newton's method or a secant method.

### Acknowledgements

We would like to thank our faculty advisors Professors R.C. Smith and H.T. Banks for their expert advice and nearly infinite patience. We would also like to thank the problem presenters from FACE International Corporation, the workshop organizers and the staff at NCSU for their time and assistance.

## References

[1] R.G. Bryant, "$LaRC^{TM}$-SI: a soluble aromatic polymide," *High Performance Polymers*, **8**, 607-615, 1996.

[2] G. Wempner, *Mechanics of Solids,* Boston: PWS Publishing, 1995.

# PROBLEM 3: MATHEMATICAL MODELING OF HIGH POWER AMPLIFIER (HPA) AND HPA LINEARIZATION THROUGH A PREDISTORTER*

Philippe Bremond[1], Danielle Bundy[2], Timothy Chevalier[3], Simone Flory[4], Brian Lewis[5], Sumithra Ramaswamy[6]

Problem Presenter:
Tien M. Nguyen
The Aerospace Corporation

**Abstract**

High Power Amplifiers (HPA's) are used by the Global Broadcasting Service (GBS) in order to boost the signal strength of transmitted data between satellite and ground-based stations. It is known that the use of these amplifiers induces distortion of the original signal before transmission. This is, in fact, a consequence of the non-linear behavior of the amplifier in the saturation region (point of maximum power transfer). In particular, both the amplitude and phase of the signal are distorted, thereby compromising the integrity of the signal. This new signal exhibits both spectral re-growth (added signal spectra outside of the desired bandwidth) and increased Bit Error Rate (BER). To compensate for these undesired effects, it is necessary to predistort the signal so that the resultant signal, when amplified, has the same power spectrum as the original signal after baseband filtering. In effect, the predistorter 'linearizes' the output of the HPA. This simply means that there is a one-to-one correlation between the input and output power. This paper describes the implementation of two-types of predistorters. The first is a non-adaptive predistorter mechanism for an HPA with known characteristics. The second is an adaptive predistorter for an HPA with unknown characteristics.

**Key words.** HPA, GBS, Predistorter, Filtering, Adaptive, Amplifier

## 1 Introduction

The Global Broadcasting Service (GBS) is involved with many aspects of our everyday lives from television broadcasts to military applications. Of major concern in the communications industry is the reliability in transferring information from one location to another. When dealing with satellite to ground based communications, for example, high power amplifiers, such as Solid State Power Amplifiers (SSPA) or Traveling Wave Tube Amplifiers (TWTA), are needed to provide the transmitted signal with ample signal strength to account for signal attenuation due to path loss. Amplifying the signal to levels required for this type of long-range transmission induces distortion in the original

*This work was supported in part by the NSF grant DMS-9704919, the Center for Research in Scientific Computation, and the Department of Mathematics at North Carolina State University.

[1]University of New Orleans
[2]University of Colorado
[3]University of New Hampshire
[4]Louisiana State University
[5]North Carolina State University
[6]Ohio State University

signal. This distortion, which are of two types and are known as AM-to-AM (amplitude distortion) and AM-to-PM (phase distortion), is one of the major concerns to communication systems engineers. These effects can cause intermodulation (IM) components and spectral regrowth, which are undesirable to system designs. For example, the intermodulation components and the spectral regrowth can cause adjacent channel interference to other services in addition to a loss in power transmission. Finally, the effects of AM-AM and AM-PM can cause the signal distortion that can degrade the Bit Error Rate (BER) performance.

Since cost is an issue in most engineering applications, a low cost mechanism for dealing with this distortion is desired. Working with the signal while it is on the satellite is simply impractical due to the logistics involved. Also, since the signal is transmitted at very high carrier frequencies (at least on the order of 106 Hertz) to keep antenna size down, any signal processing at the modulated frequency would be very costly. As a result, signal processing is performed at the ground-based stations using baseband frequencies. The actual technique used in correcting the non-linear behavior of the HPA involves signal predistortion. Predistortion involves using known properties of the HPA in order to develop an algorithm for predistorting the original signal. This is done so that when the predistorted original signal is passed through the HPA and finally transmitted, its power spectrum should look identical to that of the original signal before amplification. In essence, the predistorter corrects for the non-linear behavior of the HPA so that the resulting amplification is a linear process.

Currently, there are a few predistorter models that correct for the non-linear behavior exhibited by the HPA to varying degrees. The models studied by the GBS community include Saleh's Model, the Extended Saleh's Model, the Linear-Log Model, and the Modified Linear-Log Model.

It is important to note that the use of each of these models requires known characteristics of the HPA in order to design a successful predistorter. These schemes are deemed non-adaptive in nature because of this necessary prior knowledge of the HPA. The problem inherent with these non-adaptive schemes is that they have no way to account for changes in the characteristics of the HPA. In practical problems of interest, it is highly probable that the characteristics of the HPA may change with any number of known or unknown variables such as time or ambient temperature. As a consequence, these methods would provide erroneous results under varying HPA operating conditions.

It was the goal of this research to develop a predistorter that would not require previous knowledge of the HPA characteristics. The predistorter would use the changing characteristics of the HPA as defined by previous signal measurements and adapt to the incoming signal. Hence, it would be an "adaptive predistorter." The group was faced with the challenge of modeling a basic communication channel incorporating the elements of baseband signal filtering, modulation and amplification and then incorporating a reliable adaptive pre-distortion algorithm to account for variable HPA operating conditions.

The organization of this report is as follows. First, in Section 2, a brief background of the Saleh and the extended Saleh methods are presented. The adaptive predistorter, which is based on a variation of the Linear-Log model, is discussed in Section 3.1. Sections 3.2 and 4 discuss the implementation of the adaptive predistorter in the physical model, the Global Broadcasting Service, and in a simplified simulation model, respectively. Implementation issues and numerical results are presented in Section 5. Finally, Section 6 contains our concluding remarks.

# 2 Saleh's and Extended Saleh's Predistorter Model

## 2.1 Saleh's Model

A simplified schematic diagram for the HPA with a predistorter is shown in Figure 1. The basis behind Saleh's model is that for an HPA of known characteristics (given by a set of discrete measured

data), an interpolation technique is applied to curve-fit the data in an effort to develop a simple functional relationship between the predistorter input and output signals [1].

Let $X(t)$ denote the input signal to the predistorter (PD), $Y(t)$ denote the output of the PD, and $Z(t)$ denote the output of the HPA. The baseband representations of these signals are:

$$X(t) = p_x(t)e^{j\theta_x(t)} \tag{1a}$$

$$Y(t) = p_y(t)e^{j\theta_y(t)} \tag{1b}$$

$$Z(t) = p_z(t)e^{j\theta_z(t)}\,. \tag{1c}$$

The signal $Z(t)$ may then be written in the form

$$Z(t) = M(p_y(t))e^{j(\theta_y(t)+\varphi(p_y(t))} \tag{2}$$

where $M(p_y(t))$ and $\varphi(p_y(t))$ are the normalized AM-AM and AM-PM responses of the HPA due to the input signal $Y(t)$. The Saleh's model is extended for HPA by including four extra parameters $(a_0, a_1, b_0, b_1)$ resulting in the following equations for the normalized AM-AM and AM-PM responses:

$$M(p_y(t)) \equiv \rho_z(t) = \frac{\alpha_0 p_y(t)}{a_0 + \beta_0(p_y(t)+b_0)^2} \tag{3a}$$

$$\varphi(p_y(t)) = \frac{\alpha_1 p_y^2(t)}{a_1 + \beta_1(p_y(t)+b_1)^2}\,. \tag{3b}$$

Here, the unknown parameters, $a_0, b_0, \alpha_0, \beta_0, a_1, b_1, \alpha_1, \beta_1$ are the coefficients which will be computed to give the best fit to the measured data.

Substituting the equations (1b), (3a) and (3b) into (2), we obtain the following equation for the output of the HPA

$$Z(t) = \left[\frac{\alpha_0 Y(t)}{a_0 + \beta_0(p_y(t)+b_0)^2}\right] e^{j\left[\frac{\alpha_1 p_y^2(t)}{a_1+\beta_1(p_y(t)+b_1)^2}\right]}\,. \tag{4}$$

From equation (4), the ideal PD output for a giving input $X(t)$ is of the form:

$$Y(t) = X(t)\left[\frac{a_0 + \beta_0(p_y(t)+b_0)^2}{\alpha_0}\right] e^{-j\left[\frac{\alpha_1 p_y^2(t)}{a_1+\beta_1(p_y(t)+b_1)^2}\right]}\,. \tag{5}$$

Once again, it is important to note that these equations are only useful for an HPA with known characteristics since the curve fitting parameters depend directly on the measured data for the HPA.

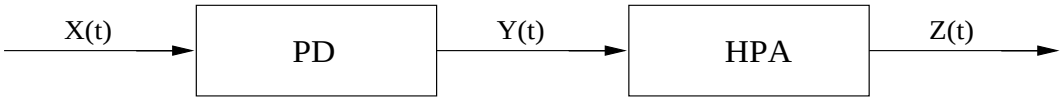



Figure 1. Simplified Flow Diagram for PD and HPA System

### 2.2 Extended Saleh's Model

A technique for modeling the PD is based on equation (5) for $Y(t)$ , which is an extension of Saleh's model. Substituting equation (1a) into equation (5) we obtain the following expressions for the amplitude and phase of the signal at the output of PD:

$$p_y(t) = \left[\frac{a_0 + \beta_0(p_y(t) + b_0)^2}{\alpha_0}\right] p_x(t) \tag{6}$$

$$\theta_{y(t)} = \theta_{x(t)} - \left[\frac{\alpha_1 p_y^2(t)}{a_1 + \beta_1(p_y(t) + b_1)^2}\right]. \tag{7}$$

Solving equation (6) for $p_y(t)$ and selecting only the negative sign for the square-root term in the numerator, we have:

$$p_y(t) = \frac{[\alpha_0 - 2b_0\beta_0 p_x(t)] - \sqrt{\alpha_0^2 - 4\beta_0 p_x(t)[a_0 p_x(t) + b_0\alpha_0]}}{2\beta_0 p_x(t)}. \tag{8}$$

Note that because of the normalized input and output relationship, the negative sign is chosen to meet the dual conditions that the square root term is real and $0 \le p_y(t) \le 1$.

If we let the amplitude and phase of the ideal PD be $p_{PD}$ and $\theta_{PD}$, respectively, then the output $Y(t)$ of the PD can be rewritten as:

$$Y(t) = X(t)p_{PD}e^{j\theta_{PD}} = p_x(t)p_{PD}e^{j(\theta_x(t)+\theta_{PD})} = p_y(t)e^{j\theta_y(t)}. \tag{9}$$

Substituting equation (7) and (8) into equation (9) and equating the amplitude and phase terms separately for the ideal PD we obtain

$$p_{PD} = \frac{p_y(t)}{p_x(t)} = \begin{cases} \frac{[\alpha_0 - 2b_0\beta_0 p_x(t)] - \sqrt{\alpha_0^2 - 4\beta_0 p_x(t)[a_0 p_x(t) + b_0\alpha_0]}}{2\beta_0 p_x^2(t)} & \text{if } p_x(t) \le 1 \\ 1 & \text{if } p_x(t) > 1 \end{cases} \tag{10}$$

$$\theta_{PD} = -\left[\frac{\alpha_1 p_y^2(t)}{a_1 + \beta_1(p_y(t) + b_1)^2}\right]. \tag{11}$$

Note that the saturation condition for HPA is implied in equation (10).

Both the Saleh Model and it's extended version are useful if the characteristics of the HPA are known. Both Saleh models provide closed form solutions for the required value of the predistorter with respect to a given input signal. However, these methods lack the versatility needed for an adaptive scheme.

## 3 Mathematical Modeling of the Adaptive Predistorter

### 3.1 The Linear-log Model

For this model, we consider the three sets of measured data that are provided for the High Power Amplifier (HPA); namely, the AM-AM curve of the normalized input and output power (both in dB), the AM-PM curve of the normalized input power (in dB), and output phase (in degrees). Using these data, we derive the corresponding AM-AM and AM-PM curves for the PD. These have to guarantee that after combining the HPA and the PD, the input signal $X(t)$ and the output signal $Z(t)$ are equal.

Let $P_{in}$(dB) be the normalized input power in the AM-AM and AM-PM curves for both HPA and PD, $P_{outHPA}(dB)$ and $P_{PD}(dB)$ be the normalized output powers in the AM-AM curves for the HPA and the PD, respectively, and $\theta_{HPA}$(degree) and $\theta_{PD}$(degree) be the output phases in the AM-PM curves for HPA and PD. Then, the PD must satisfy the following conditions:

**(i)** $P_{PD}(\text{dB}) = P_{in}(\text{dB}) - P_{outHPA}(\text{dB})$ for $P_{in}(\text{dB}) \leq 0\text{dB}$

**(ii)** $\theta_{PD}(\text{degree}) = -\theta_{HPA}(\text{degree})$

**(iii)** $P_{PD}(\text{dB}) = -P_{outHPA}(\text{dB})$ for $P_{in}(\text{dB}) > 0$ dB

Figure 2 illustrates the amplitude characteristics of the HPA from a real data set provided to us by Dr. Tien. It is noted that the graph starts out appearing to be linear. However, once the output power approaches the saturation point, the graph loses its linearity. When the input power is graphed versus the output phase (see Figure 3), the graph lacks linearity in the beginning but regains it toward the end.

Because of the non-linearity of the output power and phase, a predistorter is needed to linearize the signal. Figure 4 represents the simple block diagram for the implementation of the adaptive predistorter into a physical model, which is the Global Broadcasting Service (GBS). We note that the adaptive predistorter is receiving information from the signal before and after the amplification. Information from the signal after it has been amplified is needed so that the predistorter can adjust and be used to predistort the next incoming filtered signal. Thus, the predistorter adapts to the signals that are being processed by the HPA. For this information to be attainable the signal has to be down-converted a number of times before entering the predistorter in order for the APD to be able to use the signal effectively.

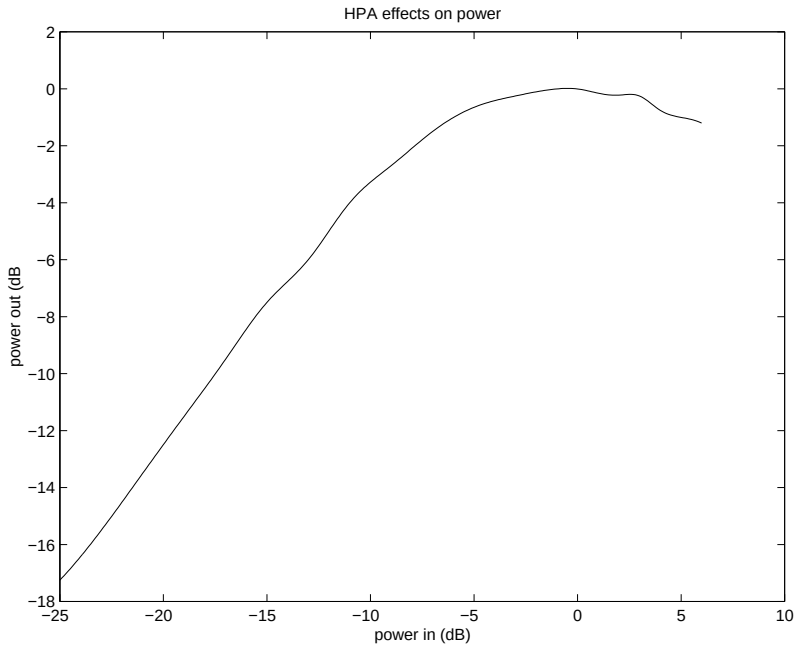



Figure 2. Input power versus output power (AM-AM).

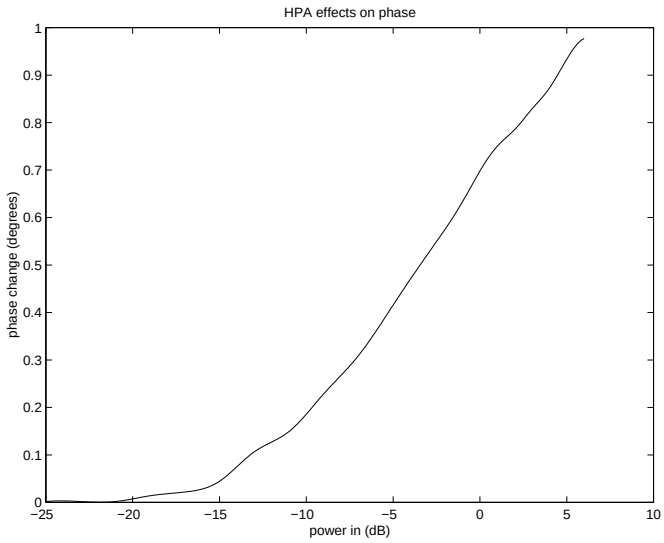



Figure 3. Input power versus output phase (AM-PM).

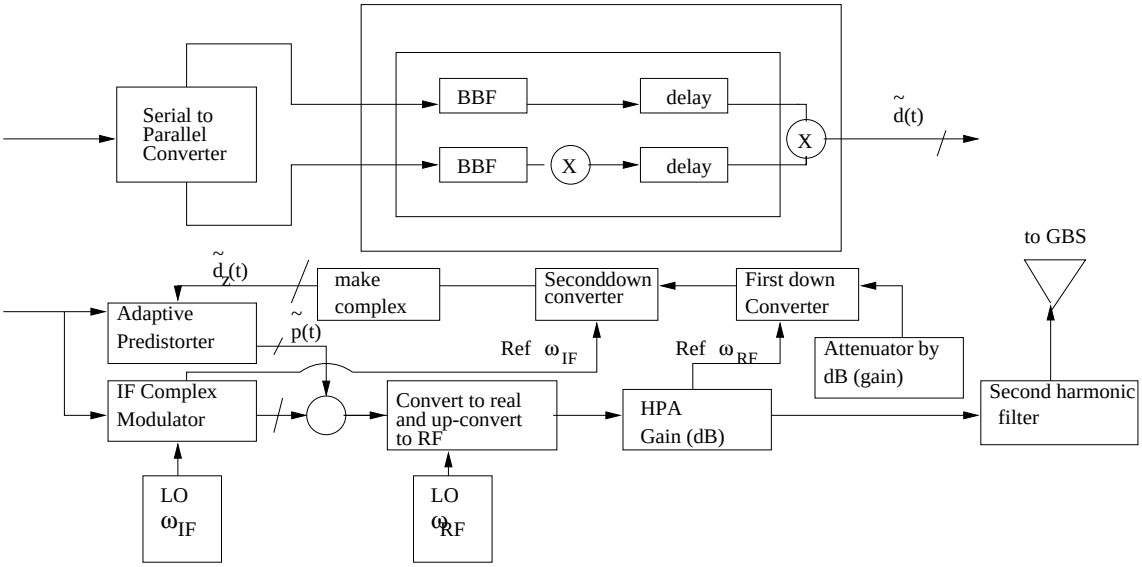



Figure 4. Simplified block diagram for the implementation of the ground transmitter with predistorter circuitry.

Figure 5 represents the flow of events surrounding the adaptive compensator. The adaptive compensator predistorts the signal using the phase and power estimators (Figure 6) and the signal is then brought back together. The input and output power estimators take the input and output power (respectively), and divide the instantaneous power with the average power. The phase estimators take the input and output signal and convert them to their respective I and Q parts (see Figure 6). Figure 6 represents what is occurring in the specific phase and power estimators. oOnce the input and output power is estimated, as well as the input and output phase, these quantities are passed on to the adaptive compensator. The processes performed by the adaptive compensator on the signals it receives are described in Figure 7.

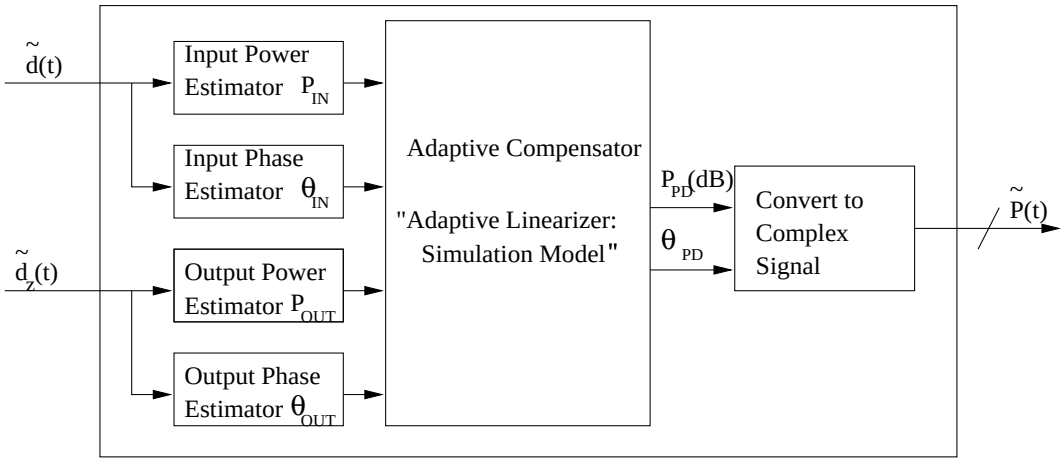



Figure 5. Simplified diagram for the adaptive predistorter.

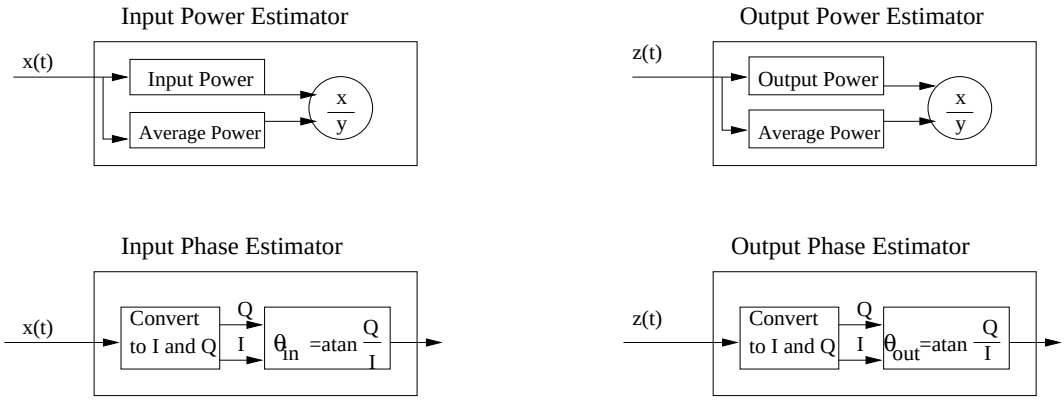



Figure 6. Block diagram for the input/output power and phase estimators.

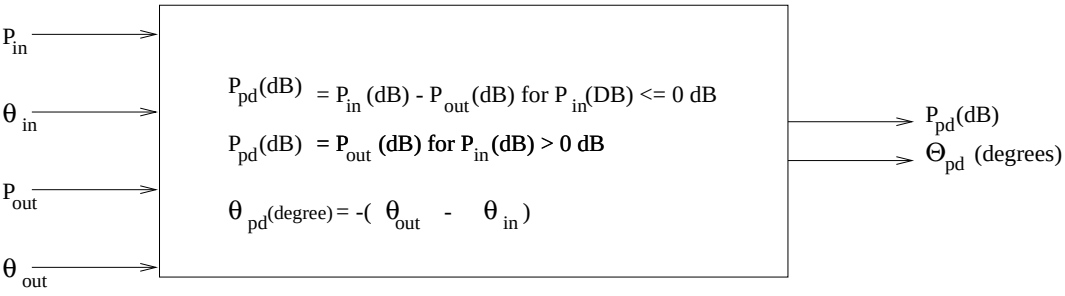



Figure 7. Description of the adaptive compensator.

### 3.2 Implementation of the Adaptive Predistorter in the GBS

The preceding diagrams in Section 3.1 give an idea of what is occurring in the predistorter. This section details the steps taken in the physical model explicitly in a step by step mathematical manner.

Step 1:

Convert the audio and (or) video signal from analog to digital and then convert the resultant digital signal from serial to parallel. This consequently improves the throughput of the channel by a factor of 2.

Step 2:

The two parallel signals: the in-phase binary signal $d_i(t)$ and the quadrature binary signal $d_q(t)$ pass through a complex baseband filter with filtering operator $\Im$. A delay $T_d$ is then created. The in-phase signal is multiplied by the cosine of the carrier frequency $\omega_c$ and the quadrature signal by its sine. Therefore, the two signals are perpendicular and do not interfere with each other. The Quadratic Phase Shift Keying is filtered to put the signal in Digital Video Broadcasting waveform. Then, they are combined to create a complex signal:

$$\Im[\tilde{d}(t)] = \Im[\tilde{d}_i(t - T_d)] + j\Im[\tilde{d}_q(t - T_d)], j^2 = -1$$

where $\tilde{d}_i(t - T_d) = d_i(t - T_d)\cos(w_c(t - T_d))$ and $\tilde{d}_q(t - T_d) = d_q(t - T_d)\sin(w_c(t - T_d))$.

Step 3:

Inside the complex modulator, the signal $\Im[\tilde{d}_{in}(t)]$ is multiplied by

$$e^{-j\omega_{if1}(t)}.$$

Then it passes through the up-converter. Let

$$S(t) = Re\Im[\tilde{d}_{in}(t)] = \Im[d_i(t - T_d)]\cos(\omega_{if1}(t)) + \Im[d_q(t - T_d)]\sin(\omega_{if1}(t))\,.$$

Because the imaginary part does not contain any information about the signal, it is filtered.

Step 4:

After passing through the up-converter with local oscillation $\omega_{if2}$ the signal is converted back to a real signal

$$S_2 = \Im[d_i(t - T_d)]\cos(\omega_{if1}(t) + \omega_{if2}(t))/2 + \Im[d_i(t - T_d)]\cos(\omega_{if1}(t) - \omega_{if2}(t))/2$$

$$+\Im[d_q(t - T_d)]\sin(\omega_{if1}(t) + \omega_{if2}(t))/2 + \Im[d_q(t - T_d)]\sin(\omega_{if1}(t) - \omega_{if2}(t))/2\,.$$

The upper side band is defined with the term $\omega_{if1}(t) + \omega_{if2}(t)$ and the lower side band with the term $\omega_{if1}(t) - \omega_{if2}(t)$. Using an image filter, only the upper side band of the signal is kept. The coefficient $1/2$ is absorbed by amplification.

This then yields the relation

$$\begin{aligned} S_T(t) &= S_2\sqrt{2Pt} \\ &= \sqrt{2Pt}\,\Im[d_i(t+T_d)]\cos(\omega_{if1}(t)) + \omega_{if2}(t)) \\ &\quad +\sqrt{2Pt}\,\Im[d_q(t+T_d)]\sin(\omega_{if1}(t)) + \omega_{if2}(t)) \\ &= \sqrt{2Pt}\,\Im[d_i(t+T_d)]\cos(\omega_{RF}(t)) + \sqrt{2Pt}\,\Im[d_q(t+T_d)]\sin(\omega_{RF}(t)) \\ &= \sqrt{2Pt}\,d_{IF}\cos(\omega_{RF}(t)) + \sqrt{2Pt}\,d_{QF}\sin(\omega_{RF}(t))\,. \end{aligned}$$

Step 5:

The signal $S_T(t)$ goes through the HPA and gains some power. The amplitude of the signal is distorted according to the characteristics of the HPA (F) and its phase is shifted by an angle $\Phi_0(t)$. That is,

$$|S_T(t)| = \sqrt{2Pt}\sqrt{d_{IF}^2 + d_{QF}^2}$$

and

$$\text{angle}[S_T(t)] = \omega_{RF}(t) + \Phi_{RF}(t),$$

where $\Phi_{RF} = \arctan\frac{d_{QF}(t)}{d_{IF}(t)}$.

Therefore the signal, $S_T(t)$, after passing through the HPA is given by

$$F|S_T(t)|e^{j[\omega_{RF}(t)+\Phi_{RF}(t)+\Phi_0(t)]}\,.$$

Step 6:

The signal, after passing through the complex modulator, the up-converter, and the HPA, has gained too much power to be filtered. Therefore, its power needs to be attenuated to be readable.

Step 7:

The signal follows its path through two down-converters. The first one has a local oscillation of $w_{if1}(t)$ and the second one has a local oscillation of $w_{if2}(t)$. After this process, the signal will be slower (lower frequency) and readable. The signal is converted to a complex signal $\tilde{d}_0(t)$. This conversion occurs to reduce the time the signal will spend inside the adaptive predistorter.

Step 8:

The signal $\tilde{d}_0(t)$ arrives at the adaptive predistorter with a distorted phase and amplitude. The adaptive predistorter follows the Log-Linear model to predistort the signal. Leaving the predistoter, $\tilde{p}(t)$ is sent back to the second up-converter with local oscillation $\omega_{if2}(t)$, so it can acquire a radio frequency.

Step 9:

This final signal goes through the HPA and the resulting signal, similar to the original one, passes through a second harmonic filter.

# 4 The Simulation Model

Due to the complexity in the GBS application and the time constraint imposed by the workshop, we implemented our idea on a simplified model, called the simulation model. Figure 8 depicts the algorithmic structure used for the simulation model:

The simulation model is a simplified version of the physical model. This is because it is not necessary to worry about modulating to a certain broadcasting frequency and the instruments can be modeled as ideal. The only 'imperfection' present in the model is the nonlinear response of the HPA. We create two random digital signals to represent the in-phase and quadrature signals that will be separated from the physical signal. These signals are filtered with a Butterworth filter, then combined and modulated into $X(t)$.

There are two steps in the simulation model: Design the HPA and the PD to check that the two distortions cancel each other out, and make the PD adaptive. The first algorithm has the original signal $X(t)$ distorted in the HPA to $Y(t)$. The PD compares both signals using the Linear-Log model and "predistorts" $Y(t)$ to become $Z(t)$. This is actually a "post-distortion" but the purpose of this model is to test that $X(t) = Z(t)$. In this model the PD should completely compensate for the distortion produced by the HPA.

In the second model the PD is truly adaptive. That is, $X(t)$ passes through the APD, which is initialized previously, then through the HPA. The APD gets $Y(t)$ from the HPA and $X(t)$, the original signal, as input, to adjust to the desired characteristic (again using the Linear-Log model). We now provide detailed descriptions of this second model.

The original signal $X(t)$ and the distorted signal $Y(t)$ are input to the APD. On the first run when $Y(t)$ is not available, a transparent PD is used and the signal is sent through the HPA to give a $Y(t)$. This isn't important to our results since we allow the predistorter to converge at the beginning. Both input signals are normalized with respect to the average of the mean power of the signal and the previous signals that have passed through. The number of signals used in the average is arbitrary (we used 10 signals for the simulation). The Linear-Log model is used to calculate the desired predistortion which is then combined with the original signal.

A consequence of the continual adaptivity of the predistorter is that there are 2 signals passing through it at any one step: One that was used to adapt the predistorter and needs to be mixed with the resulting predistortion, and one that will be used in the next step to estimate the changing characteristics of the HPA. A given signal passes through the APD, then through the HPA and back into the APD to adjust the predistorter. This predistortion must be mixed with the signal that was analyzed to maintain accuracy. However, at the same time as the APD is being adjusted, another signal passes through it at the beginning of the loop. It is important to keep these signals distinct and make sure that the correct signal is predistorted and sent out.

After the predistorter, the signal $X_{PD}(t)$ is passed to the HPA. Data for a TWTA is used for the characteristics of the HPA, but to model a changing HPA we multiply the data by a constant between .9 and 1.0, chosen randomly for each signal. The original data ranges from -25 dB to 6 dB. Below -25 dB, a linear relationship with a slope of one is assumed; the output power or phase is equal to the input for small input power. For input power between -25 dB and 0 dB (saturation), cubic spline interpolation is used to find the HPA output. Above 0 dB, the output is set to that of saturation (by definition).

For an actual broadcasting situation, the HPA should be running at or near saturation, but the input signals run at a variety of powers. To modify this, the input power is increased to saturation in the HPA. Then the difference between the original output power or phase and that of the HPA's characteristics at saturation is the actual distortion by the instrument. This distortion is added to the input signal, giving $Y_{PD}(t)$.

Recall that this signal was predistorted with the previous setting of the APD, then was used to calculate the true predistortion required. This new "predistortion" is be added to $Y_{PD}(t)$ to get the ideal signal $Z(t)$.

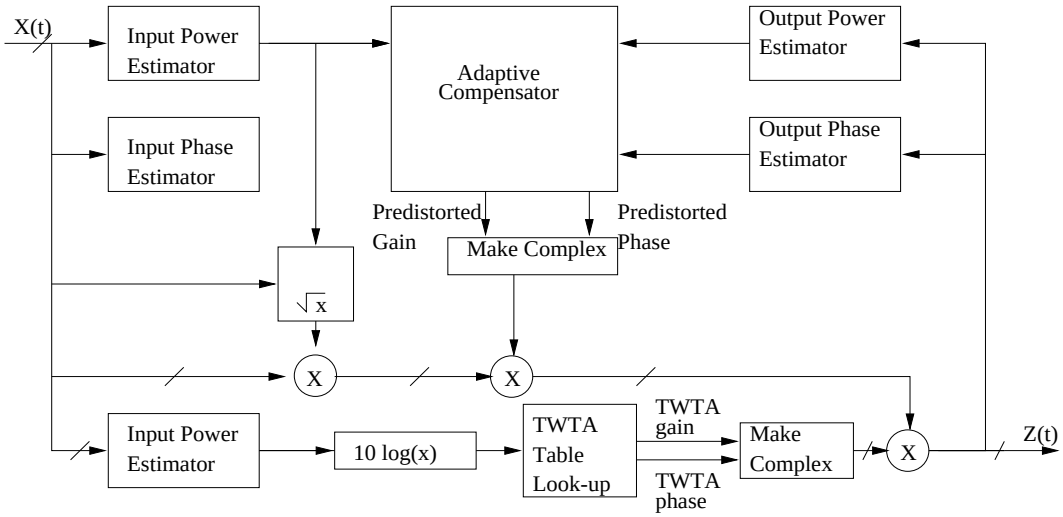



Figure 8. Simplified block diagram for the simulation model

# 5 Implementation and Numerical Results

## 5.1 Implementation

It was decided that the algorithm for the simulation model would be implemented in Matlab. The group came to this decision because of Matlab's extensive toolboxes. The program consists of a number of modules to perform the following procedures. Using the Uniform Random Number generator of Matlab, a signal is generated. This signal is then filtered using the Filter function. The filtered signal is then sent to a function that simulates the High Powered Amplifier. Conforming to the simulation model, we then passed the signal into the predistorter. This function took the original signal, along with the amplified signal and used cubic spline interpolation to "undo" what the HPA had done. The resulting signal should look exactly like the filtered signal.

## 5.2 Numerical Results

After completing the simulation program in Matlab, the group performed a number of test runs to insure that the simulation was functioning as expected. A result from the series of tests is given in Figure 9. To test, the program included statements to plot the power spectrum of the signal at various stages. As stated, the program first generates a random signal. This signal is then sent through the Matlab filter function. A plot of a sample filtered signal is given in Figure 9 titled "INPUT SIGNAL." It is then passed through the HPA. The output HPA signal is given in the graph in Figure 9 titled "Output of Signal after HPA." We note that the distortion has been introduced by the HPA. The signal passes through the predistor which uses the known numerical characteristics of the HPA and cubic spline interpolation to define the predistorter for the signal. The plot of the output signal using the predistorter is shown under the title "Output of Signal using Predistorter." This signal is the same as the original filtered signal which confirms the functionality of the predistorter.

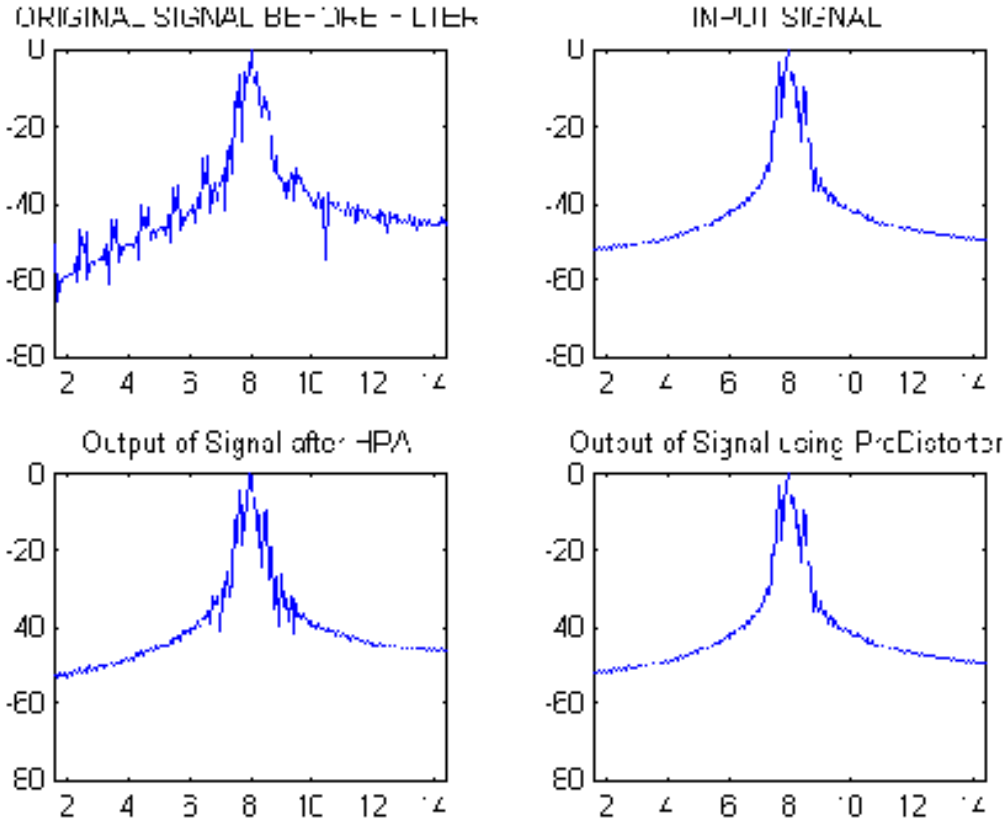



Figure 9. Numerical results of the simulation model.

## 6 Conclusions

The next natural step is to try and implement a truly adaptive simulation program. The group pondered this problem and actually tried to implement a simulation. The simulation, however, proved to be more of a task than time allotted. However, the basic outline is not so hard to explain (the implementation proved to be the difficulty). The concept is to run a randomly generated signal through the HPA a number of times (using slight variations on the values representing the HPA). During each pass, power and phase characteristics would be collected and retained about the HPA. These characteristics would then be averaged and used to predistort the next signal. This simulation will conform to the procedures described by the physical model. In the physical implementation, the characteristics of the HPA for each specific pass will be unknown by the predistorter. An average, however, will provide a sufficient set of values for the predistorter to sufficiently alter each successive signal.

### Acknowledgements

The time spent working on the project with Dr. Nguyen and Dr. Tran proved to be a very enriching experience. The group was able to see the entire process of mathematical modeling, from the algorithmic design all the way to programming a simulation. The workshop provided our group

with an excellent opportunity to explore a real world project. We would like to thank Dr. Nguyen and Dr. Tran for all of their time and help.

## References

[1] T.M. Nguyen, J. Yoh, A.S. Parker, D.M. Johnson, and H.T. Tran, "Modeling of HPA and HPA Linearization Through a Predistorter: Global Broadcasting Service Applications," *IEEE Trans. on Broadcasting,* submitted.

# PROBLEM 4: TIME DEPENDENT CONSOLIDATION OF FINE POWDERS

Karen Bliss[1], Linda Connolly[2], John Matthews[3], Shailesh Naire[4], Lakshmi Puthanveetil[5], Shannon Wynne[6]

Problem Presenter:
T. Antony Royal
Jenike & Johanson, Inc.

**Abstract**

The flow of granular materials has been a challenging mathematical endeavor since the advent of plasticity theory in the late 1950's. Jenike's original work in this field was built on foundations of continuum mechanics, plasticity theory, and soil mechanics. Advances in computer technology continually warrant a closer look at the mathematics that is the major limitation of rapid progress in this field. There are a lot of problems without adequate mathematical solutions. Some introductory material on stress and velocity fields in bins with flowing granular material will be presented as a precursor to the problem to be addressed.

The problem offered is time dependent consolidation of compressible fine powders. One application is predicting how air escapes from a powder when a bin or silo is filled and then allowed to deaerate with time. There is a parallel consolidation problem in soil mechanics that occurs in water saturated clays and a model is used to predict the settlement of foundations built on clay. For instance in Mexico City, settlement of up to 15 ft in some areas has occurred in the last 50 years due to building. Unfortunately powder/air behavior in a bin is much more complex than water/soil, so the mathematical models are very different. In large bins settlement time can be several days, in stockpiles it can be several weeks.

The original problem involves a Lagrangian frame of reference. The general three dimensional problem is too computationally intensive for a quick solution. Instead we focus on an axisymmetric problem to reduce the problem to two spatial and one time dimension. The problem will be considered first as a pseudo-2D spatial problem by introducing an approximation that effectively reduces the problem to one spatial dimension.

## 1 Introduction and Motivation

Jenike & Johanson is a specialized engineering firm which provides clients with solutions to bulk solids handling problems. One of the interests at Jenike & Johanson is computer modeling of the settlement of fine powders in bins. When aerated, fine powders behave like fluids, and so settlement properties become very important. The handling of fine powders presents some difficulties such as: flooding (uncontrolled flow), no-flow (occurs when deaerated), erratic flow, and so on. If a solid is not given enough time to settle, then flooding can occur. On the other hand, allowing solids to settle

[1]University of Missouri-Rolla
[2]University of New Orleans
[3]North Carolina State University
[4]University of Delaware
[5]University of Massachusetts-Amherst
[6]University of California-Irvine

too long may result in no-flow. The amount of time for a powder to deaerate depends on the fill rate of a silo/bin. Understanding the settlement of fine powders over time facilitates more efficient handling of fine powders.

The original problem was to model the settlement of fine powders in a bin or silo of arbitrary, but simple geometric shape (i.e., cylinder or cone). If possible, we were to model the settlement process as the bin/silo is filled. This is a three dimensional problem and quite complex. We first simplified the problem by working with a cylinder that is filled instantaneously. Using the axisymmetry of the bin, we reduced the problem to one dimension.

From physical intuition, we expect the following behaviour over time. Air escapes from the top of the bin, and the pressure decreases over time, with the pressure at the top being the atmospheric pressure. Stress on the solid increases, with zero stress at the top of the bin. Density also increases over time.

The remainder of this paper describes the model and several numerical approaches that were developed to understand the simplified one-dimensional problem. In Section 2 we define the dependent variables, independent variables, and parameters for the model. In Section 3 we describe the model, the model originally presented, and then a slightly different model. Numerical approaches are described in Section 4, and finally, numerical results are given in Section 5.

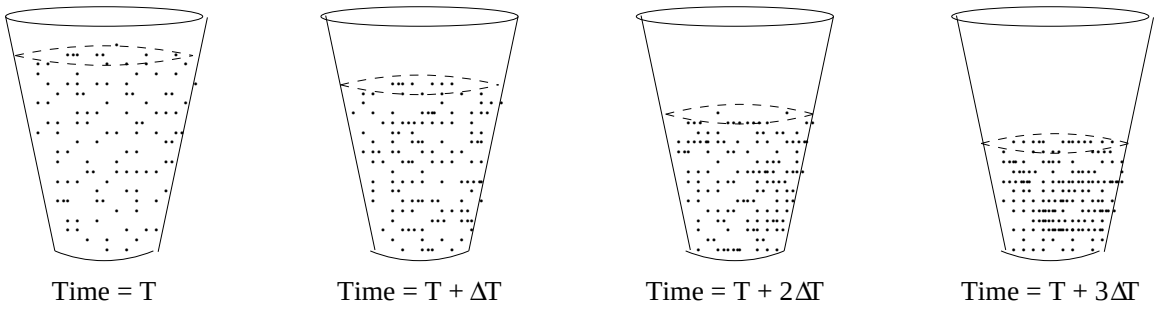



Figure 1. Compression of granular material over time

## 2 Definitions

**Independent Variables**

t = time in seconds
z = height in inches

**Dependent Variables**

p = gas pressure in interstices of granular solid measured in pounds per square inch
$\sigma$ = average vertical stress on horizontal cross-section measured in pounds per square inch
$\gamma$ = bulk weight density (weight per volume) measured in pounds per cubic inch
$u_g$ = velocity of the gas measured in inches per second
$u_s$ = velocity of the solid measured in inches per second
K = permeability of the solid
A = cross-sectional area at z
$\nu$ = porosity

**Parameters**

k = coefficient of wall friction
$\phi'$ = coefficient of internal friction
$\Gamma$ = particle density
H(t) = the height of the solid at time t measured in inches

# 3 Model Development

## 3.1 Original Model

At the beginning of the workshop, we were presented with a model where the stress equation is given in Eulerian coordinates, and the pressure equation is given in Lagrangian coordinates. The equation for stress $\sigma$,

$$\partial_z \sigma + \partial_z p + \sigma \left(A_2 - A_3\right) + \gamma = 0,$$

is a pseudo 2D equation with gas pressure gradients, which has area parameters:

$$\begin{aligned} A_2 &= \frac{1}{A}\frac{dA}{dz} \\ A_3 &= \oint k \left(\tan\theta + \tan\phi'\right)\, d\theta. \end{aligned}$$

The equation for pressure is a 1D isotropic permeability equation (similar to the diffusion equation),

$$\frac{1}{A}\partial_z \left(\frac{pAK}{\gamma}\partial_z p\right) = \nu \partial_t p - \frac{p}{\Gamma}\partial_t \gamma$$

where

$$\begin{aligned} \nu &= 1 - \frac{\gamma}{\Gamma} \\ \gamma(\sigma) &= \gamma_m \left(1 + \frac{\sigma}{\sigma_m}\right)^{\beta_m} \\ K &= K_o \left(\frac{\gamma}{\gamma_o}\right)^{-a}. \end{aligned}$$

These equation may be used to solve for the unknowns: pressure $p$ and stress $\sigma$ (or bulk density $\gamma$). This model will be referred to as model 1.

## 3.2 A Slightly Different Model

In order to have both equations in the same coordinate framework, we derived a new model that is very close to the model proposed originally. The coordinate system for this model is not material specific, but rather is an external coordinate system in which the material moves, i.e., an Eulerian framework.

The column of granular material will have a varying height given by $H(t)$, and conservation of mass in that column dictates that

$$\frac{d}{dt}\int_0^{H(t)} \gamma \, dz = 0$$

Furthermore, for a small element of the column we have the following relation involving the change in the amount of gas in that element

$$\begin{aligned} \frac{d}{dt}\int_V \left(1 - \frac{\gamma}{\Gamma}\right) \rho \, dV &= \int_S \rho \left(u_g \cdot n\right) \, dS \\ &= -\int_V \nabla \cdot \left(\rho u_g\right) \, dV. \end{aligned}$$

Since the choice of $V$ is arbitrary, we have an equation for the conservation of mass of the gas which in one dimension is

$$\partial_t \left[\left(1 - \frac{\gamma}{\Gamma}\right) \rho\right] + \partial_z \left(\rho u_g\right) = 0.$$

A similar conservation equation can be derived for the mass of the solid material,

$$\partial_t \gamma + \partial_z \left(\gamma u_s\right) = 0.$$

Combining these two partial differential equations with Darcy's law

$$u_g - u_s = -\frac{K}{\gamma} \partial_z p$$

and the stress equation

$$\partial_z \sigma + \partial_z p + \sigma \left(A_2 - A_3\right) + \gamma = 0$$

gives a system of five equations with five unknowns. Figure 2 shows the force balance, from which the stress equation was derived. Furthermore, the isothermal ideal gas relation

$$\frac{p}{p_0} = \frac{\rho}{\rho_0}$$

allows the system to be expressed in terms of pressure $p$ instead of $\rho$.

A first step in simplifying the system is to use Darcy's law to eliminate $u_s$. This results in the equation

$$\partial_t \gamma + \partial_z \left(\gamma u_g + K \partial_z p\right) = 0$$

We integrate this equation in $z$ to give

$$\begin{aligned} 0 = \int_0^z \left(\partial_t \gamma + \partial_{\tilde{z}} \left(\gamma u_g + K \partial_{\tilde{z}} p\right)\right) d\tilde{z} &= \int_0^z \partial_t \gamma \, d\tilde{z} + \left(\gamma u_g + K \partial_z p\right)_{\tilde{z}=0}^{\tilde{z}=z} \\ &= \int_0^z \partial_t \gamma \, d\tilde{z} + \gamma u_g(z) + K \partial_z p(z) \end{aligned}$$

Solving this for $u_g$ allows $u_g$ to be eliminated from the system and results in an equation similar to the original model with one extra term involving $\int_0^z \partial_t \gamma$. The new pressure equation is then

$$\left(1 - \frac{\gamma}{\Gamma}\right) \partial_t p - \frac{p}{\Gamma} \partial_t \gamma - \partial_z \left(\frac{p}{\gamma} \int_0^z \partial_t \gamma \, d\tilde{z}\right) - \partial_z \left(\frac{pK}{\gamma} \partial_z p\right) = 0.$$

For materials that are nearly incompressible, this extra term is small and may perhaps be neglected. For compressible material, this term is important. The resulting Eulerian system is as follows:

$$\left(1-\frac{\gamma}{\Gamma}\right)\partial_t p - \frac{p}{\Gamma}\partial_t\gamma - \partial_z\left(\frac{p}{\gamma}\int_0^z \partial_t\gamma\, d\tilde{z}\right) - \partial_z\left(\frac{pK}{\gamma}\partial_z p\right) = 0 \quad (1)$$

$$\partial_z\sigma + \partial_z p + \sigma\left(A_2 - A_3\right) + \gamma = 0 \quad (2)$$

with the additional assumptions,

$$\gamma = \gamma_m\left(1+\frac{\sigma}{\sigma_m}\right)^{\beta_m} \quad (3)$$

$$K = K_0\left(\frac{\gamma}{\gamma_0}\right)^{-a} \quad (4)$$

For boundary conditions we take

$$\sigma(z_{top}) = 0 \qquad \rightarrow \qquad \gamma(z_{top}) = \gamma_m, \quad (5)$$

$$p(z_{top}) = 2116 \text{ psi}, \quad (6)$$

$$\frac{\partial p}{\partial z}(z_{bottom}) = 0. \quad (7)$$

For initial conditions, we take a parabolic profile for pressure that satisfies the above boundary conditions,

$$p_0 = -(H-z)^2 + 2H(H-z)p(z_{top}). \quad (8)$$

This initial condition is somewhat arbitrary, but is physically reasonable. We can then solve the stress equation numerically, to get an initial condition for $\sigma$ satisfying the model system of equations. This model will be referred to as model 2.

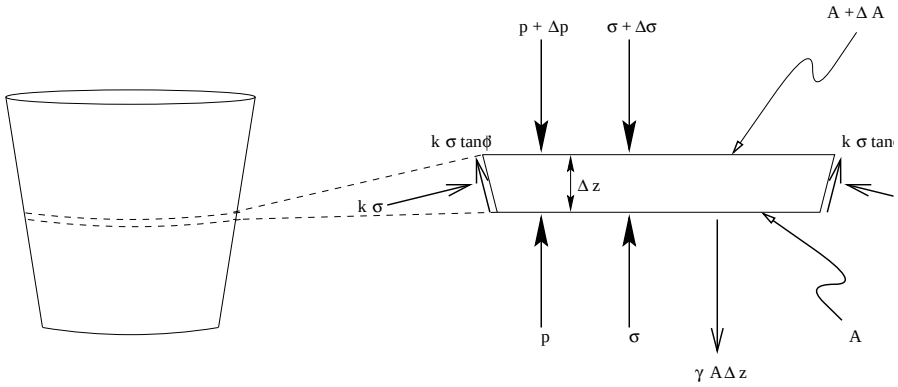



Figure 2. Force Balance

### 3.3 Remarks

In the original model, one problem came from the fact that the stress equation is in an Eulerian framework whereas the pressure equation is in a Lagrangian framework (see Figures 3 and 4). Therefore, terms containing partial derivatives with respect to z are different between the two equations

and are therefore incompatible. Also, the original model assumes that there is no change in volume over time. This causes an error in the pressure equation which is small in the incompressible case but cannot be ignored otherwise. When we derived this pressure equation with an Eulerian formulation, we get the same equation with an extra term

$$-\partial_z \left( \frac{p}{\gamma} \int_0^z \partial_t \gamma \, d\tilde{z} \right)$$

which accounts for this change in volume over time.

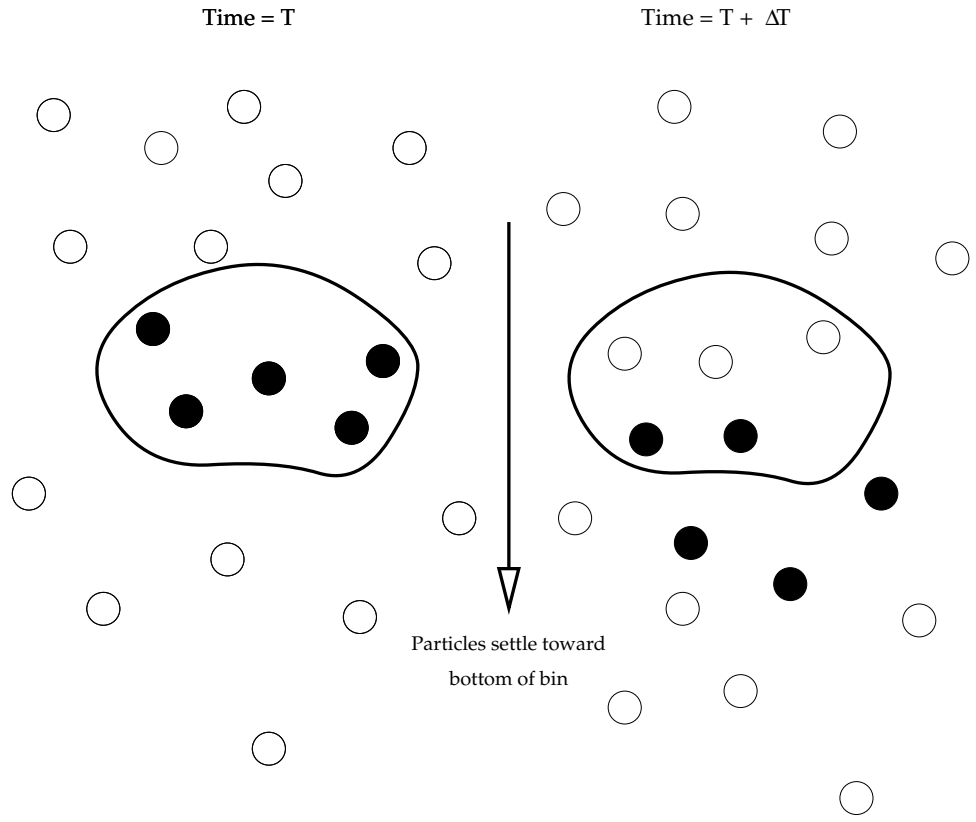



Figure 3. Eulerian Framework

## 4 Computational Algorithms

Several computational approaches were developed to solve this problem for both models. Though Model 1 is technically not correct, it was beneficial for understanding some numerical difficulties in solving this system of equations. For instance, the pressure equation is Lagrangian, and hence the grid moves with the material. So for Model 1 we allowed for time dependent, uneven grid spacing. After each time iteration, new grid points were calculated by conserving mass in each layer.

With the second model both equations are Eulerian. The grid is now fixed, however, the top boundary moves as the powder settles. In fact, the entire domain shrinks over time. A technical

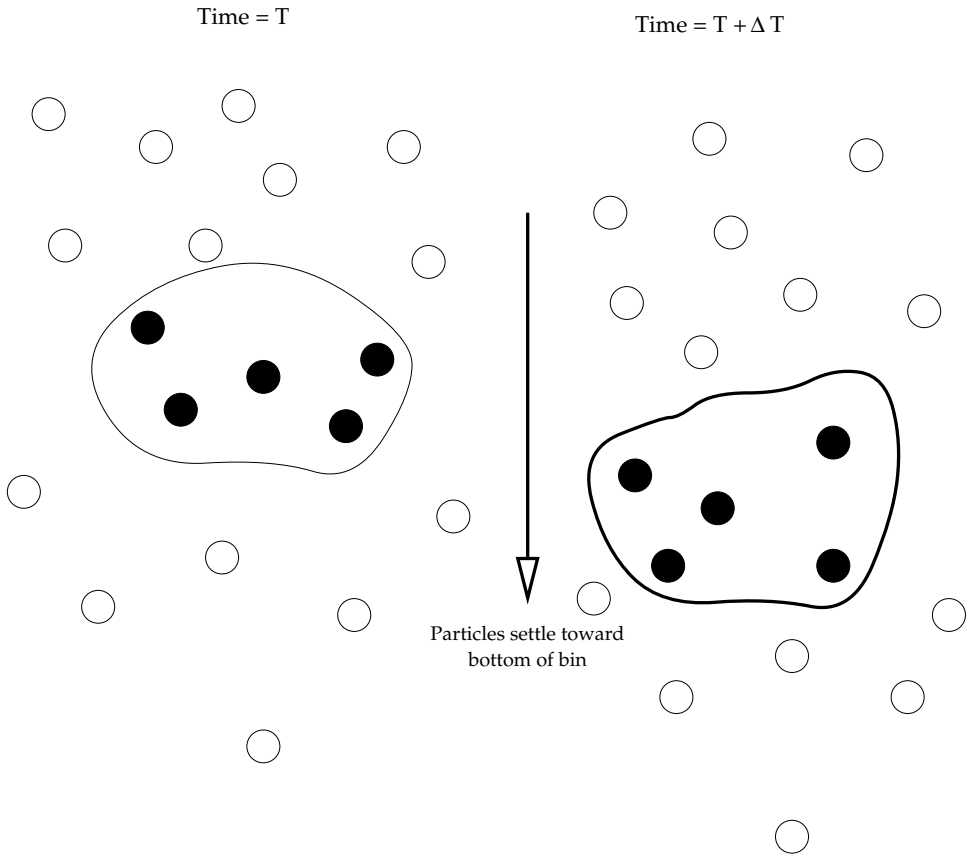



Figure 4. Lagrangian Framework

issue for Model 2 was whether to remesh the grid as the domain shrinks, or simply remove layers from the top of the bin as the mass in these layers falls below a certain tolerance.

In this section we describe three approaches, of varying complexity, to numerically solve the powder consolidation problem. The first approach is a predictor-corrector type method. The second method uses an algebraic differential solver. The third approach involves a nonlinear PDE solver.

## 4.1 Predictor-Corrector Approach

The first approach is a predictor-corrector type method, and is based on the assumption that $\partial\gamma/\partial t$ is small, at least initially. To start, we drop the $\partial\gamma/\partial t$ from the pressure equation to obtain a guess for $p$ at the next time level. This guess (prediction) is then used in the stress equation to obtain a guess for $\sigma$ at the new time level. We then have a guess for $\gamma$ at the new time level, and hence an estimate for $\partial\gamma/\partial t$. We then go back and solve the pressure and stress equations to obtain a corrected $p$, $\sigma$, and hence $\gamma$. The details of this method are given below.

### The Algorithm

Step 1:

First, we set our initial conditions. We let $p_{N+1} = 2116$, the pressure at the top boundary of the powder, and use the parabolic profile given above to set an initial $p$. The initial $p$ is used to calculate an initial value for $\sigma$, which is unknown except at the top boundary of the powder, where $\sigma_{N+1} =$

0. We calculated the initial value for $\sigma$ by solving

$$\partial_z \sigma_o + \partial_z p_o + \sigma_o (A_2 - A_3) + \gamma_o = 0$$

implicitly for $\sigma_o$. We use

$$\gamma_o = \gamma_m$$

and solve for $\sigma_o$, iterating and updating $\gamma_o$ until $\sigma_o$ converges. We used central differences to solve for $\partial_z \sigma_o$ except at the bottom boundary where we used a forward difference.

Step 2:

Next, we want to solve the pressure equation

$$\frac{1}{A}\partial_z \left( \frac{pAK}{\gamma} \partial_z p \right) = \nu \partial_t p - \frac{p}{\Gamma} \partial_t \gamma$$

for p using finite differences. We do not know the initial value of $\partial_t \gamma$. To calculate the first iteration for pressure, we assumed that the change in $\gamma$ with respect to time was negligible. Therefore, with this term essentially dropped, a guess of p was determined. A forward difference was used for the time derivative. Central differences were used for the spatial derivatives for interior points of p. Forward and backward differences were used for endpoints of p, respectively. The guess for p was used to calculate values for $\sigma$ and $\gamma$ for this time step using the same implicit scheme described in the previous step.

Step 3:

After calculating the new $\gamma$, we substituted this value into our original pressure equation to correct the value of p. Then we used this p to update $\sigma$ and $\gamma$.

Step 4:

We assumed throughout the problem that the mass in each layer is a constant. So initially we started with a mass $M$ distributed equally in $N$ layers, where the mass in each layer is given by, $m = M/N$. Also we started with a uniform grid, each layer of width $\Delta z = H/N$. After updating $p$, $\sigma$, and $\gamma$, we updated the value of $\Delta z_j$ by conserving the mass in each layer, i.e., $m = \int_{\Delta z_j} \gamma$. Using the area of a trapezoid to approximate the integral, we obtained the following equation:

$$\Delta z_j = \frac{m}{A \left[ max(\gamma_{(j-1)}, \gamma_j) - 0.5 * |\gamma_{(j-1)} - \gamma_j| \right]}$$

Thus we regrid the problem by calculating the new values of z by:

$$z_1 = 0$$

$$z_j = z_{j-1} + \Delta z_j, \qquad j = 2, 3 \ldots N+1.$$

Step 5:

Increment time step $\Delta t$.

Step 6:

To advance in time, we continued the guess-correct scheme. However, instead of initially dropping the $\partial\gamma/\partial t$ term from the pressure equation, we used the $\gamma$ values calculated previously. Then we repeated steps 3 to 6.

**Numerical Results**

Numerical results for this method were similar for both models. The model produced reasonable results as long as the material was nearly incompressible. However, we are more interested in compressible material, and in this case the predictor-corrector method does not converge. It appears that the $\partial\gamma/\partial t$ is not small initially. The error from dropping this term for an initial guess dominates the total error.

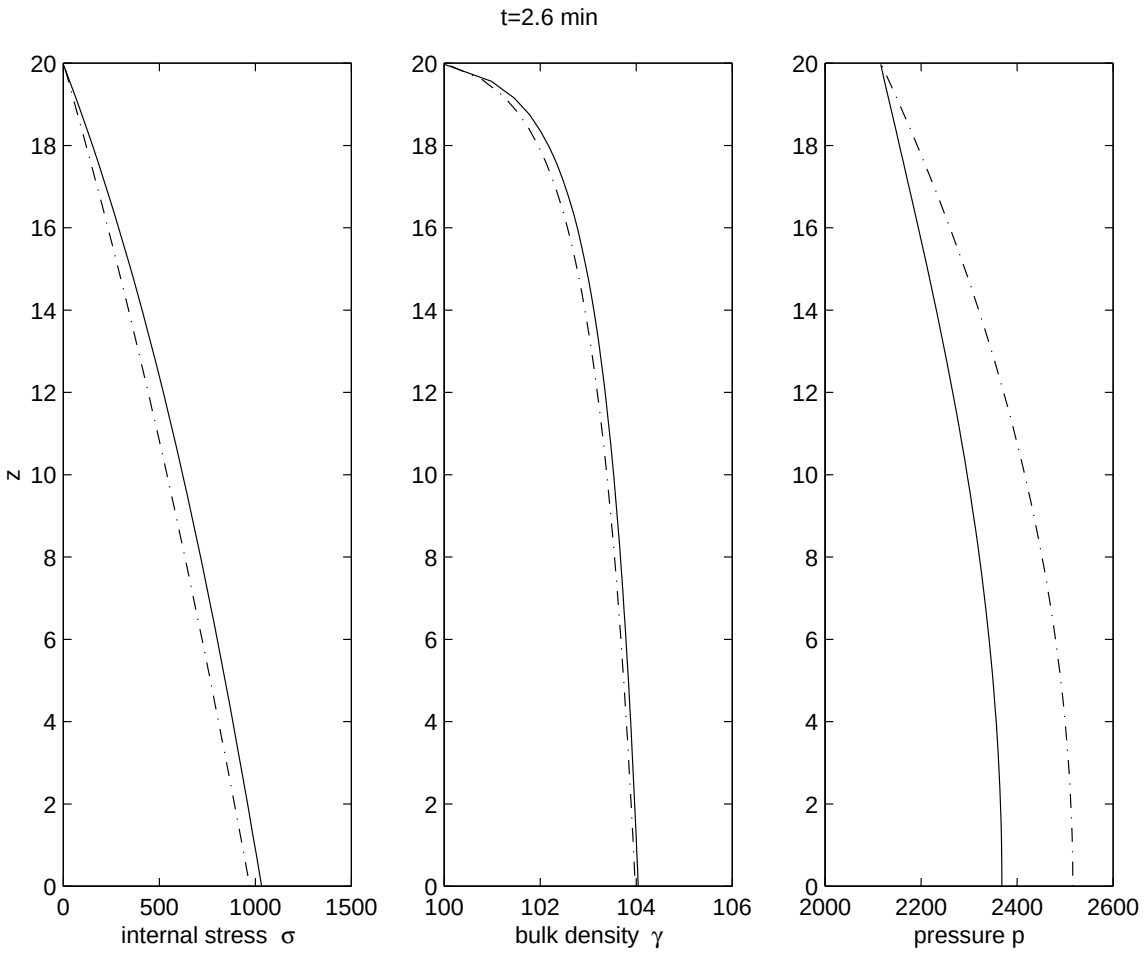



Figure 5. Dotted line is the initial data. Solid line represents the solution at time t = 2.6 minutes.

## 4.2 Differential Algebraic Approach

### The Algorithm

Step 1:

We first substitute the constitutive relations (3) and (4) into equations (1) and (2) so that the system of equations can be rewritten with $\sigma$ and $p$ as the dependent variables.

We also observe from the stress equation, i.e.,

$$\partial_z \sigma + \partial_z p + \sigma\left(A_2 - A_3\right) + \gamma_m \left(1 + \frac{\sigma}{\sigma_m}\right)^{\beta_m} = 0 \tag{1}$$

for a known pressure $p$ at any time level we can solve for the corresponding stress $\sigma$ using a Backward Euler scheme, i.e.,

$$\frac{\sigma_j^k - \sigma_{j-1}^k}{\Delta z_j} + \frac{p_j^k - p_{j-1}^k}{\Delta z_j} + \gamma_m \left(1 + \frac{\sigma_j^k}{\sigma_m}\right)^{\beta_m} + \sigma_j^k (A_2 - A_3) = 0 \tag{2}$$

The subscripts are for the spatial discretizations with respect to $z$ and the superscripts for the time discretization, $j$ =0,1...,N and $k$ =0,1.... Solving for $\sigma_{j-1}^k$ we obtain

$$\sigma_{j-1}^k = \Delta z_j \left(\sigma_j^k + (p_j^k - p_{j-1}^k) + \gamma_m \left(1 + \frac{\sigma_j^k}{\sigma_m}\right)^{\beta_m} + \sigma_j^k (A_2 - A_3)\right) \tag{3}$$

Starting from $k = 0$ (initial time) and $j = N$, using the initial pressure profile, i.e. (8), and the boundary condition for $\sigma$ at the top we can successively compute $\sigma_j^k$ for $j$ =N-1,N-2...0.

Step 2:

From Step 1 we have in principle that $\sigma = \mathcal{G}(p)$. We can use this to compute

$$\partial_t \sigma \tag{4}$$

Using the chain rule we have

$$\partial_t \sigma = \partial_p \sigma \partial_t p \tag{5}$$

Substituting this into the pressure equation results in an Ordinary Differential equation in time for the pressure $p$. We have used an Ordinary Differential equation solver to compute the updated pressure, i.e. pressure at new time level. We again use Step 4 of the previous algorithm to compute the new $\Delta z_j$.

Step 3:

The updated pressures are then used to compute the corresponding stress $\sigma$. In essence we use the pressure at a particular time level to solve for the corresponding stress which is then used to compute the pressure at the new time level.

Another way is to simultaneously solve the spatially differenced equations for the pressure and the stress. This results in an algebraic equation for the stress and an Ordinary Differential equation for the pressure, this system of equations can be simultaneously solved using DASSL (Differential-Algebraic system solver).

## 4.3 Non-Linear Algebraic Approach

### The Algorithm

Consider the two equations in the given system:

$$\begin{aligned}\left(1-\frac{\gamma}{\Gamma}\right)\partial_t p - \frac{p}{\Gamma}\partial_t\gamma - \partial_z\left(\frac{pK}{\gamma}\partial_z p + \frac{p}{\gamma}\int_0^z \partial_t\gamma ds\right) &= 0\\ \partial_z\sigma + \partial_z p + \sigma\left(A_2 - A_3\right) + \gamma &= 0\end{aligned}$$

Here is one way of approaching the problem that allows the use of a non-linear solver to attempt to find a solution. One can view $\sigma$ as a function of $p$ since if $p$ is known at a certain time $t = t_{k+1}$ then the second equation can be treated like an ODE and solved for the corresponding $\sigma$ at time $t = t_{k+1}$. Since $\gamma$ is defined to be a function of $\sigma$, it is also a function of $p$. Finally, $K$ is a function of $\gamma$ and thus a function of $p$ as well.

Consequently, one can formulate the problem using the first equation by writing

$$\left(1-\frac{G(p)}{\Gamma}\right)\partial_t p - \frac{p}{\Gamma}\partial_t(G(p)) - \partial_z\left(\frac{pK(p)}{G(p)}\partial_z p + \frac{p}{G(p)}\int_0^z \partial_t G(p) ds\right) = 0$$

where $\gamma = G(p)$ and $K = K(p)$. The function $G(p)$ can be found by various methods but in this first attempt, only a forward Euler method was used. With only one equation in $p$ one has only to discretize and attempt to solve it. To do this, one writes the above equation as $F(p^{k+1}) = 0$ and a non-linear solver can be used to find $p^{k+1}$.

The solver used here was Dr. C. T. Kelley's `nsola.m` for use in the program Matlab. An ODE solver was written by hand for finding $G(p)$ instead of relying on the built-in solver in Matlab.

One issue that had to be dealt with was the changing boundary $H(t)$. To account for this, first a new pressure $p^{k+1}$ was found on the old domain $H(t_k)$ and then using the law for conservation of mass a new height $H(t_{k+1})$ was estimated. Then the pressure that had been calculated was scaled onto the new domain. Clearly, this method is crude, but given the constraints of this project, it was a reasonable choice at the time.

### Numerical Results

Unfortunately, the results of this method were unacceptable as illustrated in Figures 6-9. The most disturbing aspect of these figures is that they show a large decrease in the height of the power in the first fraction of a second but almost no change in the remaining time.

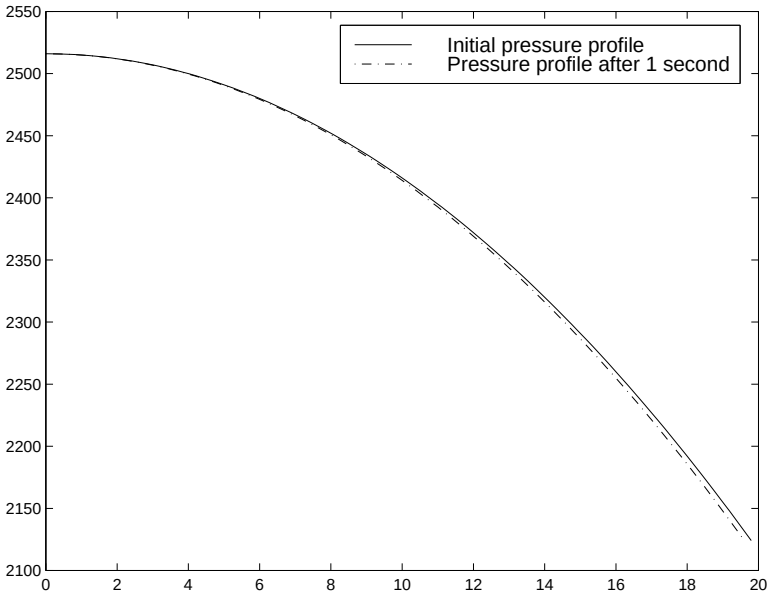



Figure 6. Comparison of pressure profiles.

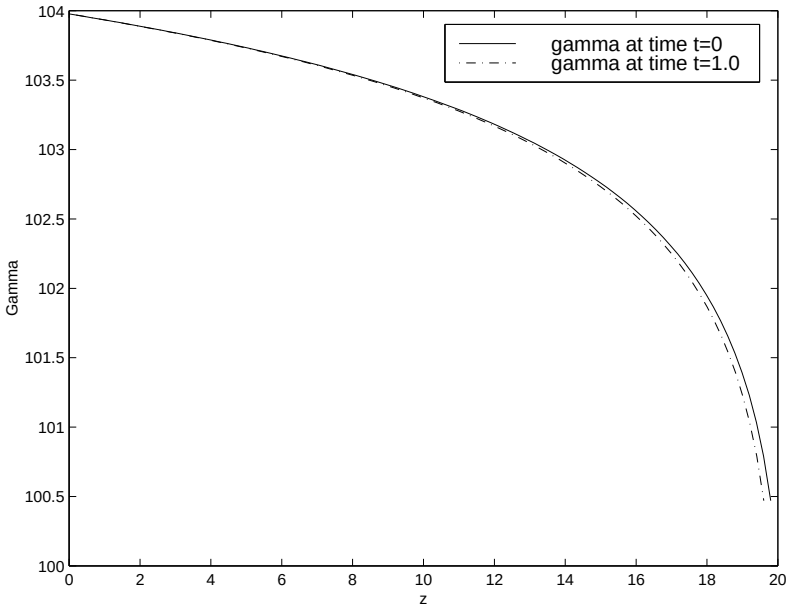



Figure 7. Comparison of gamma profiles.

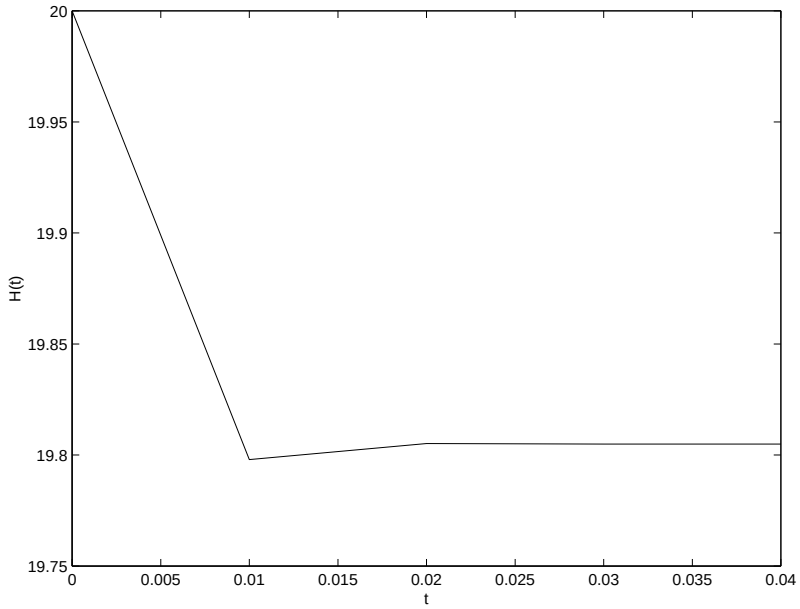



Figure 8. Height of solids as a function of time.

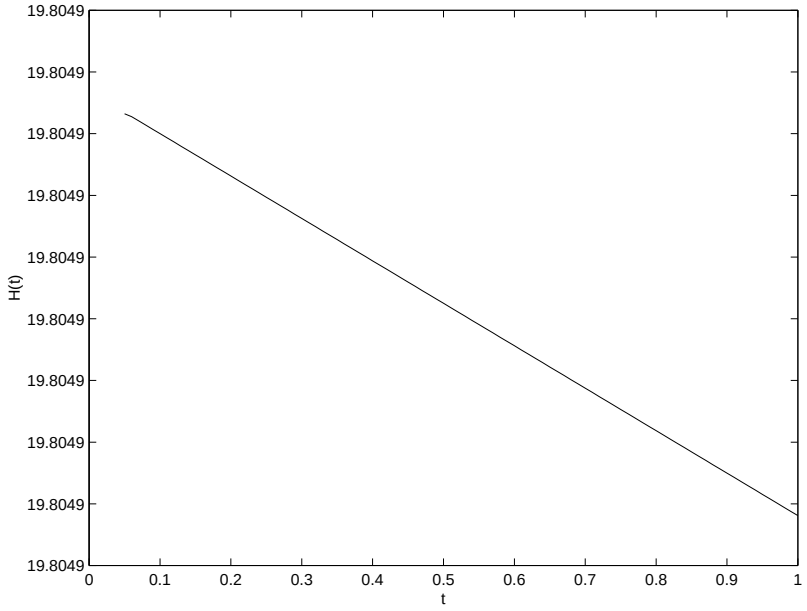



Figure 9. Height from $t = 0.05$ to $t = 1.0$.

# 5 Conclusions and Future Work

During this workshop, we developed a new model for the time dependent consolidation of fine powders. Three different approaches were applied and accurate results were obtained for the incompressible case of an instantaneously filled one-dimensional cylinder. Future work in this field would be to develop codes which model different shaped bins, filling of the bin layer by layer, and the flow of the powder out of the bin.

# 6 Appendix. Stencils used for discretizations

We have used second order accurate forward and central differencing formulas to discretize the spatial derivatives, i.e. $p_z$, $\sigma_z$ and $\partial_z(\frac{pK}{\gamma}\partial_z p)$. Special one - sided differencing formulas have been used to discretize at the bottom and top of the pile.

**Differencing for $\mathbf{p_z}$:**

We have used central difference formulas for any grid point inside the computational domain, i.e.,

$$p_z = A_j p_{j-1} + B_j p_j + C_j p_{j+1}$$

where

$$\begin{aligned} A_j &= \frac{-\Delta z_{j+1}}{\Delta z_j(\Delta z_j + \Delta z_{j+1})} \\ B_j &= \frac{\Delta z_{j+1} - \Delta z_j}{\Delta z_j \Delta z_{j+1}} \\ C_j &= \frac{\Delta z_j}{\Delta z_{j+1}(\Delta z_j + \Delta z_{j+1})} \end{aligned}$$

$j$ = 1,2...N-1. For $j$ =0, we use the boundary condition $p_z = 0$.

**Differencing for $\sigma_\mathbf{z}$:**

We have used central difference formulas for any grid point inside the computational domain, i.e.,

$$\sigma_z = A_j \sigma_{j-1} + B_j \sigma_j + C_j \sigma_{j+1}$$

where

$$\begin{aligned} A_j &= \frac{-\Delta z_{j+1}}{\Delta z_j(\Delta z_j + \Delta z_{j+1})} \\ B_j &= \frac{\Delta z_{j+1} - \Delta z_j}{\Delta z_j \Delta z_{j+1}} \\ C_j &= \frac{\Delta z_j}{\Delta z_{j+1}(\Delta z_j + \Delta z_{j+1})} \end{aligned}$$

$j$ = 1,2...N-1. We use a forward difference formula to evaluate $\sigma_z$ at $z = 0$, i.e.,

$$\sigma_z = A_0 \sigma_0 + B_0 \sigma_1 + C_0 \sigma_2$$

where

$$\begin{aligned} A_0 &= \frac{-(2\Delta z_1 + \Delta z_2)}{\Delta z_1(\Delta z_1 + \Delta z_2)} \\ B_0 &= \frac{\Delta z_2 + \Delta z_1}{\Delta z_1 \Delta z_2} \\ C_0 &= \frac{-\Delta z_1}{\Delta z_2(\Delta z_1 + \Delta z_2)}. \end{aligned}$$

**Differencing for $\partial_{\mathbf{z}}(\frac{\mathbf{pK}}{\gamma}\partial_{\mathbf{z}}\mathbf{p})$:**

We have used central difference formulas for any grid point inside the computational domain, i.e.,

$$\sigma_z = A_j(\frac{pK}{\gamma}\partial_z p)_{j-1} + B_j(\frac{pK}{\gamma}\partial_z p)_j + C_j(\frac{pK}{\gamma}\partial_z p)_{j+1}$$

where

$$\begin{aligned} A_j &= \frac{-\Delta z_{j+1}}{\Delta z_j(\Delta z_j + \Delta z_{j+1})} \\ B_j &= \frac{\Delta z_{j+1} - \Delta z_j}{\Delta z_j \Delta z_{j+1}} \\ C_j &= \frac{\Delta z_j}{\Delta z_{j+1}(\Delta z_j + \Delta z_{j+1})} \end{aligned}$$

$j$ = 1,2...N-1. We use a forward difference formula to evaluate $\partial_z(\frac{pK}{\gamma}\partial_z p)$ at $z = 0$, i.e.,

$$\sigma_z = A_0(\frac{pK}{\gamma}\partial_z p)_0 + B_0(\frac{pK}{\gamma}\partial_z p)_1 + C_0(\frac{pK}{\gamma}\partial_z p)_2$$

where

$$\begin{aligned} A_0 &= \frac{-(2\Delta z_1 + \Delta z_2)}{\Delta z_1(\Delta z_1 + \Delta z_2)} \\ B_0 &= \frac{\Delta z_2 + \Delta z_1}{\Delta z_1 \Delta z_2} \\ C_0 &= \frac{-\Delta z_1}{\Delta z_2(\Delta z_1 + \Delta z_2)}. \end{aligned}$$

# PROBLEM 5: HIGH PRECISION, HIGH ACCURACY CONTROL OF A FLUID MICRODISPENSER UNDER VARIABLE LOAD

Chongsheng Cao[1], Jerawan Chudoung[2], Jeffrey Grant[3], David Kan[4], Ilker Tunay[5]

Problem Presenter:
Jeff Sachs
D.H. Wagner Associates

**Abstract**

This report describes the mathematical modeling and simulation of a fluid microdispenser pump. The device consists of two interconnected syringe pumps with ball and socket anti-backflow valves, which work simultaneously in opposite directions to fill from a supply at atmospheric pressure and deliver to a load of unknown resistance and variable volume.

The model will be used to design a controller for the stepper-motor driven pumps to maintain a very precise output flow rate. The load flow depends on fluid compressibility and mass density. Since the properties of the fluids are variable and unknown in advance, the simulation model must incorporate these effects to test the controller under realistic conditions.

Starting from basic mass and momentum conservation laws for a Newtonian fluid, two models have been developed to describe the compressible flow. The first is a differential-algebraic model that solves for the boundary conditions between the connected elements while integrating the differential equations. This is implemented as a Matlab M-file. It has been observed that the equations are stiff and very small step sizes must be used in the integration for the solution to converge. The second one is an ordinary differential equation model which lumps some of the elements together and solves for the constraint equations in advance to arrive at a fourth order nonlinear system. A model for the valves is also incorporated. This is implemented in Simulink and the simulation proceeds without much difficulty. It has been observed that when simulated with typical displacement profiles for the pistons, the model describes the dynamics reasonably well. A more comprehensive approach should include dynamics of the stepper motors, too.

Several ideas concerning the design of a controller were proposed, but were not implemented due to lack of time. It was concluded that the controller must have some on-line parameter identification capabilities to learn the properties of the fluid being delivered and the load. Since the main variable to be controlled, i.e., load flow, cannot be measured, feedback should be obtained from the pressure transducers.

## 1 Introduction

The fluid microdispenser device under consideration consists of two interconnected syringe pumps with ball and socket anti-backflow valves. The syringes are controlled by independent stepping motors which can be given instructions to rotate through minute angles. A helical thread translates the rotational motion to longitudinal plunger motion. The rotational motions of the motors are

[1]University of California, Davis
[2]Virginia Polytechnic Institute and State University
[3]North Carolina State University
[4]University of California, Los Angeles
[5]Washington University

controlled to produce a desired flow rate. The microdispenser is equipped with two pressure sensors, one between the two pumps and one at the output port. Output flow rate must be maintained within approximately $\pm 0.1\%$ even when changing the fluid compressibility or viscosity over a wide range. Thus, the device must use information from the pressure sensors in conjunction with a control algorithm to maintain the desired flow rate even if the properties of the fluid entering the device change over time.

Fluid flow into and out of the microdispenser is controlled by the two syringe pumps. Each pump contains a pipe, inside of which is a motor controlled piston as depicted in Figure 1. Longitudinal displacement is responsible for drawing fluid into the pipe and delivering fluid out of the pipe. Increasing displacement corresponds to drawing fluid into the pipe. Likewise, decreasing displacement corresponds to delivering fluid. Thus, when piston displacement is zero there is no fluid inside the pipe. As a pump goes through successive cycles of intake and delivery, piston displacement rises and falls periodically. The shape of this "displacement profile," the piston displacement as a function of time, is dictated by the motors. Determining what profiles to use in order to maintain the desired delivery rate is of primary importance.

By design, the volume of the first pump is twice that of the second pump. To take advantage of this feature, displacement profiles should be selected so that the two pumps work together in the following manner:

1. Beginning of cycle, pump 1 fills as pump 2 delivers.
2. Pump 1 delivers as pump 2 intakes. Since pump 1 has twice the volume of pump 2, pump 2 fills while fluid continues to flow out of the device.
3. Start new cycle.

An example pair of displacement profiles is shown in Figure 2. Note that a cycle begins with piston 1 displacement at zero (pipe chamber filled) and piston 2 displacement at it's maximum value (pipe chamber filled). Also note that the, roughly speaking, out-of-phase nature of the two profiles is the key to assuring an unceasing flow of fluid out of pump 2.

Before determining displacement profiles, our team had to first construct a model that quantified the relevant physics of the microdispenser. Part of our team constructed a model in which the entire device was conceptualized as a collection of individual pieces working together in series. Another part of our team constructed a simpler model that could be implemented more quickly so that a control algorithm could be developed and displacement profiles considered. These models will be respectively referred to as Model A and Model B.

# 2 Model A

## 2.1 Governing Equations and Assumptions

We first consider the case in which a compressible fluid flows in a straight pipe with the direction of flow assumed to be solely in the longitudinal direction. The length of the pipe is taken to be $L$ and the section area of the pipe is $A$ as depicted in Figure 3. We will denote by $u(x)$ the velocity of the flow, $p(x)$ the pressure of the flow, $\rho(x)$ the density of the flow, and $W(x) = A\rho(x)u(x)$ the mass velocity of the flow.

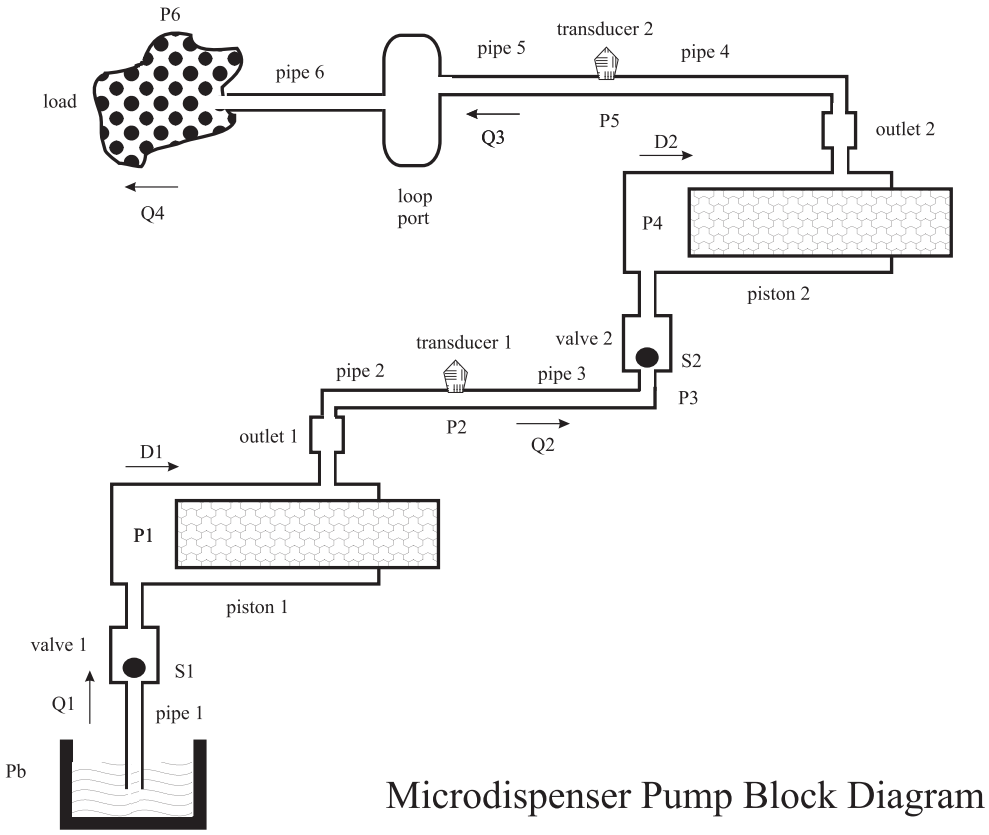



Figure 1. Microdispenser pump schematic.

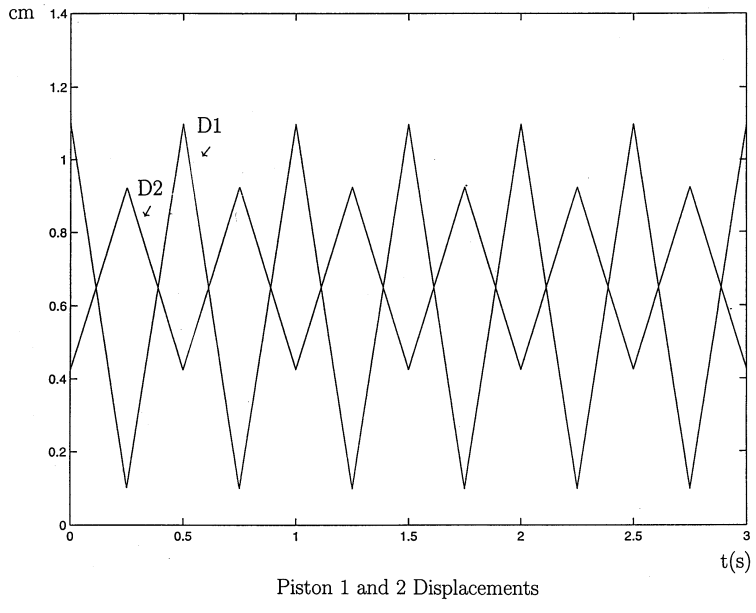



Figure 2. Representative displacement profiles.

We will use following two principles when modeling the flow.

- **Mass Conservation Law:** Conservation of mass yields

$$W(x_1) - W(x_2) = \frac{d}{dt}\left(A\int_{x_1}^{x_2}\rho(y)dy\right) \qquad 0 \le x_1 < x_2 \le L. \tag{1}$$

- **Darcy's Law:** The pressure gradient drives the flow

$$u(x) = -R\frac{\partial p}{\partial x} \qquad 0 < x < L\,. \tag{2}$$

  Here $R$ is the resistance, which depends on the viscosity of the flow, the compressibility of the flow, the length of the pipe, and the cross-sectional area of the pipe. The resistance, $R$, will be determined through a fit to data.

To simplify the derivation of the model, we make the following two assumptions. We note that both are motivated by empirical observations regarding the system being modeled.

**Assumption 1:** The density satisfies

$$\rho(x) = \rho_0(1 + \alpha p(x)) \quad \text{and} \quad \alpha << 1,$$

where $\rho_0$ is the density of the fluid when the pressure is zero and $\alpha$ is a constant depended only on the fluid.

**Assumption 2:** The velocity of the flow is linear, i.e.

$$u(x) = u(0) + \frac{u(L) - u(0)}{L}x, \qquad 0 \le x \le L.$$

The combination of Assumption 2 and equation (2) yields

$$p(x,t) = f(t) - \frac{u(0)}{R}x - \frac{u(L) - u(0)}{2LR}x^2,$$

where $f(t)$ will be determined later. Therefore, we have

$$\rho(x,t) = \rho_0\left[1 + \alpha(f(t) - \frac{u(0)}{R}x - \frac{u(L) - u(0)}{2LR}x^2)\right]. \tag{3}$$

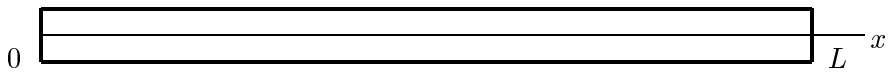



Figure 3. Schematic of the reference pipe.

## 2.2 Simple Model

Our device has ten parts comprised of three pipes, two pumps, one load, two valves and two transducers as depicted in Figure 4. One transducer is in the middle of the second pipe (Part III), and the other is in the middle of the third pipe (Part V). Each of the six pieces were treated as if they were physically shaped as pipes. It was assumed that each valve has 2 states, namely, open and closed.

The first step involved the specification of parameters. We let $(R_j, L_j, A_j),\ j = 1, \cdots, 6$ be the resistance, length, and area of each part of the device. The length and area of the two motors are denoted by $L_{m1}, A_{m1}$ and $L_{m2}, A_{m2}$, respectively. The movements of the motors are denoted by $D_1(t)$ and $D_2(t)$, and are periodic functions designed to control the output of the fluid. We will denote by $(P_j(x,t), u_j(x,t), W_j(x,t)),\ 0 \le x \le L_j$, the pressure, velocity and mass velocity of the flow in each part.

We now model the individual pieces of the device including the valves with the pumps, and model the transducers with pipes.

**PIPE:** Parts I, III, V and VI, i.e., $j = 1, 3, 5, 6$.

By applying the equations (1) and (3), we get

$$W_j(0) - W_j(L_j) = \alpha\rho_0 L_j A_j[\dot{f}_j(t) - \frac{L_j}{6R_j}(2\dot{u}_j(0) + \dot{u}_j(L_j))]$$

where $\cdot$ denotes $\frac{\partial}{\partial t}$. Noting that $W = \rho A u$, we conclude that

$$\begin{aligned} &\alpha L_j[\dot{f}_j(t) - \frac{L_j}{6R_j}(2\dot{u}_j(0) + \dot{u}_j(L_j))] \\ &\quad = (1 + \alpha f_j(t))(u_j(0) + u_j(L_j)) + \frac{\alpha L_j}{2R_j}(u_j(0) + u_j(L_j))u_j(L_j) \end{aligned} \tag{4}$$

**PUMP:** Parts II and III, i.e., $j = 2, 4$.

We will assume that the pump has the configuration shown in Figure 5. Since the motor moves, the limits of the integral will change as time progresses. In this case, the (1) can be reduced to

$$W(0) - W(L_j) = A_j \int_0^{L_j} \dot{\rho}_j(x)dx + A_{mi} \int_0^{D_i(t)} \dot{\rho}_{mi}dx + A_{mi}\dot{D}_i(t)\rho_{mi}(D_i(t))$$

where $A_{mi}$ is the area of the $i^{th}$ motor and $\rho_{mi}$ is the density of the fluid in the $i^{th}$ motor. It then follows that $i = j/2$. Furthermore, we assume that $\rho_{mi}(x) = \rho_j(0)$ so that

$$\begin{aligned} &\alpha L_j[\dot{f}_j(t) - \frac{L_j}{6R_j}(2\dot{u}_j(0) + \dot{u}_j(L_j))] + \frac{A_{mi}}{A_j}[1 + \alpha f_j(t) + \dot{f}_j(t)] \\ &\quad = (1 + \alpha f_j(t))(u_j(0) + u_j(L_j)) + \frac{\alpha L_j}{2R_j}(u_j(0) + u_j(L_j))u_j(L_j). \end{aligned} \tag{5}$$

To combine the components, we require the following assumption.

**Assumption 3:** The valve is closed when $\dot{D}(t) \ge 0$.

**CONNECTION EQUATIONS:** It is natural that we set

$$f_j(t) = f_{j+1}(t) \quad u_j(L_j) = u_{j+1}(0), \qquad j = 1, \cdots, 5. \tag{6}$$

When valve I is closed, we will replace the equations (6) of the case $j = 1$ by

$$u_1(L_1) = u_2(0) = 0. \tag{7}$$

When valve II is closed, we will replace the equations (6) of the case $j = 3$ by

$$u_1(L_3) = u_4(0) = 0. \tag{8}$$

Finally, the input and output conditions yield

$$f_1(t) = c_1 \quad u_6(L_6) = c_2. \tag{9}$$

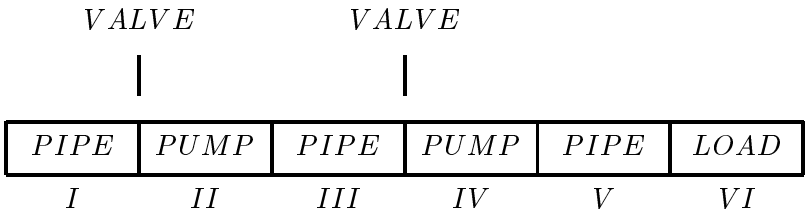



Figure 4. Components of the pump assembly.

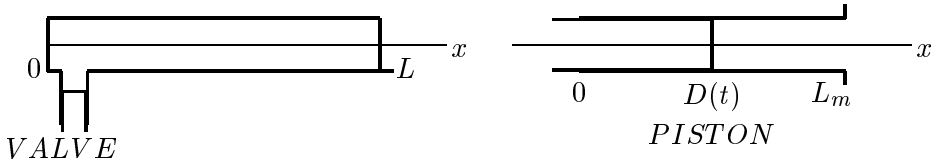



Figure 5. Valve and piston configuration in the pump.

## 3 Numerical Implementation of Model A

In this section, we outline the derivation of the mathematical model, its discretization and relevant numerical methods.

**Model**

Consider a pipe with compressible fluid. In three dimensions, the pertinent quantities are:

$$\begin{aligned}
u_L &= \text{Velocity of fluid on left} \\
u_R &= \text{Velocity of fluid on right} \\
\rho_L &= \text{density of fluid on left} \\
\rho_R &= \text{density of fluid on right} \\
A &= \text{cross-sectional area of pipe} \\
L &= \text{length of pipe} \\
V &= AL = \text{volume of pipe} \\
P_L &= \text{pressure of fluid on left} \\
P_R &= \text{pressure of fluid on right} \\
W_L &= A\rho_L u_L = \text{mass flux on left} \\
W_R &= A\rho_R u_R = \text{mass flux on right}.
\end{aligned}$$

We first assume that the flow will be laminar, rather than turbulent, since the Reynolds number is small. Next, from the standard equations for conservation of mass and balance of momentum, it follows that

$$\rho_0 AL\frac{\alpha}{2}\left[\frac{d}{dt}P_L + \frac{d}{dt}P_R\right] = A(\rho_L u_L + \rho_R u_R)$$

$$A\rho_0\frac{L}{2}\left[\frac{d}{dt}u_L + \frac{d}{dt}u_R\right] = A(\rho_L < \bar{u}_L^2 > -\rho_R < \bar{u}_R^2 > +P_L - P_R) - 4\pi L\mu(u_L + u_R)$$

Here $\alpha, \mu, \rho_0$ are physical constants.

To simplify the equations, we assume that the fluid in the pipe has negligible acceleration, which implies that $\frac{d}{dt}u$ and $\bar{u}$ vanish. With some scaling, the equations become:

$$\begin{aligned}\frac{d}{dt}P_L + \frac{d}{dt}P_R + C_1(W_L - W_R) &= 0\\ P_R - P_L - C_2\left[\frac{W_L}{\rho_L} - \frac{W_R}{\rho_R}\right] &= 0\end{aligned}$$

where $C_1 = \frac{2}{\rho_0 V}$ and $C_2 = \frac{4\pi L\mu}{A^2}$. The equations for the load and pistons are similar. The load, in addition, has variable cross-sectional area at one end. The piston has variable displacements.

With boundary conditions for the pressure and the mass flux, we have a $2N_{pipe} + 2N_{piston} + 2 + 2$ differential algebraic system of equations. (Here $N_{pipe}$ is the number of pipes, $N_{piston}$ is the number of pistons, 2 for the boundary conditions and 2 for the valves.) We noted that the system is stiff.

**Discretization**

Because the system is stiff, an implicit scheme for solving the system must be used. We used a backwards Euler scheme. For the pipe, this has the form

$$\begin{aligned}P_L^{new} + P_R^{new} + \Delta t(W_L^{new} - W_R^{new}) &= P_L^{old} + P_2^{old}\\ P_R^{new} - P_L^{new} - C_2\left[\frac{W_L^{new}}{\rho_L} - \frac{W_R^{new}}{\rho_R}\right] &= 0.\end{aligned}$$

In matrix form, this can be written as

$$\begin{pmatrix} -\Delta t C_1 & 1 & \Delta t C_1 & 1\\ -\frac{C_2}{\rho_L} & 1 & -\frac{C_2}{\rho_R} & -1\end{pmatrix}\begin{pmatrix} W_L^{new}\\ P_L^{new}\\ W_R^{new}\\ P_R^{new}\end{pmatrix} = \begin{pmatrix} P_L^{old} + P_R^{old}\\ 0\end{pmatrix}.$$

One feature of this discretization is the component-wise form it takes. That is, to add another component to the model, one adds a $2 \times 4$ sub-matrix to the solution matrix. For example, for a pipe and a load, the solution matrix is made of two $2 \times 4$ sub-matrices plus two boundary conditions, which yields a $6 \times 6$ matrix.

# 4 Model B

The model developed in the previous section requires solving a set of differential-algebraic equations to determine the behavior of the microdispenser. When combined with the dynamics of the valves, simulating these equations becomes difficult. Instead, it is possible to introduce some simplifying assumptions to arrive at a model consisting only of ordinary differential equations. This simplification becomes possible by lumping some elements of the pump and representing them by a single equation, instead of several parts which depend on each other via the boundary conditions at the junctions. In this way, the algebraic equations can be solved initially, leaving a system of a small number of simultaneous differential equations. It is then straightforward to implement this model in Matlab and obtain simulation results which resemble experimental results. This simplified model can be used to test control algorithms quickly, before taking the expense of implementing the algorithm on the actual production platform for fine tuning.

## 4.1 Simplifying Assumptions and ODE Model

We start with the conservation of mass law or continuity equation (1) for a circular pipe of length $L$ and the isothermal fluid state equation in Assumption 1. In addition to Assumption 2, we make the following assumption.

**Assumption 4.** The pressure profile along the pipe is linear, i.e.,

$$p(x) = p(0) + \frac{p(L) - p(0)}{L} x, \qquad 0 \leq x \leq L.$$

Therefore, we do not employ Darcy's Law (2). The use of (1) along the length of the pipe yields

$$W(0) - W(L) = A\rho(0)u(0) - A\rho(L)u(L) = AL\rho_0 \alpha \frac{dp}{dt} \tag{10}$$

where the average pipe pressure is defined as $p \equiv (p(0) + p(L))/2$. Next, calculating the Reynolds number for all the pipes and piston gaps, we see that it is well below 2000 for the specified range of delivery rates, indicating that all flows are laminar [1]. This considerably simplifies the conservation of momentum equation since we can neglect the inertia terms and retain only viscosity terms. Using a parabolic velocity profile across the cross-section and ignoring entry and exit losses, the pressure drop across the pipe under steady flow conditions is given by the Hagen-Poiseuille law [1]

$$p(0) - p(L) = \frac{4\pi L\mu}{A}(u(0) + u(L)) = RQ \tag{11}$$

where $\mu$ is the absolute viscosity, the line resistance is defined as $R \equiv \frac{8\pi L\mu}{A^2}$ and the average volumetric flow is $Q \equiv \frac{(u(0)+u(L))A}{2}$.

Employing equations (10) and (11) it is possible to describe the dynamics of the pump. Referring to the block diagram in Figure 1, $D_1$ and $D_2$ denote the piston displacements, $S_1$ and $S_2$ are logical variables, so that $S_i = 1$ when valve i is open, $S_i = 0$ when valve i is closed. Let us consider valve 1, inlet, piston 1 chamber, outlet 1, connection rings, pipe 2, transducer 1 and pipe 3 as a single element, with total dead volume $V_{st}$. Similarly, consider valve 2, inlet, piston 2 chamber, outlet 2 and connection rings as one element with dead volume $V_{p2dead}$. The dead volumes of pipe 4, transducer 2, pipe 5 and the loop port are augmented with the load volume, with the sum denoted by $V_{Ltot}$.

The ball-and-socket valves are modeled as follows. When the ball is seated in the socket, the valve is closed and there is no leakage. The ball is displaced from the socket and the valve is opened if the pressure across the valve exceeds a certain threshold. To close the valve again, the fluid must start flowing in the reverse direction, carrying the ball to the socket. In the interim, some amount of backflow occurs until the ball makes contact. Denoting by $b_i$ the displacement of the ball in valve i, the behavior is described by

$$\dot{b}_i = Q_i \quad , \quad S_i = \begin{cases} 1, & \Delta p > \text{threshold or } b_i > 0 \\ 0, & b_i \leq 0 \end{cases} \tag{12}$$

where the ball is allowed to move only in the limits $0 \leq b_i \leq b_{max}$. Therefore Assumption 3 is no longer necessary.

For numerical purposes it is useful to scale the pressure values so that very large and very small numbers will not appear in algebraic operations. To this end, define $P_i \equiv \alpha_0 p_i$ and select $\alpha_0$ close to the nominal value of $\alpha$. Then, using (10) for the volume $V_{st}$ combined by the volume created by the motion of piston 1, we obtain the dynamics of the first component as

$$\frac{\alpha}{\alpha_0}\left(V_{st} + A_p D_1\right)\dot{P}_1 = \left(1 + \frac{\alpha}{\alpha_0}P_1\right)\left(-A_p v_1 + Q_1 - Q_2\right) \tag{13}$$

where $A_p$ is the piston area and $v_1$ is the piston 1 velocity. Denote the pipe resistances by $R_{p_i}$ and the total resistance of piston 1 gap, pipe 2 and pipe 3 as $R_{st}$. Using (11) we get

$$Q_1 = S_1 \frac{P_b - P_1}{\alpha_0 R_{p_1}} \quad , \quad Q_2 = S_2 \frac{P_1 - P_4}{\alpha_0 R_{st}} \quad , \quad P_2 = P_1 - Q_2 R_{p_2} \alpha_0$$

where $P_b$ is the ambient pressure at the source bottle. Similar to (13), the dynamics associated with the second component, i.e., $V_{p2dead}$ and half of $V_{Ltot}$ combined by the volume created by the motion of piston 2, can be written as

$$\begin{aligned} &\frac{\alpha}{\alpha_0}\left(V_{p2dead} + A_p D_2 + \frac{V_{Ltot}}{4}\right)\dot{P}_4 \\ &\qquad = \left(1 + \frac{\alpha}{\alpha_0}P_4\right)\left(-A_p v_2 + Q_2\right) - \left(1 + \frac{\alpha}{2\alpha_0}P_4\right)Q_4. \end{aligned} \tag{14}$$

The load flow $Q_4$ and the flow into the loop port $Q_3$ are calculated as

$$\begin{aligned} Q_4 &= \frac{P_4 - P_a}{\alpha_0\left(R_L + R_{p_4}\right)} \\ Q_3 &= \left(1 + \frac{\alpha}{\alpha_0}P_4\right)^{-1}\left[\frac{V_{Ltot}}{4}\frac{\alpha}{\alpha_0}\dot{P}_4 + \left(1 + \frac{\alpha}{2\alpha_0}P_4\right)Q_4\right] \end{aligned} \tag{15}$$

where $R_L$ is the load resistance and $P_a$ is the ambient pressure at the end of the load.

### 4.2 Simulation Results

Equations (12-15) describe a fourth order dynamical system with states $b_1$, $b_2$, $P_1$, $P_4$, inputs $v_1$, $v_2$ and output $Q_3$. Combining these with an appropriate stepper motor dynamic equation with desired piston displacements as command inputs and piston displacements and velocities as outputs,

a complete simulation model of the microdispenser pump is obtained. The model is transferred into Matlab Simulink and simulated using an algorithm designed for stiff systems.

As an example, the behavior of the system to a typical input of triangular profiles in Figure 2 is shown in Figure 6. The profile aims to keep the output flow at $0.15mL/s$ and corresponds to open-loop control. As seen in the plot of the flow across valve 2, backflow phenomenon is observed. This also creates a sharp dip in the output flow. Also notice the difference between the pressure rise times between the pistons. As piston 1 starts to deliver, it has to supply fluid to a larger volume, hence the slower rise compared to piston 2.

Overall, the simplified model appears to approximate the reality sufficiently well for a first attempt. To gain more confidence, the model must be tested with several different control profiles employed in practice and the experimental measurements must be compared with the simulation results.

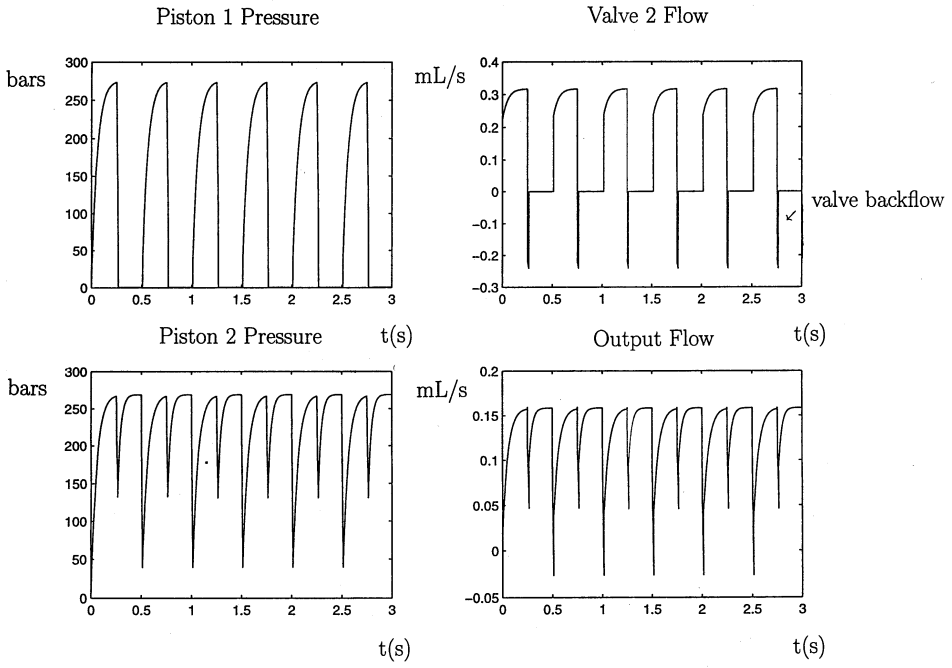



Figure 6. Pump pressure due to the triangular input depicted in Figure 2.

## References

[1] Merritt, H. E., *Hydraulic Control Systems*, John Wiley and Sons, Inc. New York, 1967

# PROBLEM 6: MASS TRANSPORT AND SURFACE REACTION OF IMMUNOASSAY

Jim Buchanan[1], Henrik Kalisch[2], Anastasios Liakos[3], Renee Yong Ran Liu[4], Shree Yvonne Whitaker[5], Igor Zhislin[6]

Problem Presenter:
Zhong Ding
Johnson & Johnson Company

**Abstract**

In clinical diagnostics product development and biochemical applications, the solid phase surface reactions and immunoassay technologies are widely used to determine the concentration of a protein in a fluid sample. In a sample, protein A with unknown concentration and protein B (a label) with known concentration bind competitively with protein C, which is coated on a solid surface. Any unbound proteins are washed out by the subsequent washing fluids. By measuring the concentration of B bound with C on the surface in a given time, the original concentration of A can be found. In this report, a fluid containing both A and B flows in a cylinder. The inner surface of a cylinder is uniformly coated with C. To assist the product design process, we need to predict the concentration of bound A on the surface as a function of time.

**Keywords:** Computational Biology, Transport/Diffusion.

## 1 Introduction.

When the human body is invaded by antigens (foreign substances that are usually protein), the immune system recognizes them and produces highly specific proteins called antibodies. Antibodies bind with antigens to form antibody-antigen complexes. The invading antigens thus become inactive and are removed from the organism later on. Although each type of antibody is highly specific for one particular type of antigen, antibodies share common structures and are different only in a region called the antigen binding site. Each antibody is a Y- shaped molecule. It has two identical antigen-binding sites, one at the end of each upper arm of the Y.

Antigen attached with a label is used in a variety of immunoassay techniques. The labeled antigen binds with the antibody and then is detected by using the label as a signal. In immunoassays, one can use the specificity of the antigen and antibody reactions to determine the existence and/or the concentration of a particular antigen. One typical immunoassay is called the *Antigen Capture Immunoassay*. In this assay, a known amount of labeled antigen ($Lb$) is added to the unlabeled sample antigen ($Ag$). Both antigens then compete in binding with the antibody ($Ab$) coated on the solid surface. The reactions taking place are as follows:

[1]University of California at Santa Barbara, buchanan@math.ucsb.edu
[3]University of Pittsburgh, alist@math.pitt.edu
[4]University of California at Berkeley, reneeliu@math.berkeley.edu
[5]North Carolina State University, sywhitak@eos.ncsu.edu
[6]Middle Tennessee State University, igor@frank.mtsu.edu

$$
\begin{array}{ccccc}
Ag & + & Ab & \overset{k_f^A}{\rightleftharpoons}_{k_b^A} & Ag:Ab \\
 & & + & & \\
 & & Lb & & \\
 & & {\scriptstyle k_b^L}\uparrow\downarrow{\scriptstyle k_f^L} & & \\
 & & Lb:Ab & &
\end{array}
$$

Any unbound proteins are washed out by the succeeding washing process. The greater the initial sample antigen concentration, the less labeled antigen is bound to antibody, and the smaller the final signal. Thus, the sample antigen's initial concentration can be determined by measuring the final signal and finding the corresponding initial concentration on the dose response curve.

In this project, our final objective is to provide a good mathematical model for the surface reaction and fluid flow in the capillary. In the model, the inner surface of a cylinder is uniformly coated with antibody. A fluid containing antigen (with unknown concentration) and an antigen/label compound (with known concentration) flows inside the cylinder. The two bind competitively with the coated antibody in the surface of the cylinder.

Our task is to predict the concentration of antigen (protein A) bound on the inner surface as a function of time along the axial direction of the cylinder. In order to get a precise result, we have to take into account the mass transport due to the movement of the fluid, the diffusion of antigen and antigen/label through the fluid, and the chemical reaction in the solid.

In the first section we describe the reaction as an isolated process, which might take place in a flask filled with a solution containing all three reactants. At this point, we assume a homogeneous distribution of the reactants, and therefore, no diffusion takes place. In the next section, we consider a model of an infinite surface which is coated with a relatively thick fluid film. The fluid contains antigen antigen/label, while antibody is stuck in the surface. At this stage, we have to include the effect of diffusion of antigen and antigen/label through the liquid, but we assume the fluid to be stationary. The reaction takes place when the antigen or antigen/label in the liquid come close enough to the surface to bind with the antibody. At last, we consider the capillary system as described above.

## 2 Kinetics in the Solution Phase.

We first consider a simple model in which we do not take into account the geometry of the physical space. We obtain a set of five ordinary differential equations for the chemical reactions.

Simplifying the notation we can rewrite our system as:

$$
\begin{aligned}
\frac{dX}{dt} &= -k_1 X Z + k_2 W, \\
\frac{dY}{dt} &= -k_3 Y Z + k_4 P, \\
\frac{dZ}{dt} &= -k_1 X Z + k_2 W - k_3 Y Z + k_4 P, \\
\frac{dW}{dt} &= k_1 X Z - k_2 W, \\
\frac{dP}{dt} &= k_3 Y Z - k_4 P,
\end{aligned}
$$

with initial conditions $X(0) = X_0$, $Y(0) = Y_0$ and $Z(0) = Z_0$. Here we have set

$$\begin{aligned} X = [Ag], \quad Y &= [Lb], \quad Z = [Ab], \\ W = [Ag : Ab], &\qquad P = [Lb : Ab], \\ k_1 = k_f^A, &\qquad k_2 = k_b^A, \\ k_3 = k_f^L, &\qquad k_4 = k_b^L. \end{aligned}$$

We can obtain a numerical approximation to the solution using a finite-difference scheme. To carry out those calculations during the preparation of this report, we used 'Matlab ODE 23s'. In Figures 1 and 2, we show some results for different initial values and reaction coefficients.

Note that depending on the values used, it may take more than 300 seconds to get near the equilibrium. To judge how close our solution is to the equilibrium, we set the time derivatives equal to zero to obtain the following relations:

$$k_{eq}^{(1)} = \frac{k_1}{k_2} = \frac{W}{XZ}$$
$$k_{eq}^{(2)} = \frac{k_3}{k_4} = \frac{P}{YZ}$$

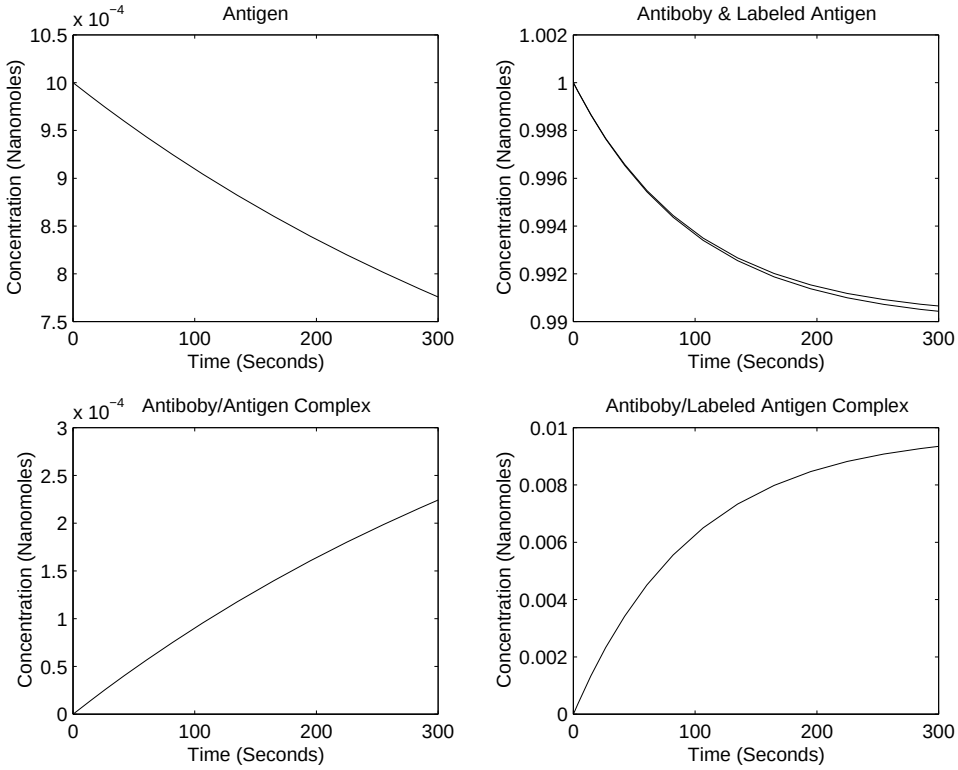



Figure 1. Low initial conditions and low coefficients

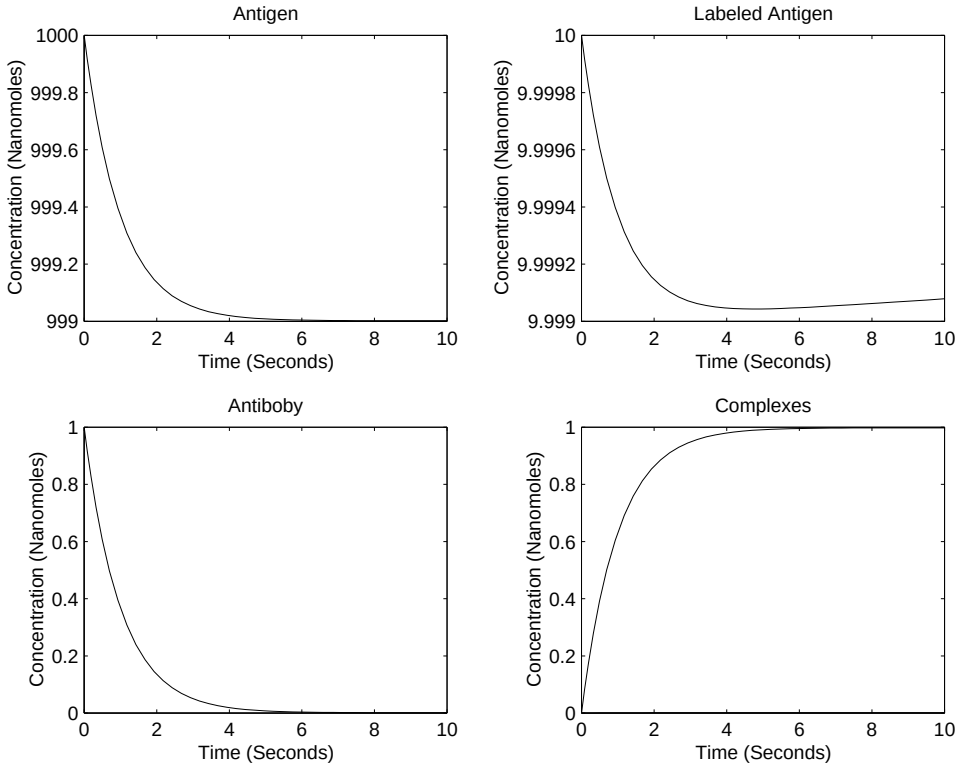



Figure 2. High initial conditions and low coefficients

In fact, we can prove mathematically that every solution within a certain range of initial values and constants will approach the equilibrium over time. First observe that we have the following symmetries in our system.

$$\begin{aligned}
\frac{dW}{dt} &= -\frac{dX}{dt}, \\
\frac{dP}{dt} &= -\frac{dY}{dt}, \\
\frac{dZ}{dt} &= \frac{dX}{dt} + \frac{dY}{dt}.
\end{aligned}$$

We integrate those equations using $X_0$, $Y_0$, and $Z_0$ as constants of integration. With this, the original system can be reduced to the following:

$$\begin{aligned}
\frac{dX}{dt} &= -k_1 X \left\{ Z_0 - X_0 + X - Y_0 + Y \right\} + k_2 \left\{ X_0 - X \right\}, \\
\frac{dY}{dt} &= -k_3 Y \left\{ Z_0 - X_0 + X - Y_0 + Y \right\} + k_4 \left\{ Y_0 - Y \right\}, \\
Z &= Z_0 - X_0 + X - Y_0 + Y, \\
W &= X_0 - X, \\
P &= Y_0 - Y.
\end{aligned}$$

Using the reduced system of two differential equations, and setting $dX/dt$ and $dY/dt$ equal to zero, we obtain a set of algebraic equations for $X$ and $Y$:

$$\begin{aligned}
0 &= -k_1 X \left\{ Z_0 - X_0 + X - Y_0 + Y \right\} + k_2 \left\{ X_0 - X \right\}, \\
0 &= -k_3 Y \left\{ Z_0 - X_0 + X - Y_0 + Y \right\} + k_4 \left\{ Y_0 - Y \right\},
\end{aligned}$$

Solving for $Y$ gives

$$\begin{aligned} Y &= \frac{k_2}{k_1}\frac{X(0)}{X} - \frac{k_2 - k_1 + C}{k_1} - X \\ Y &= -\frac{X}{2} - \frac{k_4 - k_3 + C}{2k_1} + \frac{1}{2}Y(0)\sqrt{(X + \frac{k_4 - k_3 + C}{2k_1})^2 + 4\frac{k_4}{k_3}} \end{aligned}$$

A solution of this system determines an equilibrium point of the system of ODE's in the X-Y plane. As seen in Figure 1, the first equation behaves like $1/X$ for small $X$, but like $X$ for large $X$. The second equation behaves almost like a constant, because the values of the coefficients are so small.

We can see directly that there is only one solution in the first quadrant, i.e., a solution for which both $X$ and $Y$ are positive. Moreover, we can prove that the equilibrium point so determined is stable, i.e. the real part of both eigenvalues of the Jacobian matrix of the derivative of the function associated to the differential equations are negative.

The Jacobian of the system is given by:

$$J = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

where

$$\begin{aligned} a &:= -k_2 - 2k_1X - k_1Y - k_1Z(0) + k_1X + k_1Y = -k_2 - k_1X - k_1Z \\ b &:= -k_1X \\ c &:= -k_3Y \\ d &:= -k_4 - 2k_3Y - k_3X - k_3Z(0) + k_1X + k_1Y = -k_4 - k_3Y - k_3Z \end{aligned}$$

The eigenvalues are given by:

$$\lambda_i = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad - bc)}}{2}$$

It is obvious that $a+d < 0$. It can also be shown that $ad - bc > 0$. This ensures that both eigenvalues are either negative or complex with the negative real part. Thus for any choice of coefficients and initial concentrations the equilibrium point is stable. We can also see this graphically on the plot of the vector field associated with the system of differential equations for $X$ and $Y$.

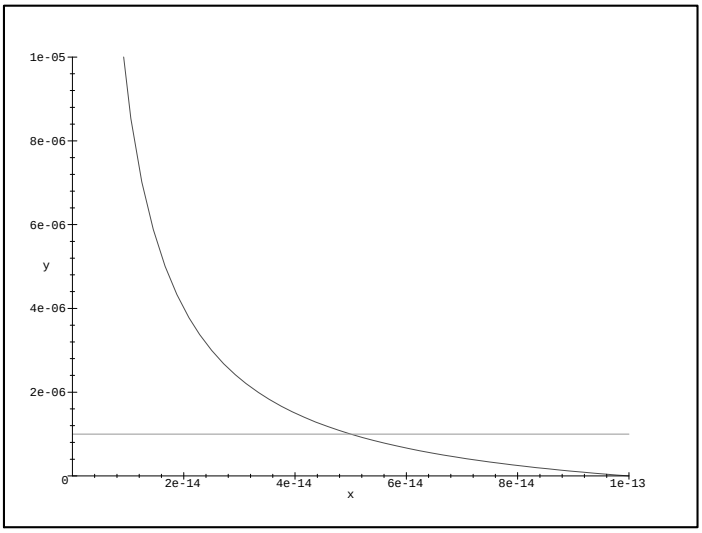



Figure 3. Nullclines for $X$ and $Y$.

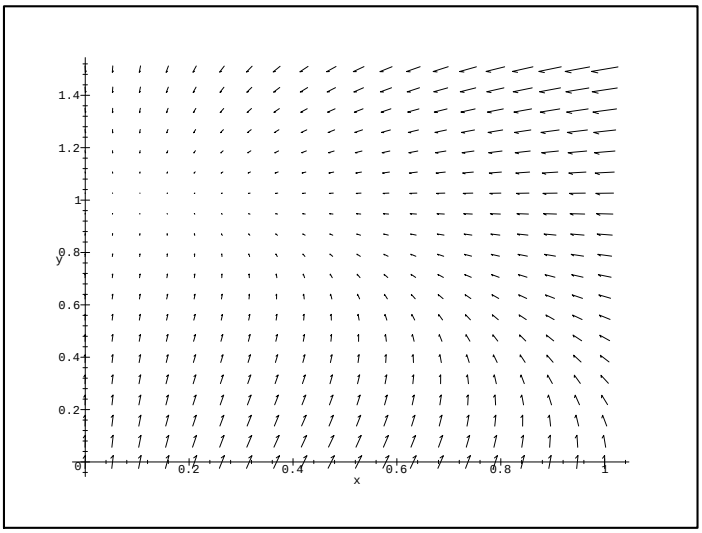



Figure 4. Stability of Equilibrium

## 3 Surface Reaction in an Infinite Planar Surface.

For this part, we consider the reaction in an idealized infinite plane coated with fluid. We assume that at any time, the distribution of the reactants in the fluid is homogeneous in any direction parallel to the plane, but not vertical to the plane. We thus have to solve a one-dimensional problem: We consider an imagined ray of fluid above a point in the surface. Because of the depletion of antigen and antigen/label in the fluid near the surface, there will be diffusion of the reactants through the fluid towards the boundary. This process is governed as usual by the diffusion equation. To derive the boundary conditions, we have to consider the transport of the reactants form the fluid into the

boundary. To this end, suppose there is a gap of size $h$ between the fluid and the surface. We then have the following equations for the concentrations of $X$ and $Y$:

$$\frac{dX_{surface}}{dt} = \frac{d_X(X_{fluid} - X_{surface})}{h} - k_1 X_{surface} Z + k_2 W$$

$$\frac{dY_{surface}}{dt} = \frac{d_Y(Y_{fluid} - Y_{surface})}{h} - k_3 Y_{surface} Z + k_4 P.$$

Letting $h$ tend to zero, we obtain

$$\frac{dX}{dt} = d_X \frac{\partial X}{\partial n} - k_1 X_{surface} Z + k_2 W$$

$$\frac{dY}{dt} = d_Y \frac{\partial Y}{\partial n} - k_3 Y_{surface} Z + k_4 P$$

for the concentrations $X$ and $Y$ in the surface. The coefficients $d_X$ and $d_Y$ appearing here are the coefficients of mass transport (see Christensen [2]). The values of $d_X$ and $d_Y$ are actually equal to the values of $D_X$ and $D_Y$, respectively, although with different units(see Frank and Kamenetskii [3]).

The system of equations for the model described above is:

$$\frac{\partial X}{\partial t} = D_X \frac{\partial^2 X}{\partial z^2}$$
$$\frac{\partial Y}{\partial t} = D_Y \frac{\partial^2 Y}{\partial z^2}$$

with boundary conditions

$$\begin{aligned}
\frac{dX}{dt} &= -k_1 X Z + k_2 W + d_X \frac{\partial X}{\partial z}, \\
\frac{dY}{dt} &= -k_3 Y Z + k_4 P + d_Y \frac{\partial Y}{\partial z}, \\
\frac{dZ}{dt} &= -k_1 X Z + k_2 W - k_3 Y Z + k_4 P, \\
\frac{dW}{dt} &= k_1 X Z - k_2 W, \\
\frac{dP}{dt} &= k_3 Y Z - k_4 P.
\end{aligned}$$

The pseudocode for obtaining a numerical solution for the partial differential equation is as follows:

1. **Initialize** concentrations of the proteins along the boundary:

$$\begin{aligned}
X_i^0 &= 1 \times 10^{-5}\text{nM}, \\
Y_i^0 &= 1 \times 10^{-5}\text{nM}, \\
Z^0 &= 1 \times 10^{-3}\text{nM}, \\
W^0 &= 0 \text{ nM}, \\
P^0 &= 0 \text{ nM}.
\end{aligned}$$

2. **While time steps remain,**

-Calculate the normal derivative (via approximation):

$$\frac{\partial X}{\partial z} \sim \frac{X_{i+1}^n - X_i^n}{\Delta z},$$
$$\frac{\partial Y}{\partial z} \sim \frac{Y_{i+1}^n - Y_i^n}{\Delta z}.$$

-Solve (by ODE solver or by forward Euler):

$$\begin{aligned}
\frac{dX}{dt} &= -k_1 XZ + k_2 W + d_X \frac{\partial X}{\partial z}, \\
\frac{dY}{dt} &= -k_3 YZ + k_4 P + d_Y \frac{\partial Y}{\partial z}, \\
\frac{dZ}{dt} &= -k_1 XZ + k_2 W - k_3 YZ + k_4 P, \\
\frac{dW}{dt} &= k_1 XZ - k_2 W, \\
\frac{dP}{dt} &= k_3 YZ - k_4 P.
\end{aligned}$$

-Solve:

$$\frac{\partial X}{\partial t} = D_X \frac{\partial^2 X}{\partial z^2},$$

using the following finite difference scheme:

$$\frac{X_i^{n+1} - X_i^n}{\Delta t} = D_X \frac{X_{i-1}^n - 2X_i^n + X_{i+1}^n}{\Delta z^2}.$$

-Similarly, solve the equation for Y.

**Repeat.**

As a first check on the numerical results, we replaced the boundary conditions with homogeneous conditions. This simplification should not alter the solution too much, since we expect the concentrations near the surface to be close to zero. Moreover, an exact solution to the problem

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial z^2}, \quad 0 < z < \infty, \quad 0 < t,$$

with

$$\begin{aligned}
C(0,t) &= 0, \quad 0 < t, \\
C(z,0) &= C_0, \quad 0 < z < \infty.
\end{aligned}$$

can be found using Fourier transforms:

$$C(x,t) = \frac{1}{\sqrt{4\pi t D}} \left\{ \int_0^\infty e^{-(z-\xi)^2/4tD} C_0 d\xi - \int_0^\infty e^{-(z+\xi)^2/4tD} C_0 d\xi \right\}.$$

The graph of $C(z,t)$ at $t = 5sec.$ is pictured in Figure 5 together with the graph obtained from the numerical approximation with zero-boundary conditions.

As can be seen, there is good agreement of the numerical solution and the exact solution. For the actual boundary conditions, our calculation also gave reasonable results. In particular, the depletion layer appears to be rather thin (see Figure 6). This will be important in the third part of our investigation. If we graph the concentrations of the reactants at the boundary over time, we observe that an equilibrium stage is reached much sooner than in the previous part (Figures 7 and 8).

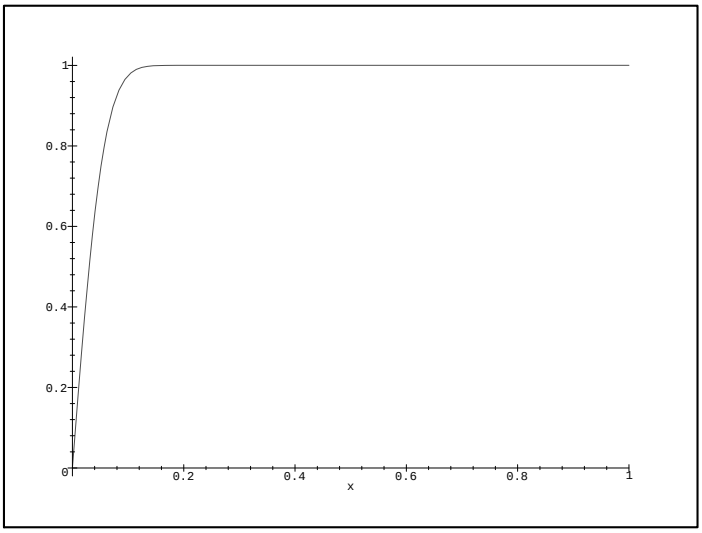



Figure 5. Exact versus numerical solutions for homogeneous boundary conditions

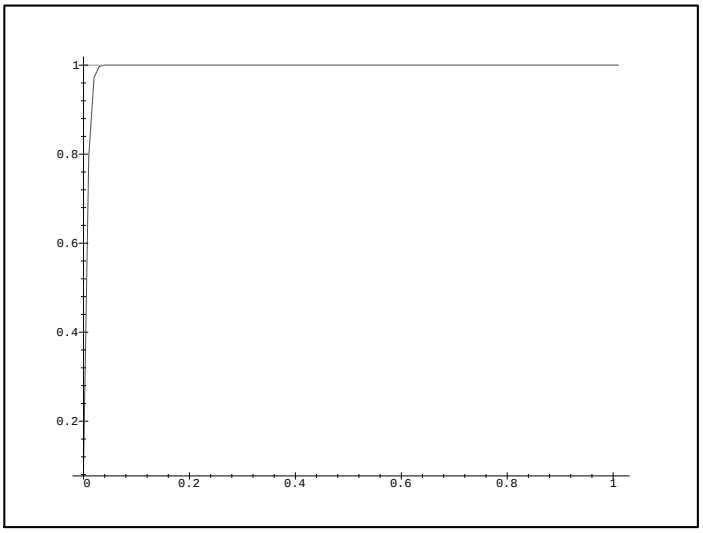
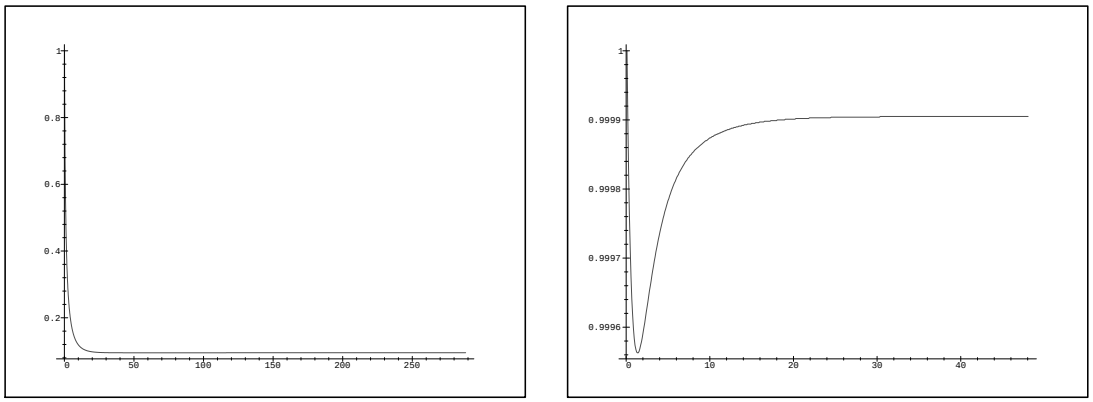
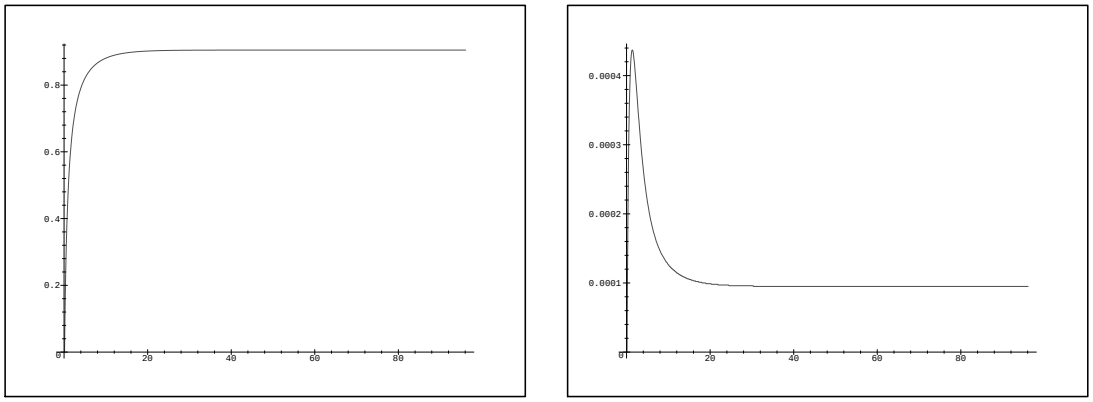



Figure 6. $X$ and $Y$ versus z at t=300 sec




(a) (b)

Figure 7. (a) $X$ versus $t$ at surface, (b) $Y$ versus $t$ at surface.




(a) (b)

Figure 8. (a) $W$ versus $t$ at surface, (b) $P$ versus $t$ at surface.

The dose-response curve which shows the bound label/antigen in the surface after 300 seconds graphed against the initial concentration of antigen in log-scale shows a nice steep decline, which warrants its practical usefulness. Inspecting the graphs of the concentrations against time during the reaction (Figures 7 and 8), we see that it would be possible to get a similar dose-response curve after 30 to 60 seconds. Of course this also depends on the particular values of the constants and initial concentrations, so it would be nice to have an estimate valid for a range of values. Using the capillary flow model it will hardly be possible to improve the results regarding the dose-response curve. The primary reason for using it is the lower cost of production.

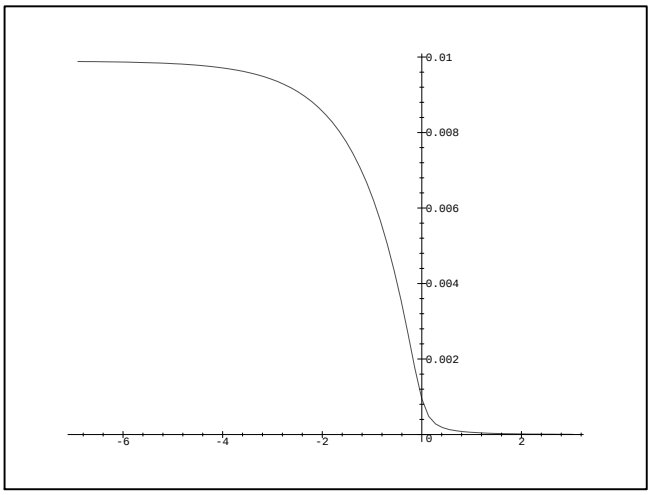



Figure 9. Dose-response curve (log scale).

## 4 Surface Reaction in a Cylinder with steady Fluid Flow

We now consider steady fluid flow in a cylinder whose inner surface is coated with antibody. The fluid, containing the antigen and antigen/label is fed into the cylinder from one side, and flows out the other. The reaction takes place in the inner surface of the cylinder. We can see directly that we have advection-diffusion equations for the concentrations of antigen and antigen/label:

$$\frac{\partial X}{\partial t} + V\frac{\partial X}{\partial z} = D_X \triangle X$$
$$\frac{\partial Y}{\partial t} + V\frac{\partial Y}{\partial z} = D_Y \triangle Y$$

The Laplacian is given in cylindrical coordinates, and we assume that the concentrations of $X$ and $Y$ are axially symmetric, so that the terms involving angle-derivatives are zero. The equations are then

$$\frac{\partial X}{\partial t} + V\frac{\partial X}{\partial z} = D_X\left(\frac{\partial^2 X}{\partial z^2} + \frac{\partial^2 X}{\partial r^2} + \frac{1}{r}\frac{\partial X}{\partial r}\right)$$
$$\frac{\partial Y}{\partial t} + V\frac{\partial Y}{\partial z} = D_Y\left(\frac{\partial^2 Y}{\partial z^2} + \frac{\partial^2 Y}{\partial r^2} + \frac{1}{r}\frac{\partial Y}{\partial r}\right)$$

On the cylinder $0 < r \leq R$ and $0 \leq z \leq 1$.

We assume the fluid to be moving strictly in the z-direction, i.e. along the axis of the cylinder. The velocity profile $V(r)$ is assumed to be quadratic:

$$V = V(r) = V_{max}\left(1 - \left(\frac{r}{R}\right)^2\right)$$

Since we observed in Section 2, that the depletion layed near the surface is rather thin, we can simplify the problem by assuming that the concentrations $X$ and $Y$ are nearly constant in the center of the cylinder. This enables us to solve the problem on the square $0 \leq z \leq 1$, $r_0 \leq r \leq R$ instead of the cylinder. The boundary conditions at $r = r_0$ are then of homogeneous Neumann type:

$$\frac{\partial X}{\partial z}(z, r_0) = 0$$
$$\frac{\partial Y}{\partial z}(z, r_0) = 0$$

Assuming steady flow of liquid through the cylinder, we also have

$$\frac{\partial X}{\partial r}(0, r) = \frac{\partial X}{\partial r}(1, r) = 0$$
$$\frac{\partial Y}{\partial r}(0, r) = \frac{\partial Y}{\partial r}(1, r) = 0$$

The fourth boundary condition is again given by the system of five ordinary differential equations, modified to take into account the transport from the fluid into the surface as in the previous section. We tried to numerically approximate a solution using a forward Euler scheme, with a second order accurate spatial discretization. Unlike in the previous case, there are great stability problems due to the terms $\frac{1}{r}\frac{\partial X}{\partial r}$ and $\frac{1}{r}\frac{\partial Y}{\partial r}$, and the fact that we have to approximate both first and second spatial derivatives. After about 40 seconds, our approximation started to blow up. We expect that by fine-tuning our code we might be able to integrate up to 60 seconds. This would probably be enough to get a decent dose-response curve. To go further in time an implicit time discretization would be advisable; however, the computation time necessary for such an approach seems prohibitive.

**Acknowledgements.**

We would like to thank Dr. Zhong Ding of Orth-Clinical Diagnostics, Inc. and Dr. Zhilin Li of North Carolina State University for their valuable contributions to this report.

## References

[1] J.R. Cannon, *The One-Dimensional Heat Equation*, Addison-Wesley, Reading, MA, 1984.

[2] Lars L. H. Christensen, "Theoretical Analysis of Protein Concentration Determination Using Biosensor Technology under Conditions of Partial Mass Transfer Limitation," *Analytical Biochemistry*, 249, pp. 153-164, 1997.

[3] D.A. Frank-Kamenetskii, *Diffusion and Heat Transfer in Chemical Kinetics*, Plenum Press, New York, 1969.