UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1998

SUMMER FACULTY RESEARCH PROGRAM FINAL REPORTS

VOLUME 5A

WRIGHT LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Colonel Jan Cerveny

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly-Clemmons

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1998

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 5 is bound in two parts, 5A and 5A. Volume 5A contains #1-17. Volume 5A contains reports #18-35. The Table of Contents for Volume 5 is included in all parts.

This document is one of a set of 15 volumes describing the 1998 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2	Armstrong Laboratory
3	Phillips Laboratory
4	Rome Laboratory
5A & 5B	Wright Laboratory
6	Arnold Engineering Development Center, United States Air Force Academy and Air Logistics Centers
	Graduate Student Research Program (GSRP) Reports
7	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10	Wright Laboratory
11	Arnold Engineering Development Center and Wilford Hall Medical Center
	High School Apprenticeship Program (HSAP) Reports
12	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A, 15B & 15C	Wright Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
'704-0188

AFRL-SR-BL-TR-00-
0780

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE
	December, 1998

4. TITLE AND SUBTITLE
1998 Summer Research Program (SRP), Summer Faculty Research Program (SFRP), Final Reports, Volume 5A, Wright Laboratory

5. FUNDING NUMBERS
F49620-93-C-0063

6. AUTHOR(S)
Gary Moore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City, CA 90230-6608

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research (AFOSR)
801 N. Randolph St.
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the SFRP participants at the Wright Laboratory.

14. SUBJECT TERMS
Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements.***

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	**PR** - Project
G - Grant	**TA** - Task
PE - Program Element	**WU** - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with....; Trans. of....; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD -	See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE -	See authorities.
NASA -	See Handbook NHB 2200.2.
NTIS -	Leave blank.

Block 12b. Distribution Code.

DOD -	Leave blank.
DOE -	Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA -	Leave blank.
NTIS -	Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self- explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

1. INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1997 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435
1997	148	98	140	427
1998	85	40	88	213

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1998 are shown in this table.

1998 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	382	85	13
(HBCU/MI)	(0)	(0)	(0)
GSRP	130	40	7
(HBCU/MI)	(0)	(0)	(0)
HSAP	328	88	22
TOTAL	840	213	42

4. SITE VISITS

During June and July of 1998, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/MIs annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1998 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1998 SRP selected participants are shown here.

1998 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	67	38	75
USAF Laboratory Funds	17	2	13
Slots Added by AFOSR (Leftover Funds)	0	0	0
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	0	0	N/A
TOTAL	84	40	88

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1998 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996	1997	1998
Faculty Members	$690	$718	$740	$740	$740	$770	$770	$793
Graduate Student (Master's Degree)	$425	$442	$455	$455	$455	$470	$470	$484
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391	$391	$400	$400	$412
High School Student (First Year)	$200	$200	$200	$200	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $52/day for faculty and $41/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and GSRP associates, 65 % claimed travel reimbursements at an average round-trip cost of $730.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-three percent (85 out of 188) of faculty associates took orientation trips at an average cost of $449. By contrast, in 1993, 58 % of SFRP associates elected to take an orientation visits at an average cost of $685; that was the highest percentage of

associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

This is the third year of using direct deposit for the SFRP and GSRP associates. The process went much more smoothly with respect to obtaining required information from the associates, about 15% of the associates' information needed clarification in order for direct deposit to properly function as opposed to 7% from last year. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1998 REPORT

The complete set of reports for the 1998 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1998 associates, as indicated below:

1998 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, USAFA, WHMC	6	11	

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. Selectees represent 44 states.

GSRP - Applicants came from 44 states. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	85
GSRP	40
HSAP	88
TOTAL	213

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	83	0	83
Master's	1	3	4
Bachelor's	0	22	22
TOTAL	186	25	109

SFRP Academic Titles	
Assistant Professor	36
Associate Professor	34
Professor	15
Instructor	0
Chairman	0
Visiting Professor	0
Visiting Assoc. Prof.	0
Research Associate	0
TOTAL	85

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	177	47
Colleague familiar with SRP	104	24
Brochure mailed to institution	101	21
Contact with Air Force laboratory	101	39
IEEE Spectrum	12	1
BIIHE	4	0
Other source	117	30
TOTAL	616	162

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	100
HSAPs	75
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B. Faculty Associates suggest higher stipends for SFRP associates.

C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1998 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

		How Many Associates Would You Prefer To Get ? (% Response)											
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
Lab	Evals Recv'd	0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
USAFA	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	**84**	**32%**	**50%**	**13%**	**5%**	**80%**	**11%**	**6%**	**0%**	**73%**	**23%**	**4%**	**0%**

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
   a. Time available for evaluation of applications:
   b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5. a. Benefits of associate's work to laboratory:
   b. Benefits of associate's work to Air Force:
6. a. Enhancement of research qualifications for LFP and staff:
   b. Enhancement of research qualifications for SFRP associate:
   c. Enhancement of research qualifications for GSRP associate:
7. a. Enhancement of knowledge for LFP and staff:
   b. Enhancement of knowledge for SFRP associate:
   c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10. a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
    b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

	AEDC	*AL*	*USAFA*	*PL*	*RL*	*WHMC*	*WL*
# Evals Recv'd	0	7	1	14	5	0	46
Question #							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1998 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 120 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1998 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13. a. Would you apply again?
    b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
    a. At Home:
    b. With Friend:
    c. On Local Economy:
    d. Base Quarters:
17. Value of orientation visit:
    a. Essential:
    b. Convenient:
    c. Not Worth Cost:
    d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1997 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	95	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1998 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1998 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 23 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from 1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

MULTIRESOLUTIONAL INFORMATION FEATURE FOR DYNAMIC CHANGE DETECTION IN IMAGE SEQUENCES

Farid Ahmed

Assistant Professor of Electrical Engineering

Pennsylvania State University, Erie

The Behrend College

School of Engineering &

Engineering Technology, Erie, PA 16563

Final Report for:

Summer Faculty Research Program

AFRL, WPAFB, Dayton, OH

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, DC

and

AFRL, WPAFB, Dayton, OH

August 1998

MULTIRESOLUTIONAL INFORMATION FEATURE FOR DYNAMIC CHANGE DETECTION IN IMAGE SEQUENCES

Farid Ahmed
Assistant Professor of Electrical Engineering
Pennsylvania State University, Erie
The Behrend College
School of Engineering &
Engineering Technology, Erie, PA 16563

Abstract

Processing of large databases of image and video sequences is becoming increasingly more challenging with the rapid growth of multimedia applications. Efficient storage and retrieval techniques necessitate the smart detection of dynamic changes in an image sequence, which may be used to form an index database. This report proposes to use a multiresolutional information metric in the determination of dynamic changes in an image sequence. In particular, scene-cut detection performance from an image sequence will be furnished. Noise tolerance of the proposed metric will also be addressed. The results developed here can be extended to the quality assessment of IR and SAR imagery.

MULTIRESOLUTIONAL INFORMATION FEATURE FOR DYNAMIC CHANGE DETECTION IN IMAGE SEQUENCES

Farid Ahmed

1 Introduction

This research originally started with the determination of the quality of an image from spatio-temporal properties of an image sequence. A number of variations of spatial and temporal information features is used in literature to determine the quality of an image. Recently, a single metric called velocital information content (VIC) has been found [1, 2] to be very effective in charting image artifacts in digital image sequences. VIC was also used to filter out some erratic frames from an IR sequence [3]. The work presented in this report is primarily motivated by these results that attempts to show the use of image quality metric in any digital image sequence for dynamic change detection. Changes in an image sequence may be sudden or slow. Scene-cut or shot-boundary is a sudden change, while zooming, panning, fading, and dissolves are examples of slow changes. A second motivation to this work is the multi-resolutional phenomenon in human visual system. We 'see' things at different resolution to extract different types of information.

In this research report, the application of a modified VIC at different resolutional levels is presented. Wavelet decomposion of image sequences is used here for the multiresolutional analysis. This report is organized as follows: an introduction to spatio-temporal feature metrics, multiresolutional analysis of metrics, simulation results with emphasis on fusion and noise tolerance, and finally comparison of different metrics.

2 Spatio-temporal metrics for Dynamic Change Detection

A sequence of images may be characterized by some spatial and temporal information content. Spatial information represents the change in an individual frame, while temporal information characterizes inter-frame changes. Histogram feature is a typically used technique for spatial information [5] in color images. Edge-based technique is also widely used in spatial feature extraction [6]. In this work, we use Sobel edge operator to exatract the spatial information.

Let us consider an image frame k, $f(i,j,k)$ having a total of P pixels. Let $s(i,j,k)$ represent the Sobel edge-operated spatial information of the frame. A metric representing spatial information content (SIC) is then given by,

$$SIC(k) = \sqrt{\frac{1}{P}\sum_i\sum_j s^2(i,j,k) - \overline{S}^2} \tag{1}$$

Here $\overline{S}$ is the mean value of spatial information of k-th frame over the total pixels.

Temporal information, on the other hand, is obtained from the frame differencing as given below. With backward difference,

$$t(i,j,k) = f(i,j,k) - f(i,j,k-1) \tag{2}$$

Temporal information content (TIC) is then given by,

$$TIC(k) = \sqrt{\frac{1}{P}\sum_i\sum_j t^2(i,j,k) - \overline{T}^2} \tag{3}$$

Here $\overline{T}$ is the mean value of temporal information of k-th frame.

As mentioned earlier, a single metric called velocital information content (VIC) which incorporates both the spatial and temporal content is found to be a better representative of change detection in image sequence.

2.1 Velocital Information Content

The velocital information is obtained from the ratio

$$v(i,j,k) = \frac{t(i,j,k)}{s(i,j,k) + a} \tag{4}$$

Here, a constant a is actually used to take care of the mathematically indeterminate situation, where $s(i,j,k) = 0$. Some statistical measure is then applied on this velocity field to come up with a single number for each frame. Below we discuss two such methods.

- Variance-based: Here, standard deviation of the velocity field of a frame is used as the metric, which is given by,

$$VIC(k) = \sqrt{\frac{1}{P}\sum_i \sum_j v^2(i,j,k) - \overline{V}^2} \tag{5}$$

  Here $\overline{V}$ is the mean value of velocital information over the total pixels.

- Entropy-based: Entropy is a widely used metric in information theory [8]. There are many different variations of entropy measurement. One of these is given by

$$VIC(k) = -\sum_i \sum_j v^2(i,j,k) \ln(v^2(i,j,k)) \tag{6}$$

  Note that, this is an un-normalized definition of entropy.

3 Multiresolutional information feature

The motivation towards using multiresolutional feature is to exploit the approximate, horizontal, vertical, and diagonal details of a dynamic change. Any 'change' in image sequence which may not be evident in one resolution or scale, may expected to be detected at some other resolution. Although, the multiresolutional analysis may incur some additional computational cost, the savings in computaion at the compressed resolution mostly offsets that.

Each resolutional decomposition demonstrates different interesting characteristics of the dynamic change which are very vital for detection. For the sake of quantitative comparison, the information features at different resolution will be combined to result in a single metric using a data fusion technique. The objective of the fusion is to obtain more discriminatory feature that can identify scene cuts from other slowly varying changes in an image sequence with higher discrimination ratio. The resulting VIC is expecterd to show sharp peaks at the location of frames corresponding to scene-cut.

3.1 Discrete Wavelet Transform for Multiresolution

Wavelet transform is a tool for decomposing a signal or an image at different resolution [9].

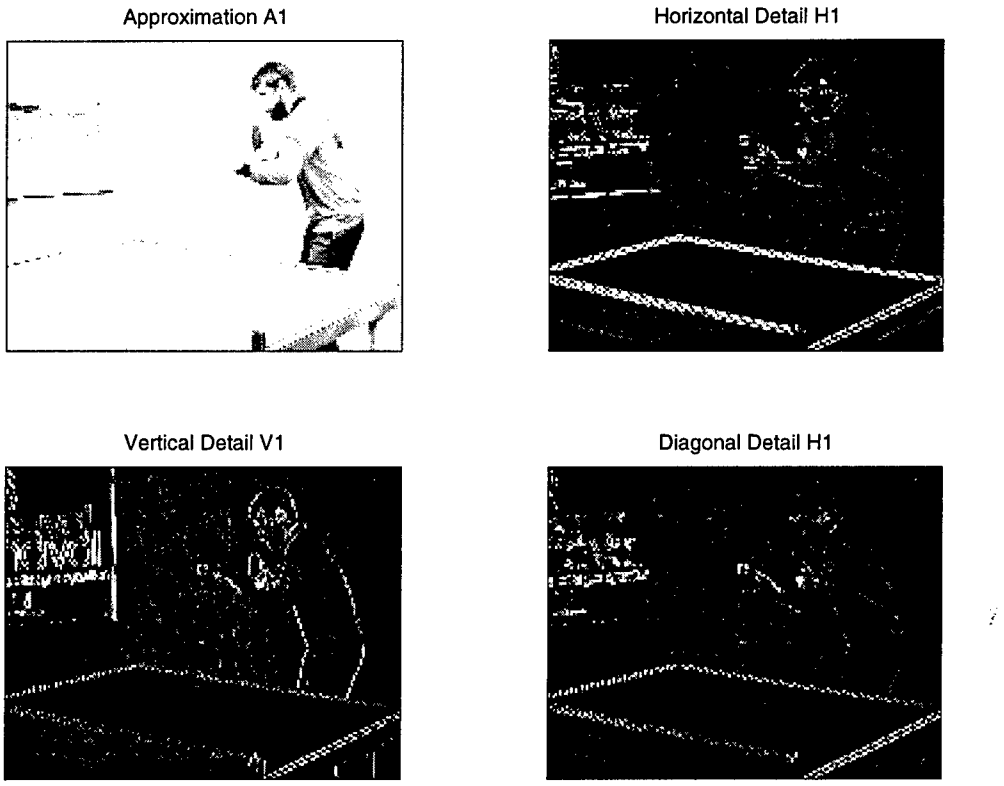


Figure 1: Decomposition of a frame into approximate and detailed coefficients

Figure 1 shows the decomposition of frame number 70 of 'Table Tennis' sequence into approximate (A1), horizontal (H1), Vertical (V1), and diagonal (D1) at resolution level 1.

Daubechies 2-tap filter was used for this analysis.

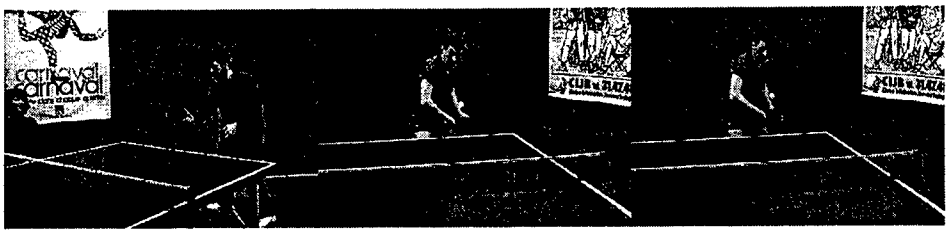


Figure 2: Frames 88, 89, and 90 (from left) of the 'Table Tennis' sequence.

4 Simulation and Results

We tested the metrics with a number of different digital image sequences. The one used in the folowing figures is the 'Table Tennis' sequence. The sequence has a total of 149 frames having a scene cut at frame number 88, as is shown from Fig. 2. Another scene-cut is at frame number 147. In order to show how well the information feature discriminates the sudden scene cut from other slowly varying dynamic changes, we define the following metric.

$$WCDR = \frac{\text{Minimum information content for scene cuts}}{\text{Maximum information content for other dynamic changes}} \tag{7}$$

Note that this is the worst case discrimination ratio.

4.1 Variance vs. Entropy-based Metrics

Fig. 3 shows the spatial, temporal, and velocital information content as formulated in Eqs. 1, 3, 5. Note that all of these formulations use the variance measurement for the metrics. Fig. 4 shows the corresponding metric values with entropy calculation as formulated in Eq. 6. It is clear from these figures that spatial information is not good for the change detection. Henceforth we shall be comparing TIC and VIC with different parameters. For both temporal and velocital metrics, gradual changes are diminished and sudden changes are enhanced with entropy measurement, as compared to the variance measurement.

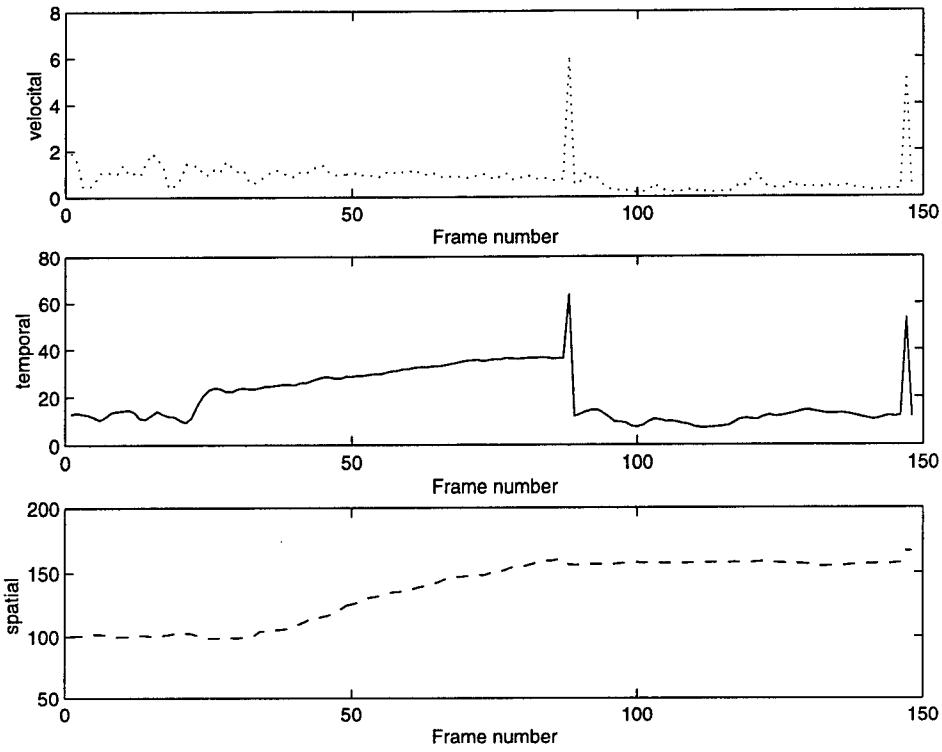


Figure 3: Spatial, Temporal, and Velocital information features from variance measurement

In other words, entropy-based metric is found to be more discriminatory in terms of sudden change detection as is clear from Fig. 5. Table 1 enumerates the corresponding WCDR values.

Table 1: Comparing variance and entropy-based metrics

Metric	WCDR	
	Velocital	Temporal
Variance-based	2.76	1.46
Entropy-based	8.53	2.13

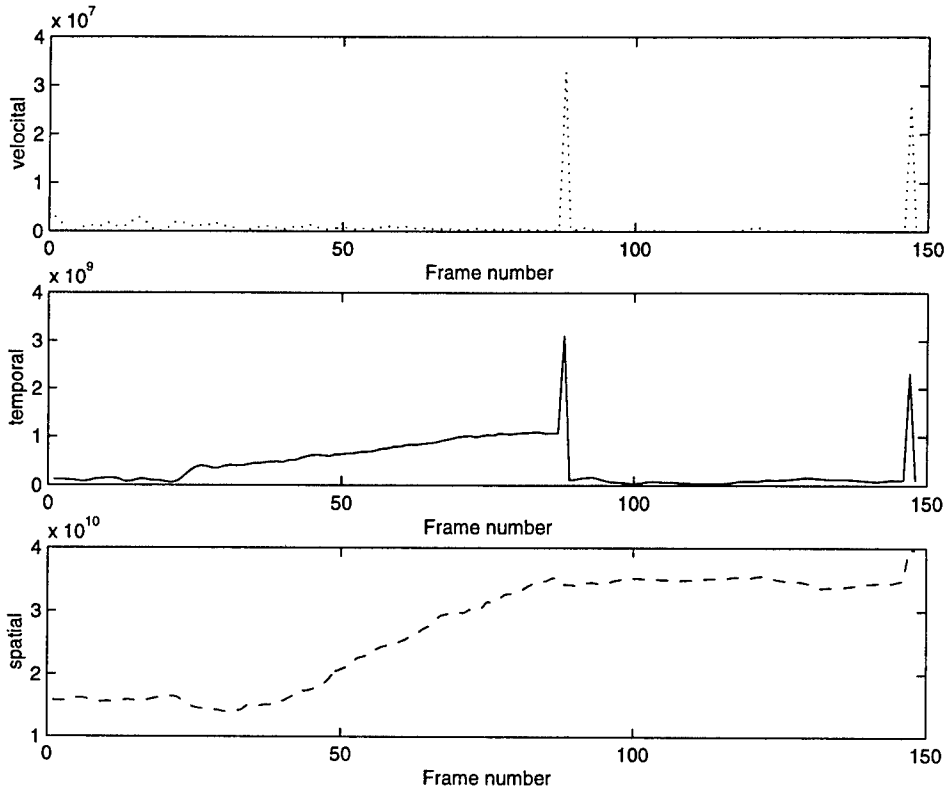


Figure 4: Spatial, Temporal, and Velocital information features from entropy measurement

4.2 Fusion of Multiresolution metrics

The frames of an image sequence are decomposed at different resolution. Original frames of the Table Tennis sequence are of 352×240. With level 1 decomposition, each of the subimages is of 176×120. Spatio-temporal metrics are computed at this compressed level. A fusion algorithm is then developed for the exploitation of the metric at different resolutions. As a simple example of fusion, suppose that the velocital information content is obtained at resolution levels A1, H1, V1, D1 by using Eq. 6, resulting in VIC_{A1}, VIC_{H1}, VIC_{V1}, and VIC_{D1}. Then the fused VIC is obtained by,

$$Fused\ VIC = VIC_{A1} \times VIC_{H1} \times VIC_{V1} \times VIC_{D1}. \tag{8}$$

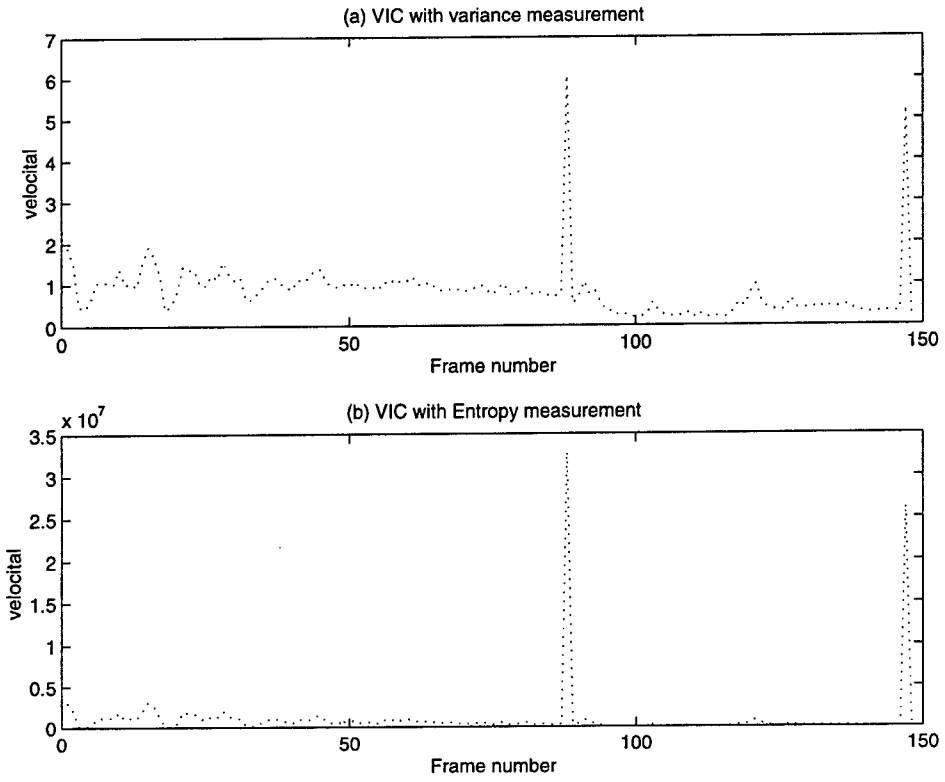


Figure 5: Velocital Information metric with variance and entropy-based calculation

4.3 Effect of different wavelet Filters

Wavelets are classified on the basis of their impulse responses, localization, regularization, orthogonality, and symmetry properties. Haar, Daubechies, Coiflets, and Symlets are orthogonal wavelets with FIR filter response. Meyer, Morlet, and Mexican Hat wavelets do not have FIR response. We here use the first type along with some biorthogonal wavelets. Haar wavelet is the smallest in support. Depending upon the order, Daubechies, Coiflet, and Symlet have different degrees of support. For example, Daubechies 2-tap filter has a support of 3, while the 10-tap has a support of 19. The general notion of support size is that wavelets with small support perform better in detecting the details.

Table 2 shows the WCDR values for both temporal and velocital information content with different wavelet filters. Note that, in general, the dicrimination ratio for VIC is much higher than that of TIC. Another observation is that the compressed wavelets (having low

Table 2: Discrimination performance for different wavelet filters

Wavelet Filter	WCDR	
	Velocital	Temporal
Haar	19.5	2.96
Daubechies 2-tap	28.12	3.03
Daubechies 3-tap	16.27	3.05
Daubechies 4-tap	9.05	3.06
Daubechies 10-tap	4.22	3.17
Biorthogonal 3.7	14.98	2.45
Coiflet 2	14.06	3.16
Coiflet 4	5.93	3.2
Symlet 3	16.3	3.05

scale and smaller support) can detect more rapidly changing details, while the stretched wavelets detect coarse features and slowly changing details. As a result information content is expected to be more with the wavelets with smaller support, which is also evident from Table 2.

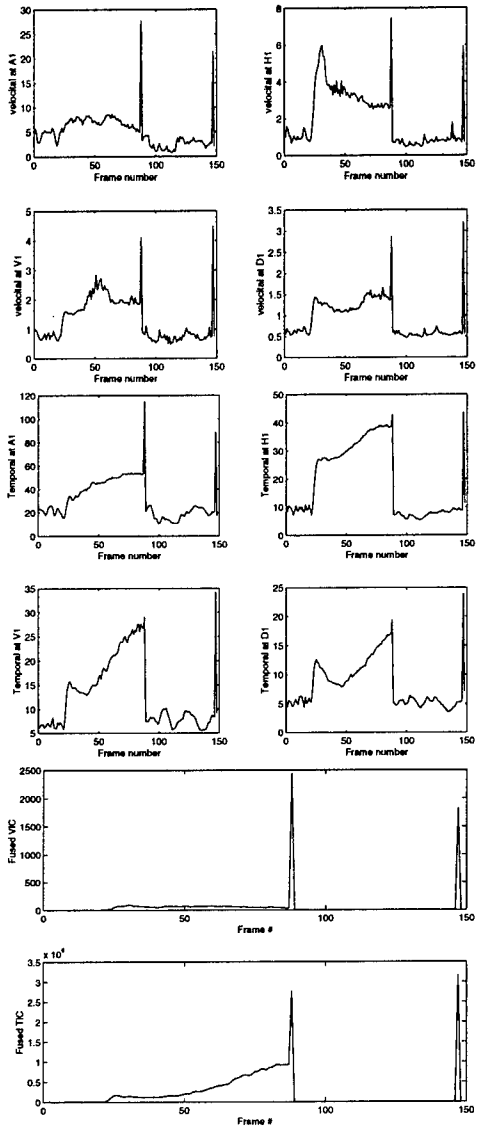


Figure 6: Multiresolutional and Fused VIC and TIC with Haar wavelet

For their better discrimination ratio, the Haar and Daubechies' 2-tap filter is used for subsequent analysis. Table 3 shows the comparison of these with entropy measurement.

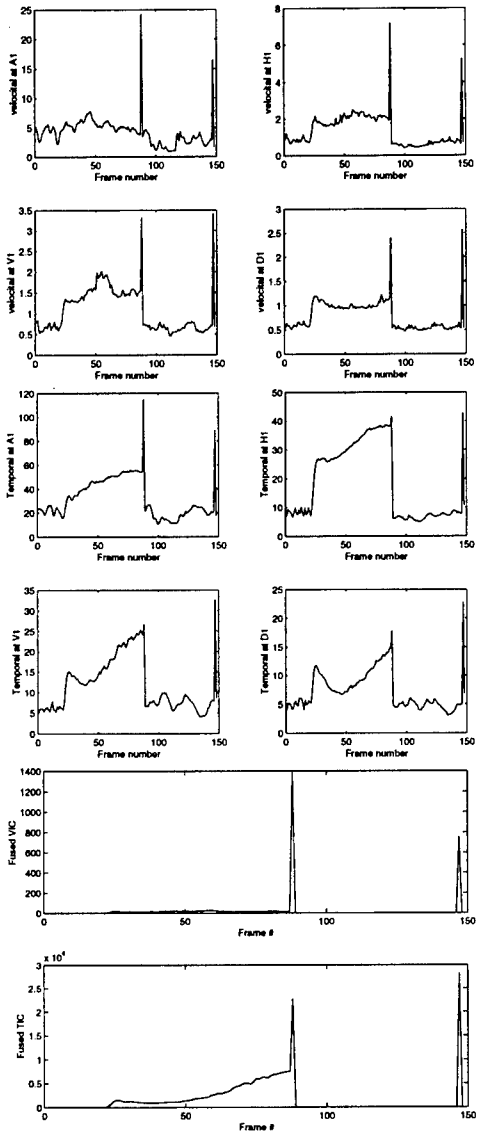


Figure 7: Multiresolutional and Fused VIC and TIC with Daubechies 2-tap wavelet (variance-based)

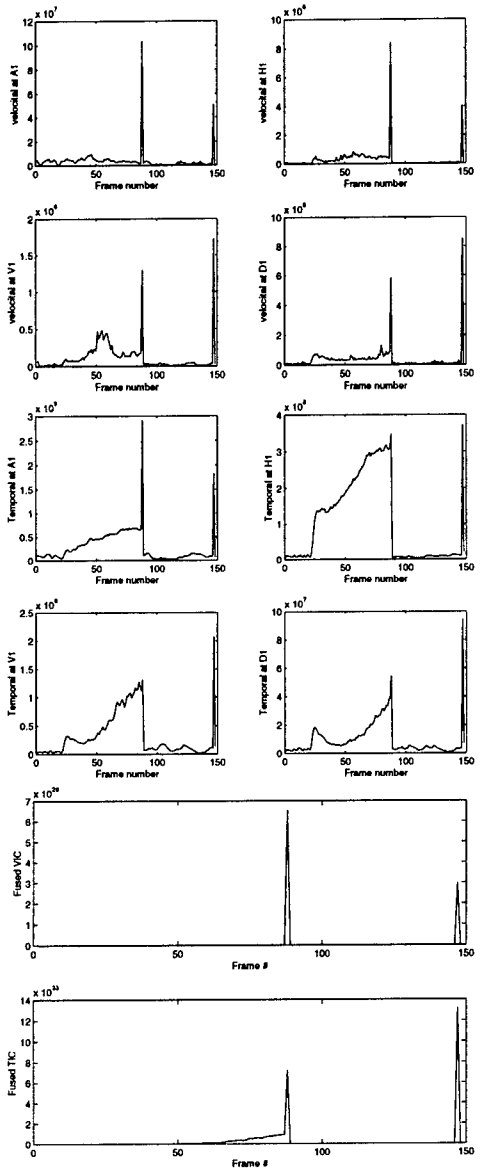


Figure 8: Multiresolutional and Fused VIC and TIC with Daubechies 2-tap wavelet (entropy-based)

Table 3: Discrimination performance with entropy measurement

Wavelet Filter	WCDR	
	Velocital	Temporal
Haar	8.6×10^4	7.77
Daubechies 2-tap	7.87×10^3	7.95

Figs. 6, 7, and 8 demonstrates the resulting VIC, TIC, and fused VIC values. Careful observation of the plots corresponding to the fused VICs will show that they are very robust in detecting the sharp changes.

Next, we obtained the information features in a little different way. Wavelet decomposition was done on the frame differences, to obtain the temporal information at different resolution. Spatial information was also obtained by applying wavelet decomposition on the Sobel-operated images. Result is shown in Fig. 9.

4.4 Noise Tolerance

Velocital information content is found to be very sensitive to noise as compared to the temporal information content. We used Daubechies' 2-tap filter for all the simulations for noise tolerance. Figure 10 shows the results.

Figure 11 shows the VIC metric and TIC metric at approximation level 1 with wavelet decomposition. Note that due to the denoising property [10] of wavelet the sensitivity of VIC to noise is reduced as is evident from a comparison with figure 10. Table 4 shows the comparisons of VIC and TIC with and without wavelet denoising for different levels of noise contents.

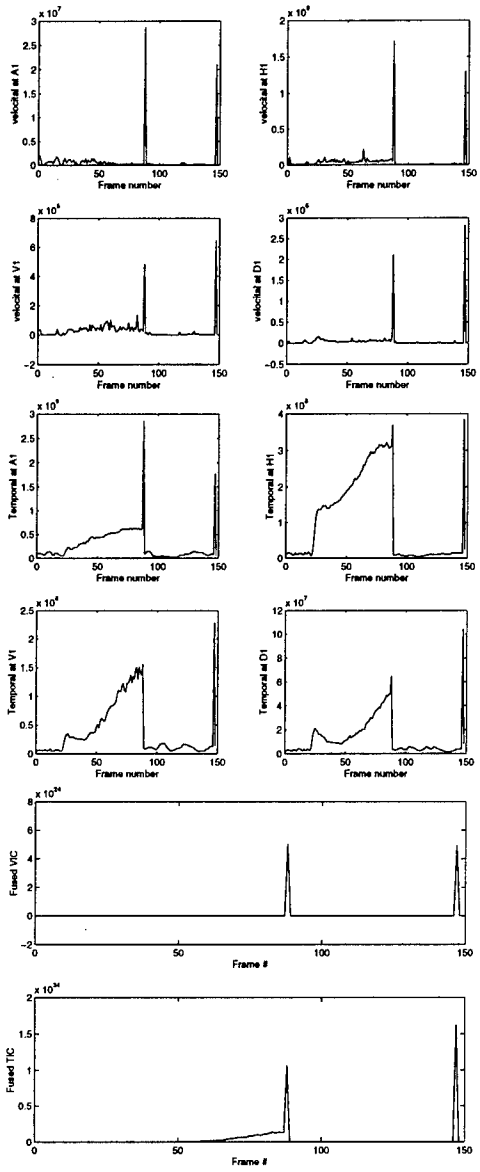


Figure 9: Multiresolutional VIC and TIC with Haar wavelet using Entropy measurement

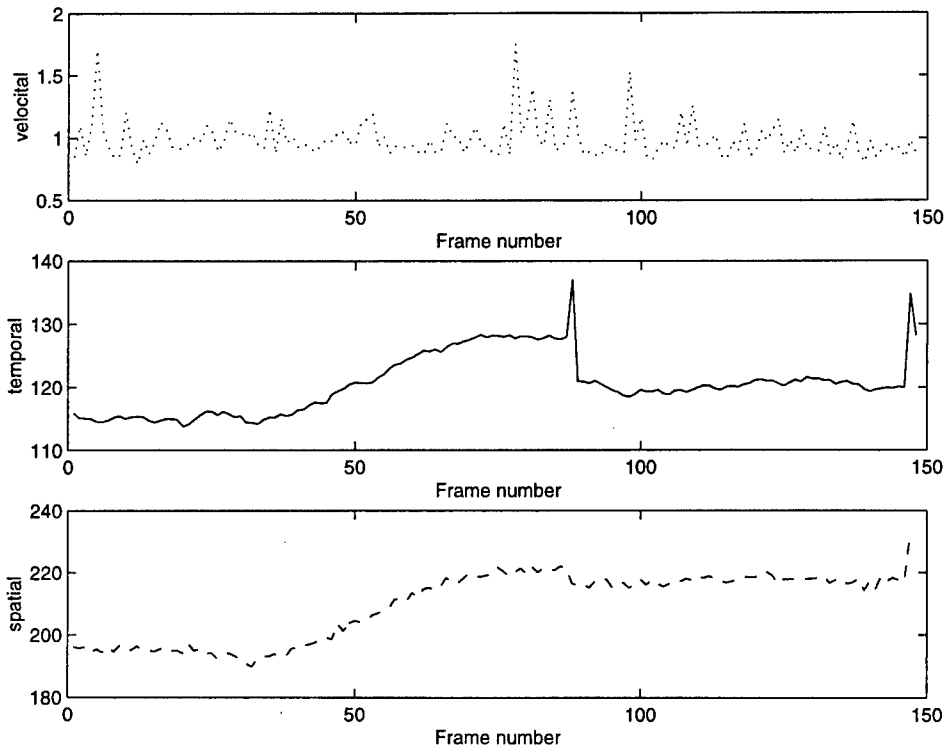


Figure 10: Spatial, Temporal, and Velocital information features with SNR=1 from variance measurement

5 Conclusion

Multiresolutional information features in terms of spatial, temporal, and velocital metrics for image sequences have been investigated. Results will find application in the dynamic change detection as well as quality assessment of digital image sequences.

REFERENCES

1. Gregory J. Power, Determining a confidence factor for automatic target recognition based on image sequence quality in *Algorithms for Synthetic Apertuire Radar Imagery V*, E. G. Zelnio *et al.* eds., Proc. SPIE 3370, Orlando, April 1998.

2. Gregory J. Power, M. A. Karim, and Farid Ahmed, Charting Image Artifacts in Digital Image Sequences using velocital information content, in *Applications of Digital*

Table 4: Discrimination performance for noisy sequence

SNR	Without wavelet denoising		With wavelet denoising	
	Velocital	Temporal	Velocital	Temporal
1	-	1.2	1.26	1.24
5	-	1.81	1.98	1.98
10	1.2	2.01	5.7	2.13

Image Processing XXI, A. G. Tescher *et al.* eds., Proc. SPIE 3460, San Diego, July 1998.

3. Gregory J. Power, M. A. Karim, Automatic Removal of Poor Quality Images from Digital Image Sequences, in *Airborne Reconnaissance XXII*, Proc. SPIE 3431, San Diego, July 1998.

4. Gregory J. Power, M. A. Karim, and Farid Ahmed, A Velocital Information Feature for Charting Spatio-Temporal Changes in Digital Image Sequences, submitted to *Journal of Electronic Imaging*, June 1998.

5. M. J. Swain and D. H. Ballard, "Color Indexing," *International Journal of Computer Vision*, Vol. 7, No. 1, pp. 11-32, 1991.

6. R. C. Gonzalez, and P. Wintz, *Digital Image Processing*, Reading, Massachussets, Addison-Wesley Publishing Co., 1987.

7. S. C. Yoon, K. Rotakonda and N. Ahuja, "Region-based Video Coding Using a Multiscale Image Segmentation," *Proc. IEEE Int. Conf. Image Processing*, pp. 510-513, 1997.

8. Claude E. Shannon, Aaron D. Wyner, N. J. A. Sloane, "Claude Elwood Shannon : Collected Papers," *IEEE Press*, 1993.

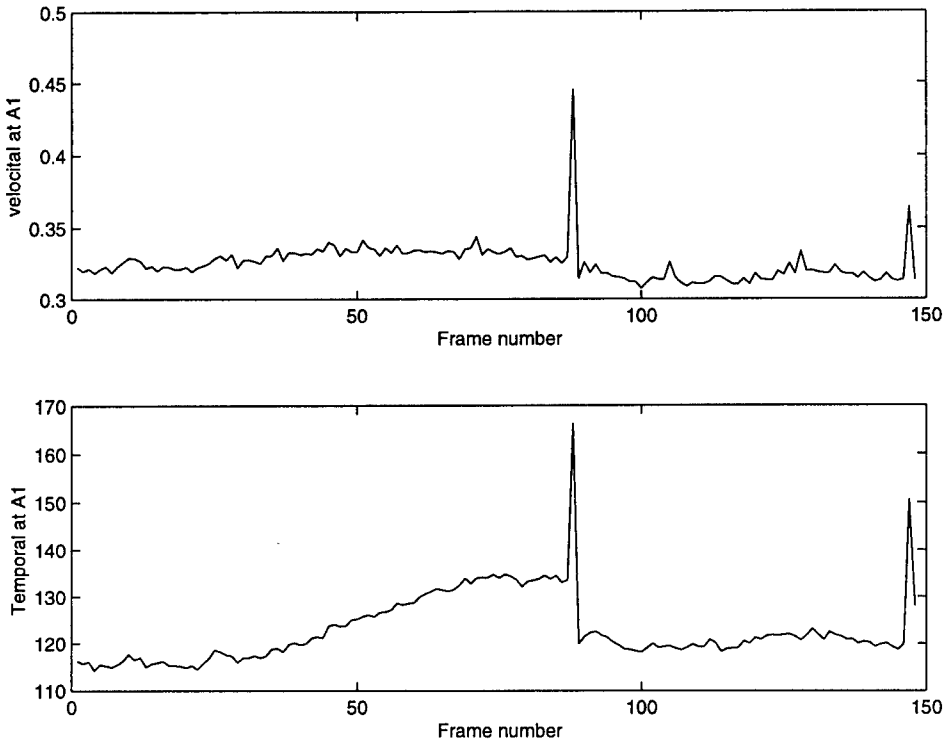


Figure 11: Temporal and Velocital information features with SNR=1 from variance measurement at approximation level 1 with denoised sequence

9. S. Mallat, "A theory for multiresolutional signal decomposition: the wavelet representation," *IEEE PAMI*, vol. 11, no. 7, pp. 674-693, 1993.

10. D. L. Donoho, "De-noising by soft thresholding," *IEEE Transaction Inf. Theory*, Vol. 41, no. 3, pp. 613-627, 1995.

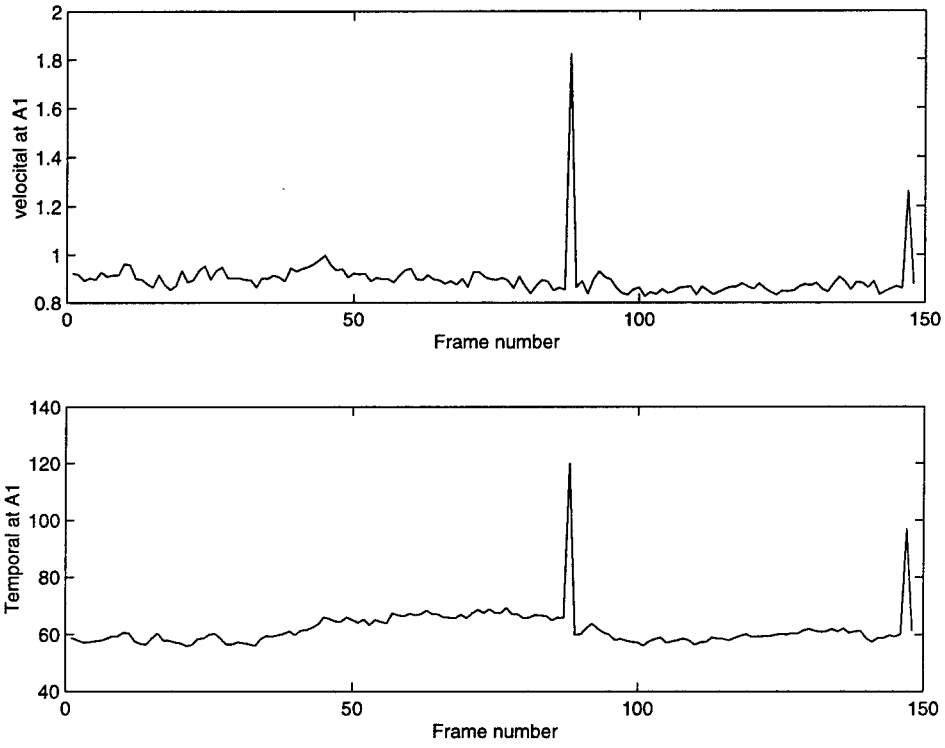


Figure 12: Temporal and Velocital information features with SNR=10 from variance measurement at approximation level 1 with denoised sequence

SYNTHESIS OF 7-BENZOTHIAZOL-2-YL-9,9-DIDECYLFLUORENE-2-YLAMINE: A VERSATILE INTERMEDIATE FOR A NEW SERIES OF TWO PHOTON ABSORBING MATERIALS & THREE NOVEL TPA DYES

Kevin D. Belfield
Associate Professor
Department of Chemistry

University of Central Florida
4000 Central Florida Blvd., P.O. Box 162366
Orlando, FL 32816-2366

Final Report for:
Summer Faculty Research Program
Air Force Research Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling AFB, Washington, DC

and

Air Force Research Laboratory (AFRL/ML)

July 1998

SYNTHESIS OF 7-BENZOTHIAZOL-2-YL-9,9-DIDECYLFLUORENE-2-YLAMINE: A VERSATILE INTERMEDIATE FOR A NEW SERIES OF TWO PHOTON ABSORBING MATERIALS & THREE NOVEL TPA DYES

Kevin D. Belfield
Associate Professor
Department of Chemistry
University of Central Florida

Abstract

The synthesis of of a new, useful intermediate, 7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine, was accomplished starting with fluorene. Regiospecific nitration of fluorene with HNO_3 and CH_3CO_2H at 85 °C led to formation of 2-nitrofluorene in 80% yield. Regiospecific iodonation of 2-nitrofluorene with I_2, $NaNO_2$, H_2SO_4, and CH_3CO_2H at 115 °C afforded 7-iodo-2-nitrofluorene in 79% yield, after recrystallization. Alkylation of 7-iodo-2-nitrofluorene was accomplished with 1-bromodecane and KOH in DMSO at room temperature, yielding 9,9-didecyl-7-iodo-2-nitrofluorene in 77% yield (after column chromatographic purification). 2-(Tri-n-butylstannyl)benzothiazole was synthesized in 90% yield (after distillation) by reaction of benzothiazole with n-BuLi at -78 °C in THF, followed by addition of tri-n-butyltin chloride. 9,9-Didecyl-7-iodo-2-nitrofluorene and 2-(tri-n-butylstannyl)benzothiazole were subjected to Stille coupling with either tetrakis(triphenylphosphine)palladium (0) or dichlorobis(triphenylphosphine)palladium (II) in toluene at 110 °C under Ar, providing 2-(9,9-didecyl-7-nitrofluoren-2-yl)benzothiazole in 61% yield. Quantitative reduction of 2-(9,9-didecyl-7-nitrofluoren-2-yl)benzothiazole with NH_2NH_2 and 10% Pd/C in EtOH/THF at 70 °C produced 7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine in 30% overall yield from fluorene.

SYNTHESIS OF 7-BENZOTHIAZOL-2-YL-9,9-DIDECYLFLUORENE-2-YLAMINE: A VERSATILE INTERMEDIATE FOR A NEW SERIES OF TWO PHOTON ABSORBING MATERIALS & THREE NOVEL TPA DYES

Kevin D. Belfield

Introduction

Multiphoton absorption can be defined as a simultaneous absorption of two or more photons via virtual states in a medium.[1] The process requires high peak power which is available from pulsed lasers. Even though multiphoton processes have been known for some time, materials that exhibit a multiphoton absorption have yet to find widespread application. The discovery of multifunctional organic materials with large multiphoton absorption cross sections has spawned a new area of research in the photonic and biophotonic fields. In particular, two photon pumped upconverted fluorescence (Figure 1) has enormous implications for multiphoton absorption-induced optical power limiting, curing of polymeric materials, stereolithography, fabrication of microelectromechanical devices, and nondestructive imaging of coating and composite interfaces.

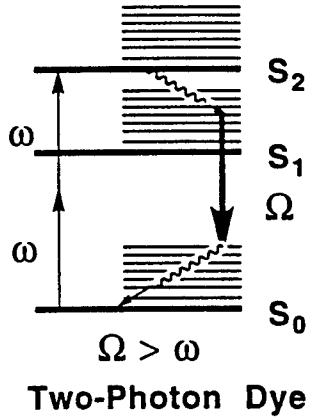


Figure 1. Illustration of upconverted fluorescence via two photon absorption.

Polymers for Optical Limiting

The optical response of an ideal optical power limiter is shown in Figure 2. The ideal optical limiter is completely transparent at low light intensities until a certain intensity level is reached. Above this threshold, the transmitted intensity remains at a constant value. Most current materials available for photonic and biophotonic applications suffer from speed, concentration, phase separation, or solubility problems. To overcome shortcomings of current materials, new, efficient two photon-absorbing chromophores could be covalently attached to polymers, affording good film forming materials which can be fabricated in a variety of configurations.

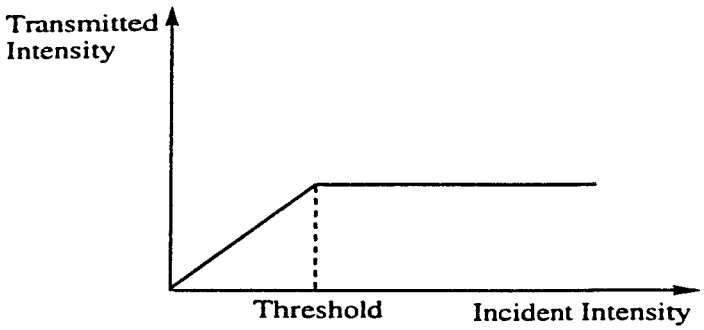


Figure 2. The optical response of an ideal optical power limiter.

Multiphoton Microscopic Fluorescence Imaging

Two photon induced fluorescence has been coupled with laser scanning microscopy to probe surfaces.[2] The major advantage of this method over single photon scanning is that the fluorescence intensity of a two photon process is quadratically dependent on the illumination intensity. This makes the fluorescence emission limited to the vicinity of the focal point and, hence, it is possible to achieve depth discrimination. Multiphoton confocal laser scanning microscopy could be a useful, nondestructive tool to study surfaces, interfaces, and fractures in polymer or glass specimens. Recently, fractures in polymer samples and polymer coatings layers (Figure 3) were imaged by this technique.[3] The images were of a methacrylate polymer matrix containing organic fluorophores. Two photon multichannel confocal microscopy was demonstrated to be useful to probe and construct images of these multilayered coatings. In principle, the dye can be covalently attached to a monomer, hence be permanently and uniformly incorporated into a polymer at high concentrations.

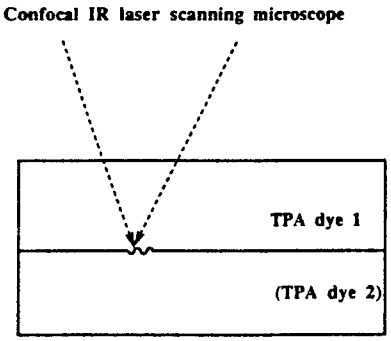


Figure 3. Interfacial imaging of coatings using two photon absorbing (TPA) coatings and confocal laser scanning microscopy.

Two Photon Curing

In many cases, thermal curing of monomers is impractical due to the size or nature of the composite, sealant, or adhesive structure, e.g., damaged aircraft parts or seams between panels. Moreover, photocuring is impractical for large objects because of the limited absorption depth of UV light. Two photon pumped upconverted fluorescence has potential for 3-D or spatially resolved photoinitiated polymerization, particularly in composites, adhesives, and sealant applications. For example, if an IR absorbing two photon dye is used, IR radiation could be used to induce two photon absorption of the dye, followed by subsequent emission of a visible photon. The visible photon could be absorbed by a photoinitiator, as illustrated in Figure 4. The two

photon absorption and fluorescence emission can be tuned by varying the molecular structure of the dye. Hence, IR radiation can be used to facilitate visible-initiated polymerization. The deep penetration of IR radiation could, potentially, be exploited to achieve deep curing of monomers, for example, in obstructed cracks and seals. With utilization of a confocal configuration, true three-dimensional (stereolithographic) and near isotropic polymerization should be possible. A logical next step would be the synthesis of monomers bearing two photon dyes and demonstration of their two photon-induced polymerization.

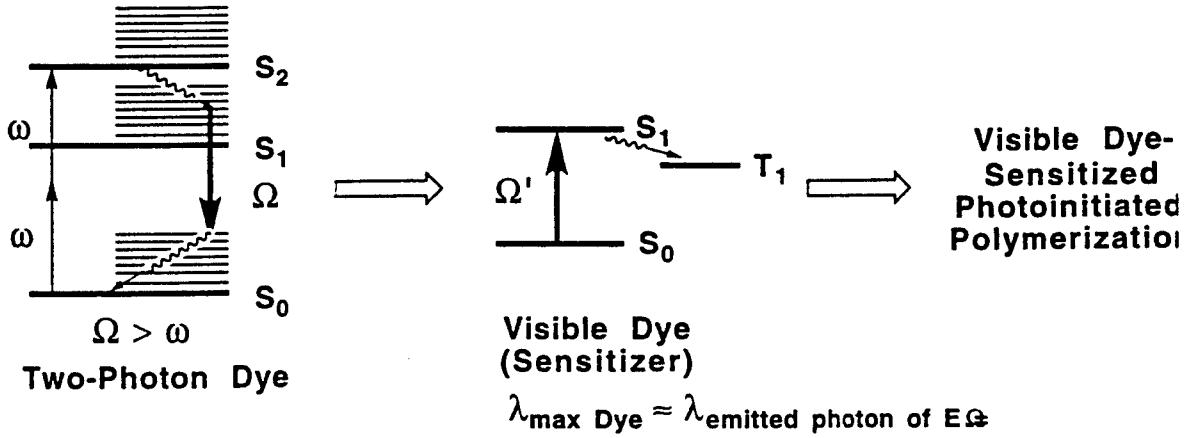
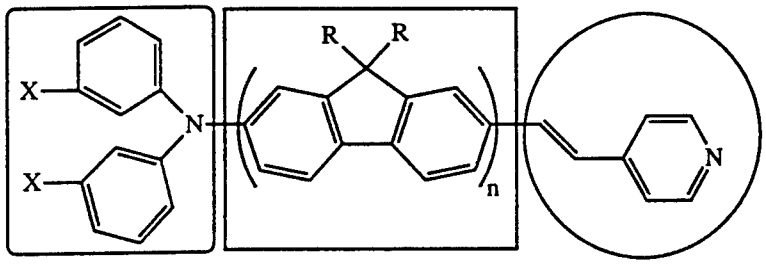


Figure 4. Schematic illustration of two photon absorption, fluorescence emission, one photon absorption, and photoinitiated polymerization.

Fluorene-based Two Photon Absorbing Chromophores

Recently, Reinhardt *et al.*[4] have prepared and characterized a number of asymmetric fluorene-based dyes that exhibit two photon absorption at 800 nm. The dyes are comprised of a polarizable π electron bridge flanked by an arylamine π electron donor functionality and a heteroarylamine π electron donor functionality and a heterocyclic π electron accepting moiety, as illustrated in Figure 5. In general, large two photon absorption cross sections are realized in such systems with strong electron-donating functionality conjugated effectively with strong electron-withdrawing functionality. Thus, the π electron bridge should be planar to facilitate effective conjugation, a condition fulfilled sufficiently well by the fluorene framework.



R = C_2H_5, C_6H_{13}, $C_{10}H_{21}$
X = H, OCH_3, OH, 2-thiophenyl
n = 0, 1

Figure 5. Generic structure of asymmetric fluorene-based chromophore.

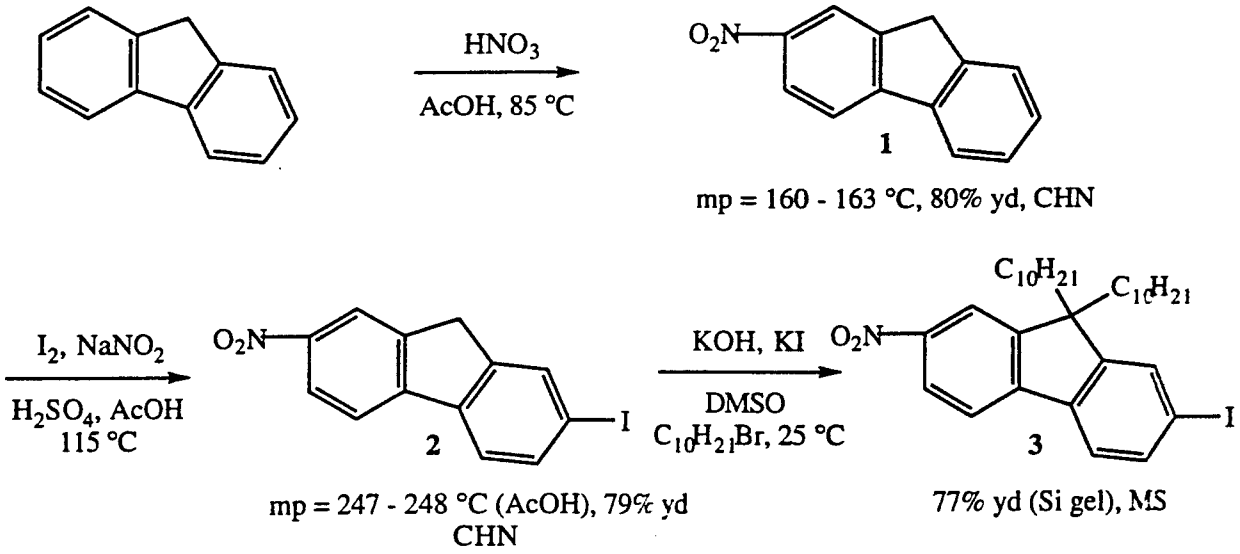
Objectives

The aim this summer was to synthesize a versatile intermediate which possessed the fluorene π electron bridge, electron-donating amine functionality, and electron-withdrawing benzothiazole moiety. In addition, long alkyl groups at the 9-position of the fluorene ring system will be incorporated to impart solubility, inhibit aggregation and crystallization, and, in suitable derivatized compounds, lead to the formation of amorphous glasses. The synthesis should be efficient, allowing for ready and economical scale-up. Functionalization of the intermediate will provide novel two photon absorbing materials. 7-Benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine (**6**) was designed and selected to be the target intermediate. In addition, three novel two photon absorbing dye motifs were designed to be prepared straight away from the versatile intermediate **6**.

Results and Discussion

Two versatile fluorene-based intermediates, 7-iodo-9,9-didecyl-2-nitrofluorene (**3**) and 7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine (**6**), were prepared in a relatively efficient multistep synthesis from readily available fluorene (Schemes 1, 5, and 6). The synthesis and characterization of these intermediates are discussed below, as are the preparations of three derivatives (**8**, **10**, and **11**) which should serve as two photon absorbing dye candidates.

Scheme 1

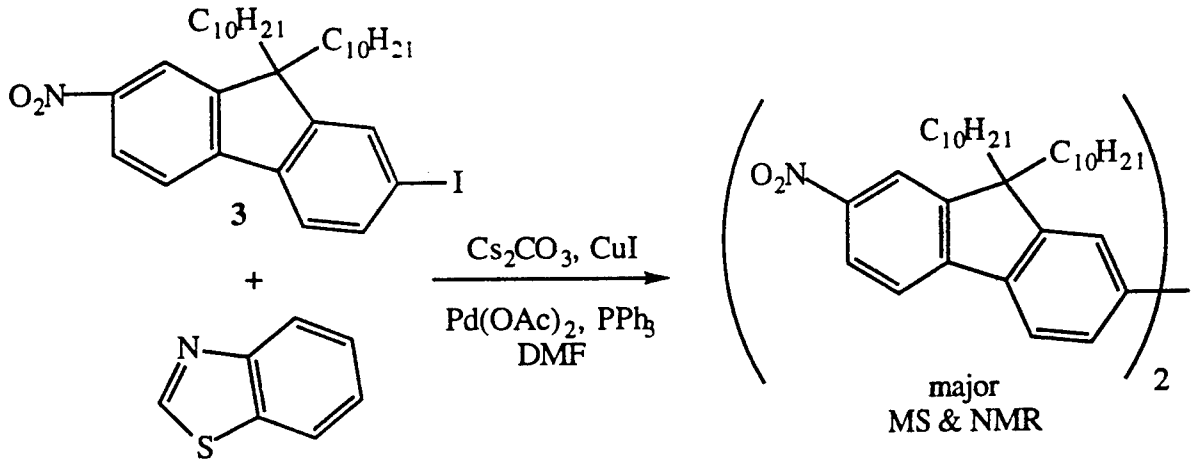


2-Nitrofluorene (**1**) was prepared in 80% yield by regiospecific nitration of fluorene with nitric acid in AcOH at 85 °C. The melting point (160-163 °C) was comparable with the literature value (157 °C),[5] while CHN analyses were in accord with calculated values. The next step in the multistep synthesis involved regiospecific iodonation of 2-nitrofluorene with iodine, sodium nitrite, AcOH and H_2SO_4 at 115 °C, affording 7-iodo-2-nitrofluorene (**2**) in 77% yield after recrystallization from AcOH. Again, the melting point (245-246 °C) was consistent with the literature value (244-245 °C),[6] while CHN analyses were in good agreement with calculated values. Dialkylation of **2** was accomplished by generation of the fluorenyl anion with KOH in DMSO and subsequent alkylation with 1-bromodecane in the presence of KI at room temperature.[4] 7-Iodo-9,9-didecyl-2-nitrofluorene (**3**) was obtained as a yellow solid in 77% isolated yield, after column chromatography. MS analysis and ^1H NMR confirmed formation of

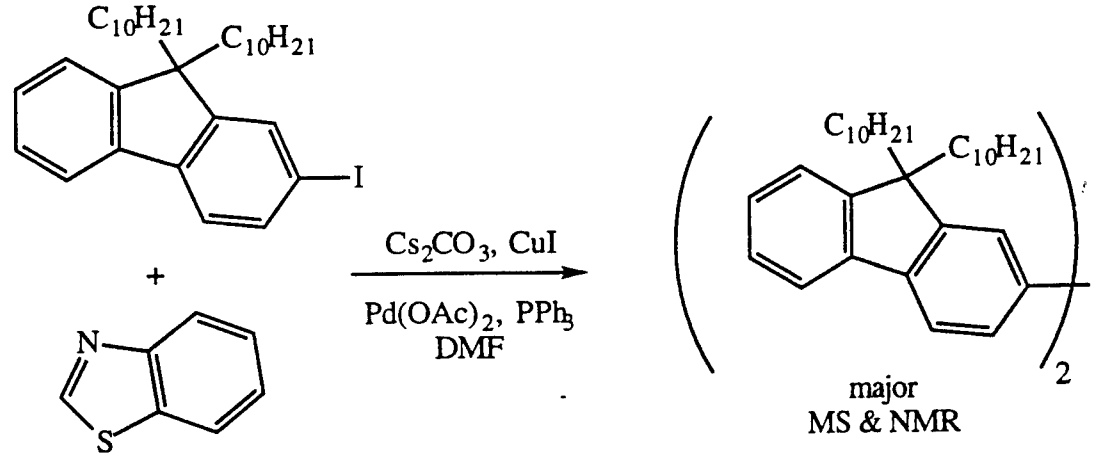
the desired product.

Next was formation of 2-(9,9-didecyl-7-nitrofluoren-2-yl)benzothiazole (**5**) from 7-iodo-9,9-didecyl-2-nitrofluorene (**3**). Although conversion of the iodide to carboxaldehyde, followed by condensation with 2-aminothiophenol was a "stand by" option, it was desirable to develop a more expeditious, direct synthetic method for the conversion of **3** to **5**. A direct coupling reaction,[7] mediated by $Pd(OAc)_2$, CuI, PPh_3, and Cs_2CO_3, was attempted between **3** and benzothiazole (Scheme 2). Surprisingly, a nitrofluorene dimer was isolated as a yellow crystalline solid as the major product, as confirmed by MS and ^1H NMR. To probe what effect the nitro group may have on the reaction, 9,9-didecyl-2-iodofluorene (prepared from decylation of 2-iodofluorene) was subjected to Pd-catalyzed coupling with benzothiazole. Consistent with the nitro analog, 9,9-didecyl-2-iodofluorene underwent reaction to form, predominantly, the dimer, as illustrated in Scheme 3.

Scheme 2



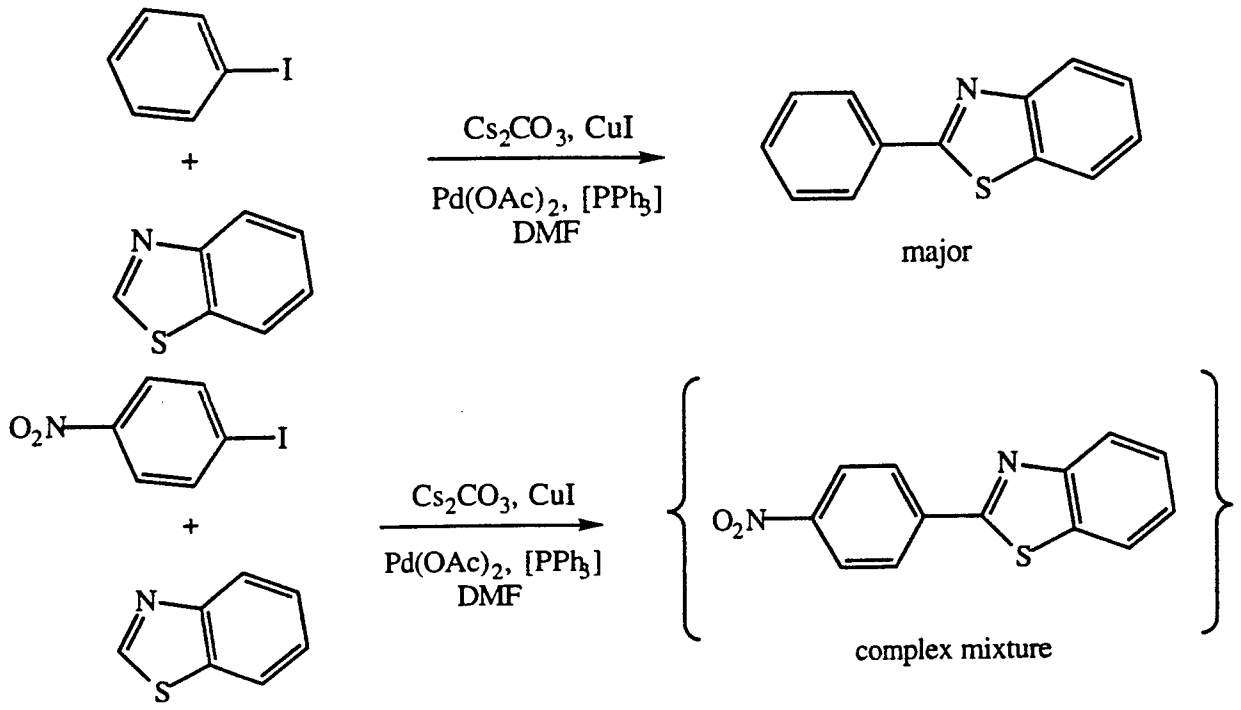
Scheme 3



To further assess the feasibility of the direct coupling, the reported example was carried out.[7] Thus, iodobenzene was reacted with benzothiazole (Scheme 4), mediated by $Pd(OAc)_2$, CuI, and

Cs_2CO_3, both in the presence and absence of PPh_3 (in one experiment tri-o-tolylphosphine was employed). In all cases, the expected coupling product, 2-phenylbenzothiazole, was secured. A similar set of experiments were conducted with 4-iodonitrobenzene, resulting in a complex mixture of products (by TLC). Hence, the direct coupling strategy was abandoned.

Scheme 4

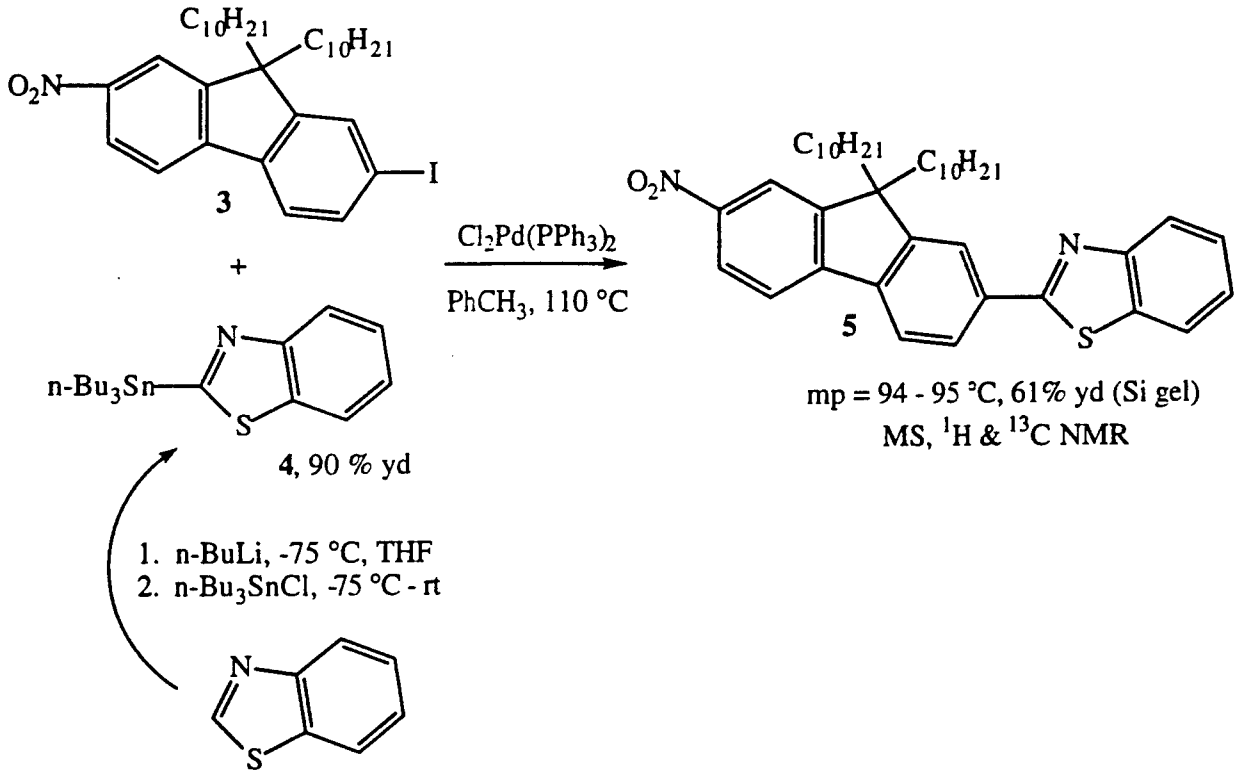


Finally, a Stille coupling procedure was devised to prepare **5**, as illustrated in Scheme 5. (2-Tri-n-butylstannyl)benzothiazole (**4**) was prepared in 90% yield by treating benzothiazole with n-BuLi in THF at -75 °C, followed by addition of tri-n-butyl tin chloride.[8] Pd-catalyzed coupling was then performed between **4** and 7-iodo-9,9-didecyl-2-nitrofluorene (**3**) in refluxing toluene with either tetrakis(triphenylphosphine)palladium (0) or dichlorobis(triphenylphosphine)palladium (II).[9] 2-(9,9-Didecyl-7-nitrofluoren-2-yl)benzothiazole (**5**) was obtained as yellow crystals in 61% yield, after column chromatographic purification (mp = 94 - 95 °C). Clearly evident in the IR spectrum were absorptions attributable to asymmetric and symmetric NO_2 stretching vibrations a 1519 and 1337 cm^{-1}, respectively. Also present was the benzothiazole C=N stretching vibration at 1589 cm^{-1}. Nitro analog **5** exhibited a UV-visible absorption ranging from 220 to 420 nm with λ_{max} = 365 nm. The diagnostic ^{13}C NMR resonance of the carbon in the 2-position in the benzothiazole ring was present at 168 ppm.[10]

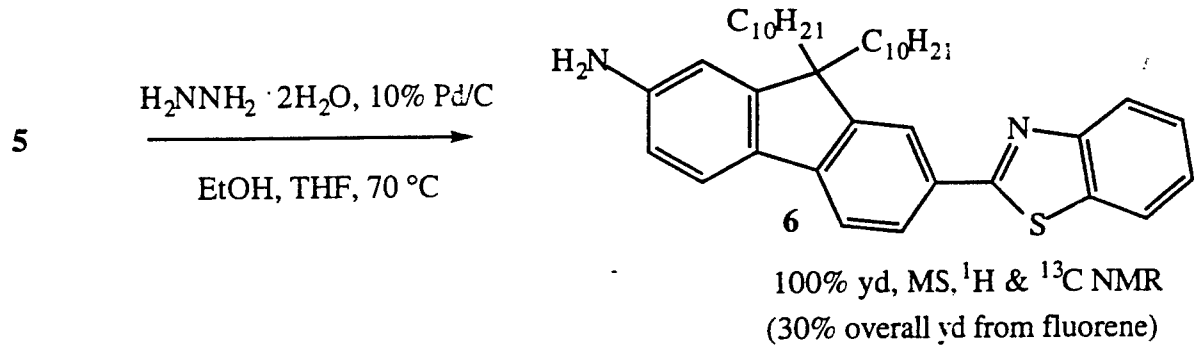
Fast, quantitative reduction of **5** was achieved using hydrazine hydrate[11] and 10% Pd/C in a 1:1 mixture of EtOH and THF at 70 °C (Scheme 6), providing 7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine (**6**) as a bright yellow sticky, viscous oil. As expected, **6** exhibited a 15 nm bathochromic shift in its UV-visible spectrum relative to **5** (λ_{max} = 380 nm, ranging from 210 - 450 nm). Characteristic stretching absorptions were observed in the IR spectrum for NH_2 (3741 and 3383 cm^{-1}) and benzothiazole C=N (1603 cm^{-1}). MS and ^1H NMR provided

additional confirmation of structure. Conclusive proof of structure was garnered by the anticipated upfield shift in the ^{13}C NMR of carbons *ortho* and *para* to the amine in the fluorene ring of **6**. *Ortho*-carbons 3 and 1 were observed at 109.8 and 114.3 ppm, respectively, compared to the corresponding carbons in **5** at 118.6 and 122.0 ppm, respectively. Even more pronounced was the chemical shift of *para*-carbon 11 of **6** at 125.0 ppm relative to that of **5** at 147.9 ppm. Amine **6** displayed a brilliant blue fluorescence in solution upon radiation with long wavelength UV radiation.

Scheme 5

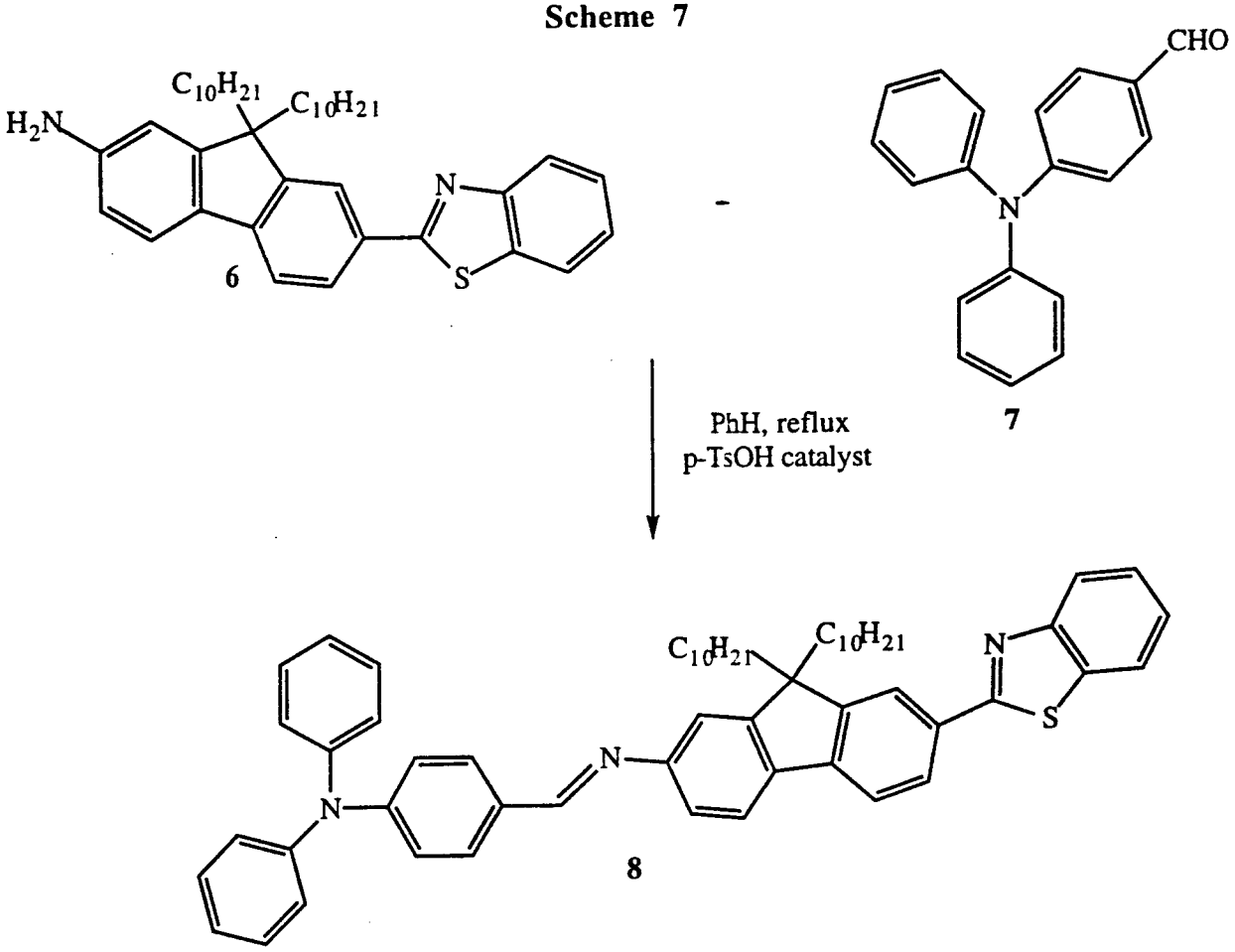


Scheme 6



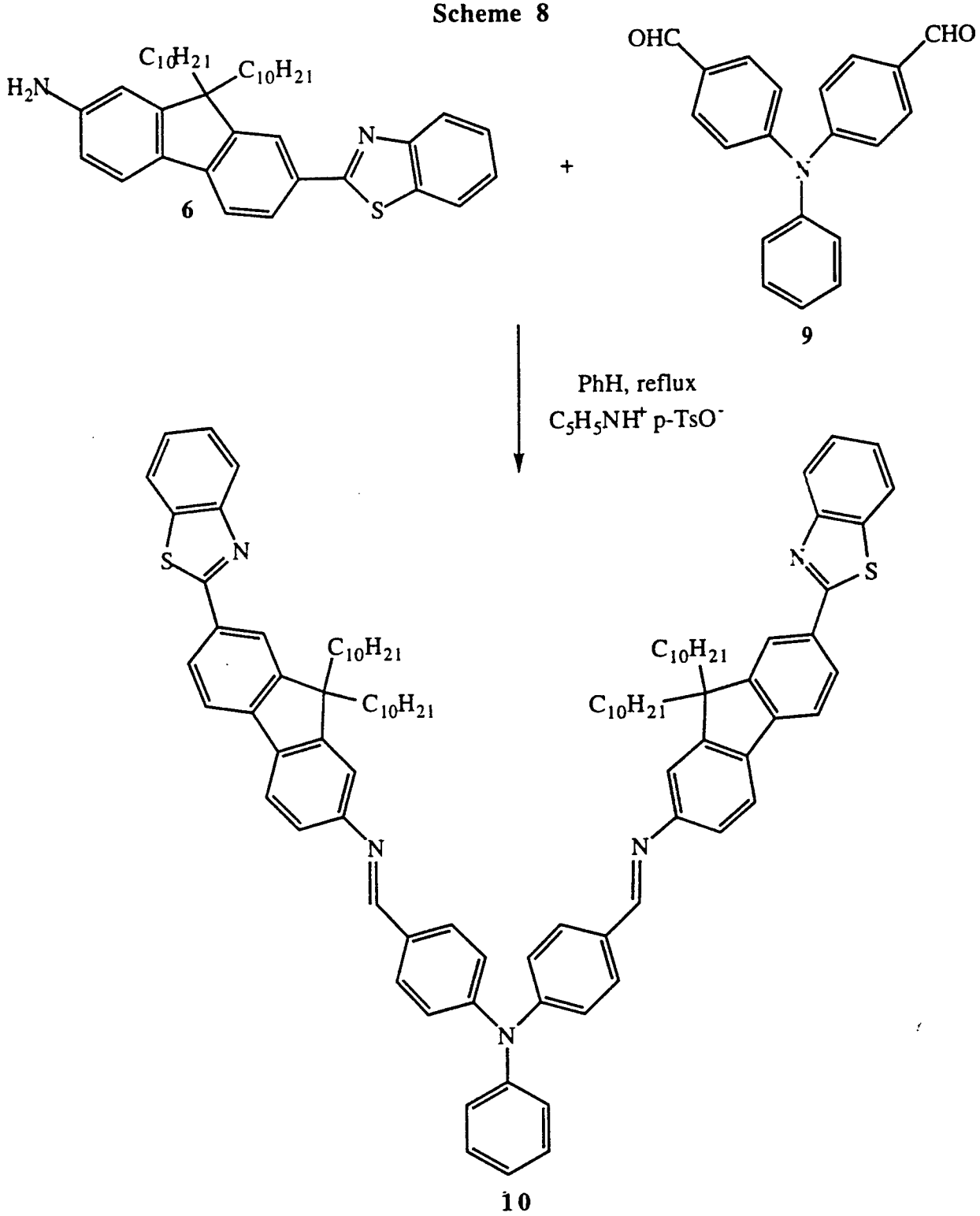
Derivatization of amine **6** was executed with three substrates, resulting in three new chromophores. First, p-TsOH-catalyzed condensation of **6** with 4-(N,N-diphenylamino)benzaldehyde (**7**, triphenylamine carboxaldehyde, prepared by Vilsmeier bisformylation of triphenylamine)[12] afforded the 4-(N,N-diphenylamino)benzaldehyde/7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine azomethine dye (**8**) as a bright yellow solid (Scheme 7).

Scheme 7



In similar fashion, **6** underwent pyridinium toluene sulfonate-catalyzed condensation with triphenylamine biscarboxaldehyde (**9**),[12] prepared by Vilsmeier bisformylation of triphenylamine, as depicted in Scheme 8. Triphenylamine biscarboxaldehyde/7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine azomethine dye (**10**) was produced as a bright orange solid.

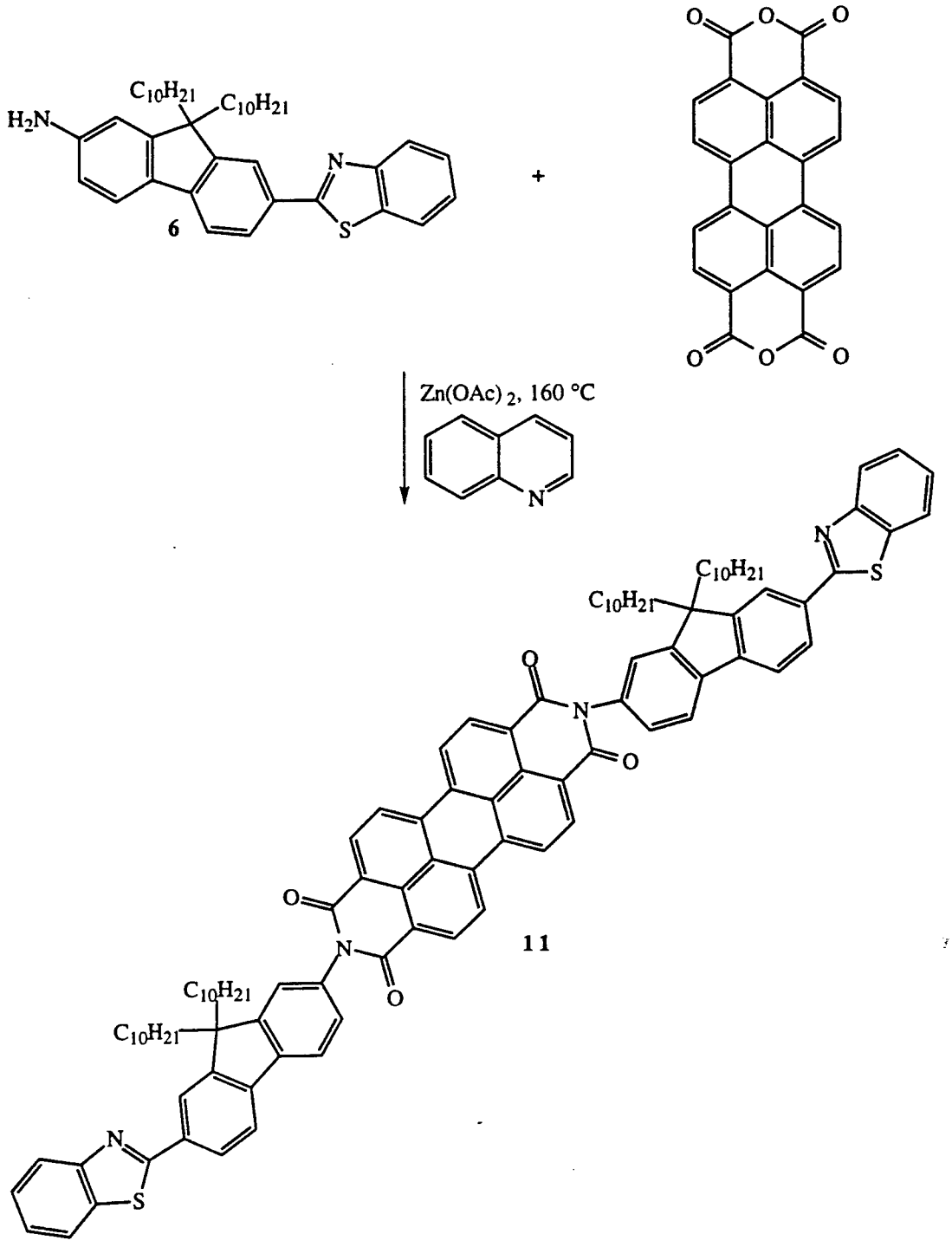
Scheme 8



In an attempt to synthesize a longer wavelength absorbing dye, **6** was subjected to $Zn(OAc)_2$-catalyzed condensation imidization with 3,4,9,10-perylenetetracarboxylic dianhydride (Scheme 9).[13] The 3,4,9,10-perylenetetracarboxylic dianhydride/7-benzothiazol-2-yl-9,9-didecylfluorene-

2-ylamine imide dye (**11**) was formed as a deep red-pink solid, after column chromatographic purification. Azomethine dye (**11**) exhibited two strong, broad absorption bands, one in the UV from 270 to 385 nm (λ_{max} = 345 nm) and another in the visible from 410 to 545 nm with λ maxima at 457, 486, and 522 nm.

Scheme 9

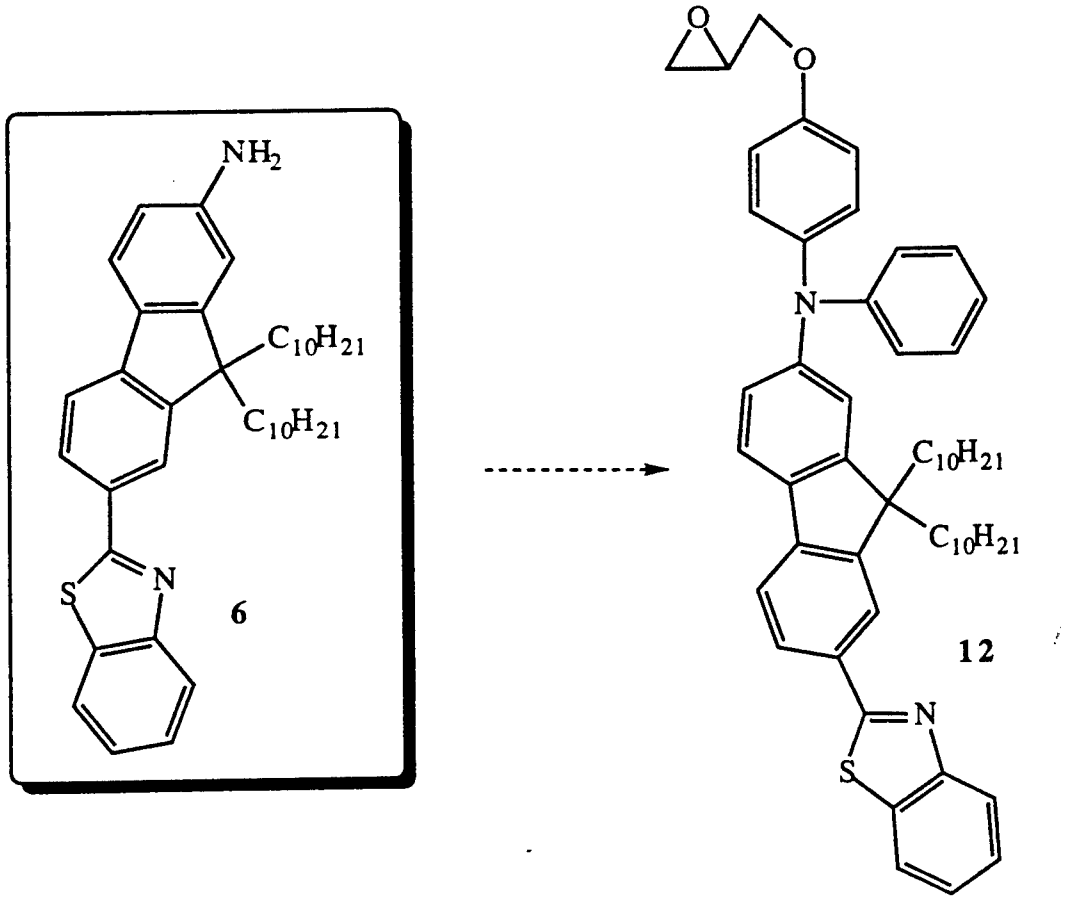


Conclusions

An efficient synthesis of two key intermediates (**3** and **6**) for the creation of new two photon absorbing dyes was accomplished, starting from readily available fluorene. All intermediates were fully characterized, including mp, UV-vis, FT-IR, MS, elemental analysis, ^{1}H and ^{13}C NMR, photoluminescence (fluorescence emission). Three novel dyes (**8**, **10**, and **11**), with varying UV-vis and fluorescence emission spectral characteristics, were prepared from the penultimate product **6** in good yield. A remarkable feature of these materials is their excellent solubility in common organic solvents, an important criterion for processing and potential device fabrication.

The efficient synthesis of aminofluorene **6** and iodofluorene **3** has opened the possibility of creating a diverse range of nonlinear optical, including multiphoton absorbing, materials, both low molar mass and polymeric. Work is underway to synthesize an epoxy derivative of **6** (Scheme 10) that will be copolymerized with an optical grade epoxide to produce a polymeric material for optical power limiting expected to have a high laser damage threshold. Importantly, monomer **12** will be a key component in two photon curing experiments to be performed in my laboratory at the University of Central Florida, in which it will serve as the two photon dye, providing upconverted fluorescence for absorption by a visible photoinitiator. Since it will then be covalently bound to the resulting polymer, it can also serve as a chromophore for subsequent two photon fluorescence imaging.

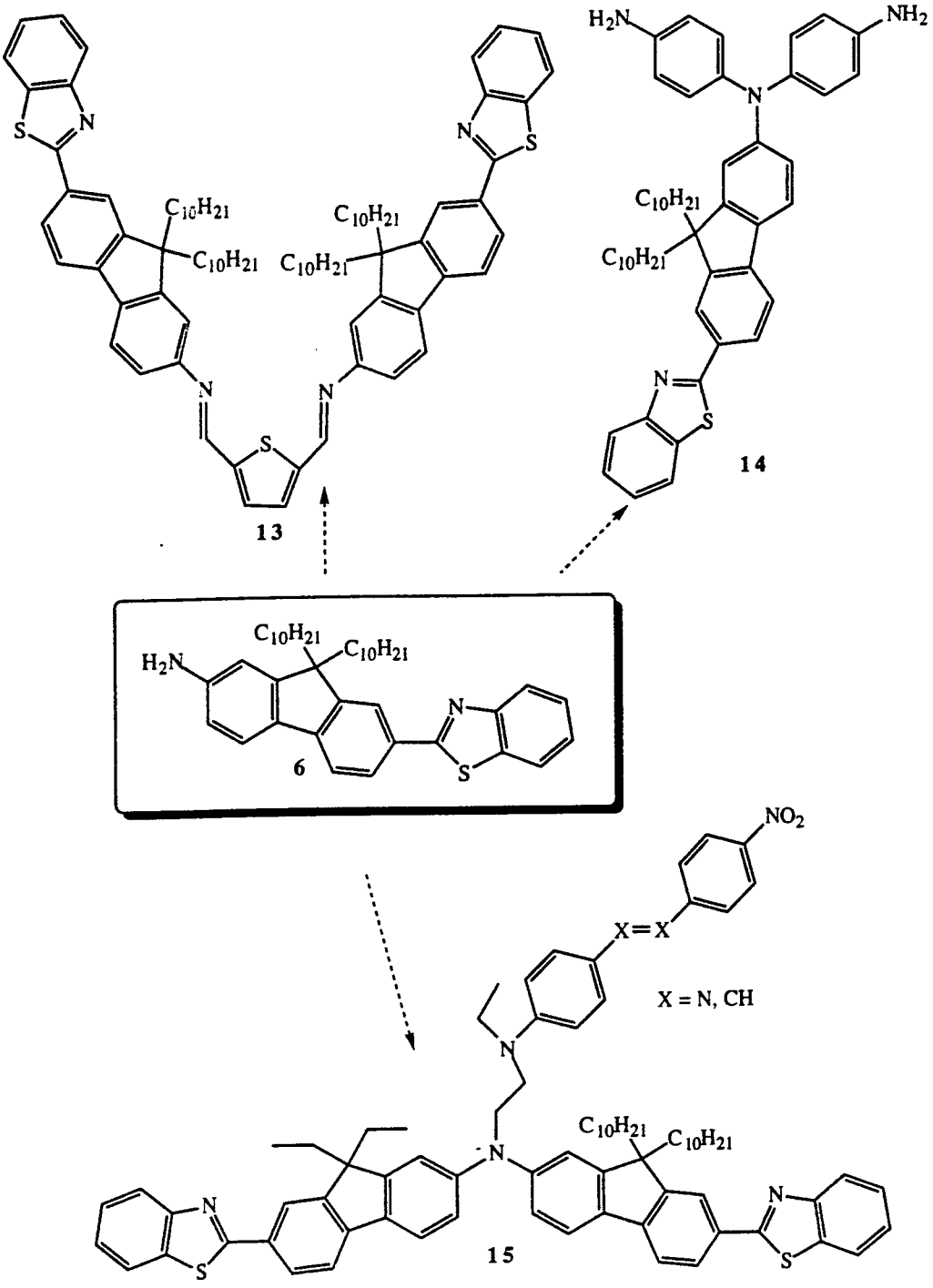
Scheme 10



Since the amine functionality of **6** can be readily functionalized, two photon absorbing dyes can be prepared whose absorption and emission properties are tailored through judicious choice of

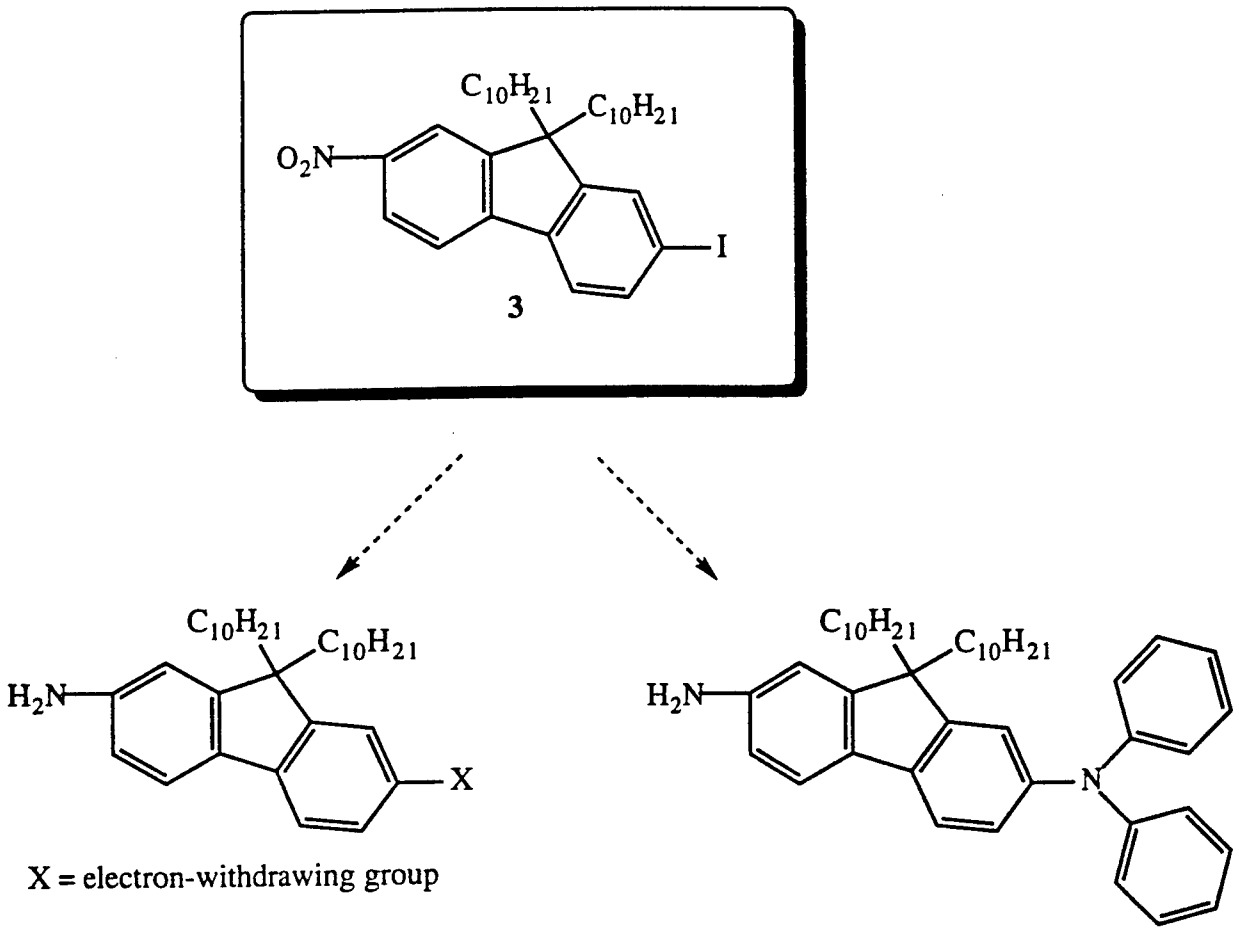
derivatization constructs (see e.g., Scheme 11). Future plans include preparation of a thiophene-based Schiff base with extended conjugation (**13**), a diamine (**14**) for the preparation of high thermal stability polyurethanes and polyimides for optical limiting, and a multifunctional compound (**15**) bearing two photon absorbing chromophores and a donor/acceptor NLO chromophore capable of being poled.

Scheme 11



Iodo derivative **3** will also serve as a key intermediate, as this can also be functionalized to provide access to wide range of multiphoton absorbing materials. A few of these possibilities are presented in Scheme 12, including conversion of the iodo functionality to either electron-withdrawing or electron-donating moieties, a compelling testament to the versatility of **3**.

Scheme 12



Experimental

General. Reactions were conducted under N_2 or Ar atmospheres. Benzothiazole was distilled under reduced pressure prior to use. THF was distilled over sodium before use. All other reagents and solvents were used as received from commercial suppliers. ^1H and ^{13}C NMR were recorded in $CDCl_3$ on Bruker AM-200 or Varian 300 NMR spectrometers, at 200 and 300 MHz, respectively, for ^1H (referenced to TMS at $\delta = 0.0$ ppm), and at 50 and 75 MHz, respectively, for ^{13}C (referenced to $CDCl_3$ at $\delta = 77.3$ ppm). UV-visible spectrophotometric measurements were recorded on a Hewlett-Packard Model 8453A spectrophotometer. Elemental analyses and MS analyses were performed at Wright-Patterson Air Force Base Materials Laboratory.

Synthesis of 2-nitrofluorene (1). Fluorene (60 g, 361 mmol) and 500 mL AcOH were heated, under N_2, to 85 °C in a 1 L 3-necked flask fitted with a mechanical stirrer, addition funnel, and thermometer. HNO_3 was added over 20 min via an addition funnel, maintaining the

temperature at 85 °C. The reaction mixture was allowed to cool to room temperature over 2 h, resulting in a yellow suspension. This was filtered, washed with 50 mL AcOH containing 1.3 g KOAc, then slurried in water and filtered. The yellow product was dried in a vacuum oven, affording 60.5 g (80% yield); mp = 160 - 163 °C (lit. 157 °C[5]). Anal. Calcd: 73.92% C, 4.29% H, 6.63% N; Found: 73.72% C, 3.72 % H, 6.52% N.

Synthesis of 7-iodo-2-nitrofluorene (2). 2-Nitrofluorene (6.4 g, 30 mmol) and 200 mL AcOH were stirred at room temperature in a 500 mL 3-necked flask fitted with N_2 inlet, condenser, stir bar, and stopper. I_2 (3.74 g, 14.7 mmol) was added and stirred for 20 min. An orange-brown liquid formed with some undissolved starting material remaining suspended. H_2SO_4 (20 mL) and 2.2 g $NaNO_2$ were then added, and the reaction mixture was heated to 115 °C for 1 h. Yellow precipitate formed. The suspension was cooled to room temperature and poured into 200 g ice, filtered, washed with water, and dried in a vacuum oven. The crude product (9.51 g) was recrystallized from AcOH (1200 mL), resulting in 8.0 g (79% yield) of yellow needles (mp = 245 - 246 °C, lit. 244-245 °C[6]). Anal. Calcd: 46.32% C, 2.39% H, 37.64 % I, 4.15% N; Found: 46.37% C, 2.18 % H, 37.96% I, 3.94% N. MS (EI, 70 eV) m/z: 337 (M+).

Synthesis of 7-iodo-9,9-didecyl-2-nitrofluorene (3). 7-Iodo-2-nitrofluorene (9.6 g, 28 mmol), 12.58 g 1-bromodecane (57 mmol), 0.49 g KI (3 mmol), and 65 mL DMSO were placed in a 250 mL 3-necked flask containing a stir bar at room temperature. Powdered KOH (6.65 g, 119 mmol) was slowly added via a solid addition funnel under N_2. The color of the reaction mixture changed from bright yellow to dark green immediately after addition of the KOH. TLC analysis (10:1 hexanes/EtOAc) indicated dialkylation was complete after 1 h. The reaction mixture was poured into H_2O and extracted with hexanes. The organic extract was washed with H_2O, dried over $MgSO_4$, and concentrated, affording 16.5 g of a viscous dark orange oil that crystallized on standing. Purification was accomplished by column chromatography (40:1 hexanes/EtOAc), providing 13.6 g of yellow solid (77% yield). MS (EI, 70 eV) m/z: 617 (M+); 477 (M - $C_{10}H_{20}$); 476 (M - $C_{10}H_{21}$); 350 (M-$C_{19}H_{40}$); 336 (M-$C_{40}H_{41}$); 85 (C_6H_{13}+); 71 (C_5H_{11}+); 57 (C_4H_9+); 43 (C_3H_7+). ^{1}H NMR (300 MHz, $CDCl_3$) δ: 8.25 (d, 1H, ArH), 8.18 (d, 1H, ArH), 7.79 (s, 1H, ArH), 7.72 (m, 2H, ArH), 7.53 (d, 1H, ArH), 1.99 (m, 4H, CH_2), 1.18 (bm, 14H, CH_2), 1.05 (bm, 14H, CH_2), 0.83 (t, 6H, CH_3), 0.56 (m, 4H, CH_2)

Synthesis of 2-(tri-n-butylstannyl)benzothiazole (4). Benzothiazole (20 g, 148 mmol) was dissolved in 150 mL dry THF and cooled to -75 °C in a 500 mL 3-necked flask, containing a stir bar and fitted with a low temperature thermometer, addition funnel, and N_2 inlet. n-BuLi (111 mL 1.6 M in hexanes) was slowly added, over 50 min, via the addition funnel, maintaining the temperature <-65 °C. During this time the reaction mixture turned from orange to brown. Tri-n-butyltin chloride (58 g, 178 mmol) was added slowly via the addition funnel, maintaining the temperature <-60 °C. The dark green reaction mixture turned yellow. After addition was complete, the reaction mixture was allowed to slowly warm to -20 °C, at which time it turned green again. It was then stirred at -10 °C for 1 h and turned yellow again. A solution of 100 mL of 3 wt% KF(aq) was added and stirred at 10 °C for 1 h. The layers were separated, in the presence of a ubiquitous polymeric gel. The organic phase was diluted with toluene, dried over $MgSO_4$, filtered, and concentrated. An orange-brown liquid resulted which was purified by vacuum distillation (bp 155-158 °C/0.2 mmHg, lit. 144-146 °C/0.15 mmHg). A yellow liquid was collected (56.3 g, 90% yield).

Synthesis of 2-(9,9-didecyl-7-nitrofluoren-2-yl)benzothiazole (5). 7-Iodo-9,9-didecyl-2-nitrofluorene (10.1 g, 16 mmol) and 8.5 g 2-(tri-n-butylstannyl)benzothiazole (20 mmol) were dissolved in 200 mL toluene and degassed under vacuum and Ar. $Pd(PPh_3)_4$ (0.47 g, 0.4 mmol) was added and degassed. The reaction mixture was heated to 110 °C under Ar. The reaction was monitored by TLC (10:1 hexanes/EtOAc) and found to be complete after about 4.5 h (reaction mixture turned black). The toluene was removed *in vacuo*, and the resulting dark orange oil was passed through a Si gel column, eluting with hexanes. An orange solid formed which was recrystallized from hexanes, affording 6.24 g of bright yellow crystals (61% yield, mp = 94-95 °C). MS (EI, 70 eV) m/z: 624 (M+); 497 (M - C_9H_{19}); 483 (M - $C_{10}H_{21}$); 357 (M - $C_{19}H_{39}$); 343 (M - $C_{20}H_{41}$); 311 (M - $C_{20}H_{43}$ - NO); 297 (M - $C_{20}H_{41}$ - NO_2); 71 (C_5H_{11}+); 57 (C_4H_9+); 43 (C_3H_7+). UV-vis (THF): λ_{max} = 365 nm (220 - 420 nm); ^{1}H NMR (200 MHz, $CDCl_3$) δ: 8.32, 8.37 (dd, 1H, ArH), 8.24, 8.18 (dd, 1H, ArH), 8.14 (m, 1H, ArH), 8.10 (m, 1H, ArH), 7.96, 7.92 (dm, 1H, ArH), 7.91, 7.88 (dd, 1H, ArH), 7.87, 7.84 (dd, 1H, ArH), 7.57, 7.38 (dd, 1H, ArH), 7.53, 7.42 (dm, 1H, ArH), 7.49, 7.46 (dd, 1H, ArH), 2.14 (m, 4H, CH_2), 1.14 (bm, 14H, CH_2), 1.04 (bm, 14H, CH_2), 0.81 (t, 6H, CH_3), 0.62 (m, 4H, CH_2); ^{13}C NMR (50 MHz, $CDCl_3$) δ: 168.0 (C2), 154.4 (C9), 153.4 (C7'), 152.9 (C13'), 147.9 (C12'), 146.7 (C10'), 141.6 (C11'), 135.3 (C2'), 134.6 (C8), 127.8 (C1'), 127.6 (C5'), 126.7 (C4'), 125.6 (C3'), 123.6 (C6), 123.5 (C5), 122.0 (C8'), 121.8 (C7), 120.6 (C4), 118.6 (C6'), 56.4 (C9'), 40.2 (Ca), 32.0 (Ch), 30.4 (Cc), 30.0 (Cd), 29.7 (Ce,f), 29.4 (Cg), 24.0 (Cb), 22.8 (Ci), 14.3 (Cj); FT-IR (KBr, cm^{-1}): 3066 (υarCH), 2924, 2852 (υalCH), 1589 (υC=N), 1519 (υNO_2 as), 1337 (υNO_2 sy).

Synthesis of 7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine (6). 2-(9,9-Didecyl-7-nitrofluoren-2-yl)benzothiazole (4.0 g, 6.4 mmol) was dissolved in a mixture of 15 mL EtOH and 15 mL THF at room temperature. To this 0.2 g 10% Pd/C was added. The mixture was heated to 70 °C under Ar. Hydrazine hydrate (2 mL, 38 mmol) was added dropwise via syringe over 20 min. The reaction mixture was stirred for 12 h at 70 °C, cooled to room temperature, and concentrated, affording a yellow-green oil. TLC (10:1 hexanes/EtOAc) indicated the reaction was quantitative. Purification was accomplished by passing a hexane solution of the amine through a Si gel plug, eluting with hexanes, providing 3.6 g of viscous, sticky yellow material (95% yield). MS (EI, 70 eV) m/z: 594 (M+); 453 (M - $C_{10}H_{21}$); 327 (M - $C_{19}H_{39}$); 313 (M - $C_{20}H_{41}$); 163 ($C_{13}H_7$+); 57 (C_4H_9+); 43 (C_3H_7+). UV-vis (THF): λ_{max} = 380 nm (210 - 450 nm); ^{1}H NMR (200 MHz, $CDCl_3$) δ: 8.09, 8.05 (dm, 1H, ArH), 8.04 (d, 1H, ArH), 7.99, 7.95 (dd, 1H, ArH), 7.90, 7.86 (dm, 1H, ArH), 7.63, 7.59, 7.56, 7.51 (dd, 1H, ArH), 7.51, 7.44, 7.39, 7.32 (dq, 1H, ArH), 7.48, 7.35 (dm, 1H, ArH), 7.19 (m, 1H, ArH), 6.67 (m, 2H, ArH), 3.82 (bs, 2H, NH_2), 1.98 (m, 4H, CH_2), 1.15 (bm, 14H, CH_2), 1.04 (bm, 14H, CH_2), 0.82 (t, 6H, CH_3), 0.66 (m, 4H, CH_2); ^{13}C NMR (50 MHz, $CDCl_3$) δ: 169.3 (C2'), 154.5 (C9'), 153.8 (C2), 150.8 (C10), 147.1 (C13), 145.1 (C12), 131.6 (C7), 130.8 (C8'), 129.2 (C8), 128.4 (C4), 127.3 (C5), 126.4 (C6), 125.0 (C11), 123.0 (C6'), 121.5 (C5'), 121.5 (C7'), 118.8 (C4'), 114.3 (C1), 109.8 (C3), 55.4 (C9), 40.8 (Ca), 32.1 (Ch), 30.4 (Cc), 30.3 (Cd), 29.8 (Ce), 29.7 (Cf), 29.5 (Cg), 24.0 (Cb), 22.8 (Ci), 14.3 (Cj); FT-IR (KBr, cm^{-1}): 3741, 3383 (υNH_2), 2926, 2855 (υalCH), 1603 (υC=N).

Synthesis of 4-(N,N-diphenylamino)benzaldehyde/7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine azomethine dye (8). Condensation of 200 mg 7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine (0.34 mmol) and 93 mg 4-(N,N-diphenylamino)benzaldehyde (7) (0.34 mmol) was conducted at reflux in 15 mL benzene with a Dean-Stark trap. After 64 h of reflux, TLC (10:1 hexanes/EtOAc) indicated presence of starting materials and product, hence, ca. 5 mg p-TsOH was added. The reaction mixture immediately turned from bright yellow to deep orange, and water droplets were observed in the condenser. Heating continued for 9 h, after which time the mixture was cooled to room temperature, passed through a Si gel plug (CH_2Cl_2), and concentrated. An orange-yellow oil was obtained (0.39 g).

Synthesis of triphenylamine biscarboxaldehyde/7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine azomethine dye (10). Condensation of 1.6 g 7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine (2.7 mmol) and 405 mg triphenylamine biscarboxaldehyde (9) (1.3 mmol) was conducted at reflux in 15 mL benzene with a Dean-Stark trap. After 64 h of relux, TLC (10:1 hexanes/EtOAc) indicated presence of starting materials and product, hence, ca. 10 mg pyridinium p-toluene sulfonate was added. The reaction mixture immediately turned from bright yellow to bright orange, and water droplets were observed in the condenser. Heating continued for 21 h, after which time the mixture was cooled to room temperature, passed through a Si gel plug (CH_2Cl_2), and concentrated. An orange oil was obtained (1.85 g).

Synthesis of 3,4,9,10-perylenetetracarboxylic dianhydride/7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine imide dye (11). Condensation of 1.0 g 7-benzothiazol-2-yl-9,9-didecylfluorene-2-ylamine (1.7 mmol) and 300 mg 3,4,9,10-perylenetetracarboxylic dianhydride (0.8 mmol), 145 mg $Zn(OAc)_2 \cdot 2H_2O$ (0.7 mmol), and 6 mL quinoline was carried out at 160 °C for 20.5 h. The reaction mixture was cooled to room temperature, diluted with EtOH, filtered, washed with MeOH, and air dried, affording 0.93 g of pink-red solid. UV-vis (THF): 270 to 385 nm (λ_{max} = 345 nm) and 410 to 545 nm with λ maxima = 457, 486, and 522 nm.

Acknowledgments

Support from the Air Force Office of Scientific Research, in the form of a Summer Research Faculty Fellowship, is gratefully acknowledged. Special thanks are due Bruce Reinhardt, the lab focal point, and Dr. Robert Evers, MLBP Branch Chief, for generous assistance, and facilitating a productive, and truly collaborative summer research experience. Monumentous credit is due to Katherine J. Schafer, a graduate student at the University of Central Florida, who also received an AFOSR Summer Fellowship and helped facilitate the progress documented in this report.

References

1. See, e.g., Bhawalkar, J. D.; He, G. S.; Prasad, P. N. "Nonlinear Multiphoton Processes in Organic and Polymeric Materials" *Rep. Prog. Phys.* **1996**, *59*, 1041-1070.
2. Denk, W.; Strickler, J. H.; Webb, W. W. "Two-Photon Laser Scanning Fluorescence Microscopy" *Science* **1990**, *248*, 73-76.
3. Bhawalkar, J. D.; Shih, A.; Pan, S. J.; Liou, W. S.; Swiatkiewicz, J.; Reinhardt, B. A.; Prasad, P. N.; Cheng, P. C. "Two-Photon Laser Scanning Fluorescence Microscopy - from a Fluorophore and Specimen Perspective" *Bioimaging* **1996**, *4*, 168-178.
4. Reinhardt, B. A.; Brott, L. L.; Clarson, S. J.; Dillard, A. G.; Bhatt, J. C.; Kannan, R.; Yuan, L.; He, G. S.; Prasad, P. N. "Highly Active Two-Photon Dyes: Design, Synthesis and Characterization Toward Application" *Chem. Mater.* **1998** (in press).
5. Kuhn, W. E. "2-Nitrofluorene and 2-Aminofluorene" *Organic Synthesis, Coll. Vol. 2*; Blatt, A. H., Ed.; Wiley: New York, 1943, pp. 447-448.

6. Marhevka, V. C.; Ebner, N. A.; Sehon, R. D.; Hanna, P. E. "Mechanism-Based Inactivation of N-Arylhydroxamic Acid N,O-Acyltransferase by 7-Substituted-N-hydroxy-2-acetamidofluorenes" *J. Med. Chem.* **1985**, *28*, 18-24.
7. Pivsa-Art, S.; Satoh, T.; Kawamura, Y.; Miura, M.; Nomura, M. "Palladium-Catalyzed Arylation of Azole Compounds with Aryl Halides in the Presence of Alkali Metal Carbonates and the Use of Copper Iodide in the Reaction" *Bull. Chem. Soc. Jpn.* **1998**, *71*, 467-473.
8. Kosugi, M.; Koshiba, M.; Atoh, A; Sano, H.; Migita, T. "Palladium-Catalyzed Coupling between Organic Halides and Organotin Compounds Involving C-N Unsaturated Bonds at the Reaction Center" *Bull. Chem. Soc. Jpn.* **1986**, *59*, 677-679.
9. Hark, R. R.; Hauze, D. B.; Petrovskaia, O.; Joullie, M. M.; Jaouhari, R.; McComiskey, P. "Novel Approaches toward Ninhydrin Analogs" *Tet. Lett.* **1994**, *35*, 7719-7722.
10. Tan, L.-S.; Srinivasan, K. R.; Bai, S. J. "New Aromatic Benzazole Polymers. I. Benzobisthiazole and Benzobisoxazole Polymers with Main-Chain Triarylamino Units" *J. Polym. Sci.: Part A: Polym. Chem.* **1997**, *35*, 1909-1924.
11. Yang, C.-P.; Lin, J.-H. "Syntheses and Properties of Aromatic Polyamides and Polyimides Based on N-Phenyl-3,3-bis[4-(p-aminophenoxy)phenyl]phthalimidine" *J. Polym. Sci.: Part A: Polym. Chem.* **1994**, *32*, 369-382.
12. Belfield, K. D.; Najjar, O. M. unpublished results.
13. Bohm, A.; Arms, H.; Henning, G.; Blaschka, P. "1,7-Disubstituted Perylene-3,4,9,10-tetracarboxylic Acids and Dianhydrides and Diimides of Said Acids" WO 97/22608, World Patent Application.

Dr. Daniel D. Bombick's report was not available at the time of publication.

RECOGNIZING LINEARITIES IN MATERIALS DATABASES

Frank M. Brown

Associate Professor

Department of Electrical Engineering and Computer Science

University of Kansas

415 Snow

Lawrence, Kansas 66045

Final Report for:

Summer Faculty Research program

Materials Process Design Branch

Materials and Manucfacturing Directorate

Air Force Research Laboratory

AFRL/MLMR

building 653, suite 13 (room 14)

2977 P Street

Wright Patterson Air Force Base

Dayton, Ohio 45433-7746

Sponsored by:

Air Force Office of Scientific Research

Bolling Air Force Base, DC

and

AFRL/MLMR

August 1998

RECOGNIZING LINEARITIES IN MATERIALS DATABASES

Frank M. Brown
Associate professor
Department of Electrical Engineering and Computer Science
University of Kansas

Abstract

The Multi-dimensional Hough Transform is proposed as a technique for searching for linearities in databases. This Transform is described and then explicitly programmed in Common Lisp. The system is illustrated by showing that it finds linearities in data representing the edge of a circle and the surface of a sphere. The system was then used to search for linearities in 21 materials datasets each involving about fifty chemical compounds and five properties. The results of one such dataset are described herein. Overall, this research suggests that the Multi-dimensional Hough Transform has potential for both organizing and simplifying data for input to various automatic classification and theory formation systems and for assessing the quality of the results of such systems.

RECOGNIZING LINEARITIES IN MATERIALS DATABASES

Frank M. Brown

1. Introduction

In recent years the Materials Laboratory of the Air Force has been investigating methods for automatically generating theories for predicting the existence and properties of unknown compounds. These methods include the Pyramidal net methods of inductive reasoning originally developed by Viktor Gladun from the then Soviet Union, Rough Set methods, and other methods.

All these methods are based on binning the property values of the objects into a small number of ranges which are then treated as a symbolic property value of that property. It is necessary to do this because these automatic classification and theory formation methods are based on symbolically matching the property values of different objects to each other in order to construct general rules relating different property values to each other. If the bins are well chosen then these methods often produce interesting and useful results, but if the bins are not well chosen then these methods are typically unable to find interesting correlations among the data. For example, [Brown,, Jacobs,Snider,Jackson&Leclair 1998] lists many rules of thumb for choosing the number of bins for Gladun's algorithm.

Given these problems, the idea arose that the binning of data should be determined from the linearities inherent in a dataset. Specifically, given a dataset represented as a spreadsheet of property values where each row represents a different object and each column represents a different property, the property values themselves may be thought of as being an m x n matrix. A linearity would simply be a hyper-plane drawn through this n dimensional hyper-space of properties containing as many datapoints as is possible. For example in the case of 2 properties (i. e. dimensions) a linearity would be a line and in the case of three properties a linearity would be a plane. The best linearities could then provide information on how to bin the values of the different properties so that the bins of different properties were chosen in a manner correlated to those linearities. For example, one method of correlated binning would be for each property to form a bin using the property values of the endpoint objects of each such linearity. In this manner the bins for different properties would be correlated to each other thereby potentially simplifying the rule systems produced by the various automatic classification and theory formation systems used at the Materials Lab.

The problem with this idea was the difficulty in finding the linearities. The obvious algorithm of creating a hyperplane from every combination of n objects being essentially m^n is exponential in n[1]. It is not so bad if there are only two properties (i.e. dimensions), for in this case it would be a quadratic algorithm: m^2 since every pair of points gives a line. Likewise, in the case of three properties (i.e. dimensions) it would be a cubic algorithm: m^3

[1]Actually mCn since the n points forming a linearity are all distinct.

since every three points defines a plane. In general however we want to find hyper-planes in n dimensional hyper-space where n is an arbitrarily large number of properties.

2. The Multi-dimensional Hough Transform System

Fortunately, Dr. Jackson (of the Materials Lab) was able to suggest an approach to solving this problem. His idea was to generalize the Hough Transform which is used in Computer Vision [Umbaugh 1988] for finding lines in two dimensional pictures. The basic idea of the Hough Transform is to choose a small number of slopes and then to compute the corresponding y intercepts for each datapoint. Such an algorithm is linear with respect to the number of datapoints. If s is the number of slopes used then the complexity is essentially: m*s. By saving the points into an array indexed by a given slope and a given y intercept range (we let the number of y intercept ranges equal the number of slopes) one can later iterate through the array searching for the cells containing the largest number of points. Since each cell contains points having the same slope and y intercept, a cell which has more than two points forms a line, and a good line has many points in its cell. This lookup process is essentially s^2, but s is intended to be a small constant much smaller than m. Thus the two steps of the Hough Transform essentially provide a linear search procedure.

In this project, we generalized the Hough Transform to n dimensions thereby dealing with n properties in the spreadsheet. The algorithm for this generalization was to choose essentially n-1 values each from a range of values of cardinallity s (representing an angle from the m-1 positive axes of the first n-1 dimensions), and then to compute the angle of the nth dimension for each of the m points. These angles then formed the index for the cells in an s^n sized array into which the points were stored. The complexity of creating the array is essentially: $m*s^{n-1}$, and the complexity of pulling the linearities out of the array is essentially s^n. Since s is intended to be a small constant this is essentially a linear algorithm in m though still exponential in n. Since s is much smaller than m this is an asymptotically significant reduction in complexity over the original naive m^n algorithm. For example, with 1000 data elements and 10 properties the naive algorithm has a complexity of 1000^{10} (i.e. 1 followed by 30 zeros) an astronomical complexity beyond current computer capabilities, whereas the Multi-dimensional Hough algorithm with 10 slots per property (i.e. s=10) has the complexity of about $1000*10^{10}$ (1 followed by 13 zeros) which though large is perhaps not beyond the realm of computer capabilities.

This generalization of the Hough Transform to n-dimensions was prototyped in the Logistica Programming Language, an advanced Artificial Intelligence Programming Language whose development was supported by the Air Force Materials Laboratory. In order to eliminate trivial linearities, a linearity defined by a cell is displayed only if the number of points in the cell is greater than the dimension of the hyperspace. Thus if the dimension is 2 then 3 points are needed for a displayed linearity and if the dimension is 3 then at least 4 points are needed for a displayed linearity. In these examples the cardinallity of s was set to 8 and the boundaries were set to be the slopes (i.e. tangents of the angle) obtained from 8 equally spaced angles from 0 to π.

3. Experiments with the Multi-dimensional Hough Transform System

The Multi-dimensional Hough Transform System was initially tested on two groups of generated data, namely the edges of circles and the surfaces of spheres. Circles and Spheres were chosen because it was thought that their apparent non-linearity and their symmetry would provide particularly difficult cases with which to deal. All numbers produced as output were rounded to the nearest small fraction for presentation purposes.

The first group of datasets, the circle datasets, were generated by producing 2^k datapoints on the edge of a unit circle as k ranged from 0 to 4. The datapoints generated and the linearities found are listed in Table 1 below. No non-trivial linearities occurred in the first three datasets with 1, 2, or 4 points evenly spaced around the circle. The linearities found in the fourth dataset with 8 points involved two tangents of the circle and two diagonal diameters through the circle. Other tangents and diagonal linearities were not found presumably due to the asymmetry in handling the last dimension (which is computed) and the inherent approximation nature of choosing the boundaries for s. There were a number of linearities found in the fourth dataset with 16 datapoints. These involved both tangents and diameters through the circle. The best two linearities found were:

((-3/8 -11/12) (-5/7 -5/7) (-11/12 -3/8) (0 1) (3/8 11/12)(5/7 5/7) (11/12 3/8))

((11/12 -3/8) (5/7 -5/7) (3/8 -11/12) (0 -1) (-11/12 3/8)(-5/7 5/7) (-3/8 11/12))

These linearities each involved 7 points, or almost half the dataset. Each was a diagonal diameter across the circle.

k	2^k	data points	linearities
0	1	#(1 0)	none
1	2	#(1 0) #(-1 0)	none
2	4	#(1 0) #(0 1) #(-1 0) #(0 -1)	none
3	8	#(1 0) #(5/7 5/7) #(0 1) #(-5/7 5/7) #(-1 0) #(-5/7 -5/7) #(0 -1) #(5/7 -5/7)	((-5/7 -5/7) (0 1) (5/7 5/7)) ((5/7 -5/7) (5/7 5/7) (1 0)) ((-5/7 -5/7) (-1 0) (-5/7 5/7)) ((5/7 -5/7) (0 -1) (-5/7 5/7))
4	16	#(1 0) #(11/12 3/8) #(5/7 5/7) #(3/8 11/12) #(0 1) #(-3/8 11/12) #(-5/7 5/7) #(-11/12 3/8) #(-1 0) #(-11/12 -3/8) #(-5/7 -5/7) #(-3/8 -11/12) #(0 -1) #(3/8 -11/12) #(5/7 -5/7) #(11/12 -3/8)	((5/7 -5/7) (3/8 -11/12) (0 -1) (-3/8 11/12)) ((3/8 -11/12) (0 -1) (0 1)) ((3/8 -11/12) (0 -1) (0 1) (3/8 11/12)) ((0 -1) (0 1) (3/8 11/12)) ((-3/8 -11/12) (0 1) (3/8 11/12) (5/7 5/7)) ((-3/8 -11/12) (-5/7 -5/7) (-11/12 -3/8) (0 1) (3/8 11/12) (5/7 5/7) (11/12 3/8)) ((11/12 -3/8) (-5/7 5/7) (1 0)) ((11/12 -3/8) (5/7 -5/7) (1 0)) ((11/12 -3/8) (5/7 -5/7) (3/8 11/12) (1 0)) ((11/12 -3/8) (5/7 -5/7) (5/7 5/7) (11/12 3/8) (1 0)) ((3/8 -11/12) (5/7 5/7) (11/12 3/8) (1 0)) ((5/7 5/7) (11/12 3/8) (1 0)) ((-5/7 -5/7) (11/12 3/8) (1 0)) ((0 1) (3/8 11/12) (5/7 5/7)) ((5/7 -5/7) (3/8 -11/12) (0 -1)) ((-11/12 -3/8) (-1 0) (5/7 5/7)) ((-5/7 -5/7) (-11/12 -3/8) (-1 0)) ((-5/7 -5/7) (-11/12 -3/8) (-1 0) (-3/8 11/12)) ((-5/7 -5/7) (-11/12 -3/8) (-1 0) (-11/12 3/8) (-5/7 5/7)) ((-3/8 -11/12) (-1 0) (-11/12 3/8) (-5/7 5/7)) ((-1 0) (-11/12 3/8) (-5/7 5/7)) ((5/7 -5/7) (-1 0) (-11/12 3/8)) ((-3/8 -11/12) (-5/7 -5/7) (0 1) (3/8 11/12)) ((3/8 -11/12) (0 -1) (-5/7 5/7) (-3/8 11/12)) ((11/12 -3/8) (5/7 -5/7) (3/8 -11/12) (0 -1) (-11/12 3/8) (-5/7 5/7) (-3/8 11/12))

Table 1: Linearities found in the Circle Datasets

The second group of datasets, the sphere datasets, were generated by producing datapoints on the surface of a unit sphere by iterating through spherical coordinates by an increment equal to π/k as as k ranged from 1 to 3 The datapoints generated and the linearities found are listed in Table 2. No non-trivial linearities occurred in the first dataset with 2 points at the north and south poles. In the second dataset with 6 evenly spaced points on the Cartesian coordinate axises of the sphere one linearity, a plane through the center of the sphere parallel to the equator, was found. There were many linearities found in the third sphere dataset which had 14 points. The best such linearity was the last one found namely:

((0 0 -1) (5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (-9/14 4/7 -1/2)

(6/7 0 -1/2) (1/9 -6/7 1/2) (-10/13 -2/5 1/2) (-9/14 4/7 1/2))

This linearity of eight points is a plane through the sphere running diagonally from the positive x and negative y quadrant to the negative x and positive y quadrant. In addition the plane is tilted on the z axis so as to include the south pole but not the north pole.

k	#	data points	linearities
1	2	#(0 0 1) #(0 0 -1)	none
2	6	#(0 0 1) #(1 0 0) #(0 1 0) #(-1 0 0)#(0 -1 0) #(0 0 -1)	((0 0 -1) (0 -1 0) (0 1 0) (0 0 1))
3	14	#(0 0 1) #(6/7 0 1/2) #(3/10 9/11 1/2) #(-9/14 4/7 1/2) #(-10/13 -2/5 1/2) #(1/9 -6/7 1/2) #(5/6 -1/5 1/2) #(6/7 0 -1/2) #(3/10 9/11 -1/2) #(-9/14 4/7 -1/2) #(-10/13 -2/5 -1/2) #(1/9 -6/7 -1/2) #(5/6 -1/5 -1/2) #(0 0 -1)	((0 0 -1) (5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (6/7 0 -1/2) (1/9 -6/7 1/2) (-9/14 4/7 1/2)) ((1/9 -6/7 -1/2) (-9/14 4/7 -1/2) (5/6 -1/5 1/2) (1/9 -6/7 1/2)(6/7 0 1/2) (0 0 1)) ((0 0 -1) (5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (6/7 0 -1/2)(1/9 -6/7 1/2)) ((1/9 -6/7 -1/2) (5/6 -1/5 1/2) (1/9 -6/7 1/2) (3/10 9/11 1/2)(6/7 0 1/2) (0 0 1)) ((0 0 -1) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (1/9 -6/7 1/2) (0 0 1)) ((0 0 -1) (1/9 -6/7 -1/2) (1/9 -6/7 1/2) (0 0 1)) ((0 0 -1) (1/9 -6/7 -1/2) (1/9 -6/7 1/2) (3/10 9/11 1/2) (0 0 1)) ((0 0 -1) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (1/9 -6/7 1/2)) ((1/9 -6/7 -1/2) (1/9 -6/7 1/2) (3/10 9/11 1/2) (0 0 1)) ((0 0 -1) (5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (6/7 0 -1/2)) ((0 0 -1) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (1/9 -6/7 1/2)(3/10 9/11 1/2) (0 0 1)) ((5/6 -1/5 1/2) (1/9 -6/7 1/2) (3/10 9/11 1/2) (6/7 0 1/2) (0 0 1)) ((0 0 -1) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (6/7 0 -1/2)) ((0 0 -1) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (3/10 9/11 1/2)) ((0 0 -1) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (3/10 9/11 1/2) (0 0 1)) ((0 0 -1) (3/10 9/11 -1/2) (1/9 -6/7 1/2) (3/10 9/11 1/2) (0 0 1)) ((3/10 9/11 -1/2) (1/9 -6/7 1/2) (3/10 9/11 1/2) (0 0 1)) ((1/9 -6/7 1/2) (3/10 9/11 1/2) (6/7 0 1/2) (0 0 1)) ((0 0 -1) (3/10 9/11 -1/2) (6/7 0 -1/2) (3/10 9/11 1/2)) ((0 0 -1) (3/10 9/11 -1/2) (3/10 9/11 1/2) (0 0 1)) ((3/10 9/11 -1/2) (3/10 9/11 1/2) (6/7 0 1/2) (0 0 1)) ((-10/13 -2/5 -1/2) (-9/14 4/7 -1/2) (5/6 -1/5 1/2) (1/9 -6/7 1/2)(3/10 9/11 1/2) (6/7 0 1/2) (0 0 1)) ((0 0 -1) (-10/13 -2/5 -1/2) (3/10 9/11 -1/2) (6/7 0 -1/2)(-10/13 -2/5 1/2) (3/10 9/11 1/2)) ((0 0 -1) (-10/13 -2/5 -1/2) (3/10 9/11 -1/2) (-10/13 -2/5 1/2)(3/10 9/11 1/2)) ((0 0 -1) (-10/13 -2/5 -1/2) (3/10 9/11 -1/2) (-10/13 -2/5 1/2) (3/10 9/11 1/2) (0 0 1)) ((-10/13 -2/5 -1/2) (3/10 9/11 -1/2) (-10/13 -2/5 1/2) (3/10 9/11 1/2) (0 0 1)) ((-10/13 -2/5 -1/2) (3/10 9/11 -1/2) (-10/13 -2/5 1/2) (3/10 9/11 1/2) (6/7 0 1/2) (0 0 1)) ((-10/13 -2/5 -1/2) (3/10 9/11 -1/2) (5/6 -1/5 1/2) (3/10 9/11 1/2)(6/7 0 1/2) (0 0 1)) ((5/6 -1/5 -1/2) (6/7 0 -1/2) (5/6 -1/5 1/2) (-9/14 4/7 1/2)) ((5/6 -1/5 -1/2) (-9/14 4/7 -1/2) (5/6 -1/5 1/2) (6/7 0 1/2)) ((5/6 -1/5 -1/2) (3/10 9/11 -1/2) (6/7 0 -1/2) (5/6 -1/5 1/2)) ((5/6 -1/5 -1/2) (5/6 -1/5 1/2) (3/10 9/11 1/2) (6/7 0 1/2)) ((5/6 -1/5 -1/2) (6/7 0 -1/2) (5/6 -1/5 1/2) (3/10 9/11 1/2)) ((5/6 -1/5 -1/2) (3/10 9/11 -1/2) (6/7 0 -1/2) (5/6 -1/5 1/2)(3/10 9/11 1/2) (6/7 0 1/2)) ((5/6 -1/5 -1/2) (3/10 9/11 -1/2) (5/6 -1/5 1/2) (6/7 0 1/2)) ((5/6 -1/5 -1/2) (6/7 0 -1/2) (5/6 -1/5 1/2) (6/7 0 1/2)) ((1/9 -6/7 -1/2) (6/7 0 -1/2) (1/9 -6/7 1/2) (6/7 0 1/2)) ((-10/13 -2/5 -1/2) (-9/14 4/7 -1/2) (-10/13 -2/5 1/2) (-9/14 4/7 1/2)) ((-10/13 -2/5 -1/2) (3/10 9/11 -1/2)(-10/13 -2/5 1/2) (3/10 9/11 1/2)) ((-10/13 -2/5 -1/2) (-9/14 4/7 -1/2) (3/10 9/11 1/2) (0 0 1)) ((0 0 -1) (3/10 9/11 -1/2) (-10/13 -2/5 1/2) (-9/14 4/7 1/2)) ((-10/13 -2/5 -1/2) (-9/14 4/7 -1/2) (1/9 -6/7 1/2) (0 0 1)) ((0 0 -1) (1/9 -6/7 -1/2) (-10/13 -2/5 1/2) (-9/14 4/7 1/2)) ((1/9 -6/7 -1/2) (-9/14 4/7 -1/2) (1/9 -6/7 1/2) (-9/14 4/7 1/2)) ((0 0 -1) (5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (3/10 9/11 -1/2) (6/7 0 -1/2) (-10/13 -2/5 1/2) (-9/14 4/7 1/2)) ((5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (-9/14 4/7 -1/2) (5/6 -1/5 1/2) (1/9 -6/7 1/2) (-9/14 4/7 1/2) (0 0 1)) ((5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (-9/14 4/7 -1/2) (5/6 -1/5 1/2)(1/9 -6/7 1/2) (-9/14 4/7 1/2)) ((0 0 -1) (5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (-9/14 4/7 -1/2) (5/6 -1/5 1/2) (1/9 -6/7 1/2) (-9/14 4/7 1/2)) ((0 0 -1) (5/6 -1/5 -1/2) (1/9 -6/7 -1/2) (-9/14 4/7 -1/2) (6/7 0 -1/2) (1/9 -6/7 1/2) (-10/13 -2/5 1/2) (-9/14 4/7 1/2))

Table 2: Linearities found in the Sphere Datasets

The Circle and Sphere experiments showed the following:

(1) The Multi-dimensional Hough Transform could be used to extract some interesting linearities from even the most non linear and symmetric data. For example, of particular interest are the 7 element diagonal diameter linearities in the circle example with 16 points and the 8 element linearity in the 14 element sphere example.

(2) The Multi-dimensional Hough Transform would not produce all linearities due to: (a) the fact that the s number of slopes form arbitrary boundaries for defining a line and (b) the fact that the last dimension is treated differently from the other dimensions since it is computed while the other ones were produced by incrementing through some angle. For example, in the circle example with eight points the points on the axes occur less often than do the other points. Likewise, since the sphere dataset with 6 points is perfectly symmetrical, having found a plane through the equator we would have desired to have found two more linearities, namely two planes running through the north and south poles, one through the x axis and one through the y axis.

Accepting the approximation and asymmetry limitations of point (2), we were encouraged by point (1) to proceed by further developing the system to apply to real materials data.

To this end we transliterated the system into Common LISP and adapted it to run with the spreadsheet formats used as input to the theory formation system based on Gladun's Algorithm that is currently being used at the Materials Laboratory. The computer code written in Common LISP for the final Multi-dimensional Hough Transformation System has been delivered to the Materials Lab.

4. Construction of the Materials Databases for testing the Hough Transform System

At this point the Materials lab had received Materials data suggesting that certain sets of compounds of three elements would be logically grouped into a number of distinct classes. These groupings were based on a computerized analysis of five properties of the atomic elements which we shall call: VE, EN, ZUNGER, MTEMP, and Z. For each atomic element the property values of each of these properties is given in Table 3. An example grouping (dataset #28) is listed in Table 4.

Dr. Leclair asked us to try to determine the linearities inherent in 21 of these groupings.[2]. Since 5 properties each of 3 elements per compound equaled 15 dimensions (i.e. n=15) and the number of objects was on the order of 50 per group (e.g. 57 in dataset #28), the complexity of applying the Multi-dimensional Hough Algorithm on such a group would be roughly: $50*s^{15}$. Since we intended to keep s equal to 8 as in our previous examples this gives: $50*8^{15}$. Given limited computer resources we reduced this complexity to essentially $50*8^5$ by averaging the property values of a given property for each of the three atomic elements in each compound. Thus, for example, the VE property of the compound Dy_Ge_Ir would be the average of the VE property of each of its three elements.

[2] Experiments were carried out for the following twenty-one Datasets:
28, 30, 41, 47, 53, 56, 57, 58, 59, 62, 63, 68, 83, 86, 99, 102, 105, 136, 137, 138, 149

Element	VE	EN	ZUNGE	MTEP	Z	Element	VE	EN	ZUNGE	MTEM	Z
Ac	3	1.1	3.12	1323	89	N	5	2.85	0.54	63	7
Ag	11	1.07	2.375	1235	47	Na	1	0.89	2.65	371	11
Al	3	1.64	1.675	934	13	Nb	5	2.03	2.76	2741	41
Am	3	1.3	4.89	1267	95	Nd	3	1.2	3.99	1283	60
As	5	2.27	1.415	1090	33	Ni	10	1.76	2.18	1726	28
At	7	2.64	1.83	575	85	Np	3	1.3	4.93	913	93
Au	11	1.19	2.66	1338	79	O	6	3.32	0.465	55	8
B	3	1.9	0.795	2573	5	Os	8	1.85	2.65	3318	76
Ba	2	1.08	3.402	998	56	P	5	2.32	1.24	317	15
Be	2	1.45	1.08	1560	4	Pa	3	1.5	4.96	1873	91
Bi	5	2.14	1.997	544	83	Pb	4	1.92	2.09	601	82
Br	7	2.83	1.2	266	35	Pd	10	2.08	2.45	1827	46
C	4	2.37	0.64	3925	6	Pm	3	1.15	3.99	1353	61
Ca	2	1.17	3	1112	20	Po	6	2.4	1.9	527	84
Cd	12	1.4	2.215	594	48	Pr	3	1.1	4.48	1204	59
Ce	3	1.1	4.5	1071	58	Pt	10	1.91	2.7	2045	78
Cl	7	2.98	1.01	172	17	Pu	3	1.3	4.91	914	94
Cm	3	1.2	3.6	1818	96	Ra	2	0.9	3.53	973	88
Co	9	1.72	2.02	1768	27	Rb	1	0.8	4.1	312	37
Cr	6	2	2.44	2130	24	Re	7	2.06	2.68	3453	75
Cs	1	0.77	4.31	302	55	Rh	9	1.99	2.52	2239	45
Cu	11	1.08	2.04	1357	29	Ru	8	1.97	2.605	2583	44
Dy	3	1.15	3.67	1682	66	S	6	2.65	1.1	386	16
Er	3	1.2	3.63	1795	68	Sb	5	2.14	1.765	904	51
Eu	3	1.15	3.94	1095	63	Sc	3	1.5	2.75	1812	21
F	7	3.98	0.72	54	9	Se	6	2.54	1.285	490	34
Fe	8	1.67	2.11	1808	26	Si	4	1.98	1.42	1683	14
Fr	1	0.7	4.37	300	87	Sm	3	1.2	4.14	1345	62
Ga	3	1.7	1.695	303	31	Sn	4	1.88	1.88	505	50
Gd	3	1.1	3.91	1584	64	Sr	2	1.13	3.21	1042	38
Ge	4	1.99	1.56	1211	32	Ta	5	1.94	2.79	3269	73
H	1	2.1	1.25	14	1	Tb	3	1.2	3.89	1633	65
Hf	4	1.73	2.91	2500	72	Tc	7	2.18	2.65	2445	43
Hg	12	1.49	2.41	234	80	Te	6	2.38	1.67	723	52
Ho	3	1.2	3.65	1743	67	Th	3	1.3	4.98	2023	90
I	7	2.76	1.585	387	53	Ti	4	1.86	2.58	1933	22
In	3	1.63	2.05	430	49	Tl	3	1.69	2.235	577	81
Ir	9	1.87	2.628	2683	77	Tm	3	1.2	3.6	1818	69
K	1	0.8	3.69	336	19	U	3	1.7	4.72	1405	92
La	3	1.35	3.08	1193	57	V	5	2.22	2.43	2163	23
Li	1	0.9	1.61	453.7	3	W	6	1.79	2.735	3683	74
Lu	3	1.2	3.37	1929	71	Y	3	1.41	2.94	1796	39
Mg	2	1.31	2.03	922	12	Yb	3	1.1	3.59	1097	70
Mn	7	2.04	2.22	1517	25	Zn	12	1.44	1.88	693	30
Mo	6	1.94	2.72	2890	42	Zr	4	1.7	2.825	2125	40

Table 3: Properties of the Elements

Dy_Ge_Ir		Gd_Os_Sn		Ge_Os_Tm
Dy_Ge_Os		Ge_Ho_Ir		Ge_Os_U
Ag_U_W		Ge_Ho_Os		Ge_Os_Yb
Dy_Os_Sn		Ge_Ho_Re		Ho_Os_Sn
Er_Ge_Ir		Ge_Ir_Lu		La_Os_Sb
Er_Ge_Os		Ge_Ir_Nd		Lu_Os_Sn
Er_Os_Sn		Ge_Ir_Pr		Ce_Fe_Re
Eu_Ge_Ir		Ge_Ir_Sm		Ce_Fe_Ta
Eu_Ge_Os		Ge_Ir_Tb		Ce_Ge_Ir
Eu_Os_Sb		Ge_Ir_Th		Ce_Ge_Os
As_Ce_Os		Ge_Ir_Tm		Ce_Os_Sb
As_La_Os		Ge_Ir_Yb		Nd_Os_Sb
As_Nd_Os		Ge_La_Os		Os_Pr_Sb
As_Os_Pr		Ge_Lu_Os		Os_Sb_Sm
As_Os_Th		Ge_Lu_Re		Os_Sn_Tb
Ga_Os_U		Ge_Nd_Os		Os_Sn_Th
Gd_Ge_Ir		Ge_Os_Pr		Os_Sn_Tm
Gd_Ge_Os		Ge_Os_Sm		Os_Sn_Y
Gd_Ge_Re		Ge_Os_Tb		Ru_Sn_Th

Table 4: Dataset Grouping #28 (coherence= .95 with 57 compounds)

Thus, from Tables 3 and 4 we constructed a dataset with 57 compounds and only 5 properties of the Compounds. The constructed dataset for the original dataset #28 is illustrated in Table 5. Each constructed dataset became the input to the Multi-dimensional Hough Algorithm.

Compound	AVERAGE VE	EAVERAGE N	AVERAGE ZUNGER	AVERAGE MTEMP	AVERAGE Z
Dy_Ge_Ir	5.333333333333333	1.67	2.6193333333333335	1858.6666666666667	58.333333333333336
Dy_Ge_Os	5.0	1.6633333333333333	2.626666666666667	2070.3333333333335	58.0
Ag_U_W	6.666666666666667	1.5200000000000002	3.276666666666667	2107.6666666666665	71.0
Dy_Os_Sn	5.0	1.6266666666666667	2.733333333333333	1835.0	64.0
Er_Ge_Ir	5.333333333333333	1.6866666666666668	2.606	1896.3333333333333	59.0
Er_Ge_Os	5.0	1.68	2.6133333333333333	2108.0	58.666666666666664
Er_Os_Sn	5.0	1.6433333333333333	2.72	1872.6666666666667	64.66666666666667
Eu_Ge_Ir	5.333333333333333	1.67	2.7093333333333334	1663.0	57.333333333333336
Eu_Ge_Os	5.0	1.6633333333333333	2.716666666666667	1874.6666666666667	57.0
Eu_Os_Sb	5.333333333333333	1.7133333333333336	2.785	1772.3333333333333	63.333333333333336
As_Ce_Os	5.333333333333333	1.7400000000000002	2.855	1826.3333333333333	55.666666666666664
As_La_Os	5.333333333333333	1.8233333333333335	2.3816666666666664	1867.0	55.333333333333336
As_Nd_Os	5.333333333333333	1.7733333333333334	2.685	1897.0	56.333333333333336
As_Os_Pr	5.333333333333333	1.7400000000000002	2.848333333333333	1870.6666666666667	56.0
As_Os_Th	5.333333333333333	1.8066666666666666	3.015	2143.6666666666665	66.33333333333333
Ga_Os_U	4.666666666666667	1.75	3.0216666666666665	1675.3333333333333	66.33333333333333
Gd_Ge_Ir	5.333333333333333	1.6533333333333333	2.6993333333333336	1826.0	57.666666666666664
Gd_Ge_Os	5.0	1.6466666666666665	2.706666666666667	2037.6666666666667	57.333333333333336
Gd_Ge_Re	4.666666666666667	1.7166666666666668	2.716666666666667	2082.6666666666665	57.0
Gd_Os_Sn	5.0	1.61	2.813333333333334	1802.3333333333333	63.333333333333336
Ge_Ho_Ir	5.333333333333333	1.6866666666666668	2.6126666666666667	1879.0	58.666666666666664
Ge_Ho_Os	5.0	1.68	2.6199999999999997	2090.6666666666665	58.333333333333336
Ge_Ho_Re	4.666666666666667	1.75	2.6300000000000003	2135.6666666666665	58.0
Ge_Ir_Lu	5.333333333333333	1.6866666666666668	2.5193333333333334	1941.0	60.0
Ge_Ir_Nd	5.333333333333333	1.6866666666666668	2.7260000000000004	1725.6666666666667	56.333333333333336
Ge_Ir_Pr	5.333333333333333	1.6533333333333335	2.8893333333333335	1699.3333333333333	56.0
Ge_Ir_Sm	5.333333333333333	1.6866666666666668	2.776	1746.3333333333333	57.0
Ge_Ir_Tb	5.333333333333333	1.6866666666666668	2.692666666666667	1842.3333333333333	58.0
Ge_Ir_Th	5.333333333333333	1.72	3.0560000000000005	1972.3333333333333	66.33333333333333
Ge_Ir_Tm	5.333333333333333	1.6866666666666668	2.596	1904.0	59.333333333333336
Ge_Ir_Yb	5.333333333333333	1.6533333333333335	2.5926666666666667	1663.6666666666667	59.666666666666664
Ge_La_Os	5.0	1.7299999999999998	2.43	1907.3333333333333	55.0
Ge_Lu_Os	5.0	1.68	2.526666666666667	2152.6666666666665	59.666666666666664
Ge_Lu_Re	4.666666666666667	1.75	2.5366666666666666	2197.6666666666665	59.333333333333336
Ge_Nd_Os	5.0	1.68	2.733333333333334	1937.3333333333333	56.0
Ge_Os_Pr	5.0	1.6466666666666665	2.896666666666667	1911.0	55.666666666666664
Ge_Os_Sm	5.0	1.68	2.783333333333333	1958.0	56.666666666666664
Ge_Os_Tb	5.0	1.68	2.6999999999999997	2054.0	57.666666666666664
Ge_Os_Tm	5.0	1.68	2.6033333333333335	2115.6666666666665	59.0
Ge_Os_U	5.0	1.8466666666666667	2.9766666666666666	1978.0	66.66666666666667
Ge_Os_Yb	5.0	1.6466666666666665	2.6	1875.3333333333333	59.333333333333336
Ho_Os_Sn	5.0	1.6433333333333333	2.7266666666666666	1855.3333333333333	64.33333333333333
La_Os_Sb	5.333333333333333	1.78	2.4983333333333335	1805.0	61.333333333333336
Lu_Os_Sn	5.0	1.6433333333333333	2.6333333333333333	1917.3333333333333	65.66666666666667
Ce_Fe_Re	6.0	1.61	3.0966666666666662	2110.6666666666665	53.0
Ce_Fe_Ta	5.333333333333333	1.57	3.133333333333333	2049.3333333333335	52.333333333333336
Ce_Ge_Ir	5.333333333333333	1.6533333333333333	2.8960000000000004	1655.0	55.666666666666664
Ce_Ge_Os	5.0	1.6466666666666665	2.9033333333333338	1866.6666666666667	55.333333333333336
Ce_Os_Sb	5.333333333333333	1.6966666666666665	2.971666666666667	1764.3333333333333	61.666666666666664
Nd_Os_Sb	5.333333333333333	1.7299999999999998	2.801666666666667	1835.0	62.333333333333336
Os_Pr_Sb	5.333333333333333	1.6966666666666665	2.9650000000000003	1808.6666666666667	62.0
Os_Sb_Sm	5.333333333333333	1.7300000000000002	2.8516666666666666	1855.6666666666667	63.0
Os_Sn_Tb	5.0	1.6433333333333333	2.8066666666666666	1818.6666666666667	63.666666666666664
Os_Sn_Th	5.0	1.6766666666666667	3.17	1948.6666666666667	72.0
Os_Sn_Tm	5.0	1.6433333333333333	2.7099999999999995	1880.3333333333333	65.0
Os_Sn_Y	5.0	1.7133333333333332	2.4899999999999998	1873.0	55.0
Ru_Sn_Th	5.0	1.7166666666666666	3.155	1703.6666666666667	61.333333333333336

Table 5: Dataset Grouping #28 with Compound Properties

set to be the average the constituent Atomic Element Properties

5. Results of applying the Multi-dimensional Hough Transform System to the constructed datasets

The Multi-dimensional Hough Transform System was applied to the constructed datasets corresponding to the 21 datasets we were asked to analyze. For each such dataset the system returned a list of linearities that it found. This list was ordered so that the linearities containing more points (i.e. compounds) appeared first and the linearities that contained fewer points appeared last. Thus the best linearity found is always at the front of the list.

Each linearity is represented by a pair of lists. The first element (i. e. CAR in LISP) of each linearity is a list of ranges corresponding to the properties of the objects in the order presented. For dataset #28 the first element of

each linearity is a list of 5 ranges corresponding to the five properties of the compounds in the order presented (i.e. Average-VE, Average-EN, Average-ZUNGER, Average-MTEMP, Average-Z). Each range consists of two objects (i. e. chemical compounds) separated by three periods representing the object with the lowest and highest (in that order) property values for the given property in that linearity. The remaining elements (i.e. CDR in LISP) of each linearity is the list of objects (i.e. compounds) of the given dataset that occur in that linearity.

For example, the best linearity found for the constructed dataset #28 was:

```
( ((Ge_Lu_Re ... Ag_U_W)
   (Ag_U_W ... Ge_Os_U)
   (As_La_Os ... Ag_U_W)
   (Ce_Ge_Ir ... Ge_Lu_Re)
   (Ce_Fe_Ta ... Os_Sn_Th))
Ru_Sn_Th    Os_Sn_Y     Os_Sn_Tm    Os_Sn_Th    Os_Sn_Tb    Os_Sb_Sm
Os_Pr_Sb    Nd_Os_Sb    Ce_Os_Sb    Ce_Ge_Os    Ce_Ge_Ir    Ce_Fe_Ta
Ce_Fe_Re    Lu_Os_Sn    La_Os_Sb    Ho_Os_Sn    Ge_Os_Yb    Ge_Os_U
Ge_Os_Tm    Ge_Os_Tb    Ge_Os_Sm    Ge_Os_Pr    Ge_Nd_Os    Ge_Lu_Re
Ge_Lu_Os    Ge_La_Os    Ge_Ir_Yb    Ge_Ir_Tm    Ge_Ir_Th    Ge_Ir_Tb
Ge_Ir_Sm    Ge_Ir_Pr    Ge_Ir_Nd    Ge_Ir_Lu    Ge_Ho_Re    Ge_Ho_Os
Ge_Ho_Ir    Gd_Os_Sn    Gd_Ge_Re    Gd_Ge_Os    Gd_Ge_Ir    Ga_Os_U
As_Os_Th    As_Os_Pr    As_Nd_Os    As_La_Os    As_Ce_Os    Eu_Os_Sb
Eu_Ge_Os    Eu_Ge_Ir    Er_Os_Sn    Er_Ge_Os    Er_Ge_Ir    Dy_Os_Sn
Ag_U_W      Dy_Ge_Os    Dy_Ge_Ir)
```

The second best linearity was:

```
(((Ge_Lu_Re ... Ag_U_W)
  (Ag_U_W ... Ge_Os_U)
  (As_La_Os ... Ag_U_W)
  (Ce_Ge_Ir ... Ge_Lu_Re)
  (Ce_Fe_Ta ... Os_Sn_Th))
Ru_Sn_Th    Os_Sn_Y     Os_Sn_Tm    Os_Sn_Th    Os_Sn_Tb    Os_Sb_Sm
Os_Pr_Sb    Nd_Os_Sb    Ce_Os_Sb    Ce_Ge_Os    Ce_Ge_Ir    Ce_Fe_Ta
Ce_Fe_Re    Lu_Os_Sn    La_Os_Sb    Ho_Os_Sn    Ge_Os_Yb    Ge_Os_U
Ge_Os_Tm    Ge_Os_Tb    Ge_Os_Sm    Ge_Os_Pr    Ge_Nd_Os    Ge_Lu_Re Ge_Lu_Os
        Ge_La_Os        Ge_Ir_Yb        Ge_Ir_Tm        Ge_Ir_Th        Ge_Ir_Tb
Ge_Ir_Sm    Ge_Ir_Pr    Ge_Ir_Nd    Ge_Ir_Lu    Ge_Ho_Re    Ge_Ho_Os
 Ge_Ho_Ir   Gd_Os_Sn    Gd_Ge_Re    Gd_Ge_Os    Gd_Ge_Ir
As_Os_Th    As_Os_Pr    As_Nd_Os    As_La_Os    As_Ce_Os    Eu_Os_Sb
Eu_Ge_Os    Eu_Ge_Ir    Er_Os_Sn    Er_Ge_Os    Er_Ge_Ir    Dy_Os_Sn
Ag_U_W      Dy_Ge_Os    Dy_Ge_Ir)
```

The range of property values (i.e. the lowest and highest values) for each property of a linearity may be computed from two objects (i.e.. chemical compounds) in each range in the list of ranges which is returned (as the CAR) for each linearity by looking up that property values in the input dataset (which in this case is given in Table 5). Thus the ranges for the above two linearities are:

Average-VE	(Ge_Lu_Re ... Ag_U_W)	4.67 to 6.67
Average-ZUNGER	(Ag_U_W ... Ge_Os_U)	1.52 to 1.846
Average-EN	(As_La_Os ... Ag_U_W)	2.3816 to 2.536
Average-MTEMP	(Ce_Ge_Ir ... Ge_Lu_Re)	1655 to 2197.6
Average-Z	(Ce_Fe_Ta ... Os_Sn_Th)	52.3 to 72

It will be recalled from the introduction that the original purpose of this system was to produce ranges for the various bins. The above linearity suggests for example that the Average-VE bin should have a bin with boundaries from 4.67 to 6.67 and that this bin would be highly correlated with the corresponding bin boundaries given for the other properties in this linearity. If this dataset were part of a much larger dataset with many other linearities all these linearities would suggest many correlated bin boundaries which may then simplify and improve the formation of rules about such data.

All 35 linearities from the constructed dataset 28 are listed in Appendix A.

The results of applying the system to the other 20 datasets are too numerous to present here, but all have been delivered in electronic media to the Materials Laboratory for assessment by the Materials Scientists of the Materials Laboratory.

6. Preliminary Analysis of the results of the constructed Dataset #28

The result that the best linearity includes all 57 datapoints suggests that all 57 points are linearly related to each other to the extent that a true linearity is approximated by the cardinallity of slopes s.

Since there were $8^5=2^{15}=32{,}000$ cells in the Hough Array, it appears extraordinary that all such points should fall within the same cell. This suggests that the best linearity found is a true linearity and not just an artifact of using too few cells.

The second best linearity is identical to the best linearity except that it omits a single point (i.e. the compound Ga_Os_U). What is going on here is that an s angle is being changed to create the next cell, but that new angle does not effect things much since 56 of the points still fall within the new cell. This suggests that linearities contained within other linearities are essentially redundant and could be omitted or alternatively one could decrease the cardinallity of s.

The difference between the best and second best linearities could also suggest that Ga_Os_U is an outlier of the class of elements. A more definitive answer in this respect might be obtained by examining all the linearities for differences from the best linearities.

7. Conclusion

The Multi-dimensional Hough Transform appears to have potential for both organizing and simplifying data for input to various automatic classification and theory formation systems and for assessing the quality of the results of such systems. In this regard we believe that further research along following lines may be productive:

(1) The slopes s should be calibrated on known examples both respect to its cardinallity and division points. In this research s has been given a cardinallity of 8 evenly spaced angles. However, in various cases the results suggest that s in too large or too small. Furthermore equal increments in angle do not translate into equal changes in slope.

(2) The system should be tested on the database of all compounds too see if the linearities it recognizes correlates with the groups of three element compounds that have been suggested as being "similar" by other theory formation methods. This should be feasible because the complexity of the Multi-dimensional Hough Transform lies basically in the number of properties involved, not in the number of compounds examined. The number of compounds is only a linear factor in the overall complexity.

(3) The system should be tied into one or more automatic classification or theory formation systems so as to provide those systems with an automated correlated binning capability.

(4) The linearities from the 21 constructed datasets used herein should be examined for significance by a Materials Scientist. Such a Scientist may be able to suggest ways to automatically assess or help to assess the significance of these results.

References

Brown, Frank M., Jacobs, G., Snider, J. S., Jackson, A. G. , Leclair, S. R., An Experimental Investigation of Gladun's Theory of Automatic Classification, Ninth Midwest Artificial Intelligence and Cognitive Science Coherence, March 20-22, Wright State University, Dayton, Ohio., AAAI publisher, 1998.

Umbaugh, Scott E., Computer Vision and Image Processing: A Practical Approach to Using CVIPtools, Prentice Hall, 1998.

Appendix A: Results of applying the Multi-dimensional Hough Transform System to the constructed dataset #28

The Linearities found for Dataset #28 are the elements of the following list. This list is sorted so that the linearities with the most members occur earlier. Each linearity is represented by a pair of lists. The first element (i.e. CAR in LISP) of each linearity is a list of five ranges corresponding to the five properties. Each element of a range is the average of the corresponding property value of each of the three atomic elements in the given compound. Thus each range is the compound with the lowest and highest average in the given linearity. The remaining elements (i.e. CDR) of each linearity is the list of compounds of the given dataset that occur in that linearity.

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ce_Fe_Re) (Ce_Fe_Ta ... Ge_Os_U) (As_La_Os ... Os_Sn_Th) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (Ge_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (Ge_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re

Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... As_La_Os) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Os_Sb_Sm) (Ce_Fe_Ta ... Ge_Os_U) (As_La_Os ... Os_Sn_Th) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Os_Sn_Y ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Ag_U_W)) Os_Sn_Y Os_Sn_Tm Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Os_Sb_Sm) (Gd_Os_Sn ... Ge_Os_U) (As_La_Os ... Ge_Ir_Th) (Ce_Ge_Ir ... Ge_Lu_Re) (Os_Sn_Y ... Ge_Os_U)) Os_Sn_Y Os_Sn_Tm Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Dy_Ge_Os Dy_Ge_Ir)

(((Os_Sn_Y ... Ag_U_W) (Ag_U_W ... As_La_Os) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Os) (Ce_Fe_Ta ... Ag_U_W)) Os_Sn_Y Os_Sn_Tm Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Os Gd_Ge_Ir As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Os_Sb_Sm) (Gd_Os_Sn ... Ge_Os_U) (As_La_Os ... As_Os_Th) (Ce_Ge_Ir ... Ge_Lu_Re) (Os_Sn_Y ... Ge_Os_U)) Os_Sn_Y Os_Sn_Tm Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Dy_Ge_Os Dy_Ge_Ir)

(((Ru_Sn_Th ... Ag_U_W) (Ag_U_W ... As_Os_Th) (Ge_Ir_Lu ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Os) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn Ho_Os_Sn Ge_Os_Yb Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Os Gd_Ge_Ir As_Os_Th As_Os_Pr As_Nd_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (Os_Sn_Y ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ho_Re Ge_Ho_Os Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Dy_Os_Sn Ag_U_W Dy_Ge_Os)

(((Ge_Lu_Re ... Os_Sb_Sm) (Ce_Fe_Ta ... Ge_Os_U) (Ge_La_Os ... Os_Sn_Th) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Lu_Os_Sn Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ho_Re Ge_Ho_Os Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Dy_Os_Sn Dy_Ge_Os)

(((Ge_Lu_Re ... Os_Sb_Sm) (Gd_Os_Sn ... Ge_Os_U) (As_La_Os ... Ga_Os_U) (Eu_Ge_Ir ... Ge_Lu_Re) (Os_Sn_Y ... Ge_Os_U)) Os_Sn_Y Os_Sn_Tm Os_Sn_Tb Os_Sb_Sm Nd_Os_Sb Ce_Ge_Os Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Dy_Ge_Os Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... As_La_Os) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Re) (Ce_Fe_Re ... Ag_U_W)) Os_Sn_Y Os_Sn_Tm Os_Sb_Sm Nd_Os_Sb Ce_Ge_Ir Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_Tm Ge_Os_Tb Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Ge_Os Gd_Ge_Ir As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Os_Sn_Tm ... Ag_U_W) (Ag_U_W ... La_Os_Sb) (La_Os_Sb ... Ag_U_W) (Ce_Ge_Ir ... Ce_Fe_Re) (Ce_Fe_Ta ... Ag_U_W)) Os_Sn_Tm Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_Pr Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Os Gd_Ge_Ir As_Os_Pr As_Nd_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Ir Er_Os_Sn Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Ir)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... As_La_Os) (As_La_Os ... Ag_U_W) (Eu_Ge_Ir ... Ge_Lu_Re) (Os_Sn_Y ... Ag_U_W)) Os_Sn_Y Os_Sn_Tm Os_Sb_Sm Nd_Os_Sb Lu_Os_Sn La_Os_Sb Ge_Os_Yb Ge_Os_Tm Ge_Os_Tb Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Ge_Os Gd_Ge_Ir As_Os_Th As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Ge_Os Er_Ge_Ir Ag_U_W Dy_Ge_Os Dy_Ge_Ir)

(((Ga_Os_U ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (Ge_Ir_Tb ... Ag_U_W) (Ce_Ge_Ir ... As_Os_Th) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Ho_Os_Sn Ge_Os_U Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Ir_Th Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Gd_Os_Sn Gd_Ge_Os Gd_Ge_Ir Ga_Os_U As_Os_Th As_Os_Pr As_Ce_Os Eu_Os_Sb Eu_Ge_Os Eu_Ge_Ir Er_Os_Sn Dy_Os_Sn Ag_U_W)

(((Ge_Lu_Re ... Ag_U_W) (Ag_U_W ... Ge_Os_U) (La_Os_Sb ... Ag_U_W) (Ge_Ir_Yb ... Ge_Lu_Re) (Gd_Ge_Re ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Sb_Sm Os_Pr_Sb Nd_Os_Sb Ce_Os_Sb Lu_Os_Sn La_Os_Sb Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Lu_Re Ge_Lu_Os Ge_Ir_Yb

Ge_Ir_Tm Ge_Ir_Th Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Ga_Os_U As_Os_Th Eu_Os_Sb Er_Os_Sn Er_Ge_Os Er_Ge_Ir Dy_Os_Sn Ag_U_W Dy_Ge_Os)

(((Ge_Lu_Re ... Os_Pr_Sb) (Ce_Fe_Ta ... Ge_Os_U) (Ge_Lu_Re ... Os_Sn_Th) (Ga_Os_U ... Ge_Lu_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Tm Os_Sn_Th Os_Sn_Tb Os_Pr_Sb Ce_Os_Sb Ce_Ge_Os Ce_Fe_Ta Lu_Os_Sn Ho_Os_Sn Ge_Os_Yb Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_Lu_Re Ge_Ir_Th Ge_Ho_Re Ge_Ho_Os Gd_Os_Sn Gd_Ge_Re Gd_Ge_Os Ga_Os_U As_Os_Th Eu_Ge_Os Er_Os_Sn Er_Ge_Os Dy_Os_Sn Dy_Ge_Os)

(((Os_Sn_Y ... Ag_U_W) (Ag_U_W ... As_La_Os) (As_La_Os ... Ag_U_W) (Ce_Ge_Ir ... Ge_Lu_Os) (Ce_Fe_Re ... Ag_U_W)) Os_Sn_Y Os_Sb_Sm Nd_Os_Sb Ce_Ge_Ir Ce_Fe_Re La_Os_Sb Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Ge_Ir_Lu Ge_Ho_Ir Gd_Ge_Ir As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Os_Sb Eu_Ge_Ir Er_Ge_Ir Ag_U_W Dy_Ge_Ir)

(((Ge_Lu_Re ... La_Os_Sb) (Lu_Os_Sn ... As_La_Os) (As_La_Os ... Gd_Ge_Re) (Ge_Ir_Yb ... Ge_Lu_Re) (Os_Sn_Y ... Lu_Os_Sn)) Os_Sn_Y Lu_Os_Sn La_Os_Sb Ge_Os_Yb Ge_Os_Tm Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Lu Ge_Ho_Re Ge_Ho_Os Ge_Ho_Ir Gd_Ge_Re As_Nd_Os As_La_Os Er_Ge_Os Er_Ge_Ir Dy_Ge_Os Dy_Ge_Ir)

(((Os_Sn_Y ... Ce_Fe_Re) (Ce_Fe_Ta ... As_La_Os) (As_La_Os ... Ce_Fe_Ta) (Ce_Ge_Ir ... Ce_Fe_Re) (Ce_Fe_Ta ... Dy_Ge_Ir)) Os_Sn_Y Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Ge_Os_Sm Ge_Os_Pr Ge_Nd_Os Ge_La_Os Ge_Ir_Tb Ge_Ir_Sm Ge_Ir_Pr Ge_Ir_Nd Gd_Ge_Ir As_Os_Pr As_Nd_Os As_La_Os As_Ce_Os Eu_Ge_Os Eu_Ge_Ir Dy_Ge_Ir)

(((Ge_Lu_Re ... As_Os_Th) (Eu_Ge_Os ... Ge_Os_U) (As_La_Os ... Os_Sn_Th) (Ga_Os_U ... Ge_Lu_Re) (Os_Sn_Y ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Y Os_Sn_Th Ge_Os_U Ge_Os_Tm Ge_Os_Tb Ge_Os_Sm Ge_Nd_Os Ge_Lu_Re Ge_Lu_Os Ge_La_Os Ge_Ho_Re Ge_Ho_Os Gd_Ge_Re Ga_Os_U As_Os_Th As_La_Os Eu_Ge_Os Er_Ge_Os Dy_Ge_Os)

(((Gd_Ge_Re ... Ce_Fe_Re) (Ce_Fe_Ta ... Ge_Os_U) (Gd_Ge_Re ... Os_Sn_Th) (Ce_Ge_Ir ... Ce_Fe_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Th Os_Sn_Tb Os_Pr_Sb Ce_Os_Sb Ce_Ge_Os Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Ho_Os_Sn Ge_Os_U Ge_Os_Sm Ge_Os_Pr Ge_Ir_Th Ge_Ir_Pr Gd_Os_Sn Gd_Ge_Re Ga_Os_U Er_Os_Sn Dy_Os_Sn)

(((Gd_Ge_Re ... Os_Pr_Sb) (Ce_Fe_Ta ... Ge_Os_U) (Gd_Ge_Re ... Os_Sn_Th) (Ga_Os_U ... As_Os_Th) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Th Os_Sn_Tb Os_Pr_Sb Ce_Os_Sb Ce_Ge_Os Ce_Fe_Ta Ge_Os_U Ge_Os_Sm Ge_Os_Pr Ge_Ir_Th Gd_Os_Sn Gd_Ge_Re Ga_Os_U As_Os_Th)

(((La_Os_Sb ... Ag_U_W) (Ag_U_W ... As_La_Os) (As_La_Os ... Ag_U_W) (Ge_Ir_Yb ... Ce_Fe_Re) (Ce_Fe_Re ... Ag_U_W)) Ce_Fe_Re La_Os_Sb Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Lu Ge_Ho_Ir As_La_Os Er_Ge_Ir Ag_U_W Dy_Ge_Ir)

(((Ru_Sn_Th ... Ag_U_W) (Ag_U_W ... Ge_Ir_Th) (Ge_Ir_Pr ... Ag_U_W) (Ce_Ge_Ir ... Ce_Fe_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Th Os_Pr_Sb Ce_Os_Sb Ce_Ge_Ir Ce_Fe_Ta Ce_Fe_Re Ge_Ir_Th Ge_Ir_Pr Ag_U_W)

(((Ge_Lu_Re ... La_Os_Sb) (Os_Sn_Y ... Ge_Os_U) (As_La_Os ... Ga_Os_U) (Ga_Os_U ... Ge_Lu_Re) (Os_Sn_Y ... Ge_Os_U)) Os_Sn_Y La_Os_Sb Ge_Os_U Ge_Lu_Re Ge_La_Os Ge_Ho_Re Gd_Ge_Re Ga_Os_U As_La_Os)

(((Ge_La_Os ... La_Os_Sb) (Ge_Ir_Yb ... As_La_Os) (As_La_Os ... Dy_Ge_Ir) (Ge_Ir_Yb ... Ge_Ir_Lu) (Ge_La_Os ... La_Os_Sb)) La_Os_Sb Ge_La_Os Ge_Ir_Yb Ge_Ir_Tm Ge_Ir_Lu Ge_Ho_Ir As_La_Os Er_Ge_Ir Dy_Ge_Ir)

(((Ge_Lu_Re ... Ru_Sn_Th) (Os_Sn_Th ... Ge_Os_U) (Ge_Lu_Re ... Os_Sn_Th) (Ga_Os_U ... Ge_Lu_Re) (Gd_Ge_Re ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Th Ge_Os_U Ge_Lu_Re Ge_Ho_Re Gd_Ge_Re Ga_Os_U)

(((Ga_Os_U ... Ag_U_W) (Ag_U_W ... Ga_Os_U) (Ga_Os_U ... Ag_U_W) (Ga_Os_U ... Ce_Fe_Re) (Ce_Fe_Ta ... Os_Sn_Th)) Ru_Sn_Th Os_Sn_Th Ce_Fe_Ta Ce_Fe_Re Ge_Ir_Th Ga_Os_U Ag_U_W)

end

CHARACTERIZATION OF ACOUSTIC SOURCES FOR HYPERSONIC RECEPTIVITY RESEARCH

Gregory A. Buck
Assistant Professor
Mechanical Engineering Department

South Dakota School of Mines and Technology
501 E. St. Joseph Street
Rapid City, SD 57701

Final Report for Summer Faculty Research Program
Wright Patterson AFB
Aircraft Aerodynamics Division

Sponsored by:
Air Force Office of Scientific Research
Bolling AFB, Washington, DC

and

Wright Patterson Research Lab

September, 1998

CHARACTERIZATION OF ACOUSTIC SOURCES FOR HYPERSONIC RECEPTIVITY RESEARCH

Gregory A. Buck
Assistant Professor
Mechanical Engineering Department
South Dakota School of Mines and Technology

Abstract

Transition from a laminar state of fluid motion to a turbulent flow remains one of the most complex, as well as one of the most important, unsolved problems in fluid mechanics. In the design of supersonic and hypersonic aircraft, prediction of boundary layer transition is of paramount importance because the turbulence dramatically increases both heat transfer and skin friction drag, with attendant increases in weight and cost, and decreases in aerodynamic performance and flight range.

Although a number of theories currently exist to predict transition, they rest heavily on knowledge of initial disturbance amplitudes in the boundary layer, which are coupled in some manner to the freestream fluctuations. The process by which freestream disturbances generate instabilities in the boundary layer is referred to as *receptivity*, and plays a pivotal role in transition. Despite several decades of intense study, receptivity mechanisms are inadequately understood at present, particularly for acoustic disturbances in supersonic boundary layers, and the need for additional experimental confirmation of theory is generally recognized. In order to conduct experimental studies of the receptivity of hypersonic boundary layers to acoustic disturbances, a controllable and measurable source of disturbances is required.

In this study, candidates for acoustic sources were identified and used to introduce disturbances into the test section of a Mach 2.8 wind tunnel. These included: 1) a simple rectangular cutout or cavity in the wind tunnel floor, 2) a rectangular cavity periodically excited by ultrasonic piezoceramic transducers tuned to 25 kHz and 40kHz, and 3) a spark source generated by repeatedly applying a high voltage across two tungsten electrodes placed in a cylindrical cavity in the wind tunnel floor. Velocity disturbances induced by the various sources were sensed using both a hot-film anemometer and a silicon diaphragm piezoresistive dynamic pressure sensor (Kulite). Schlieren photography was also used to visualize density differences in the flow and to document probe positions.

The results indicate that by means of appropriate source excitation, signal sampling and averaging, disturbances from both the excited cavity and the spark source can be successfully identified within the disturbance Mach cone, despite the presence of large amplitude acoustic noise radiated from the turbulent side wall boundary layers of the wind tunnel. Accounting for differences in sensor frequency response, both hot-film and Kulite measurements qualitatively duplicate the dynamic characteristics of the sources; in general, the Kulite probe, although physically larger, provides a higher sensitivity. However, the low signal to noise ratio and the presence of multiple sources of noise, give rise to a number of operational restrictions; continued source and sensor development is warranted.

Introduction and Background

Among other benefits, hypersonic flight improves weapon survivability and response time. Boundary layer transition to turbulence is central to hypersonic vehicle design primarily because the turbulence increases heat transfer to the vehicle. Higher heat transfer generally requires higher-performance thermal protection, at the price of increased weight and cost. Transition also impacts engine and aerodynamic performance due to large increases in drag from turbulent skin friction, particularly important for vehicles with large wetted areas and extended flight times. These factors place a premium on understanding the basic mechanisms underlying the transition process, which are not well understood at present, as evidenced by NASA's decision to postpone further work on the National Aerospace Plane until this phenomenon can be more reliably predicted and (ultimately) controlled.

Although a number of theories currently exist to predict transition, they rest heavily on knowledge of initial disturbance amplitudes in the boundary layer, which are coupled in some manner to the freestream fluctuations. The process by which freestream disturbances generate instabilities in the boundary layer is referred to as receptivity, and plays a pivotal role in transition. Freestream disturbances will have energy distributed over some range of wavelength and frequency (or equivalently wavelength and phase velocity), and the wavelength of a freestream disturbance will not, in general, match the wavelength of a boundary layer instability of the same frequency. For example, a critical issue in low-speed receptivity theory is the mismatch between acoustic wavelengths and Tollmein-Schlichting (TS) wavelengths, which may be two orders of magnitude smaller.[1] Also, the amplitude distribution of the disturbance in the boundary layer will not necessarily match the eigenfunction of a boundary layer instability. Details of this evolution from freestream disturbance to boundary layer instability, including the amplitude of the engendered boundary layer instability, are a desired outcome of receptivity study. The review of low speed receptivity by Saric, et al.[1] is an excellent starting point for the study of receptivity. Choudhari and Streett[2] review receptivity in supersonic boundary layers.

Freestream disturbances in a compressible flow are often represented as three independent modes, distinguished by the principal physical mechanisms governing their propagation through the compressible medium. They are: 1) an entropy mode, 2) a vorticity mode, and 3) an acoustic mode. The entropy mode in a wind tunnel, for example, is typically described as "temperature spottiness", i.e. a fluctuation in the fluid total temperature and density due to nonuniform upstream heating. It convects at the freestream velocity and has a pressure equal to the freestream. The vorticity mode arises from sources of vorticity in the flow (turbulence, wakes, boundary layers, etc.) and also has constant pressure and freestream convection velocity. The acoustic mode is distinguished from the vorticity and entropy modes in that it consists of isentropic pressure, density and temperature fluctuations which propagate at some velocity not necessarily equal to the freestream velocity. A point on a sound wave created by a stationary source propagates normal to the wavefront at the sound speed, a, and convects with the freestream velocity, U. The points on the wavefront directly upstream and downstream of a stationary source in supersonic flow propagate downstream at U-a (slow acoustic wave) and U+a (fast acoustic wave), respectively.

Of special interest to supersonic and hypersonic transition are sound waves created by moving sources, specifically, acoustic radiation from "eddy Mach waves" [3] created by turbulent eddies in adjacent boundary layers. These eddies may be thought of as wavy walls moving at some convection velocity c_r, typically less than the freestream U. For convection velocities which are supersonic relative to the freestream sonic condition, i.e. $M_{rel} = (U-c_r)/a > 1$, Mach waves traveling at c_r and inclined at the relative Mach angle ($= \text{Arcsin}(1/M_{rel})$) will be created.

A notable difference between low-speed and supersonic receptivity is the close correlation between the phase velocities of the instabilities and the freestream acoustic disturbances.[2] Amplified first and second mode disturbances have dimensional phase velocities between $c_0 = U - a$ (the slow acoustic mode, $c_0/U = 1 - 1/M$ in non-dimensional terms), and U.[5] The wavenumbers of the freestream acoustic disturbances are thus much more closely tuned to the boundary layer instabilities. Phase velocities between U and U-a correspond to disturbances traveling subsonically relative to the freestream. As noted by Laufer,[4] as the edge Mach number of a turbulent boundary layer increases, more of the radiated acoustic energy is concentrated in disturbances traveling supersonically relative to the freestream ($c_r < U - a$). The most intense eddy Mach radiation will thus have phase velocities less than the amplified boundary layer instabilities, and some mechanism would be required to transfer this energy to higher phase velocities. Some acoustic radiation with phase velocities between U and U-a will, exist, however, and no tuning mechanism is required for these disturbances.

Until the early 1990's, linear stability theory formed the basis for the most advanced tool for hypersonic boundary layer transition prediction, the "e^N" method.[5] The e^N method, however, is still fundamentally a correlation method. Recent advances such as Parabolized Stability Theory,[6] Direct Navier Stokes Simulation,[7] and Compressible Linear Navier Stokes,[8] open the possibility of computing finite-amplitude disturbance growth from receptivity through breakdown. Since the receptivity process is intrinsic to all transition processes, further experimental work on receptivity is warranted, not only to obtain new physical insight, but also to provide data and validation for computational benchmarking. Successful experimental studies rest on the development and characterization of a good acoustic disturbance source.

General Requirements for Acoustic Sources

The ideal acoustic source should produce disturbances with a well-defined and independently variable frequency and wave number spectrum, and independently variable amplitude, from the linear to non-linear range. Maslov, et al.[9], have conducted experiments at Mach 2 using a glow discharge to provide a harmonic acoustic source. In the first version of this experiment, a discharge in a flat splitter plate created acoustic radiation, which impinged on the test boundary layer developing on an adjacent flat plate. In a second version of this experiment, the test flat plate was rolled 180 deg., so that the test boundary layer was now on the side opposite to the acoustic source on the splitter plate.[10] The test boundary layer was thus shielded from direct radiation from the source. The leading edge of the test flat plate thus appeared as a line acoustic source to the test boundary layer, in emulation of previous theoretical studies in the Russian literature. Oblique input waves inclined 20 to 40 degrees to the freestream were found to give maximum boundary layer response.

The second experiment was repeated at Mach 6 using point and line glow discharge sources on the splitter plate.[11] Input disturbances of 31.6 and 50 kHz were created. The largest boundary layer oscillations in this case were created by oblique waves inclined approximately 60 degrees to the freestream. The glow-discharge used in this experiment tended to produce a rather complex three-dimensional wave number spectrum due to sound created by traveling waves in the splitter-plate boundary layer.

Acoustic frequencies of interest in supersonic and hypersonic flows can be well into the ultrasonic range. A typical non-dimensional frequency $F = 10^{-4}$ would coincide with a dimensional frequency of 11.4 kHz for freestream Mach 4, with ambient stagnation conditions.

For example, under Mach 8 conditions at Arnold Engineering Development Center von Karman Facility Tunnel B, the "most dangerous" second mode frequencies are of the order of 100 kHz.

A ubiquitous problem in stability experiments is determination of the required acoustic amplitude. The amplitude of the disturbance must be large enough to be reliably detected, but not so large as to introduce non-linear effects (if comparison with linear stability theory is sought). Nishioka and Morkovin,[12] often cited in low-speed receptivity work, state that a threshold for nonlinear effects is 95 dB (SPL=1.1 Pa). Saric has shown that the threshold for nonlinearity depends on the background vorticity level, and has shown linear response up to 110 dB (SPL=6.3 Pa) for narrow band acoustic input.[13] Extrapolation of these levels to compressible flow is questionable. One avenue is to look at the order-of-magnitude of sound levels in conventional supersonic facilities, where acoustic noise is known to influence transition locations. Hot-wire measurements by Laufer showed broadband pressure fluctuations of approximately 0.04-0.06% of freestream dynamic pressure in the JPL supersonic tunnel over a Mach range of 1.6 to 5.0. For conditions typical of the Mach 4 Quiet Flow Ludweig Tube at Purdue University, this would be a sound pressure level of 70 to 105 Pa (131 to 134 dB).

A primary difficulty in supersonic and hypersonic acoustic receptivity experiments is generating sufficiently high-amplitude, high-frequency noise in a low-density environment. For a planar acoustic wave, the sound pressure level, p', may be related to the fluid velocity after the wave passage, U',via the acoustic impedance of the medium, ρa, where ρ and a are the undisturbed density and sound speed, respectively, as $p' = \rho a U'$. At a loudspeaker sound source for example, U' is velocity of the speaker surface. For the same speaker in a low-density environment, the sound pressure level will thus scale on ρa.

One drawback of continuous stationary sources in supersonic flow is that they create a pattern of constructive/destructive interference between the fast and slow portions of the acoustic waves emanating from the source. [14] Therefore any receptivity study using such a source would need to correct for this interference, and would also require some technique to isolate the two waves. Impulsive sources, such as blast waves and sparks are not faced with these problems and have also been studied as viable candidates.

Identification of Candidate Acoustic Sources

In subsonic receptivity experiments, acoustic disturbances have typically been generated using speakers. In Saric's experiments, [1,13] the speakers were located around the circumference of the plenum of the wind tunnel and were phased so as to generate planar acoustic waves normal to the freestream velocity. In studies of traveling disturbances in 3-dimensional flows, Buck and Takagi[15] used a speaker coupled to a small hole to introduce a point source of disturbances into a rotating disk boundary layer. Several methods were considered for introducing acoustic disturbances in a supersonic wind tunnel settling chamber, based on older work. Sparks and double diaphragm shock tubes were investigated in the early 1960's as sources for dynamic blast-loading effects.[16,17] Several factors make introduction of acoustic waves in the settling chamber difficult. One issue is that only fast waves can pass the sonic throat. In addition, sound levels are attenuated through the throat passage[18] and for these reasons the present study considers only acoustic sources placed in the test section.

Ultrasonic transducer/receiver pairs were also explored as a means of generating acoustic disturbances. Piezoceramic transducers available commercially from APC Corporation in both 25 kHz and 40 kHz models, were found to offer acceptable amplitude output in a reasonably

compact source package. Previous testing of the 40 kHz model in a vacuum chamber down to a pressure of 0.2 psia, produced promising results when the pressure disturbance was detected using the paired ultrasonic receiver [14]. However the diameter of this receiver sensor is 16 mm, somewhat restrictive with regard to the localized boundary layer measurements that are desired for receptivity experiments. Hence the ultrasonic transducer source was considered a viable candidate for the present study; however alternative means to detect the disturbances were examined, as described below.

Two options are available for impulsive sources. The first is a spark discharge and the second is laser-induced breakdown.[19] In studies performed in the Purdue Quiet Flow Ludweig Tube, the acoustic wave from a laser-induced breakdown was resolved using interferometry. However, the acoustic wave could not be resolved using a hot-wire anemometer.[20]

A spark discharge creates an impulsive source of heat, light and sound as a high voltage (typically greater than 8000 volts) is applied across two tungsten electrodes, separated by a small gap, ionizing the gas between the electrodes, causing a spark to jump across the gap. The light duration is typically less than 80 nanoseconds and for the present work, a preliminary study was conducted in the laboratory to examine the acoustic structure of a typical spark. The source was a Xenon Corporation Model 437-B Nanopulse ™ System. Two schlieren photographs taken at 50 and 100 microseconds after spark firing are shown in Fig. 1. The complex multiple wave structure is believed to be due to the spark gap geometry. The shock velocity decreases as the shock radius increases, which is typical of spherical shocks. A rough estimate of the shock Mach number for the 50 microsecond delay is obtained by dividing the measured shock radius by the time delay. This gives an average Mach number of approximately 1.3. The shock velocity for the second case is estimated by dividing the difference in the measured radii between the two cases by 50 microseconds. This gives an average Mach number of 1.2. Based on these preliminary observations, the spark source was regarded to be a good candidate for further study.

Of additional concern are disturbances associated with the orifice where sound is introduced. Any distortion of the acoustic wave due to the flowfield around the source (including the boundary layer) must be assessed. The orifice must be placed above or below the model so that extraneous vorticity disturbances do not interact with the model flowfield. This makes it difficult, if not impossible, to create plane waves at normal incidence with sound sources in the test section.

Another candidate that was identified for study was a cavity, or cut-out in the wind tunnel floor. Simple rectangular cavities have been well-known to produce acoustic radiation since at least the 1950's[21] and this technique was suggested by Schneider[22] in a previous communication as a potential source. A similar arrangement has been used to provide high-amplitude forcing of a free shear layer.[23] One benefit of such an arrangement is that the cut-out should produce cylindrical sound waves. A drawback is that such cavities usually generate sound with a large number of harmonics.

In summary, based on preliminary studies and previous experiments, three candidate acoustic sources were identified as most promising for the present study. They were 1) a simple rectangular cutout or cavity in the wind tunnel floor, 2) a rectangular cavity periodically excited by ultrasonic piezoceramic transducers tuned to 25 kHz and 40kHz, and 3) a spark source generated by repeatedly applying a high voltage across two tungsten electrodes placed in a cylindrical cavity in the wind tunnel floor. The development of the experimental procedure for using these sources is discussed below.

Experimental Apparatus and Procedure

The three sources identified above were used to generate freestream disturbances in the Mach 2.8 supersonic wind tunnel at the Air Force Institute of Technology (AFIT). Figure 2 shows a schematic of the test section (2.5 inch by 2.5 inch cross section) geometry as viewed from the side, with the flow direction from left to right. Photographs of the wind tunnel apparatus and test section are shown in Figures 3 and 4, respectively. The facility is of the blowdown type and typical stagnation conditions of 30 psig and 25 °C were used for the runs reported herein, although the stagnation pressure was varied earlier in the experiment to study its influence on disturbance characteristics.

Since future receptivity experiments require measurements of disturbances within the boundary layer as well as in the freestream, the velocity disturbances induced by the various sources were sensed using both a hot-film anemometer and a silicon diaphragm piezoresistive dynamic pressure sensor (Kulite). The hot-film probe was a TSI Model 1218-20 curved boundary layer probe, with the sensing element located upstream of the centerline shown in Figure 2 (probe body centerline) by 0.5 inch. The hot-film probe was used with a 1:1 bridge provided in the TSI Model IFA100 Intelligent Flow Analyzer for higher frequency response. The hot-film overheat ratio was set as high as possible to improve signal sensitivity; the resistance ratio of film operating resistance to film cold resistance (taken at 0°C) was between 2 and 2.2 for all runs reported herein. The Kulite dynamic pressure probe was held in the flow with the sensing surface 1.125 inch upstream of the centerline indicated on Figure 2, and was supplied with an 18 VDC bridge excitation. The 0-5 psid series XCW-093 Kulite sensor was referenced to local atmospheric pressure and was specified to provide 50 mV full scale output for 15 VDC bridge excitation. The bridge output voltage was amplified by a gain of 100 using an external conditioning circuit. Both hot-film and Kulite signals were AC coupled to a LeCroy model 9384 Digital Oscilloscope, which was used to trigger and record data samples as discussed in the procedure below. The vertical positions of each of the probes were varied during the experiment and the spanwise positions were as follows:

Hot-film probe-	Centered across the span
Kulite probe-	0.5 inch to the right of the span centerline(facing downstream)

Schlieren photography was also used to visualize density differences in the flow and to document vertical probe positions. A Cooke Corporation PalFlash 500 spark light source with a vertical slit was used with two 60 inch focal length spherical mirrors; the first to collimate the source and the second to image the test section on 4 x 5 inch Type 87 Polaroid film. The knife edge was placed at the focal point of the second spherical mirror and, unless otherwise stated, was fixed in a vertical orientation to emphasize streamwise density gradients in a vertical plane at the center of the span. The relative position of the schlieren window to the other components in the test section is shown in Figure 2, for reference on the photographs. The lower tangent point of the window drops below the floor of the test section by 0.125 inch, as will be apparent from the photographs as well. Originally it was hoped to use the schlieren system to visualize disturbance wavefronts from the various sources as was done by Krishnamurty [21] for acoustic waves from cavities at Mach 0.8 and 1.3. Although the schlieren system was able to clearly visualize disturbances in still air from the Xenon Corp. 437-B Nanopulse spark source, as previously discussed (Figure 1), distinct wavefronts from cavity sources at Mach 2.8 (either excited or unexcited) were not observable in the present study.

At the onset of this experiment it was recognized that acoustic noise radiated from the turbulent sidewall boundary layers would be a predominant concern. Electrical noise induced by the high voltage discharge from the spark source was also observed, particularly in the hot-film signal. Ensemble averaging, a common practice in stability experiments, was also employed in this study to

allow identification of the signal from the background noise. Guided by a previous numerical experiment of Kimmel [24] where the number of ensemble averages was varied parametrically on a sine burst signal superposed with random noise of 100 times the amplitude of the signal, 862 ensemble averages were employed for all runs reported herein. Based on Kimmel's results, a larger number of ensemble averages might be desirable; the chosen value of 862 averages of a 500 point sample was a memory limitation of the A/D sampling hardware (LeCroy 9384 scope).

Tied to the issue of ensemble averaging, are the issues of acoustic source excitation and sample triggering, which were observed to be critical to the successful identification of disturbances in the supersonic freestream. Since the LeCroy 9384 computes ensemble averages in the time domain, sampling must be triggered at the same point in the phase of the input signal driving the acoustic source in order to preserve the identity of the measured signal from either sensor (Kulite or hot-film), during the averaging process. In addition, the Nanopulse spark source firing rate was practically limited to 50 Hz, so it was decided to modulate the harmonic excitation of the piezoceramic sources at 50 Hz to produce repetitive sine bursts of either 25 kHz or 40 kHz (depending on which source was being used) that could be used as the ensembles. A Wavetek Model 75 function generator was used to produce the 50 Hz, 10V p-p square wave which was adopted for both purposes, namely to trigger the spark source, and to attenuate the high frequency sinusoidal excitation of the piezoceramic sources (provided by an HP Model 3312A function generator supplied with the external modulating square wave). In the former role, the square wave was also used to trigger the data sampling which was set at a rate of 250 kHz for both Kulite and hot-film sensors in order to resolve acoustic frequencies up to 100 kHz or so in the freestream. In the latter role, when the piezoceramic sources were used (with 10 V p-p sine burst excitation), data sampling (again at a rate of 250 kHz) was triggered by the ANDing of two events using the LeCroy "smart trigger" feature, namely the square wave above a threshold voltage of 5.6 V AND the sine excitation above an arbitrary threshold of 3.5 V, thus phase locking the sampling to the same point of the excitation signal. Alternatively, the excitation could be produced with a phase lock so that each of the sine bursts occurring at 50 Hz is initiated at the same point in the sine phase, so that triggering on the square wave alone would be sufficient. During the development of this experimental procedure, it was also observed that when using the spark source, a 5X BNC attenuator should be employed between the LeCroy trigger input and the input trigger circuit of the Xenon Nanopulse spark source, in order to avoid feedback from the high voltage spark electrical spike and inadvertent triggering of the sampling.

With the sampling conducted as described above, 862 samples of 500 points each were taken from both the Kulite and the hot-film signals and stored to the LeCroy 9384. The raw data as well as the time averaged signals and their spectra were logged for each of the cases described below. Fortuitously, the 862 ensemble averages thus triggered at 50 Hz and sampled at 250 kHz upon each trigger, consumed a total sample time of 862/50 = 17.2 seconds which did not exceed the blowdown time of the wind tunnel. The sampling time could be shortened for the piezoceramic transducers by modulation of the excitation signal with a higher frequency square wave; for this study however, it was desired to keep identical sampling for both spark and piezoceramic sources, so that side-by side comparisons of the two sources could be made.

During the development of the procedure described above, considerable effort was expended in simply finding a technique that would allow identification of the acoustic disturbance from the background noise. This development examined a number of ideas, including various cavity length to depth ratios, a splitter plate to introduce disturbances into the freestream through a thinner boundary layer, a spark source with narrower gap width, and various excitation and triggering schemes. However, this report documents only the final series of runs that can be used to characterize the acoustic sources and considers only one gap width for the spark source (0.075 inch) and one cavity length (streamwise direction) to depth (vertical direction) ratio, L/d = 1. Despite these difficulties

early on, additional study should precede any final conclusions. For example, it was observed in the splitter plate runs that the support sting cross section was too large and created significant tunnel blockage; therefore it would be worthwhile redesign the plate with a smaller support sting with less flow blockage and the disturbances so introduced re-examined. In addition, a more thorough study of various cavity L/d's is also warranted, as is the development of a spark source designed for application as an acoustic disturbance source, rather than a light source. These issues will be discussed more fully in the Conclusions section of this report.

The following table summarizes the final runs and the conditions associated with each. For the hot-film measurements of disturbances in the freestream, the probe was placed at a vertical position within the Mach cone of the disturbances where the output from the sensor was large (typically in the lower Mach wave from the cavity trailing edge). A complete study of other positions within the Mach cone and the influence of probe position on disturbance amplitude was not conducted. The vertical probe positions are documented in the schlieren photographs for each case listed below (provided in a separate document), and as stated above, all cavity runs are for L/d = 1, and sampling for each ensemble was at 250 kHz with smart triggering (except for the spark source runs which only require the simple trigger), unless otherwise stated. The stagnation temperature is designated as T_0, stagnation pressure P_0 and the ambient barometric pressure P_b.

Case Number	Conditions	T_0	P_0	P_b
		°C	psig	in Hg
803-1(no photo)	Flow off baseline, no cavity, simple 50 Hz trig.	23.6	NA	29.09
803-2	Flow on baseline, no cavity, simple 50 Hz trig.	23.4	30	29.09
803-3	Unexcited cavity, simple 50 Hz trigger	23.5	30	29.09
803-4	Unexcited cavity, probes moved, simple 50 Hz	23.3	30	29.09
803-5	Unexcited cavity, simple 50 Hz trigger	23.4	30	29.04
803-6	Repeat of 803-5	23.4	30	29.04
803-7	Excited cavity, 25 kHz, simple 50 Hz trigger	23.9	30	29.07
803-8	Repeat of 803-7	23.5	30	29.07
804-1(no photo)	Unex cav., 500 kHz sample, simple 50 Hz trig.	22.9	30	29.11
804-2	Unex cav.,	22.5	30	29.11
804-3	Excited cavity, 25 kHz, simple 50 Hz trigger	22.6	30	29.11
804-4(no photo)	Flow off, excit. on at 25 kHz, smart trigger	22.6	NA	29.11
804-5	Excited cavity, 25 kHz	23.0	30	29.11
804-6	Excited cavity, 25 kHz	23.0	30	29.11
804-7(no photo)	Flow off baseline, excitation off	23.0	NA	29.10
804-8	Flow off baseline, excitation on 25 kHz	23.0	NA	29.10
804-9	Excited cavity, 25 kHz, Kulite probe moved	23.0	30	29.10
804-10(no photo)	Flow off baseline, excitation on 25 kHz	23.0	NA	29.10
804-11	Excited cavity, 25 kHz	22.8	30	29.10
804-12	Unexcited cavity (unplugged from 804-11)	22.6	30	29.09
805-1(no photo)	Flow off baseline, excitation on 40 kHz	22.4	NA	29.10
805-2	Excited cavity, 40 kHz	22.5	30	29.10
805-3	Repeat of 805-2	22.4	30	29.10
805-4	Unexcited cavity (unplugged from 805-3)	22.5	30	29.10
805-5	Excited cavity, 40 kHz	23.0	30	29.10
805-6	Unexcited cavity (unplugged from 805-5)	22.5	30	29.08
805-7	Excited cavity, 40 kHz	22.7	30	29.08
805-8	Excit cav., 40 kHz, probes up out of Mach cone	22.5	30	29.07
805-9	Repeat of 805-8	22.5	30	29.07
805-10	Excited cavity, 40 kHz, probes back in Mach	23.4	30	29.07

805-11	Excited cavity, 25 kHz, probes in lower Mach	25.0	30	29.05
805-12	Repeat of 805-11, Kulite moved down slightly	23.5	30	29.05
805-13	Excited cavity, 25 kHz, probes up out of Mach	23.5	30	29.05
806-1	Flow off tuning run, excit. on 25 kHz	22.8	NA	29.02
806-2	Flow off tuning run, piezo unplugged	22.8	NA	29.02
806-3	Flow off tuning run, blue valve on	22.8	NA	29.02
806-4	Excited cavity, 25 kHz, probes up out of Mach	23.0	30	29.03
806-5(new HW)	Excit. cav., 25 kHz, HW in Mach, Kulite out	23.5	30	29.01
806-6	Excited cavity, 25 kHz, adj. HW	23.8	30	28.99
806-7	Unexcited cavity	23.0	30	28.98
806-8	Excit cav, 25 kHz, HW low Mach, Kulite out	23.5	30	28.98
806-9	Excit cav., 25 kHz, both probes in Mach cone	23.5	30	28.97
806-10(no photo)	Flow off, spark excitation at 50 Hz	24.5	NA	28.97
806-11	Spark excitation at 50 Hz	24.0	30	28.97
806-12	Repeat of 806-11	24.0	30	28.97
806-13(no photo)	Flow off, spark excitation off baseline	24.5	NA	28.97
806-14	Flow on, spark excitation off baseline	24.0	30	28.97
806-15	Flow off, spark excitation on, probes up	24.4	NA	28.97
806-16	Flow on, spark excitation on, probes up/out	23.8	30	28.97
807-1	Flow off, spark excitation on, 1000 avgs	23.9	NA	29.06
807-2	Same as 807-1 except tape blocked			
807-3	Flow off, spark excit on	23.9	NA	29.06
807-4	Flow on, spark on, probes up/out of Mach cone	23.1	30	29.06
807-5	Flow on, spark on, probes in Mach cone	23.3	30	29.05
807-6	Flow on, spark off, probes in Mach cone	23.0	30	29.05
807-7	Flow on, spark on, HW in Mach, Kulite up	23.5	30	29.05
807-8(no photo)	Flow off, spark on, same probe pos. as 807-9	NR	NR	NR
807-9	Flow on, spark on, Kulite in Mach, HW out	22.8	30	29.05
807-10	Flow on, spark on, Kulite in Mach, HW out	22.0	30	29.05

Included in the above, several runs were also made to supply additional confirmation that acoustic signals were indeed being measured by the Kulite and hot-film probes. These involved computing a number of ensemble averages of each signal both with, and without, a physical block (duct tape) between the source and the sensors. Although not recorded to disk, similar runs were also made earlier with the piezoceramic sources as well. These runs are discussed in the Results and Discussion section below.

Experimental Results and Discussion

The unexcited cavity results for both Kulite and hot-film probes are given in the spectrum plots of Figures 5 and 6, respectively, where the spectra of the signals with, and without the cavity are shown. Baseline, flow-off noise for each sensor is also displayed on these plots, and the difference in scale between the Kulite and the hot-film amplitude axes should be noted. These plots clearly show a peak at approximately 28 kHz due to the presence of the cavity, which was also subsequently confirmed in the spectrum of the signal obtained by subtracting the no-cavity data from the cavity data. The Kulite probe was located in the upper Mach wave (from the cavity leading edge) and the hot-film probe in the lower Mach wave (from the cavity trailing edge), as shown in the photograph for Case 803-5. Depending on the dominant mode expected, this finding is not inconsistent with previous experimental correlations and theory.

Disturbances in the freestream produced by the cavity excited with the piezoceramic source at 25 kHz are represented by Case 804-11 and are shown in Figure 7. The ensemble averaged time signals (trace A: hot-film, trace C: Kulite) and corresponding spectra for the averaged signals (trace B: hot-film, trace D: Kulite), indicate detectable peaks at the input frequency of 25 kHz in each sensor, although the Kulite is a much higher level signal. This is a general observation for all runs reported herein, and for this reason, the hot-film measurement of disturbances was considerably more difficult, largely due to this lower signal to noise ratio. Figure 7 also documents the presence of considerable broadband, high-frequency noise in the freestream, which compounds the difficulty of the measurement. Confirmation of these measurements was made not only by repeating the same conditions (Cases 804-5, 804-6, 806-5, 806-6, 806-8, 806-9) but also by examining the signals with 1) flow off and excitation on as shown in Figure 8, representing Case 804-10, 2) flow on and excitation off as shown in Figure 9 representing Case 804-12, and 3) flow on with excitation on, but the probes moved up out of the Mach cone for the disturbances as shown in Figure 10, representing Case 805-13. These figures corroborate the expected behavior for acoustic sources and confirm the physical validity of the measurements. Similar results from the 40 kHz excited cavity are shown in Figures 11, 12, and 13 and lead to similar conclusions. These results also show that the 25 kHz source produces a signal somewhat larger than the 40 kHz source (about 10 dB on average), so that for the L/d =1 cavity it would be the preferred source. Another observation worthy of mention is the extreme sensitivity of the piezoceramic source output (as detected by either probe, but more readily observable with the Kulite) to input driving frequency. Frequency drift, either in the HP 3312A function generator output or the required input tuned frequency of the piezo source, resulted in a drop in signal amplitude, during a 1 or 2 hour period. Retuning to a slightly different frequency restored the signal, but could be a potential problem in future experiments.

As mentioned above, additional confirmation of the measurements was obtained by a comparison of the signals in still air, with and without a physical block (duct tape) over the source cavity. These results are shown in the time traces of Figures 14 and 15, for the 25 kHz excited cavity. For this comparison, both Kulite and hot-film signals were sampled at 10 MHz and averaged for 200 ensembles of 500,000 points each. The 25 kHz burst excitation is shown on trace 2 of these figures; trace C is the ensemble averaged Kulite and trace A is the ensemble averaged hot-film. These results clearly indicate that the acoustic disturbances are not detected when blocked by the tape.

Results from the spark source are best studied by first examining the baseline flow off, spark on condition as represented by Case 806-10, shown here as Figure 16. Trace A shows the ensemble averaged hot-film time response and trace B the corresponding spectrum, while trace C shows the Kulite average time response and trace D the corresponding Kulite spectrum. The time per division and frequency per division are indicated as the second items in the left hand boxes. In this run, the hot-film and Kulite probes are approximately 2.4 inches from the spark source and the still air temperature is 24.5 C. The figure shows a measured time delay for the disturbance to reach the sensor of about 0.17 msec, which would indicate an average Mach number of 1.04, in qualitative agreement with the schlieren measurements taken earlier which indicated a disturbance Mach number of 1.2. The lower value might be attributable to the weakening of the spark under rapid fire conditions, but this conclusion requires further study. At any rate the measured delay is reasonable for an acoustic disturbance and it is duplicated by both sensors. The spark source also displays a characteristic ringing at about 8.5 kHz, again consistent in both sensors, which may be due to internal wave reflections in the spark gap, or the tunnel wall cavity. Also apparent in the Kulite signal is a higher frequency oscillation at about 30 kHz, which could be due to the dynamic response of the Kulite itself, although the manufacturer published natural frequency is considerably higher (about 150 kHz). Figure 17 (Case 806-11) shows the same measurements with the flow on and shows a number of interesting observations. First the time delay for the spark disturbance to reach the sensors is shortened to about 0.1 msec, again duplicated by the two independent sensors. It is difficult to

estimate what would be expected for this elapsed time since initially the disturbance is propagating in the cylindrical cavity of the tunnel wall (where the flow is certainly not easily predictable), then into the subsonic boundary layer and finally into the supersonic boundary layer and freestream Mach cone. The shortening is consistent with this physics, however. Secondly, there is an apparent shift upward in the ringing of the source to about 9.5 kHz, possibly due to complex flow in the tunnel wall cavity interacting with the acoustic wave. Lastly, the 30 kHz oscillations in the Kulite probe are damped completely when the flow is on. The larger mean pressure on the transducer face probably acts to inhibit the free oscillations of the diaphragm under the impulsive acoustic excitation. Figure 18 (Case 806-14) shows the baseline flow on, excitation off condition, indicating that no acoustic signal is present in either sensor, as expected. Figures 19 and 20, representing Cases 806-15 and 806-16 respectively, show runs made with both probes moved up so that with flow on, they would be outside the Mach cone of the disturbances. Figure 19 is the flow off case, indicating a longer time for disturbances to reach the sensors, as expected; the signals are somewhat reduced in amplitude, particularly noticeable in the hot-film measurement. Figure 20 verifies the acoustic nature of the disturbances as they are all swept downstream, confined to the zone of influence for the Mach 2.8 flow, and are unable to reach either sensor. Although additional research on the influence of gap spacing and geometry are certainly warranted, these results offer promise that a spark may provide a detectable, impulsive acoustic source for receptivity studies.

Conclusions and Recommendations

Although all three candidate sources produced measurable disturbances in the freestream, the unexcited cavity generates a disturbance that is probably too weak to be very useful in subsequent receptivity studies. The 28 kHz peak in the spectrum for the Mach 2.8 flow is difficult to identify from the background noise, particularly using the hot-film sensor. Other L/d ratios could be studied to see if they produce larger amplitude disturbances, but the biggest drawback is that there is no convenient way to trigger phase locked data sampling with the unexcited cavity source.

Both piezoceramic sources used to excite the L/d =1 cavity provide a controllable source of disturbances which is readily detectable with the ensemble averaging technique described above. For this cavity aspect ratio, it appears that the 25 kHz source may provide a higher amplitude freestream disturbance, although there is a longer sustained ringing in the Kulite signal for this frequency, possibly related to some natural frequency in the sensor itself or in the cantilevered support sting. A longer time between bursts could be studied to see if this ringing will damp significantly, although this will increase the total sample time, if the same number of ensemble averages is used. It would also be worthwhile to consider alternative hardware for producing the modulated sine bursts, since the HP 3312A was observed not to produce a hard zero between bursts; a function generator with digital output frequency control may also help to alleviate the problem with frequency drift described above.

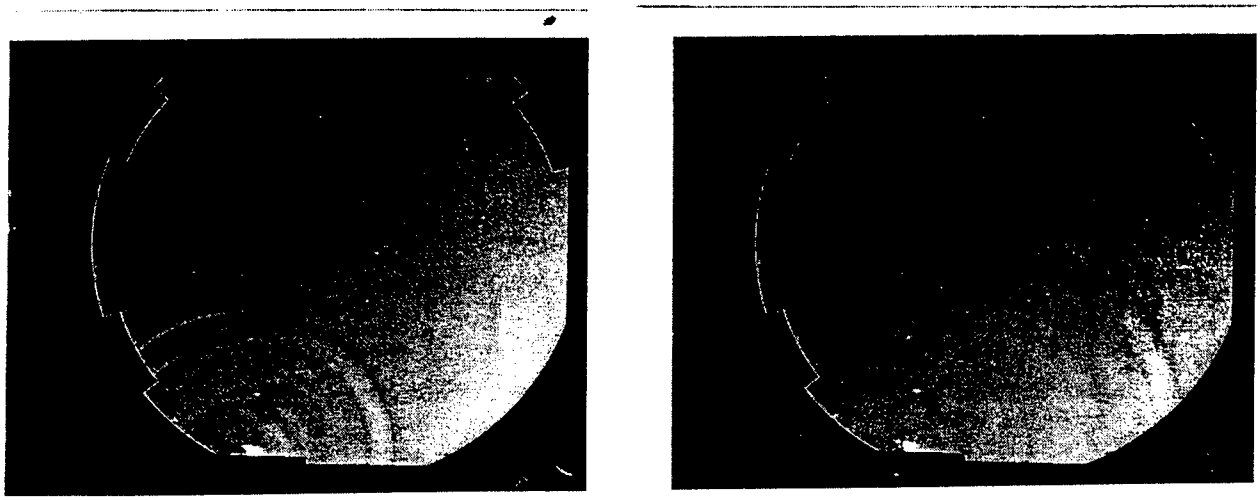
The spark source used in the present study was originally designed as a light source for rapid event capture but also produces measurable acoustic disturbance levels. However the high voltage discharge creates a large spike of electrical noise at each firing, which is picked up by the hot-film sensor, as shown in Figures 16 through 20. The disturbance rises above this background noise and is readily detectable, and it although may be impossible to remove or shield the sensor from this induced noise, the electrical charging and firing circuits should be examined to see if some design improvements related to this are possible. A source with higher firing rate would also be desirable to shorten total sampling time. Another observation worthy of note is the conspicuous decrease in light intensity from the spark source upon evacuation of the wind tunnel prior to blowdown, presumably also weakening the acoustic disturbance produced. This results from the difficulty in ionizing the

low density air between the electrodes, a condition that also results with supersonic flow in the test section. A source with a larger gap and higher available voltage might produce higher intensity acoustic disturbances under these low pressure conditions, but such adjustment is not readily attainable with the commercial spark light source. Other techniques for producing impulsive blasts with the desired control (and perhaps even wavefront geometry) could be examined and the development of a source tailored specifically to this application would probably be worthwhile.

Related to the sensors themselves and the sampling techniques, several issues are worthy of discussion. The Kulite sensor generally shows greater sensitivity (with the external gain of 100), but is physically larger and may not be acceptable for the localized boundary layer measurements required in receptivity studies. The hot-film sensor is suitable for boundary layer measurements but should be operated at as high an overheat as possible. The runs conducted after Case 806-4, used a different probe operating at a somewhat lower overheat (2.0 compared to 2.2 for the previous probe) and for these runs a decreased sensitivity was observed. In addition, specialized signal conditioning such as bandpass filtering and amplification for the hot-film (also employed in the experiments of Maslov, et al.) is worthy of further study for future experiments. Also, although there is always a trade-off between maximum sampling rate, memory, and A/D discretization error, a 12 bit A/D sampling will provide resolution of 1 part in 4096 compared to the 1 part in 256 presently employed with the 8 bit A/D conversion. On a 50 mV range, for example, a signal change of 0.1 mV would be undetectable with the 8 bit machine, whereas 12 bit A/D conversions will sense any changes greater than 12 μvolts. In concert with a higher number of ensemble averages, this should make small time dependent disturbances superposed on a mean AC fluctuation arising from time dependent noise more readily detectable.

Acknowledgements

The author wishes to acknowledge the Air Force Office of Scientific Research for their support of this work under the Summer Faculty Research Program. Dr. Roger Kimmel, Dr. John Schmissuer, and Dr. Jon Poggie of the Air Vehicles Directorate at Wright Patterson Laboratory provided considerable technical and hands-on input and were instrumental in the success of this undertaking, as was the support of the Air Force Institute of Technology faculty (Dr. William Elrod) and technical staff (Mr. Jay Anderson and Mr. Andy Pitts). The competent assistance of the AFIT Model Shop, who graciously provided fabrication and machine work under the tight time constraints of the Summer Faculty Program, is also gratefully acknowledged.



50 μsec delay

100 μsec delay

Figure 1 Spark induced disturbances in still air at 25 °C

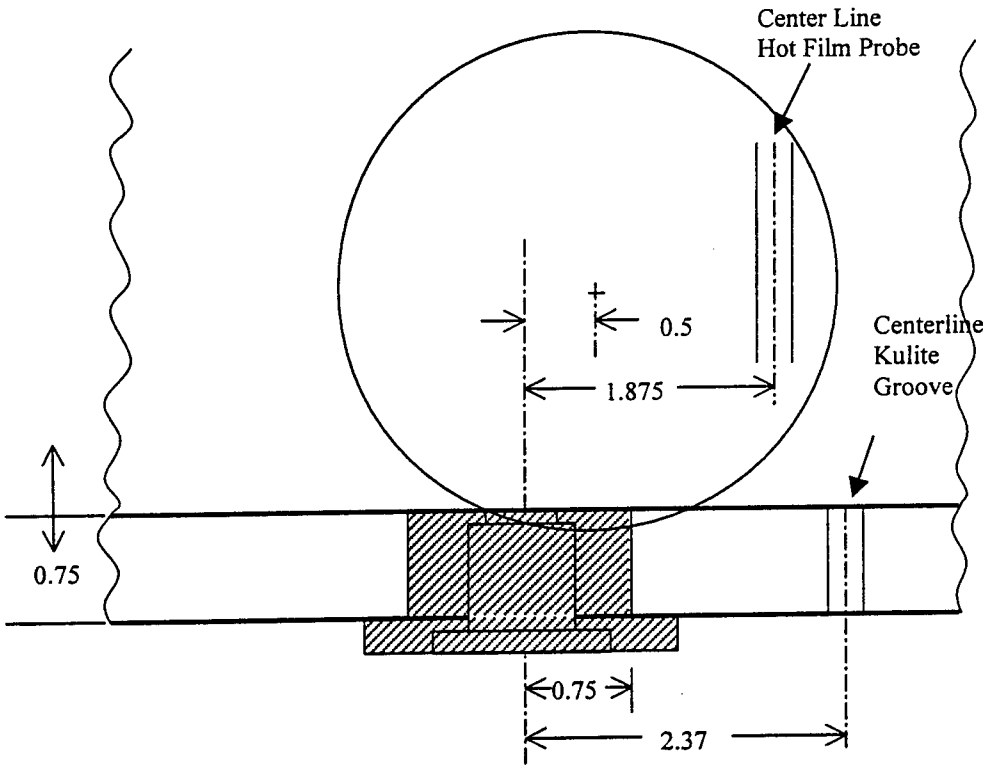


Figure 2 Schematic of Wind Tunnel Test Section Geometry

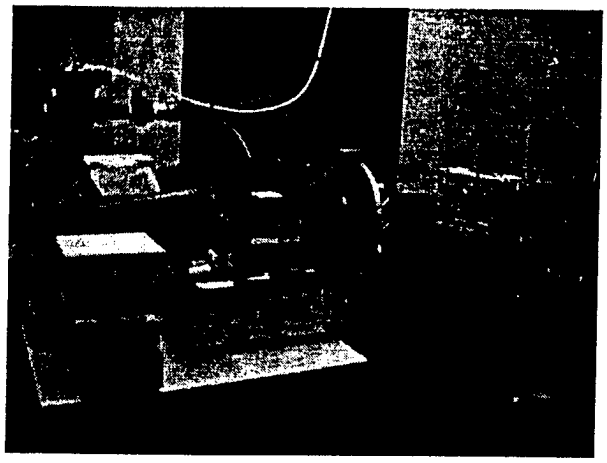


Figure 3 AFIT Mach 2.8 Wind Tunnel Apparatus

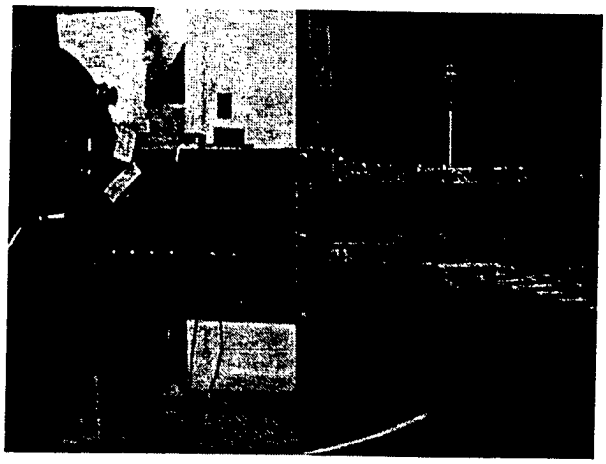


Figure 4 AFIT Mach 2.8 Wind Tunnel Test Section (1-side removed)

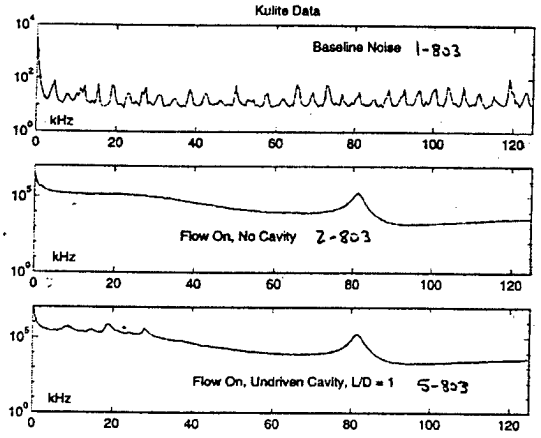


Figure 5 Simple cavity – Kulite data

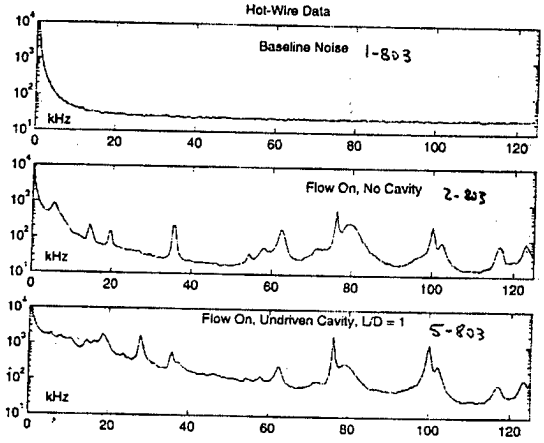


Figure 6 Simple cavity – Hot-film data

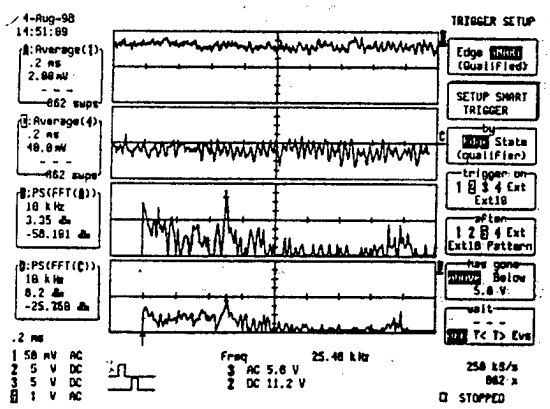


Figure 7 Excited cavity-25 kHz

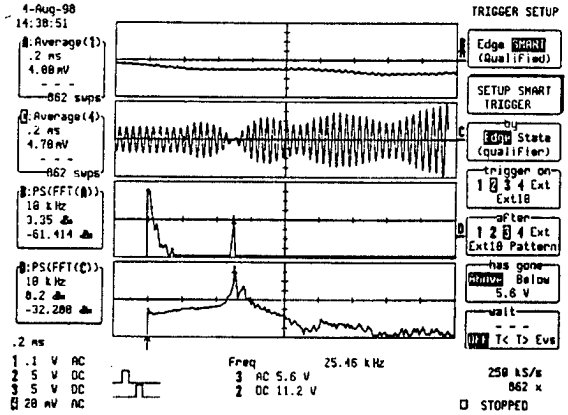


Figure 8 Flow off-25 kHz excitation on

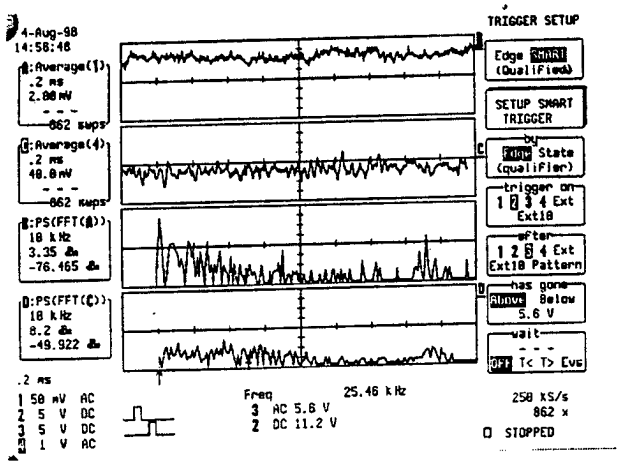


Figure 9 Flow on, excitation off

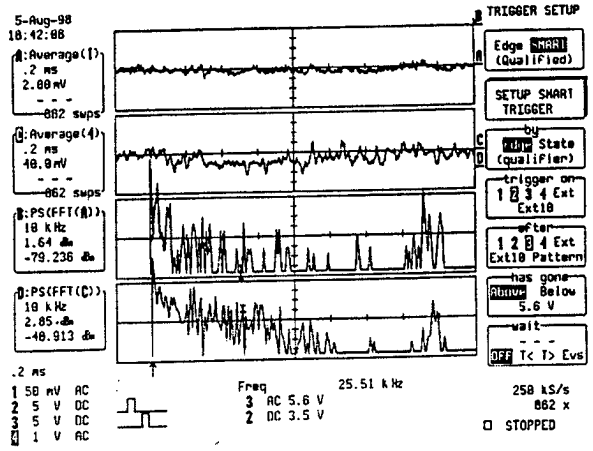


Figure 10 Excited cavity-25 kHz Probes moved up

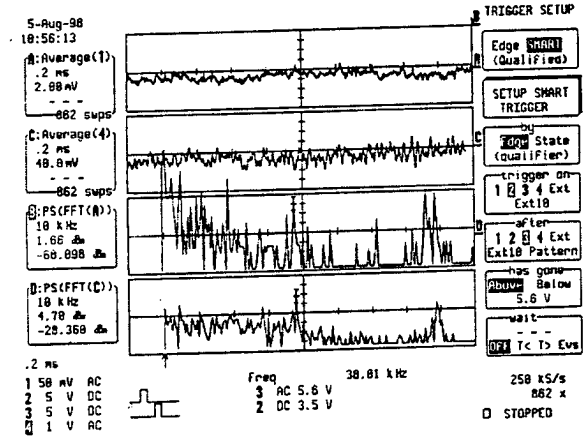


Figure 11 Excited cavity-40 kHz

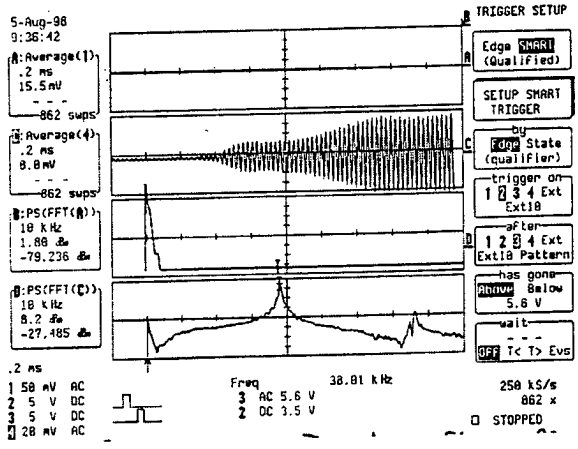


Figure 12 Flow off 40 kHz excitation

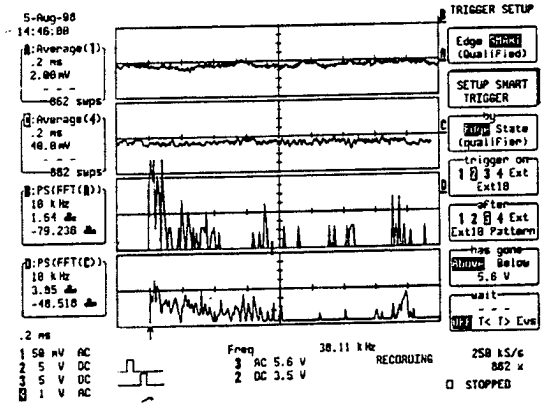


Figure 13 Excited cavity-40 kHz Probes moved up

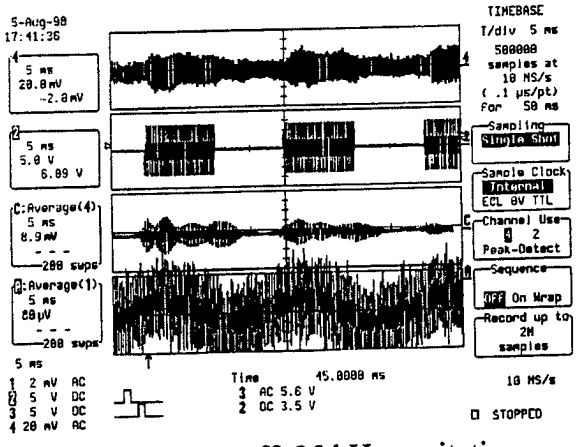


Figure 14 Flow off, 25 kHz excitation Unblocked

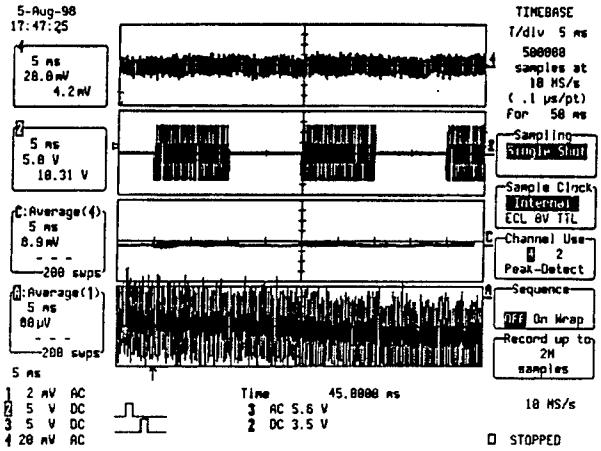


Figure 15 25kHz excitation – tape blocked

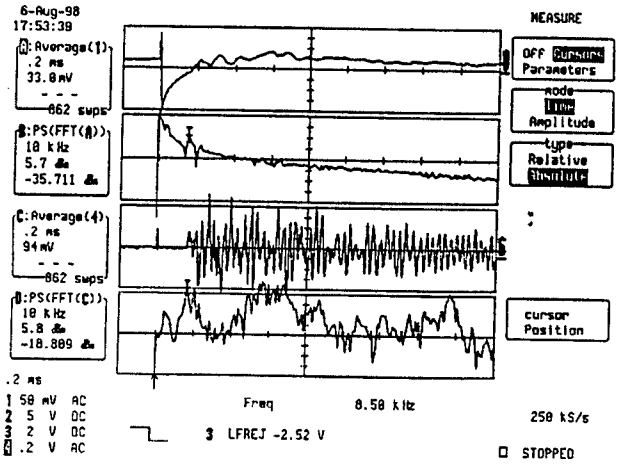


Figure 16 Flow off, spark excitation

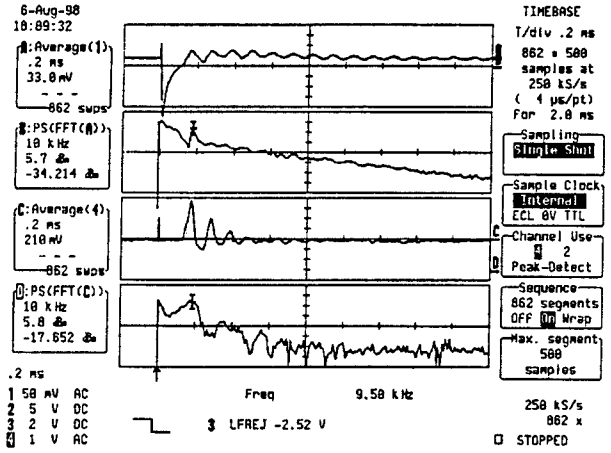


Figure 17 Flow on, spark excitation

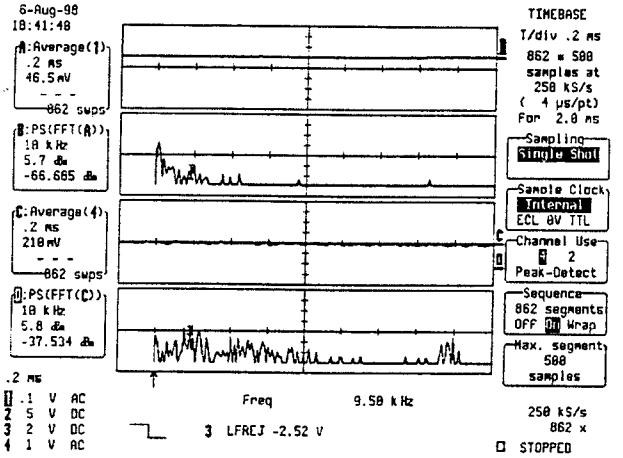


Figure 18 Flow on, spark off

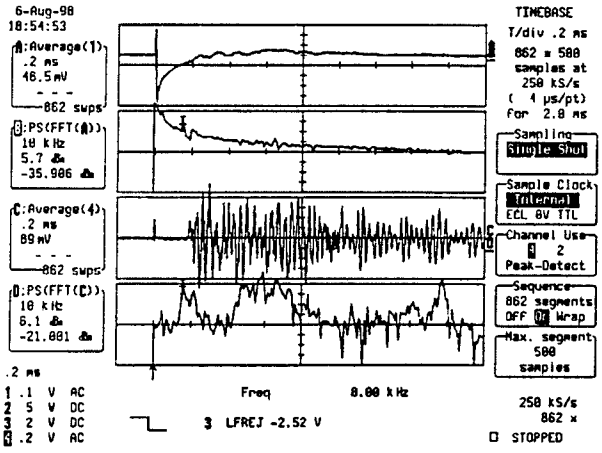


Figure 19 Flow off, spark excitation
Probes moved up

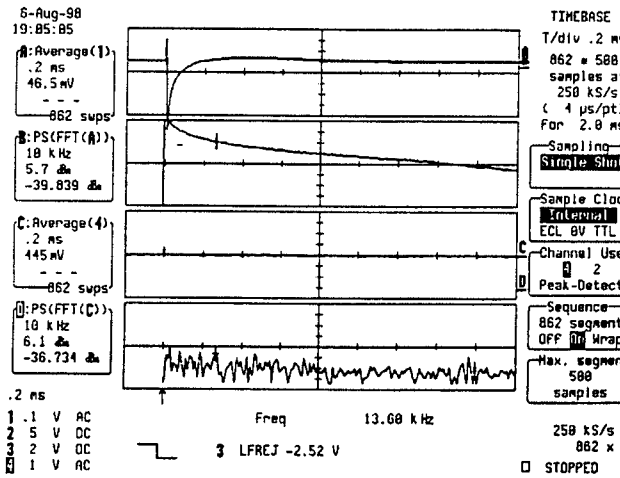


Figure 20 Flow on, spark excitation
Probes moved up

References

[1] Saric, W. S., Reed, H. L., and Kerschen, E. J., "Leading Edge Receptivity to Sound: Experiments, DNS and Theory," AIAA 94-2222, June, 1994.

[2] Choudhari, M. and Streett, C., "Boundary Layer Receptivity Phenomena in Three-Dimensional and High-Speed Boundary Layers," AIAA 90-5258, Oct. 1990.

[3] Phillips, O.M. , "On the generation of sound by turbulent supersonic shear layers", Journal of Fluid Mechanics, vol. 9, 1960, pp. 1-28.

[4] Laufer, J., "Some Statistical Properties of the Pressure Field Radiated by a Turbulent Boundary Layer," *The Physics of Fluids*, vol. 7, no. 8, August 1964, pp. 1191-1197.

[5] Mack, L. M., "Boundary-Layer Stability Theory," *Special Course on Stability and Transition of Laminar Flow*, edited by R. Michel, AGARD Report No. 709, pp. 3-1 to 3-81, 1984.

[6] Herbert, Th., Stuckert, G. K., and Lin, N., "Method for Transition Prediction in High-Speed Boundary Layers," Air Force Wright Laboratory Technical Report WL-TR-93-3097.

[7] Zhong, X., "Direct Numerical Simulation of Hypersonic Boundary-Layer Transition Over Blunt Leading Edges, Part I: A New Numerical Method and Validation," AIAA 97-0755, January 1997.

[8] Streett, C. L., "Direct Harmonic Linear Navier-Stokes Methods for Efficient Simulation of Wave Packets," AIAA 98-0784, January 1998.

[9] Maslov, A. A., and Semenov, N. V., "Excitation of Natural Oscillations in a Boundary Layer by an External Acoustic Field," translated from *Izvestiya Akademii Nauk SSSR, Mekhanika Zhidkosti I Gaza*, no. 3, pp. 74-78, May-June, 1986.

[10] Semionov, N. V., Kosinov, A. D., and Maslov, A. A., "Experimental Investigation of Supersonic Boundary-Layer Receptivity," *Transitional Boundary Layers in Aeronautics*, Proceedings of the colloquium "Transitional Boundary Layers in Aeronautics," Amsterdam, 6-8 December 1995, Henkes, R. A. W. M., and van Ingen, J. L., ed, Royal Netherlands Academy of Arts and Sciences, P.O. Box 19121, 1000 GC Amsterdam, the Netherlands, 1996, pp.413-420

[11] Maslov, A. A., Shiplyuk, A. N., Sidorenko, A. A., and Arnal, D., "Leading Edge Receptivity of the Hypersonic Boundary Layer on a Flat Plate," Preprint no. 1-98, Institute of Theoretical and Applied Mechanics, Russian Academy of Sciences Siberian Branch, Novosibirsk, 1998.

[12] Nishioka, M., and Morkovin, M. V., "Boundary-Layer Receptivity to Unsteady Pressure Gradients: Experiments and Overview," *Journal of Fluid Mechanics*, vol. 171, 1986.

[13] Saric, W. S., and White, E. B., "Influence of High-Amplitude Noise on Boundary-Layer Transition to Turbulence," AIAA 98-2645, June, 1998.

[14] Kimmel, R.L., AFOSR Progress Report Dec. 97 to Jan 98.

[15] Buck, G.A. and Takagi, S., "Observations of Traveling Disturbances from a Point Source in Rotating Disk Flow", AIAA 97-2299, June, 1997.

[16] Lemcke, B., "Double-Shock Shock Tube for Simulating Blast Loading in Supersonic Flow," *AIAA J.* vol. 1, no. 6, June 1963, pp. 1417-1418.

[17] Miller, H. R. "Shock-on-Shock Simulation and Hypervelocity Flow Measurements with Spark-Discharge Blast Waves," *AIAA J.* vol. 5, no. 9, September 1967, pp. 1675-1677.

[18] Candel, S. M., "Acoustic Conservation Principles and an Application to Plane and Modal Propagation in Nozzles and Diffusers," Journal of Sound and Vibration, vol. 41, no. 2, 1975, pp., 207-232.

[19] Salyer, T. R., Randall, L. A., Collicott, S. H., and Schneider, S. P., "Use of Laser Differential Interferometry to Study Receptivity on a Hemispherical Nose at Mach 4," AIAA 98-0238, January 1998.

[20] Schmisseur, J. D., "Receptivity of the Boundary Layer on a Mach-4 Elliptic Cone to Laser-Generated Localized Freestream Perturbations," Doctoral Dissertation, Purdue University School of Aeronautics and Astronautics, December 1997.

[21] Krishnmurty, K., "Sound Radiation From Surface Cutouts in High Speed Flow," Doctoral Dissertation, California Institute of Technology, Pasadena, California, 1956.
[22] Schneider, S., P., Purdue University School of Aeronautics and Astronautics, 1997.
[23] Sato, N., Imamura, R., Shiba, S., Takahashi, S., Tsue, M., and Kono, M., "Advanced Mixing Control in Supersonic Airstream with a Wall-Mounted Cavity, AIAA 96-4510-CP, 1996.
[24] Kimmel, R.L., Private Communication, July 1998.

SOME CRITICAL ISSUES OF THE NEXT GENERATION TRANSPARENCY PROGRAM

Joe G. Chow
Associate Professor
Department of Industrial and Systems Engineering

Florida International University
University Park Campus
Miami, FL 33199

Final Report for:
Summer Faculty Research Program
Wright Research Site

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

And

Wright Research Site

September 1998

SOME CRITICAL ISSUES OF THE NEXT GENERATION TRANSPARENCY PROGRAM

Joe G. Chow
Associate Professor
Department of Industrial and Systems Engineering
Florida International University

Abstract

Low-pressure injection molding of polycarbonate is a new promising technology for manufacture of fighter aircraft transparencies. Compared to the current method, this approach has the benefits of reducing the manufacturing cycle time, eliminating the metal frame assembly, and forming geometrically complex shapes. A program, called Next Generation Transparency (NGT), has been initiated to reduce risk for the development of an injection molded frameless transparency for the Jointer Strike Fighter. The focus of this summer research is to investigate two critical issues of the NGT program: through-canopy escape by explosive severance and fabrication of optical quality NGTs.

Current escape clearance techniques are either canopy fracturing using explosive cords or rocket jettison of canopies. For the NGT program, the former approach is preferred since it is faster, lighter, and more efficient. However, the injection-molded polycarbonate, chosen to resist bird strike, is very fracture resistant. Little is known concerning how to design and configure an explosive severance system for this type of material. Explosive severance tests performed up to this point have achieved some success in fracturing ploycarbonate, but significant testing will be required to develop a reliable through-canopy escape system. This summer, the principal investigator studied the previous test results thoroughly and, based on his understanding, he proposed to perform 3-D modeling and analysis for this problem at his institution to complement future experiments that will be conducted.

Accurately machined and polished mold surfaces are required to produce optically acceptable transparencies. The Confirmation Frameless Transparencies (CFTs) produced during the previous Frameless Transparency Program exhibited waviness in certain regions of the optical area. This surface waviness, caused by excessive manufacturing errors in the mold, will create unacceptable optical distortion for the pilot. To better understand the cause of this problem, the principal investigator visited Delta Tooling Inc., the contractor for fabricating CFT and NGT molds. During his visits, he studied the mold fabrication technology and compared the technology that was employed to fabricate the CFT tool to the new high speed machining technology that will be used in manufacturing the NGT tool. To conduct an extensive study on this issue, the principal investigator submitted a proposal to Manufacturing Technology Directorate (AFRL/ML) in July 1998 to seek funding to perform this research at his university.

SOME CRITICAL ISSUES OF THE NEXT GENERATION TRANSPARENCY PROGRAM

Joe G. Chow

1. Introduction

The AFRL/VACE at Wright Patterson Air Force Base is responsible for the research and development of advanced aircraft transparency systems. One of their most recent efforts is to develop next generation transparency (NGT) for an advanced strike fighter. The purpose of the NGT program is to demonstrate and validate the ability of injection-molded frameless transparency technology to be incorporated in an affordable and integrated design that meets the mission requirements for the Jointer Strike Fighter.

The NGT is based on the Frameless Direct-Form Transparency Program (FTP) that was initiated by the US Air Force in 1985. The direct-form process is derived from the injection molding process. During molding, the polycarbonate material is injected into the mold using a low pressure and long cycle process. The edge of the mold cavity is thickened to facilitate direct attachment of the canopy to the aircraft frame. Inserts for latching and hinges are placed within the mold and the material injected around them. This process was validated during several contracted efforts since the start of the FTP.

Compared to the current bent-from-sheet fabrication process, the advantage of the direct form process is low cost, potential increase in life cycle, improved material properties, thickness tailoring, design freedom, and forming process repeatability. In addition, it is possible to make a frameless transparency system which would eliminate a large number of parts, limit thermally induced stresses, and reduce replacement time and manpower required.

Since the NGT program is based on a new technology that is significantly different from the current method, it has many important unresolved issues. During his summer research, the principal investigator has concentrated on the following two issues:

- How to explosively sever a thick polycarbonate transparency ?
- How to fabricate a large-scale optical quality NGT ?

The description of each problem and the principal investigator's research efforts on these problems are provided in the following sections.

2. Through-Canopy Escape by Explosive Severance

The Joint Striker Fighter will be designed to be used by several customers: US Air Force, US Navy, US Marine Corps, British Air Force and British Royal Navy. From canopy's point of view, each customer has different requirements and needs. For the Air Force, bird strike resistance is an important requirement while, for the US Navy and Marine, fast emergency escape (less than 0.3" second) is a primary concern. To resist bird strike's impact, stronger and tougher material, such as 0.75" thick polycarbonate, should be used. For faster escape in an emergency situation, through-canopy by explosive severance (instead of jettisoning the entire canopy assembly) should be adopted.

The goal of this project is, therefore, to develop a reliable explosive severance system that can fracture a polycarbonate canopy completely at various operating temperatures so that the pilot can escape through it.

During the past few years, the following work has been conducted to investigate the feasibility of using explosive to sever polycarbonate aircraft transparencies:

- Augmented Shock Wave Severance at NASA Langley
- Optimized NGT Solution at Ensign Bickford Company
- NGT Explosive Severance Test at Navy's Indian Head Division

A brief review of their work and summary of their results are presented as follows.

2.1 Augmented Shock Wave at NASA Langley (1992-1995)

This work, conducted primarily by Bement at NASA Langley, utilized a new concept (augmented shock wave) to sever or weaken a polycarbonate transparency [1,2]. This technique employs embedding explosive cords into parallel grooves that are cut into a surface of the polycarbonate. The cords are initiated simultaneously to produce shock waves that progress toward the centerline between the cords and the lower surface of the material. Intersecting incident and reflected waves augment at the centerline to fail or weaken the material in tension.

Some of the key parameters they studied are:

- Setup (Imbedded explosive vs. Using polyurethane as buffer)
- Number of Cords (Single vs. Dual)
- Sample Size (6" x 6", 18" x 18", F-16 panels)
- Temperature (-65, 70 and 165 °F)
- Sample Thickness (0.45" and 0.89")
- Distance Between the Cords

The primary findings from Bement's studies are:

- Buffered setup works better than the imbedded setup.

  The buffer material, polyurethane layer, enhances the coupling of explosive pressure impulses into the polycarbonate. This layer is so effective that it remains intact after the explosive, and no cratering is induced into the polycarbonate. As a result, this setup can fracture thicker material than simply embedding explosive in the polycarbonate material.

- Double cords are more effective than single cord.

  To fracture the same polycarbonate panel, more explosive is required for single cord than double cords. Also, double cords produce no inboard debris while single cord may.

- Distance between two cords is critical.

  If the distance between two cords is too small, spalling is likely to occur. If the distance is too large, the panel may not be severed effectively.

Some of the results he reported, however, are contradictory to common sense. For example, data in Reference [1] showed that at high temperature (165 °F), 0.895" thick polycarbonate coupon could be fractured without spalling while the same material couldn't be severed at a much lower temperature (-65 °F). He attributed this unexplainable result to the fact that the history of the polycarbonate materials he used was unknown. For this reason, this result could not be applied to the NGT.

2.2 Tests Performed by Ensign Bickford Corporation

This work began from 1997 and temporally suspended in April 1998. The engineers at Ensign Bickford started with investigating different explosive methods and then decided to concentrate on augmented shock wave technique. To implement this principle in their testing, they used the buffered setup, as suggested by Bement' results. Their test setup is shown in Figure 1.

Similar to Bement's work, they also conducted an extensive study on different parameters. Some of the main parameters they investigated are:

- Number of Cord (Single vs. Double)
- Cord Layout (Straight line vs. Corner)
- Type of Explosives (MDF and PBXN)
- Coupon Thickness (0.50" and 0.75")
- Temperature (-65, 70 and 200 °F)
- Type of Polycarbonate (Extruded, Injection Molded, Commercial Grade, Military Grade)

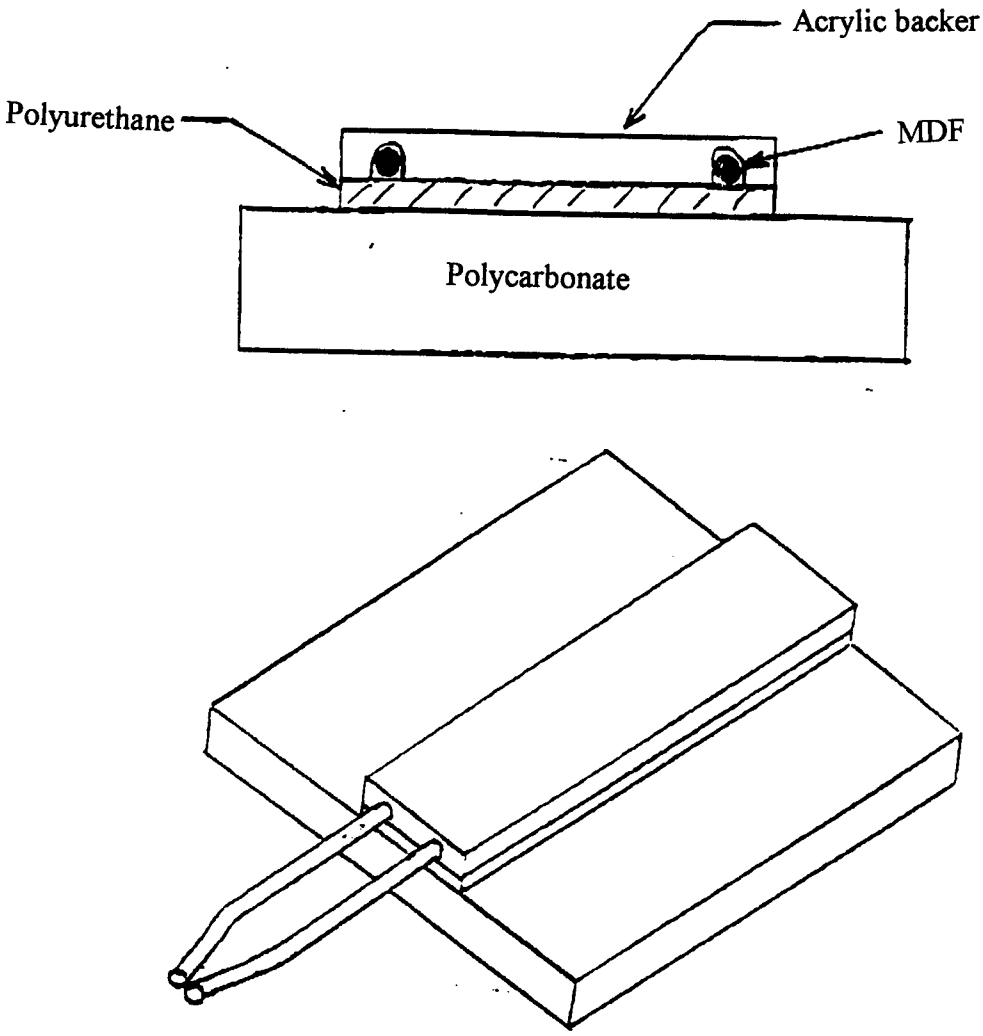


Figure 1.Explosive Severance Setup at Ensign Bickford Corporation

The conclusions for Ensign Bickford's work are:

- At room temperature, polycarbonate coupon can be fractured.
- At 160 degree F, polycarbonate coupon may fracture.
- At 200 degree F, polycarbonate coupon is tough to fracture.

During the most recent testing they conducted, they also took some high speed films. The goal of taking high speed films on the explosive severance test was twofold: (1) capture explosive detonation and resulting damage of the polycarbonate and (2) achieve a better understanding of crack initiation and propagation. A sample of their high speed film testing results is shown in Figure 2.

2.3 Navy's Work at Indian Head Division

Recently, the Cartridge Actuated Device (CAD) Research and Development at Indian Head Division conducted a test to determine the suitability of utilizing current US Navy canopy explosive concepts to weaken the NGT material during an emergency escape situation [3]. The two concepts selected are:

- Mild Detonating Cord (MDC, 5.5 grains/ft, single cord)
- Flexible Linear Shape Charge (FLSC, lead sheathed HNS-II loaded to 13.5 grains/ft)

The test setups for these two configurations are shown in Figure 3.

A windscreen comprised of CFT polycarbonate was first sectioned into 3" x 6" coupons. The thickness of the coupon was reduced from 0.75" to approximately 0.375". The thickness of the modified coupons was measured at each end of the coupons tested, since it is usually not uniform. Once this had been completed, the explosive lines were installed. The coupons, with explosive lines, were placed in groups of five. There were a total of six groups, three for the MDC configuration at three different temperatures (-65, 70, and 225 °F) and another three for the

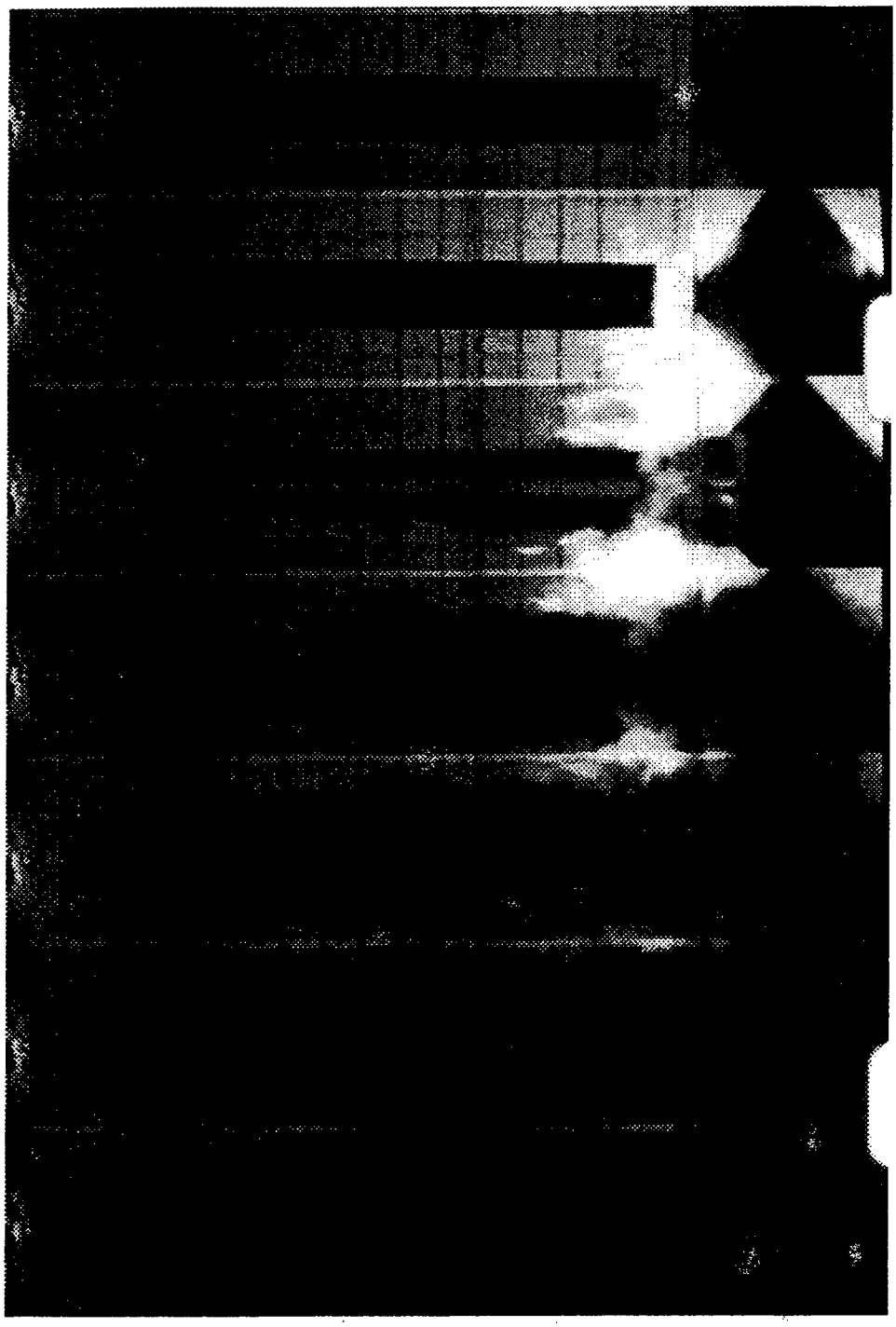
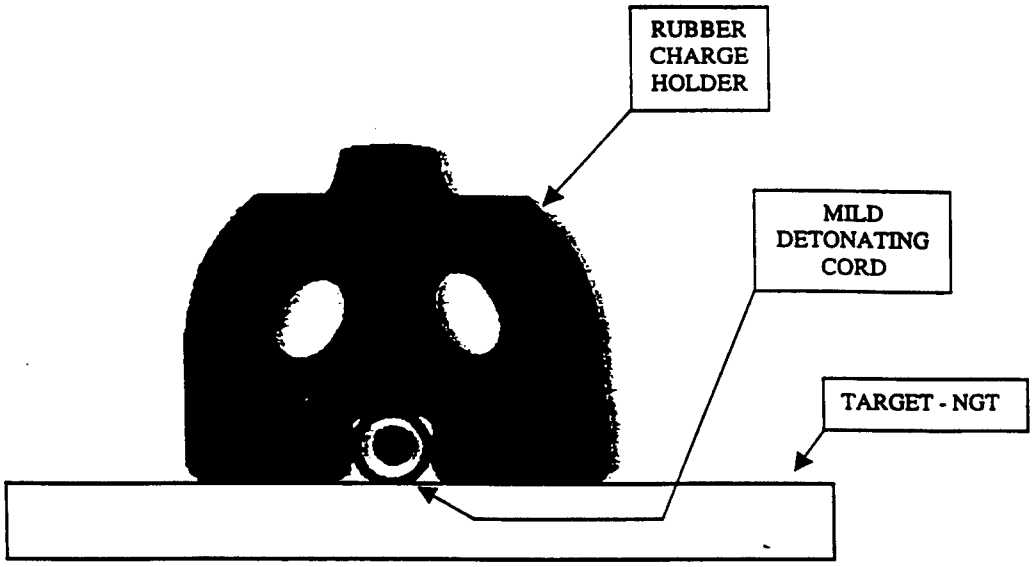
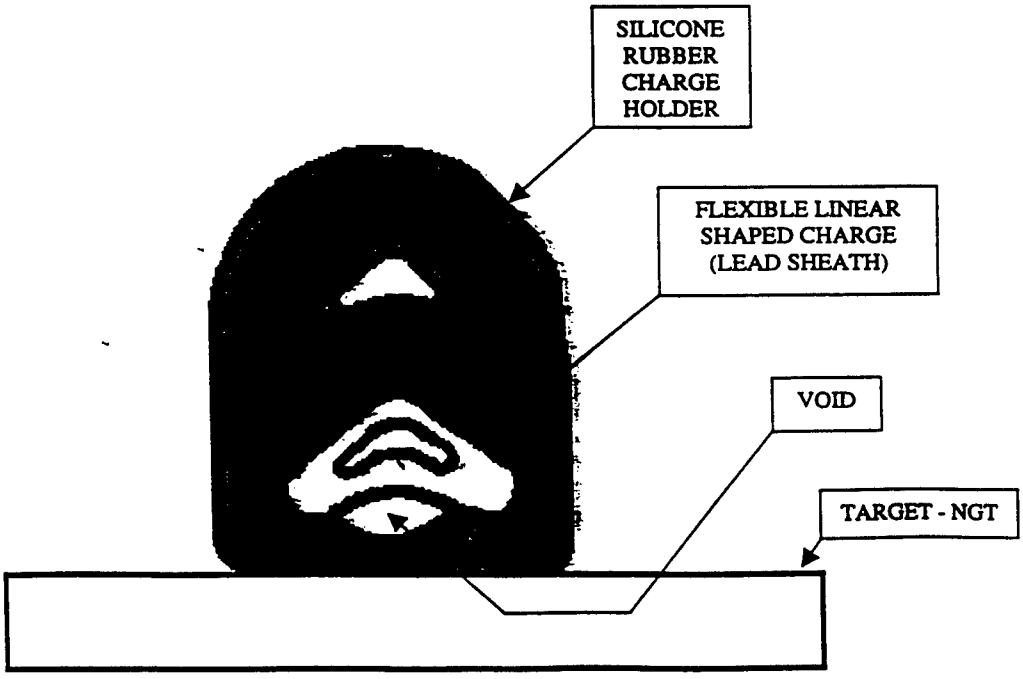


Figure 2. Sample of Ensign Bickford's High Speed Film Results



MF16 KIT DETONATING CORD ASSEMBLY
Currently Used on the S-3 Series Aircraft



MW68/69 FLEXIBLE LINEAR SHAPED CHARGE ASSEMBLY
Currently Used on the S-3 Series Aircraft

Figure 3. Test Setups for Navy's Explosive Severance Configurations

FLSC configuration tested at the same temperatures.

When the explosive tests had been completed, each coupon was sectioned into 3-sub coupons. The depth of penetration of the shock wave/FLSC jet was measured at each end of the smaller sub-coupons. These sub-coupons were then shear strength tested along the area where the explosive lines were installed to determine a weakened state for each. Their results demonstrated current US Navy explosive severance concepts are viable for the NGT materials when the thickness is limited to 0.500" or less. Encouraged by the success of this preliminary test, they planned to conduct another round of testing with larger amount of explosive in the near future.

2.4 The Principal Investigator's Research Plan

Due to the complicated nature of this problem, the principal investigator plans to conduct this project in two phases. The emphasis of each phase for the explosive severance project is:

1. Acquire an in-depth understanding of the problem during Summer Research Program.

The principal investigator had studied the work performed by various organizations and discussed with the personnel who has been involved in this project.

2. Develop a 3-D predictive model through Summer Research Extension Program funding ($25,000).

After reviewing the work that has been performed, the following conclusions can be drawn:

- No reliable explosive severance system has been identified for the NGT. Thus, significant testing will still be required in the future.

- The experiments conducted so far have been very inefficient since they were done on a trial and error basis.

To complement future experiments, the investigator principal intends to perform 3-D analysis to provide some insights and directions for the testing efforts. The principal investigator will seek funds from the follow-on Summer Research Extension Program (SREP) to support this effort.

To perform 3-D modeling and analysis for the explosive severance project, the commercial software DYNA3D software will be used. DYNA3D is a general purpose, explicit three dimensional finite element application used to analyze the large deformation response of inelastic solids and structures. Because of its general purpose design and its ease of use, DYNA3D has become a standard in a number of application areas and industries including automotive, aerospace, metal manufacturing, plastics and defense.

Applications for DYNA3D simulations and analysis include: airplane crash, armor penetration, automotive crash, glass forming, metal forming, sheet metal stamping, and train crash.

To enhance the success of this modeling effort, Dr. Cesar Levy, a faculty member in the Department of Mechanical Engineering at Florida International University, has been invited and agreed to collaborate on this research endeavor.

3. Fabrication of Optical-Quality NGTs

Injection molding is typically used to fabricate small parts in large quantities. High optical quality injection molded parts in particular are very small in size (eyeglasses). Therefore, the aircraft transparency is significantly larger than typical injection molded parts and it has stringent optical requirements. The CFTs produced during the previous FTP exhibited surface waviness in the optical area, which may result in unacceptable optical distortion. The cause of this problem can be attributed to geometric inaccuracy of the CFT tool.

To identify a potential solution of this problem, the principal investigator visited Delta Tooling in Auburn Hills MI, the primary contractor for the CFT and NGT mold fabrication, to observe the process of fabricating an injection mold. While he was there, he made a special effort to compare the technology that was used to fabricate the CFT tool and the technology that will be used to produce the NGT tool. What follows are his findings from his visits to Delta Tooling.

3.1 Major Steps for Fabricating An Injection Mold

After Delta Tooling receives the part geometry from their customer, they will go through the following steps to fabricate an injection mold:

- Surfacing : generate mold surfaces and 2-D drawings (hydraulics, etc.)
- Machining : create core and cavity surfaces

  - Rough-Mill
  - Semifinish-Mill
  - Finish-Mill

- Benching (hand grinding) : remove cutter marks
- Spotting : press two mold halves together to ensure critical surfaces match well
- Polishing : significantly improve surface finish of the mold surfaces
- Inspection : check the mold dimensions using a coordinate measuring machine

In this manufacturing process, machining and benching are the two most critical steps in determining the geometric accuracy of the tool. For machining, tool wear is a main concern since the mold material, P20 steel, is fairly hard. Excessive tool wear tends to degrade the geometric accuracy and surface finish of the mold surfaces. For a large mold, several days are usually required to complete finishing-milling operations. Therefore, to ensure good surface finish and geometric accuracy after finishing operations, frequent tool change is required. Benching error is also a source of geometric errors, since hand grinding is not as controlled as a programmed toolpath on a machine tool.

3.2 Major Improvements for Mold Fabrication

Since the fabrication of the CFT tool, significant improvements have been made in the mold fabrication technology. The most significant improvement is the availability of high-speed die and mold machining. The main benefit of machining an injection mold at higher speeds is to improve the mold's surface finish and geometric accuracy so that benching can be reduced or eliminated. To support high speed machining, several key technological improvements have been made. These improvements are discussed as follows.

1. The Machine

The new type of machine tool usually has high structural rigidity with a high level of damping. They also have a better design. As an example, the X and Y movements for the newer machine are performed by the spindle, instead of by the machine table. Th advantage of moving the spindle, instead of the table and the work piece, is that the spindle weight is fixed and it is usually much lighter than the work piece.

2. High-Speed Spindles

The spindle for high speed machine is precise, rigid, and can run reliably at 8,000 to 30,000 rpm for longer periods with minimal growth from heat. In addition, the spindle rotation error is usually a fraction of conventional machines to minimize tool eccentricity.

3. The Control

The new machine control is capable of maintaining an accurate cutter path during fast moves. Since the size of the part program is huge because step over is so small, the control typically uses a 32-bit microprocessor, with high-speed interpolation and block processing and substantial look-ahead capability. The long look-ahead capability that comes with high speed machine tools would allow the controller to anticipate path shifts and avoid accidents at high speeds.

4. Cutting Tools

New types of cutting tools have far better heat and wear resistance capabilities than conventional carbides. A superior cutting tool for P20 steel is TICN coated carbides. With their better wear resistance than conventional uncoated and TIN coated carbides, the tool life of newer inserts can be significantly extended. Other advanced cutting tool materials, such as Cubic Boron Nitride (CBN), have also emerged. According to Delta' in-house tool life evaluation results, CBN can last much longer than the best coated carbides in machining P20 steels (a ratio of 10 to 1).

3.3 Fabrication Technology for the CFT and NGT Tools

To ensure the highest quality for the NGT, Delta Tooling will apply the state-of-the-art technologies to the NGT mold fabrication. The following table shows the difference between the equipment and procedure used for the CFT and NGT tools.

Table 1. Comparison Between the Fabrication Technology for CFT Tool and NGT Tool

	CFT Tool	NGT Tool
Machine Tool	Droop and Rein	Mecof
Maximum RPM	3,000	12,000
Maximum Feed Rate (ipt)	120	600
Stepover (in)	0.030	0.008
Cutting Tool	TIN coated carbide	TICN coated or CBN
XY Movement	By table	By spindle
Benching	Six sizes of stones	One or two sizes
Polishing	Three grades	Three grades

By using the most advanced technologies, Delta Tooling expects the geometric accuracy of the NGT tool (on the order of 0.001") will be significantly better than that of the CFT tool (on the order of 0.004").

3.4 The Principal Investigator's Research Plan

High speed machining has emerged as a promising solution for producing highly accurate injection molds and optical quality NGTs. One of the key components in implementing this new technology is the use of CBN cutting tool since conventional carbides will have a very short tool life due to accelerated wear at high speeds. However, commercial CBN milling inserts have just become available recently and they are not in wide use in molding industry yet. Therefore, to better understand their effects on the optical quality of injection molded parts, extensive study is required. The principal investigator recently submitted a proposal to AFRL/ML to request funds to investigate the effect of high speed machining using CBN tools on the manufacture of NGTs.

4. References

1. Bement, L., Schroeder J., and Schimmel, M., "Explosive Fracturing of Aircraft Canopy Materials", Presented at the 1992 SAFE Symposium, Las Vegas, Nevada, November 2-3, 1992.

2. Bement, L. and Schimmel M., "Augmented Shock Wave Severance of Materials", Presented at the 31st AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, San Diego, California, July 10-12, 1995.

3. Blachowski, T., "Next Generation Transparency Baseline Fragilization Test Results", Technical Report, Naval Sea Systems Command, 1998

Documentation of the Airflow Patterns within an Aircraft Engine Nacelle Simulator

Peter J. Disimile
Associate Professor
Department of Aerospace Engineering

The University of Cincinnati
Cincinnati, Ohio 45221-0070

Final Report for:
Summer Faculty Research Program
Wright-Patterson Research Site

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright-Patterson Research Site

September 1998

Documentation of the Airflow Patterns within an Aircraft Engine Nacelle Simulator

Peter J. Disimile
Associate Professor
Department of Aerospace Engineering
The University of Cincinnati
Cincinnati, Ohio 45221-0070

Abstract

Airflow patterns within the F-22 Aircraft Engine Nacelle Fire Test Simulator located at Wright-Patterson Air Force base were studied. By passively tagging the nacelle airflow with smoke particles and illuminating the flow field with either laser or halogen light, the flow patterns within the nacelle were recorded. The experimental results indicate that the inlet screen produced multiple adjacent jet flows that strongly interacted with one another, resulting in the gross instability of the inlet flow. The nature of the flow pattern within this first nacelle compartment effectively acts as a cross flow, thereby forcing each of the multiple jet structures to take on a kidney shape. It was further observed, that within this kidney structure two counter-rotating streamwise vortices occurred.

Introduction

The use of fire suppression systems in modern engine nacelles is of paramount importance to the survivability of the engine and crew. The fire suppressant currently used is Halon 1301. However, in light of the Montreal protocol, production of this fire suppression agent is now banned. Although numerous Halon replacement candidates have been proposed, many of them lack an equivalent chemical activity and as a result, their efficiency as a fire suppressant has been reduced. To compensate for this reduction in agent effectiveness a more robust means of agent distribution and mixing is desired without the use of additional hardware.

One of the key aspects of fire suppression within an engine is concerned with the complex flow pattern that occurs between the engine nacelle and the engine core. In this case the primary sections include the compressor, combustor, and turbine assemblies. It is in these regions that once a fire has occurred the immediate need for fire suppression is of great importance. To evaluate the ability of a fire suppression system to extinguish engine nacelle fires, which may occur as a result of component failure, accidental or otherwise, a unique facility was constructed, under the Halon replacement program, for the Air Force Research Laboratory at the Wright-Patterson Air Force base. This facility, known as the Aircraft Engine Nacelle Fire Test Simulator, referred to as the AENFTS, is used to test various fire extinguishment systems and agents for their ability to suppress fire. In this case such fire tests can be performed within an engine nacelle under simulated operating conditions. A photograph of the simulated engine core, fabricated for utilization in the current nacelle simulator test is presented on the next page in Figure 1, and shows the rear and side regions of the test core.

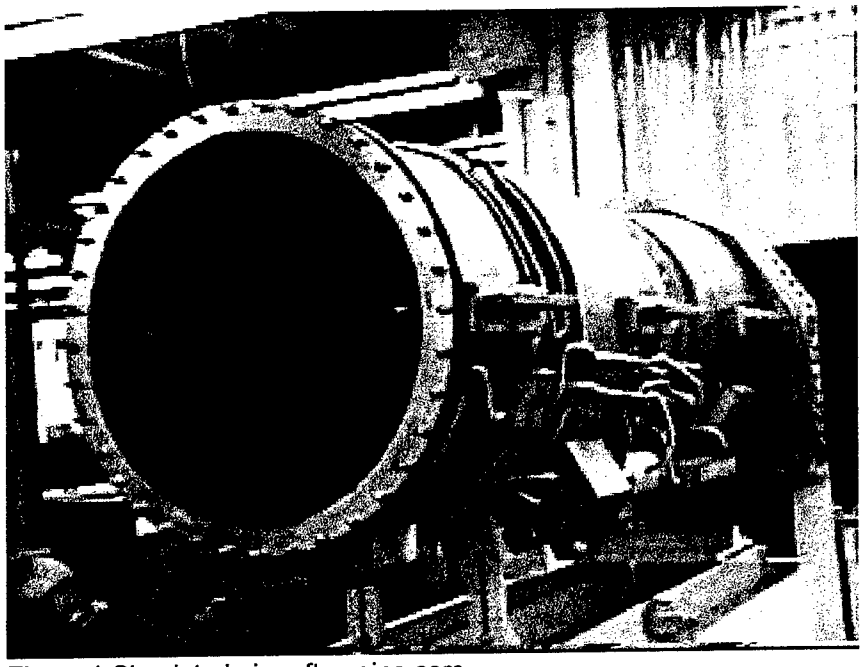


Figure 1 Simulated aircraft engine core

In addition to the major engine components considered, there are numerous bluff objects that are found within an engine nacelle. Without doubt, each aircraft configuration has a unique nacelle design containing the engine core and varying amounts of clutter in the form of tubes, boxes and enclosures. A close-up image of the simulated core is shown in Figure 2, emphasizing the major sources of clutter contained within an actual F-22 nacelle. On close inspection of the lower half of this image a region containing several boxes and enclosures is observed. Furthermore, a cylindrical container centrally located within this region depicts the area known as Fire Zone 1.

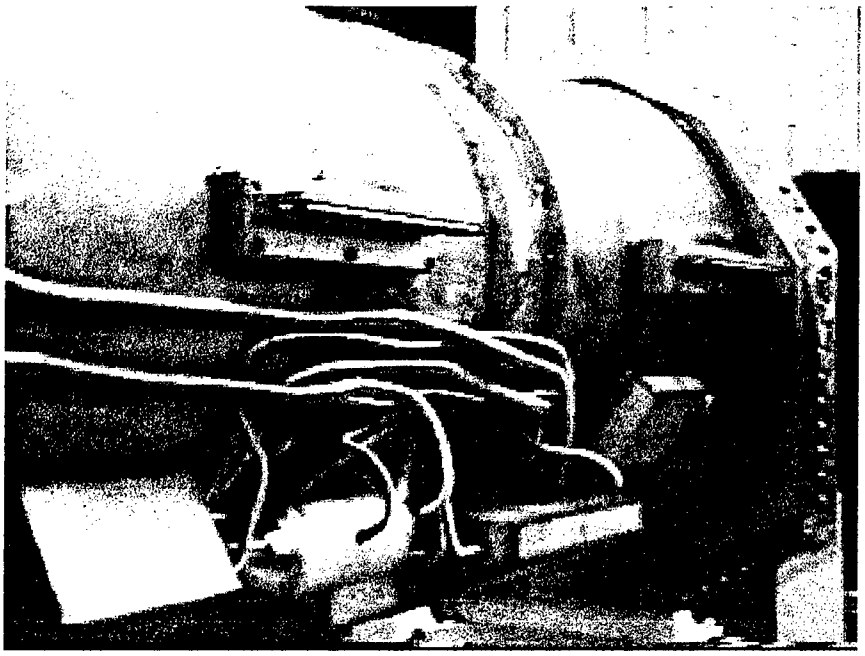


Figure 2 Close up of the clutter components

In an actual operating engine some of the tubes represented by the clutter would carry jet engine fuel or hydraulic fluid. Therefore, these tubes may serve as a fuel source for a potential fire, while other tubes may enclose electrical wiring that can act as a potential ignition source. Once a failure occurs, tubes carrying volatile or combustible fluids may leak and be ignited by an ignition source. The leaking fluid may exit the tube in the form of a spray jet or a droplet forming a pool of fuel. For comparison purposes the spray jet scenario was selected and will be discussed at a later point in the text.

Although engine nacelle fires may be caused by either fuel being sprayed and ignited or from pool-fires a turbulent diffusion flame is produced. In either type of fire, flame stability is enhanced through the formation of a recirculation zone behind any of the multitude of obstructions found to exist in the nacelles airflow path. The obstacles present in the airflow path, referred to earlier as bluff bodies, act as flame holders, thereby stabilizing and protecting a potential flame from being attenuated by the main nacelle airflow or suppressed by the fire extinguishing system. The importance of agent entrainment into the recirculation (or combustion) zone of an obstacle stabilized flame is paramount and determines the amount of agent required for fire extinguishment, as well as the time necessary for the safe attenuation of a fire.

In order to have effective fire extinguishment a sufficient amount of suppression agent is required to be transported into the fire zone. However, because of aircraft weight considerations, the minimum amount of agent necessary to achieve fire suppression is always desired. Unfortunately, agent distribution is confined to only a few locations, thereby making the dynamics of the agent discharge jet and nacelle air flow, crucial to flame stability and fire extinguishment. This is further complicated by the effect of engine core surface temperature, surrounding air temperature, and air pressure within the nacelle.

Therefore, to minimize damage sustained to an aircraft, after a fire has been detected, both a sufficient quantity and an effective distribution of a fire extinguishing agent is required. To accomplish this goal an understanding of both the nacelle airflow and agent discharge is necessary. To this end an experimental program was undertaken to document the air flow pattern within the F-22 engine nacelle simulator under the condition of low nacelle airflow.

Strategy

To gain insight into the flow pattern within the F-22 engine nacelle, the AENFTS located at the Wright-Patterson site was selected for evaluation. Specifically, the airflow patterns within the nacelle simulator were investigated and subsequently documented. This was accomplished by using direct injection flow visualization. In this technique the generation of small smoke particles, typically between one and 10 microns was achieved using a commercial vaporizing unit. These small particles were essential to ensure that the trajectory of the particles would follow the complex internal nacelle flow field. To enable visualization these smoke particles were illuminated by either a halogen lamp or a laser light sheet. Light scattered from these tiny particles enabled the trajectory of the smoke particles to be made visible. The following photograph (Figure 3) depicts the nacelle simulator (AENFTS) highlighting five critical regions. The first two regions depict the locations of the Top Front and Rear Viewing Ports, with the third region located at the Fire Zone. The forth and fifth locations are at the air inlet duct and the Agent Discharge Port. It should be noted that the inlet airflow enters the nacelle in an asymmetric manner, through an inlet screen close to the agent discharge port.

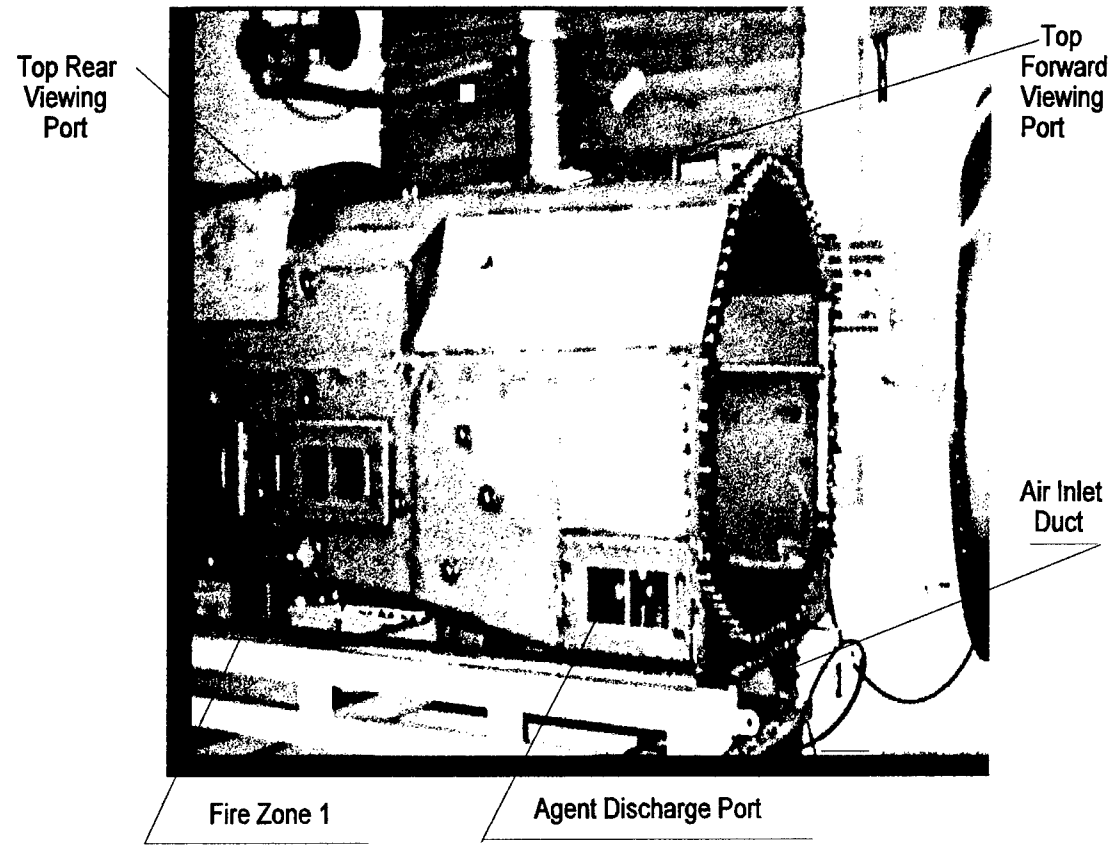


Figure 3 View of the nacelle with critical locations labeled

As previously mentioned, the smoke was produced using a commercial smoke generator. This was made possible by the vaporization of a liquid whose formulation was similar to that of mineral oil. The vaporized fluid was then introduced into the main nacelle airflow via two methods:

I) direct injection into the nacelle through the flooding of inlet air duct, just upstream of the inlet screen, and

ii) the direct injection of the smoke through eight 0.25 inch plastic tubes inserted and fixed to the nacelle inlet screen.

Since the inlet screen passages were nominally oriented at 45 degrees, each tube was inserted in the screen and aligned accordingly, making sure the end of the tubes were approximately flush with the screen's outer surface. The tubes were arranged such that their locations represented a "V dot". That is, one smoke tube was positioned at the vertex and three tubes on each side of the inlet screen forming a V pattern. The eighth and last tube which completed the "V dot" was placed parallel to the centerline of the simulated engine core. In other words, the tube was placed streamwise along the nacelle, above the V and in line with the vertex (see Figure 4).

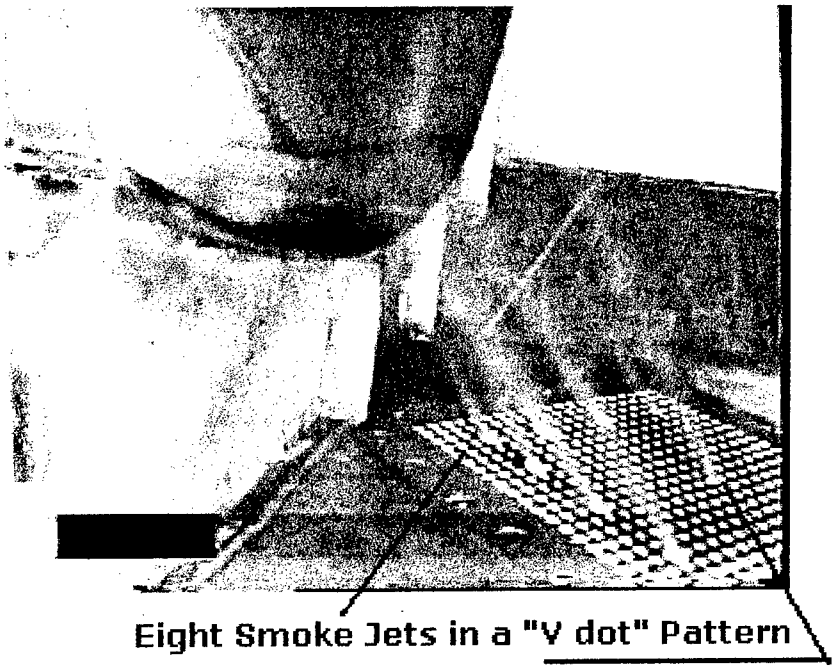


Figure 4 Image emphasizing the positions of smoke tubes

Although the major portion of the flow visualization tests were performed with the tubes arranged in a "V dot" pattern, several early tests were conducted with the smoke tubes placed in two parallel lines, approximately side by side and this will discussed later in the text.

The passively tagged nacelle airflow was illuminated in the areas of interest with either laser or halogen light. When a macroscopic or three-dimensional view of the flow structure was desired, a halogen source was used for illumination. If detailed or two-dimensional features of a complex flow regime was

desired a laser light was used for the illumination source. Due to the coherent nature of a laser, the laser beam could be expanded to form a relatively thin laser light sheet. In the present work laser beams were expanded with the appropriate optics to form a laser sheet approximately 1.5 mm in thickness. Both an Argon-Ion gas laser (producing a blue light) and a Helium-Neon laser (producing a red light) were utilized. Scattered light was then collected using miniature CCD color cameras and recorded on VHS video tape for post processing.

Results

As smoke was injected into the nacelle air inlet duct and forced through the inlet screen a complex swirling motion was produced. This fully three-dimensional motion resembles a classical horseshoe vortex, similar to the flow pattern that occurs in a snow storm when blowing snow moves past a tree. The following image (Figure 5) shows the initiation of this three-dimensional swirling motion, whose rotation is in the clockwise direction.

Unfortunately, this image is presented in a two-dimensional format and therefore, the three-dimensional nature of this motion may be difficult to appreciate. However, this highly three-dimensional flow was easily realized in real time, during the performance of the experiment. As previously noted this motion may be described as a horseshoe vortex, whose two tails extend downstream into the nacelle between the first bulkhead and the engine core/clutter (visible in the left hand side of the image). Complicating this motion further is the fact that the individual jets that make the swirling motion contain counter rotating pairs of streamwise vortices.

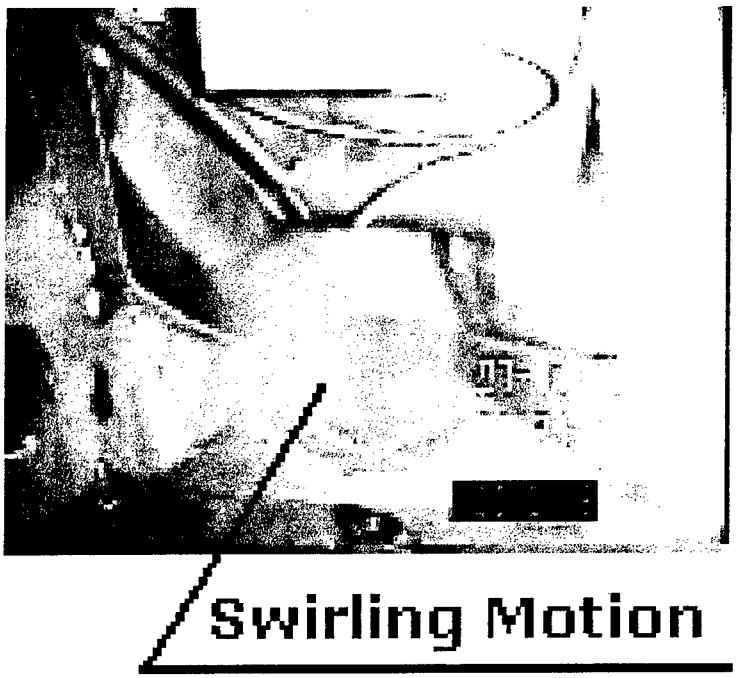


Figure 5 Initiation of the three-dimensional clockwise swirling motion that exists in the first bay of the nacelle simulator.

As previously noted the airflow enters the nacelle through the inlet duct that is located at the bottom of nacelle forward compartment (see Figure 3). After which the airflow is redirected upward through the inlet screen. The inlet screen consists of a large number of circular passages that have been drilled at a 45 degree angle (as previously seen in Figure 4) and act as a number of individual jets. The orientation of these jets are such that the flow is directed downstream into the nacelle. Unfortunately, a large portion of the incoming air exiting the screen and impinges on the clutter and the underside of the simulated engine core. This effectively blocks and redirects the airflow and causes it to split and swirl before entering the nacelle. Viewing this process, in the time mean sense, suggests that the split in the airflow is even. However, closer inspection of the video, or examination of single frames, clearly indicates the oscillating nature of the stagnation point.

Figure 6 A laser lightsheet positioned at 45 degrees to the exiting jet flow originating at the nacelle inlet screen

From the analysis of a number of single frames, of which Figure 6 is a single frame randomly selected from the VHS video tape, three features are worthy of mentioning. The first feature provides evidence of the jet flows impinging on the underside of the engine core. In the second case, the flow shows asymmetry as it approaches the engine core, and finally, the swirling nature of the flow shows how the stagnation point oscillates back and forth. Also apparent in the video is a circulation region positioned on each side of the nacelle. These are the tails of the complex swirling motion observed in Figure 5.

The unsteadiness of the stagnation point is to a large degree a result of the perturbations caused by the interaction of the adjacent jets. This can be seen in Figure 7. Here the structures produced by the individual jets are shown to interact. However, it must be noted that the smoke streams are originating from two parallel rows of jets separated by approximately two diameters. Therefore as the jet spacing is reduced the interaction can be expected to be more violent and unsteady.

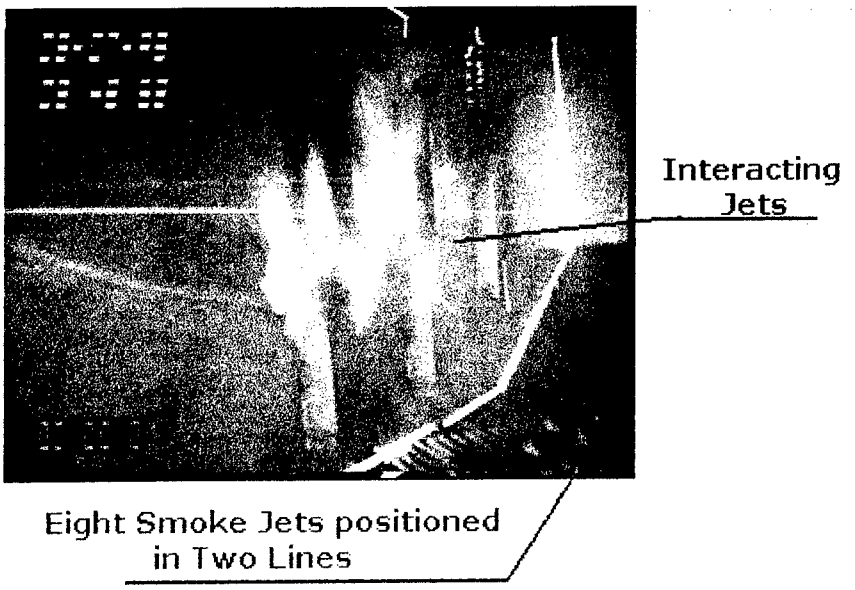


Figure 7 Strong interaction of the kidney shaped smoke jets are observed within the nacelle

The physical extent of the inlet airflow can be observed in Figure 8. In this figure the smoke tubes are again positioned in the "V dot" pattern and the laser light sheet oriented, normal to the jets, at 45 degrees. This enabled the light sheet to be positioned perpendicular to the smoke filaments axis in order to allow isolation of the individual jet structures. Also readily apparent in this figure is the kidney shaped cross section of the three smoke structures that make up the lower portion of the jet. Such a structure is reminiscent of vortical structures observed in 'jet in a crossflow' studies, in which a jet is injected into the cross stream. As the jet penetrates into the flow it also undergoes a bending, aligning itself with the freestream and transforming its vortical structure. A characteristic of these types of flow fields are the kidney shaped structures that contain two counter rotating streamwise vortices.

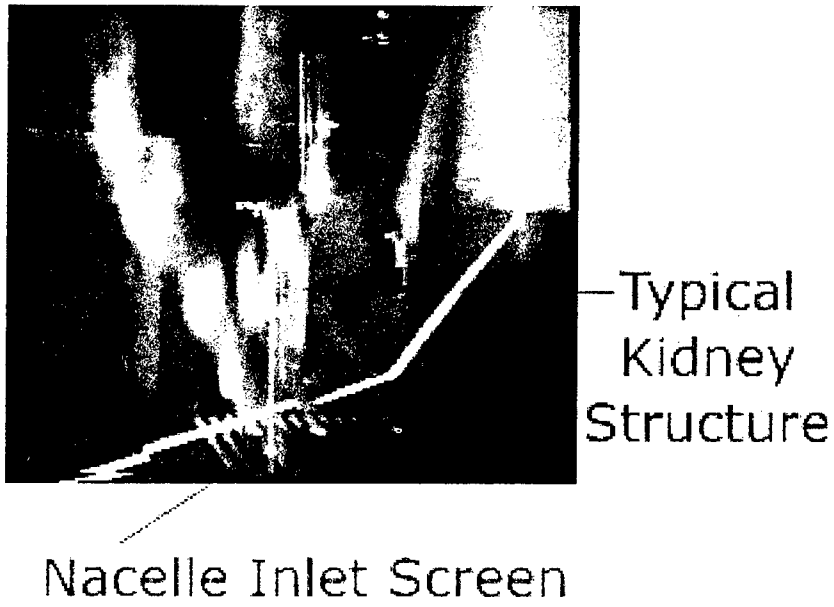


Figure 8 Kidney shaped vortical structures are observed

Under flow conditions containing favorable pressure gradients these streamwise vortical structures will intensify and become highly coherent, thereby delaying their breakdown. This increased lifetime can result in this type of vortical pair being observed many diameters from their point of origin. Although most of these streamwise vortical pairs would interact with the engine clutter, the flow along the bottom of the nacelle appeared not to impede their coherence. These vortical structures were observed downstream in Fire Zone 1 (Figure 9).

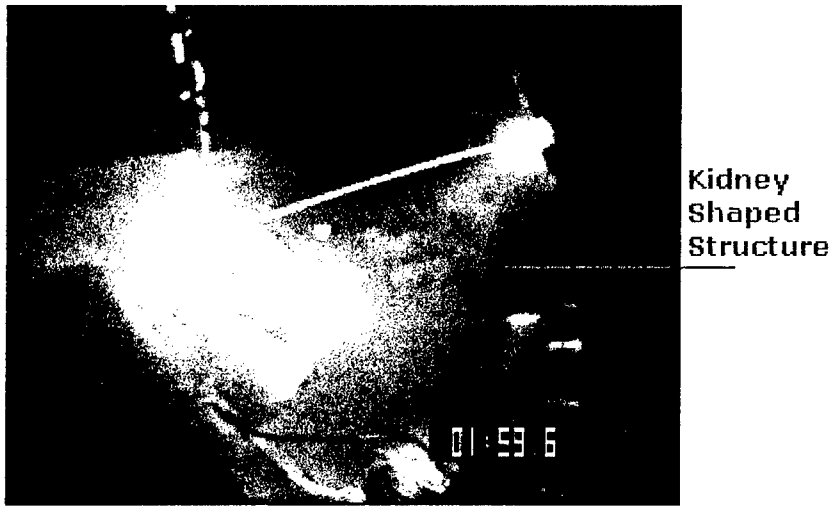


Figure 9 Observation of a counter rotating pair of streamwise vortices in Fire Zone 1.

An examination of past nacelle fire tests was performed in order to search for additional support for the existence of organized flow structures within a fire zone. Specifically, evidence of counter-rotating streamwise vortices was sought within the fire zone, while spray combustion was in progress. To achieve this goal video tapes of simulated AENFTS fire tests during the Halon replacement program were examined and Figure 10 is a typical photograph. In this particular phase of the Halon program a fuel jet spraying JP-8 was intentionally ignited and allowed to reach a steady thermal state. At that point in time the fire extinguishing system was energized and the amount of fire extinguishing agent needed to suppress the fire was recorded. For the purpose of the present argument the following figure has been included in this report and clearly shows the presence of the kidney shaped structures. Furthermore, this figure demonstrates the ability of these vortical structures to alter the fire pattern by separately entraining fuel and redirecting a flame above the existing fuel spray.

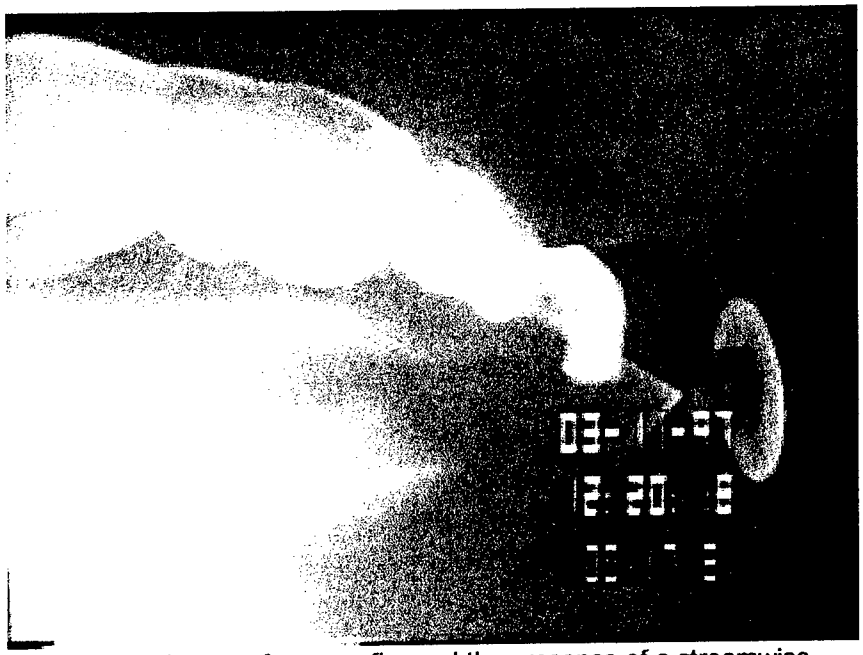


Figure 10 An image of a spray fire and the presence of a streamwise vortical structure redirecting a portion of the flame.

Summary

The overall effect of nacelle airflow on agent discharge still needs to be further examined. However, the present study clearly shows the existence of a complex three-dimensional swirling motion just upstream of the first bulkhead at the agent discharge nozzle location. Through the use of laser light scattering from small smoke particles, it was determined that jets which exited the inlet screen interacted with one another in a violent manner and became unstable. It was also observed that they would impinge on the underside of the engine core and clutter components present within the nacelle. This resulted in a redirection of the flow and an interaction with the individual jets as they exited the screen. This interaction manipulated the jet structure, in such a way as to take on a kidney shape, and strongly resembled that of a jet in a crossflow. The resulting kidney structures were made up of counter-rotating pairs of streamwise vortices and exhibited a high degree of coherence. These kidney structures were observed downstream in the nacelle at Fire Zone 1 in both cold flow tests and during fire tests performed under the Halon replacement program.

Defining and understanding the mean flow is only one part of the problem. The complexities of the recirculation and stagnation zones coupled with the inherent instability of interacting jets must be appreciated to fully understand the affect of the internal nacelle airflow on agent distribution. Only after a sufficient understanding of these problems can modeling of the nacelle airflow and agent distribution be successful.

SENSORS FOR FOCAL PLANE ARRAY PASSIVE MILLIMETER-WAVE IMAGING

Numan S. Dogan
Associate Professor
Department of Electrical Engineering

North Carolina A&T State University
539 McNair Hall
1601 East Market Street
Greensboro, NC 27411

Final Report for:
Summer Faculty Research Program
AFRL/MNGS

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

and

AFRL/MNGS, Eglin Air Force Base, FL

August 1998

SENSORS FOR FOCAL PLANE ARRAY PASSIVE MILLIMETER-WAVE IMAGING

Numan S. Dogan
Associate Professor
Department of Electrical Engineering
North Carolina A&T State University

Abstract

Sensors for focal plane array passive millimeter-wave imaging have been investigated. Passive Millimeter-Wave (PMMW) imaging systems open a new window in the electromagnetic spectrum for imaging through fog, clouds, dry snow, smoke, and sandstorms. A PMMW camera capable of generating real time display of the imaged scene, similar to video camera, has been developed by TRW[1]. In order to achieve reasonable price/performance tradeoff, low cost sensors for PMMW imaging are required. We have investigated sensors for direct detection radiometers, superheterodyne radiometers, and microbolomers. Results show that monolithic superheterodyne receivers have price/performance advantage compared to other technologies.

1. Introduction

The smart tactical autonomous guidance systems (STAG) program was initiated at Eglin Air Force Base in 1990 under the sponsorship of the Armament Directorate of Wright Laboratory. The purpose of the STAG program is to investigate and support passive millimeter-wave activities. As part of this program, a mobile test bed referred to as MAPS (Millimeter-Wave Analysis of Passive Signature) has been developed [2]. The MAPS provides the data base of passive millimeter-wave information concerning terrain, atmospheric and target signatures needed to support the critical development technologies associated with the STAG program. Passive Millimeter-Wave(PMMW) imaging systems open a new window in the electromagnetic spectrum for imaging through fog, clouds, dry snow, smoke and sandstorms.

Applications benefiting from such a capability are numerous. Foremost, in terms of market entry, is the low visibility autonomous landing of aircraft. Its impact on civil and military aviation is quite significant both from an operational and a cost/benefit point of view. Military intelligence, surveillance and reconnaissance benefit from a new all-weather asset enabling easier target detection and cueing due to the unique signature provided by metals and metal-like objects in the scene. All-weather perimeter surveillance and navigation on the ground or at sea are also missions enabled by PMMW imaging systems. A growth market in PMMW imaging will generate applications from fire fighting, emergency vehicles, search and rescue, to perimeter surveillance at correctional and other critical facilities.

The key component of the PMMW camera/radiometer is the Monolithic Microwave/Millimeter-Wave Integrated Circuit (MMIC) receiver chip. Having such a receiver chip in large quantities at low cost is very important for the development efforts in PMMW imaging. This report comparatively examines the capabilities and performances of various technologies.

2. Background

Every physical body not at absolute zero emits electromagnetic radiation at all frequencies, by virtue of the thermal energy the body possesses. Planck presented a radiation law relating the brightness, β, of a blackbody at a temperature T to frequency f. The brightness is the amount of power arriving at a unit area, over a unit solid angle, and in a given frequency interval; this can be stated as

$$\beta = \frac{2hf^3}{c^2}\frac{1}{e^{hf/kt}-1}(Wm^{-2}rad^{-2}Hz^{-1})$$

where

h = Planck's constant (= 6.55×10^{-34} joule-s),

f = source frequency (Hz),

c = velocity of light (= 3×10^8 m/s),

k = Boltzmann's constant (= 1.38×10^{-23} joules/K),

T = source temperature (K).

For the observed frequencies that are few hundred gigahertz (GHz) and temperatures at thousands of degrees Kelvin, hf<<kT. The brightness, with this approximation, reduces to the Rayleigh-Jeans law of the form

$$\beta = \frac{2f^2kT}{c^2}$$

PMMW imaging/radiometer systems use sensors that detect and quantify the emitted radiation from objects. Figure 1 illustrates a radiometer with a bandwidh (B) and a gain (G). Usually, the received noise energy densities radiated by a blackbody are very small and the collecting area is limited; thus, measurements can be made only indirectly on amplified versions of the signals. The combination of antenna, amplifier, and the display is the PMMW imager.

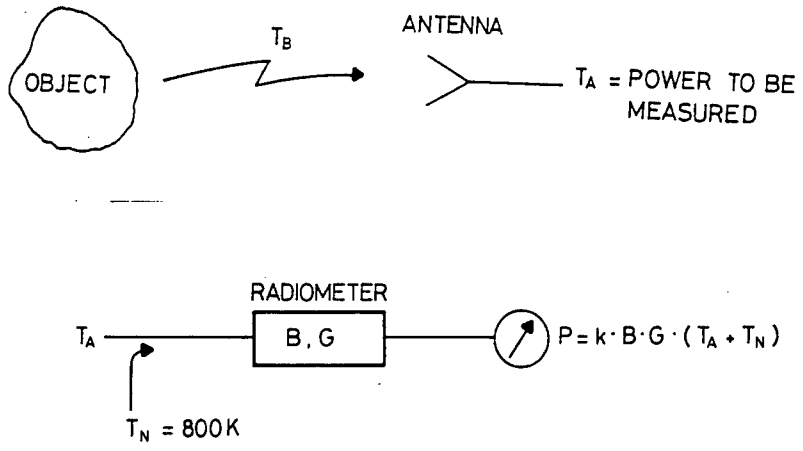


Figure 1 Radiometer

Every component in the PMMW imager generates electrical noise power that contributes to the noise output of the system. In practice, the system noise output can be many orders of magnitude greater than the signal being measured. One of the most important figure of merit is *the radiometer sensitivity* which is the smallest change in system noise input power that can be detected in the system output. This figure is taken to equal the uncertainty in measuring the off-source noise power output referred to the system input; expressed as [3]

$$\Delta T = KT_s / \sqrt{Bt}$$

where

T_s is the system noise temperature

B is the bandwidth before detection

t is the time spent integrating the measurement.

K is a factor close to unity that is set by the radiometer system.

ΔT given in the above expression is a basic limitation to the sensor performance of PMMW imager. The PMMW imager design goal is to obtain a performance as close as possible to the theoretical ΔT.

Earlier developments in first millimeter-wave imagers/radiometers date back to the 1950s. These systems were bulky, had low thermal sensitivity and poor spatial resolution. Since these early developments there have been considerable advances in semiconductor solid state devices and MMIC technology which have allowed the size and weight of imagers to be reduced.

3. Radiometers

Most of the radiometers reported to date fall into one of the following three categories; (1) Direct detection or tuned-radio frequency (TRF) radiometers, (2) Superheterodyne radiometers, and (3) Microbolometers. In the following, each of these radiometers are reviewed.

3.1 Direct Detection Radiometers

Figure 2 shows the block diagram of a direct detection/TRF radiometer. A Passive millimeter-wave camera based on this principle with the capability of generating real time displays of the imaged scene, similar to video cameras, has been under development at TRW. The TRW camera uses a focal-plane array of direct detection MMIC receivers based on 0.1 micron T-gate passivated pseudomorphic AlGaAs/InGaAs/GaAs HEMT technology. The on-wafer temperature sensitivity)T is ~0.4K with a 10 msec time constant. The camera utilizes quasi-optical imaging through a plastic lense (18" diameter aperture), focusing the incoming MMW radiation through a high-pass filter, off of an image oversampling mirror, and onto the 1040 (40Hx26V) receiving antennas of the FPA. The MMW radiation is amplified and rectified in the 1040 MMIC chips, and the resultant 4160 (4x1040 due to the image oversampling) video signals are passed through various electronic and image processing steps which prepare the image for final display

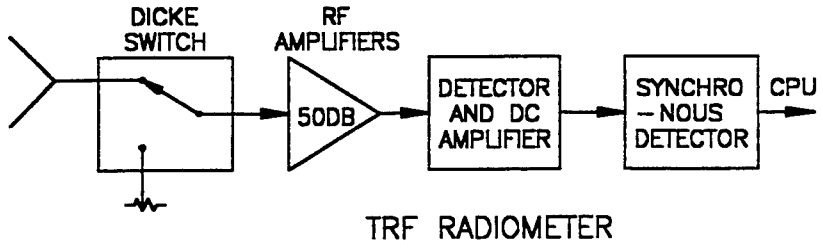


Figure 2. Direct Detection/TRF Radiometer

The Transistor gain fluctuation which is quite severe for short channel transistors such as the ones used in TRW camera degrade the performance of the radiometer. To overcome this problem, Weinreb [4] uses pilot-signal stabilization. At present, W-band MMIC LNA's used in the TRW cameras are expensive. They require MBE grown wafers and e-beam lithography. . The major advantages and disadvantages of direct detection radiometers can be summarized as: Advantages: (1) No need for Local Oscillator (LO) power, (2) Simple circuit design and topology, and (3) Good temperature sensitivity. Disadvantages: (1) Expensive process ($30/receiver chip), (2) Excessive Transistor gain fluctuations, and (3) Limited number of suppliers.

3.2 Superheterodyne Radiometers

Figure 3 shows the block diagram of a superheterodyne radiometer. Superheterodyne radiometer utilizes a mixer, an IF amplifier, and a detector. One of the enabling technologies used at millimeter wave frequencies is the Finite Ground Coplanar (FGC) transmission line. FGC transmission line is a modified Coplanar Waveguide (CPW) that provides improved performance at millimeter-wave frequencies.

Millimeter-wave monolithic detectors and Varactor multipliers have been demonstrated using FGC lines as circuit elements and transmission lines[5]. FGC technology is inexpensive and utilizes the technology developed for conventional GaAs MESFETs and HBTs. Unlike the direct detection radiometer which requires 0.1 micron T-gate passivated pseudomorphic AlGaAs/InGaAs/GaAs HEMT technology, the lithography requirement for superheterodyne receiver based on FGC technology is not critical. The first stage which handles the incoming high-frequency millimeter-wave signal is the mixer which down converts the frequency to IF. Typical published results show that acceptable noise figures and conversion losses can be achieved using conventional MMIC technology. Amplification of the down converted signal at IF can easily be handled by mature low noise MESFET technology. As the following results show, the degradation due to lack of a preamplifier before the mixer is minimal.

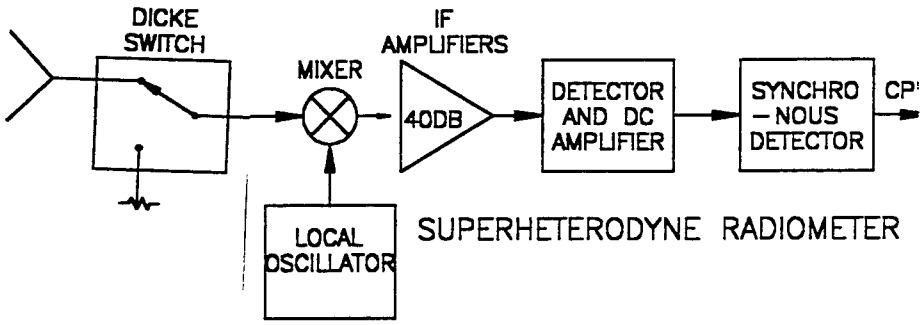


Figure 3. Superheterodyne Radiometer

Using the published results in the literature[6-10] for MMIC mixers and low-noise amplifiers, we can evaluate the noise figure or system temperature of the radiometer. The following parameters are typical values for the conversion loss and noise temperature of W-band mixers and low noise IF amplifiers;

L_c = 6 dB , conversion loss of a W-band mixer

T_{mix} = 1000 K, noise temperature of a W-band mixer

F_{IF} = 1.5 dB, noise figure of an IF low noise amplifier

The overall noise figure of the system including the mixer and the preamplifier is expressed as

$$F = L_c + 10 \log (N_R + F_{IF} - 1)$$

where N_R is the noise ratio of the mixer expressed as $N_R = T_{mix} / T_o$. T_o is the ambient temperature of the system in degree K. Using the values shown above, the overall noise figure of the system is given by

$$F = 6 + 10\log (3.33 + 1.5 - 1) = 11.83 \text{ dB}$$

The system temperature is expressed as

$$T_{syst} = T_o (F-1)$$
$$= 300 \text{ x } (11.83-1) = 3{,}250 \text{ K.}$$

Using the estimated system temperature we can calculate the expected radiometer temperature sensitivity for a bandwidth of 10 GHz and 10 ms integration time;

$$\Delta T = T_{syst} / (B.t)^{0.5}$$
$$= 3{,}250 / (10^{10} \text{x} 10^{-2})^{0.5}$$
$$= 0.325 \text{ K.}$$

The temperature sensitivity of a superheterodyne radiometer is certainly comparable to the temperature sensitivity of a direct detection radiometer. The major advantages of a superheterodyne radiometer include: (1) MMIC receiver can be fabricated using conventional GaAs IC technology, (2) Transistor gain fluctuations are very small for MESFET and HBTs resulting in better amplifier gain stability. The major disadvantage of a superheterodyne radiometer is that it requires a local oscillator(LO). LO power can be coupled using quasi optical techniques and may not increase the system design costs significantly.

Due to the explosive growth in the cellular communication market, GaAs technology has matured and cost per wafer is very reasonable. It is expected that W-band MMIC receivers can be fabricated for $1-$3 per chip in large quantities.

3.3 Microbolometers

Metal-film bolometers have been used for a long time in various radiative thermal detections [11]. Recent developments in bolometers use micromachining techniques which made the detector performance approach the radiation limit. Great attention is also being paid currently to develop monolithic focal plane-array (FPA) bolometer to meet the demand of robust, economical, and uncooled thermal imagers in related commercial markets, such as fire or surveillance purpose. A review article by P.L. Richards [12] discusses the basic principles of a bolometer. Figure 4 shows a microbolometer with an integrated dipole antenna intended for millimter-wave imaging applications[13]. The millimeter-wave signal is received by the integrated dipole antenna and coupled to the microbolometer as a current that causes a change in the temperature of the sensing material. Sensing material that is frequently used in microbolometers include Tellurium, Bismuth, Niobium, Vanadium oxide, and YBaCuO[14-17]. Typical value for the temperature coefficient of resistance for these materials is 1-5 %. One of the most important parameters of a bolometer is the voltage responsivity that is expressed as

$$S(f) = \Delta V / \Delta P$$

$$= (\eta \alpha V_b) / (G(f)+j2\pi fC) \quad (V/W)$$

where

η = optical coupling efficiency

α = (1/R)(dR/dT), temperature coefficient of resistivity

V_b = DC bias voltage

G = Thermal conductance

C = Heat capacity.

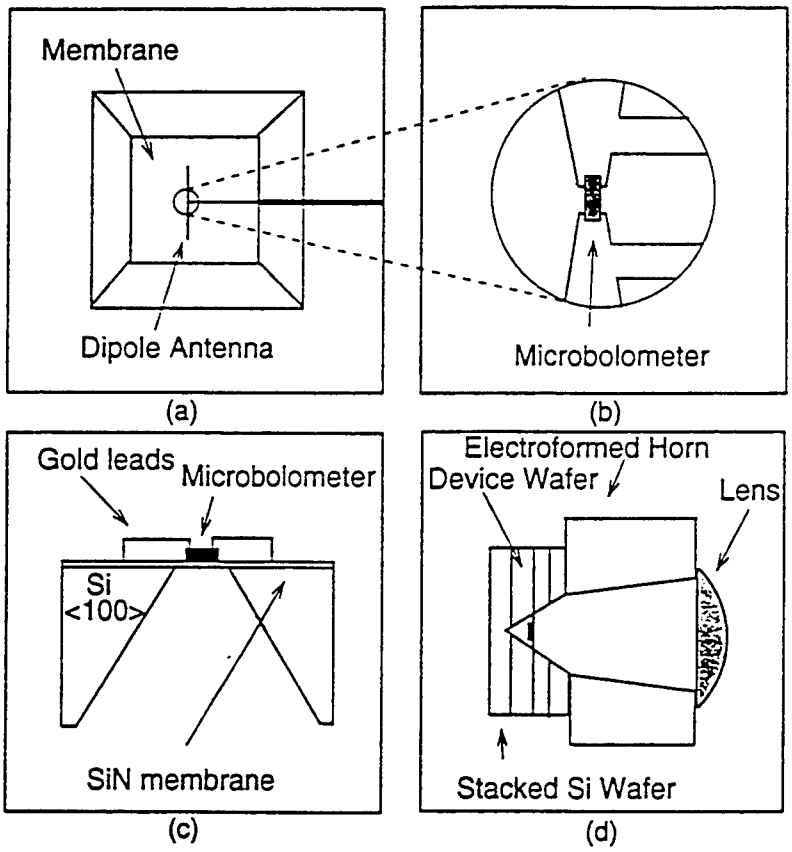


Figure 4. Micromachined Microbolometer (a) top view, (b) close up view of the microbolometer, (c) side view, (d) microbolometer with integrated horn

The response time of the bolometer is expressed as;

$$\tau = C/G \text{ (sec.)}.$$

The other important parameter of a bolometer is the noise equivalent power (NEP) which is expressed as;

$$NEP = \{4kT^2G + (4kTR/S^2) + [e^2 + (iR)^2]/S^2 + (Ef^2/S^2)\}^{0.5} \qquad (W/Hz^{0.5})$$

The first and second terms in the equation are the phonon and Johnson noise, respectively. The third and the last terms are the amplifier noise and 1/f noise of the bolometer, respectively. Typical values published to date for these parameters are;

$S \sim$ 10-1,000 v/W

$NEP \sim 10^{-10}$ W/Hz$^{0.5}$

$\tau \sim$ 1-10 ms.

Although significant progress has been made in microbolometers, their performance still lacks the performance of the direct detection and superheterodyne radiometers. The other main disadvantage of microbolometers is that most of the work is carried out in the research labs and no commercial source is available for mass production.

4.0 Conclusion

This report presents the results of a comparative study of sensors for focal plane array (FPA) passive millimter-wave imaging. Since large number of sensors are required for FPA imaging, the cost of sensor element (receiver chip) is a critical issue. Direct detection radiometers, superheterodyne radiometers, and microbolometers have been considered in this study. The results indicate that the superheterodyne radiometer is a viable candidate for low cost production with good sensor performance characteristics.

5. REFERENCES

[1] Larry Yujiri et al., "Passive Millimeter-Wave Camera", Proceedings of the SPIE Conference on Passive Millimeter-Wave Imaging Technology , Vol. 3064, , pp 15-22, Orlando, FL, April 1997.

[2] Doc Ewen, Roger Smith, Bryce Sundstrom, Byron Belcher, and Keith Trott, "Millimeter-Wave Analysis of Passive Signatures (MAPS)", Proceedings of the SPIE Conference on Passive Millimeter-Wave Imaging Technology , Vol. 3064, , pp 2-14, Orlando, FL, April 1997.

[3] G. Evans and C.W. McLeish, "RF Radiometer Handbook", Artech House, 1977.

[4] Sander Weinreb, "Millimeter-Wave Integrated Circuit Radiometers", Proceedings of the SPIE Conference on Passive Millimeter-Wave Imaging Technology , Vol. 3064, , pp 80-89, Orlando, FL, April 1997.

[5] G.E. Ponchak, S. Robertson, F. Brauchler, J. East and L.P.B. Katehi, "Finite Ground Coplanar Waveguide for Microwave and Millimeter-Wave Integrated Circuits," ISHM 1996, Proc. of the Int. Symp. on Microelectronics, Minneapolis, Minnesota, Oct. 1996, pp. 517-521.

[6] S. Raman and G. Rebeiz, "A 94 HGz Uniplanar Subharmonic Mixer," 1996 IEEE MTT-S Digest, pp 385-388.

[7] U. Guttich, J.-M. Dieudonne, A. Klaassen, H. Sledzik, K.-E. Schmegner, U. Goebel, and M. Boheim, "A Low Noise Heterodyne 89 GHz MMIC Module for the Multifrequency Imaging Microwave Radiometer (MIMR)," 1996 IEEE MTT-S Digest, pp 21-24.

[8] R.S. Virk, S.A. Maas, M.G. Case, M. Matloubian, P. Lawyer, H.C. Sun, C. Ngo, and D.B. Rensch, "A Low-Cost W-Band MIC Mixer Using Flip-Chip Technology," IEEE Microwave and Guided Wave Letters, Vol. 7, No. 9, September 1997, pp. 294-296.

[9] S. Gearheart, "Integrated Millimeter-Wave and Submillimeter-wave Antennas and Schottky-Diode Receivers," Ph.D. Thesis, The University of Michigan, Ann Arbor, Michigan, 1994.

[10] S. Gearheart and G. Rebeiz, " A Monolithic 250 GHz Schottky-Diode Receiver," IEEE Transactions on Microwave Theory and Techniques, Vol. 42, No. 12, pp 2504-2511, December, 1994.

[11] J.S. Shie, Y.M. Chen, M.O. Yang, and B.C.S. Chou, "Characterization and Modeling of Metal-Film Microbolometers," IEEE Journal of Microelectromechanical Systems, Vol. 5, No.4, December 1996, pp. 298-306.

[12] P.L. Richards, "Bolometers for Infrared and Millimeter Waves," Journal of Applied Physics, 1 July 1996, pp. 1-24.

[13] Arifur Rahman, Erik Duerr, Gert de Lange, and Qing Hu, "Micromachined Room-Temperature Microbolometer for mm-Wave Detection and Focal-Plane Imaging Arrays", Proceedings of the SPIE Conference on Passive Millimeter-Wave Imaging Technology , Vol. 3064, , pp 122-133, Orlando, FL, April 1997.

[14] M.E. MacDonald and E.N. Grossman, "Niobium Microbolometers for Far-Infrared Detection," IEEE Trans. On MTT 43, pp 893-896, 1995.

[15] S.W. Wentworth, D.P. Neikirk, "Composite Microbolometers with Tellurium Detector Elements," IEEE Trans. Microwave Theory and Techniques," Vol. 40, No. 2, Feb. 1992, pp. 196-201.

[16] H. Jerominek, F. Picard, and D. Vincent, "Vanadium Oxide Films for Optical Switching and Detection," Optical Engineering, Vol. 32, No. 9, September 1993, pp. 2092-2099.

[17] C.M. Travers, A. Jahanzeb, D.P. Butler, Z.Celik-Butler, "Fabricatio of Semiconducting YBaCuO Surface-Micromachined Bolometer Arrays," IEEE Journal of Microelectromechanical Systems, Vol. 6, No. 3, September 1997, pp. 271-276.

MECHANICAL STRENGTH MODELING OF PARTICLE STRENGTHENED NICKEL-ALUMINUM ALLOYS STRENGTHENED BY INTERMETALLIC γ' (Ni_3Al) PRECIPITATES

James M. Fragomeni
Assistant Professor
Department of Mechanical Engineering

Ohio University
251 Stocker Engineering Center
Athens, Ohio 45701-2979

Final Report for:
Summer Faculty Research Program
Wright Research Site

Sponsored by:
Air Force Office of Scientific Research
Wright-Patterson Air Force Base, Ohio

and

Wright Research Site

August 1998

MECHANICAL STRENGTH MODELING OF PARTICLE STRENGTHENED NICKEL-ALUMINUM ALLOYS STRENGTHENED BY INTERMETALLIC γ'(Ni_3Al) PRECIPITATES

James M. Fragomeni
Assistant Professor
Department of Mechanical Engineering
Ohio University

Abstract

A nickel-aluminum alloy strengthened by the γ' (Ni_3Al) intermetallic ordered coherent precipitates was used as a demonstration material to develop the method of modeling to predict strengthening behavior during plastic deformation as a consequence of the γ' particles acting as obstacles to the dislocations and thus impeding their glide motion through the alloy. The two most dominate strengthening mechanisms in the Ni-Al system are order strengthening when the particles were smaller than the critical looping radius, and Orowan strengthening when the particles were larger than the looping radius. In the overaged condition when the particles are large in size, the dislocations bypass and loop the particles by the Orowan mechanism. In the underaged to peak aged conditions where the particles are usually smaller than the looping radius, the dislocations shear the precipitates during deformation. The overall approach was determined from the particle coarsening kinetics, dislocation mechanics, thermodynamics, resolved shear stress and texture, as well as the dislocation particle interaction mechanisms. The single crystal strength for particle strengthening was used as the basis for estimating the total polycrystalline strength. Thus, the total precipitation strengthening response of the demonstration alloy was determined from the active microstructural strengthening mechanisms in the underaged, peak-aged, and overaged conditions. The total polycrystalline yield strength included contributions from the intrinsic lattice strength, the solid solution strengthening, grain size strengthening, and particle strengthening which included the order hardening and Orowan strengthening contributions. The total mechanical yield strength for a Ni-6.27wt.%Al alloy was predicted for the peak-aged condition, and was found to be in good agreement with the experimental peak yield polycrystalline strength data for Ni-6.27wt.%Al.

The overall goal of this research was not only to develop a strength model for Ni-Al, but to provide the framework and basis for mechanical strength modeling of high temperature superalloys and high temperature discontinuously reinforced aluminum alloys that are strengthened by γ' (Ni_3Al). The theoretical work described in this report provides a solid foundation for future mechanical behavior modeling research for much more complex alloy systems, such as various high temperature superalloys, and high temperature discontinuously reinforced aluminum alloys, and would incorporate the microstructural parameters and processing variables of alloys with very unique microstructures. Most of the theoretical modeling would apply to these other alloys, however, the extent of the interactions and contributions of the various strengthening mechanisms would be different.

MECHANICAL STRENGTH MODELING OF PARTICLE STRENGTHENED NICKEL-ALUMINUM ALLOYS STRENGTHENED BY INTERMETALLIC $\gamma'(Ni_3Al)$ PRECIPITATES

James M. Fragomeni

Introduction

This report discusses the topic of strengthening of various materials which obtain most of their strengthening from various types of obstacles to dislocation motion such as dispersoids, precipitates, and grain boundaries. This study of the various mechanisms of strengthening from different kinds of particles in the microstructure is important for accurately modeling the mechanical strengthening of a metal or alloy. The mechanical strength of alloys is usually modeled in terms of the deformation behavior of single crystals. A nickel-aluminum alloy strengthened by the Ni_3Al intermetallic phase by was selected as a demonstration material to develop the method of modeling presented for predicting the mechanical strength. The long term objective of this investigation is to extend the current modeling effort to much more complex alloy systems such as some high temperature superalloys, such as nickel base superalloys, and high temperature discontinuously reinforced aluminum alloys.

The property that defines the plastic deformation or yielding in single crystals is the critical resolved shear strength, which is the stress at which dislocations glide freely through the single crystal. During deformation, the precipitate particles are either sheared/cut or bypassed by the gliding dislocations, depending on their sizes. The larger particles will be bypasses or looped by the dislocations and the smaller particles will be sheared by the dislocations. Thus, the particles inherently impede the motion of the dislocations since they act as obstacles or barriers to the dislocation glide motion. Some of the technologically important particle-hardened alloys such as precipitation strengthened Ni-Al, Ni-Ti, Ti-Al, Fe and Ni base superalloys, and Al-Li, are primarily strengthened by ordered $L1_2$ intermetallics, such as γ'-Ni_3Al, γ'-Ni_3Ti, Ti_3Al, and δ'-Al_3Li, and achieve good strength because the intermetallic strengthening precipitates effectively impede the dislocation glide motion during plastic deformation. Ordered coherent precipitates are important microstructural constituents in these and various other precipitation-hardenable alloy systems. These precipitates are sheared by the dislocations in the underaged and peak-aged conditions and looped by the dislocations in the overaged conditions. Strengthening by ordered coherent precipitates occurs when a matrix dislocation shears an ordered precipitate and creates an antiphase boundary on the slip plane on the particle that is being sheared. As long as the ordered precipitates are sheared during plastic flow, the dislocations responsible for plastic deformation travel in pairs, with the trailing dislocation restoring the disorder on the {111} slip planes created by the leading dislocation. Therefore there is no net change in order behind the dislocation pairs. The leading and trailing dislocations are always separated by a region of antiphase boundary. The energy of the antiphase boundary always connects the two dislocations to a dislocation pair. The antiphase boundary energy per unit area on the slip plane, γ_{abp}, represents the force per unit length opposing the motion of the dislocation as it penetrates the particle {11,21}. The separation distance between the leading and trailing dislocations is a result of the equilibrium between the attractive force of the antiphase boundary and the repulsive

force of the two dislocations of equal sign {22}. Antiphase boundaries can lie on any plane, but the two most important for Ni-Al are the {111} and {100} planes.

Theoretical Approach

Dislocation-Particle Bypassing of Large Particles

The dependence of strengthening on the interparticle spacing was first proposed by Orowan {1} who showed that the stress necessary to force a dislocation between two particles was inversely proportional to their spacing. High temperature discontinuously reinforced aluminum alloys are strengthened primarily by two different size distributions of the strengthening phase so therefore it becomes necessary to model the strengthening response of the dispersion strengthened alloy in terms of a bimodial particle size distribution. The theoretical method being taken will be derived to be applicable to precipitation hardened alloys with more than one strengthening phase or with a bimodal particle size distribution of the same strengthening phase. For the high temperature discontinuously reinforced aluminum alloys the primarily strengthening of the dispersion strengthened alloy results from the hard intermetallic particles. The strengthening results from these particles impeding the dislocation glide motion through the Orowan mechanism via. dislocation particle bypassing/looping. The dislocations do not enter the particles, but rather bypass them leaving dislocation loops around the particles. The dislocation loops which surround the particles are usually referred to as Orowan loops. The nature of the particles does not effect the observed strengthening from bypassing, provided the particles are strong enough to withstand the local stress that the dislocation exerts on them without shearing or fracturing. The nature of the particles can be very simple such as GP zones, which are solute rich regions with the same structure as the matrix, or very complicated such as intermetallic compounds or complex carbides. The version of the Orowan expression currently accepted and used for spherical dispersoids is given by {1}

$$\Delta\tau = 2\ \{G_m b/((4\pi)(1-\nu)^{0.5})\}\{1/\lambda\}\{\ln[d_p/r_o\} \tag{1}$$

where G_m is the shear modulus of the matrix, b is the Berger's vector of the matrix, ν is Poisson's ratio, λ is the interparticle spacing, d_p is the mean planar particle diameter, r_o is the core radius of dislocations often taken as 2b. However, a much more simplier form of the Orowan model can be expressed as

$$\Delta\tau = G_m b/\lambda \tag{2}$$

Various models have been developed for estimating the interparticle spacing λ in terms of the volume fraction, average particle size, and number of particles intersected per unit area of slip plane. In the case of spherical particles of uniform diameter d, which are distributed in a triangular array on the slip plane the interparticle spacing is given by {2}

$$\lambda = L_p - d_p = 1.075/(N_A)^{0.5} - \pi d/4 \tag{3}$$

where L_p is the average center-to-center distance between particles, N_A is the number of particles intersected per unit area of slip plane, and f is the volume fraction of the dispersoids. Or in terms of the particle diameter d and volume fraction, λ is given by {2}

$$\lambda = 0.538d(2\pi/3f_v)^{0.5} - \pi d/4 \quad (4)$$

where d is the average particle size diameter, and f_v is the volume fraction of the dispersoid phase.
Kocks {3} and also Forman and Makin {4} determined a similar estimate of the interparticle separation given by

$$\lambda = 1.25r\,(2\pi/3f_v)^{0.5} - 1.633r \quad (5)$$

where r is the average particle radius of the dispersoid distribution, and f_v is the volume fraction of dispersoids.
The free space between particles has also been approximated by Ardell {11} by the so-called square lattice spacing, L_s, which is given by

$$L_s = (2\pi/3f_v)^{0.5}r \quad (6)$$

The number of dispersoid particles that intersect a given microstructural plane at various random intersections of the particle can be estimated in terms of the volume fraction of dispersoids and average size radius of the dispersoids. This parameter n_{total} can be expressed as {5-9}

$$n_{total} = (3\,f_v)/(2\,\pi\,r^2) \quad (7)$$

In addition, the total number of particles per unit volume can be expressed as {5-10}

$$N_v = (6f_v)/(\pi d^3) \quad (8)$$

where N_v is the total number of dispersoid particles per unit volume and can be related to the total number of particles per unit area by the expression {5-10}

$$n_{total} = N_v d \quad (10)$$

Determining the total strength for the combined strengthening effect of two (or more) different distributions of particles is necessary since both distributions have a complex nonlinear additive effect on the total strength of the dispersion strengthened alloy. Thus the appropriate strengthening addition model must be chosen for incorporating the contributions to the strengthening from the two different particle size distributions in the microstructure.

In a given microstructure, the dislocations do not shear or bypass the precipitate particles through the particle diameter, but at various sections or chords through the precipitates. Some or the precipitates may be sheared or looped by the dislocations through the center and the other particles at random chord lengths. This is because not

only is a real distribution of particle sizes not monodisperse where all the particles have the same size, but the distribution of particle sizes will intersect any given glide plane at various chord lengths of the individual particle sizes. Thus, the average planar particle radius is a more accurate estimate of the actual area of interest than the average particle size of a given distribution of dispersoids. The average planar particle radius is related to average particle size through the expression given by {11}

$$r_{pl} = \pi r/4 \tag{11}$$

where r_{pl} is the average planar radius of the precipitate particles. The average planar particle size can be used instead of the average particle size to give more accurate predictions of sheared and looped areas. Also, utilizing the average planar particle size will result in the average planar particle cross sectional area, which would be given by the expression {11}

$$A_{pl} = 0.61685\ A_{particle} \tag{12}$$

The value for the total cross sectional area on any given microstructural plane can be expressed from the matrix cross-sectional area and the planar particle cross sectional area and expressed as

$$A_{total} = A_{matrix} + A_{pltot} \tag{13}$$

where A_{pltot} is the total planar particle cross-sectional area on the given microstructural plane, and A_{matrix} is the total area of the matrix of the given microstructural plane. In addition, the volume fraction of precipitates in the microstructure can be shown to be approximately equal to the area fraction of precipitates on a given microstructural plane. The volume fraction of precipitates of any shape can be rigorously determined from the following ratios {5,6,7}

$$f_v = V_{particle}/V_{total} \tag{14}$$

where $V_{particle}$ is the total volume of all the particles and

$$V_{total} = V_{particle} + V_{matrix} \tag{15}$$

Thus, if the total planar particle cross sectional area on a given microstructural plane and the volume fraction are known, then the total cross sectional area of a given plane can be determined from the relationship

$$A_{total} = A_{pl}/f_v \tag{16}$$

where A_{total} is the total planar particle cross sectional area on the given microstructural plane.

Complex Strengthening: Particle Summing Strength Rules

It is not unusual to have more than one type of strengthening phase in a given alloy {15,17,18,19,25,26}. There have been various expressions developed to account for the contributions to the strengths of particles of two or more strengths. Considering two different types of particles having single crystal critical resolved shear strengths of τ_{c1} and τ_{c2}, and assuming the flow stress of the matrix to be zero, the total single crystal critical resolved shear stress, τ_c, for dispersion/particles strengthening can be estimated from a superposition rule. The most common expressions of the superposition rules for the estimating strength from two different types or strengths of particles, are given by Brown and Ham {12}, and Labusch {13} as

$$\Delta\tau_c = \Delta\tau_{c1} + \Delta\tau_{c2} \tag{17a}$$

$$(\Delta\tau_c)^2 = (\Delta\tau_{c1})^2 + (\Delta\tau_{c2})^2 \tag{17b}$$

$$\Delta\tau_c = \Delta\tau_{c1}\,(X_1)^{0.5} + \Delta\tau_{c2}\,(X_2)^{0.5} \tag{17c}$$

$$(\Delta\tau_c)^{1.5} = (\Delta\tau_{c1})^{1.5} + (\Delta\tau_{c2})^{1.5} \tag{17d}$$

where X_1 and X_2 are areal fractions of the types of obstacles and are defined by the relationships:

$$X_1 = n_{s1}/n_s \tag{18a}$$

$$X_2 = n_{s2}/n_s \tag{18b}$$

where n_{s1} and n_{s2} are the number of particles per unit area on the given microstructural plane of particle strength or type 1 and particle strength or type 2 respectively. The number of particles per unit area is sometimes approximated in terms of the so-called square lattice spacing, L_s, defined by {11}

$$n_s = 1/(L_s)^2 \tag{19}$$

A general expression has been proposed for writing the total strengthening due to two types of precipitates simultaneously and given by {14,15}

$$(\Delta\tau_c)^q = (\Delta\tau_{c1})^q + (\Delta\tau_{c2})^q \tag{20}$$

where q is an adjustable parameter exponent usually ranging from 1 to 2. Equation (17a) is a special case of equation (20) corresponding to a linear (q = 1) superposition rule. Equation (17b) is also a special case of equation

(20) which corresponds to the pythagorean (q=2) superposition rule, first proposed by Koppenaal and Kuhlmann-Wilsdorf {16}. Equation (17c) is a law of mixtures. The linear superposition rule, for most cases, is a poor approximation, except for a mixture of a few strong obstacles among many weak ones {12}. An example of this is the case of weak precipitates in a solid solution. Pythagorean superposition is most accurate when the strengths of the two types of particles are approximately equal. The law of mixtures given by equation (17c) has been reported to be almost as good as the pythagorean law, except for obstacles of very different strengths {12}. However, Ardell {11} has shown that the law of mixtures is generally not in good agreement of the computer simulation experiments of Foreman and Makin {20}, except for narrow distributions of strong obstacles. The general superposition rule with a variable q, provides excellent agreement with the results of computer simulation experiments of dislocations through an array of obstacles by Foreman and Makin {20}. Ardell {11} has shown that the best agreement occurs when q is approximately equal to 1.4 for obstacles of very different strengths, and equals approximately 1.7 to 1.8 for obstacles of more similiar strengths. Huang {35} used a value of q = 1.5 when adding the strengthening from δ' and T_1 precipitates in an Al-Li-Cu-Zr alloys. Thus, using the above expressions for obstacles of different strengths, the contributions can be added from Orowan strengthening and particle shearing.

Dislocation Particle-Shearing for Small Particles

For a given distribution of particle sizes, some of the particles of the distribution may be sheared by the dislocations while other bypassed or looped by the dislocations, depending on their sizes relatively to the Orowan looping radius. The mechanism(s) by which the particles are sheared by the dislocations are different that the Orowan mechanism, and strengthening of the alloy increases as the particle size increases since it requires more force for the dislocations to shear through larger particles than the smaller particles. With the Orowan mechanism, the strengthening decreases as the particle size increases. The strengthening that results from dislocation-particle shearing is related to the particle size, the volume fraction of the precipitates, the Burger's vector, the antiphase boundary energy, etc. Strengthening of a particle hardened alloys having ordered coherent precipitates occurs when a matrix dislocation shears an ordered precipitate and creates an antiphase boundary on the slip plane of the ordered particle. Various alloys strengthened by ordered coherent precipitates include nickel-base superalloys, aluminum-lithium alloys, nickel-aluminum alloys, titanium-aluminum alloys, and stainless steels strengthened by γ' precipitates. These alloys have a crystal structure of the $L1_2$ type i.e., the Cu_3Au structure. It is characteristic of alloys strengthened by ordered coherent precipitates that the dislocations travel in groups or pairs. The dislocations typically travel in pairs because the passage of a pair of matrix dislocations (b=a<110>/2) through a precipitate particle restores perfect order on the {111} slip plane. The Burgers vector can be taken as approximately 0.254 nm for Ni-base alloys {68}. In $L1_2$-long-range ordered alloys such as Ni-Al, perfect dislocations have a_o<110> Burgers vectors; such dislocations dissociate into two $(a_o/2)$<110> partial dislocations with an antiphase boundary in between them on the slip plane of the particle. In disordered face centered cubic materials, the Burgers vector of perfect dislocations is of the type (a_o)<110>, where a_o is the lattice constant. Thus, when a dislocation passes through an ordered precipitate, an antiphase boundary (APB) is formed, of surface energy γ_{apb} per unit area.

However, in Ni-Al alloys, since the dislocations travel in pairs, the leading dislocation is impeded since it creates the APB, and the trailing dislocation is assisted since it restores the order. In Ni-Al, screw dislocations cross-slip from {111} planes, where they are mobile, to {010} planes, where they are immobile by the Kear-Wilsdorf {67} mechanism. An estimate of the maximum force exerted by a spherical particle may be obtained by assuming the dislocation lies along a diameter, in which case the maximum force of interaction is {23}

$$F = 2\, r\, \gamma_{apb} \tag{21}$$

and the corresponding critical resolved shear strength for particle strengthening is given by {11}

$$\tau_o = (\gamma_{apb}/b)(3\pi^2 \gamma_{apb}\, r f_v/32\Gamma)^{0.5} \tag{22}$$

where Γ is the dislocation line tension, τ_o is the critical resolved shear strength based on order strengthening, γ_{apb} is the antiphase boundary energy, and f_v is the volume fraction of the precipitate phase. This equation explains the increase of the strength on ageing in terms of the increase in r for a given volume fraction, and indicates why order hardening is thought to be a very important mechanism in nickel alloys in which large volume fractions of the ordered γ' phase (Ni_3Al, Ni_3Ti, or Ni_3Al,Ti) may be present {23}. Raynor and Silcock {27} derived a similar expression for the order strengthening mechanism of order hardened alloys given by

$$\tau_o = (\gamma_{apb}/2b)[(3\pi^2 \gamma_{apb}\, r f_v/32\Gamma)^{0.5} - f_v] \tag{23}$$

Brown and Ham [12,24] developed a expressions to incorporate the effect of the second dislocation of the pair on the critical resolved shear strength. It is difficult to determine the effect because it depends upon the statistical interaction between the trailing dislocations and the particles that have already been sheared by the leading dislocations. The second dislocation is pulled forward by the antiphase domain boundary remaining in the precipitates which it intersects. As τ_o is increased from zero, the first dislocation bends forward more, the second straightens out, and forward stress on the first dislocation is substantially increased by the interaction of the second dislocation with disordered particles {23}. Depending on the particle size r, Brown and Ham {12} have derived two equations for order strengthening for paired dislocation movements given by

$$\tau_o = (\gamma_{apb}/2b)[(4\, \gamma_{apb}\, r f_v/\pi\Gamma)^{0.5} - f_v] \tag{24}$$

for $\pi\Gamma/4\gamma_{apb} < r < \Gamma/\gamma_{apb}$. However, when $r > \Gamma/\gamma_{apb}$ the expression for order strengthening becomes

$$\tau_o = (\gamma_{apb}/2b)[(4\, f_v/\pi)^{0.5} - f_v] \tag{25}$$

where r is the average particle size radius for a distribution of particle sizes, f_v is the corresponding volume fraction of the particle distribution, and γ_{apb} is the antiphase domain boundary energy. Martin {23} shows that in general the mechanisms of shear strength is simply related to $f_v^{0.5}$ and $r^{0.5}$. Starke {30} and Sanders {29} determined a general form for the shear strength, τ, due to an array of shearable particles with an average particle size r, and a volume fraction f_v given by

$$\tau = C\, f_v^m\, r^n \tag{26}$$

where C, m, and n are all constants, with m an n are usually taken to be around 0.5. Sanders {29} applies this expression for aluminum-lithium alloys, in which order hardening dominates, and Starke {30} applies this expression valid for 7000 series aluminum alloys.

Predicting the Total Mechanical Strengthening:

The total mechanical strengthening contribution of a dispersion hardened or precipitation hardened alloy system is related to the single crystal strengthening contribution from the particles distributed in the matrix, as well as the contribution from other factors such as the solid solution strengthening contribution, and any contributions from the intrinsic lattice strength or from the grain size. The total mechanical strengthening contribution can be represented as the yield strength or proof strength, and given by the expression

$$\sigma_{total(proof)} = M\Delta\tau_{particle} + \Delta\sigma_q \tag{27}$$

where M is the Taylor factor, $\Delta\tau_{particle}$, is the total particle strengthening contribution, $\Delta\sigma_q$ is the as-quenched strength (the 0 hour artificial aging time strength), and $\sigma_{total(proof)}$ is the total mechanical strengthening of the alloy. The total particle strengthening contribution, $\Delta\tau_{particle}$, of the total mechanical strength represents the all the strengthening contribution from all the particles such as order, Orowan, etc. The as-quenched strength is given by

$$\Delta\sigma_q = \Delta\sigma_{ss} + \Delta\sigma_{gs} + \Delta\sigma_i \tag{28}$$

where $\Delta\sigma_i$ is the baseline intrinsic lattice strength i.e., the Peierls-Nabarro strengthening contribution, $\Delta\sigma_{ss}$ is the solid solution strengthening contribution, and $\Delta\sigma_{gs}$ is the grain size strengthening contribution. The Peierls-Nabarro strengthening is the inherent lattice resistance to dislocation glide and represents the minimum stress necessary to move a rigid dislocation by one lattice constant. The Peierls stress is negligibly small, $\Delta\sigma_i \approx 10^{-5}G$, for face centered-cubic (FCC) and close-packed hexagonal (CPH) metals with their close packed, 'smeared out' atoms {28}. Close packed planes are the most widely separated and thus shear most easily and thus have the lowest Peierls stress. Fleischer {62} developed a model to estimate the amount of solid solution hardening based on the interaction

of a screw dislocation with a substitutional impurity atom. Fleischer {62} estimated the critical shear stress τ_c from solution hardening to be equal to

$$\tau_c = \{K_{max}^{1.5} C_F^{0.5}\}/\{b(6E_L)^{0.5}\} \quad (29)$$

where K_{max} is the maximum obstacle strength, C_F is the areal concentration of solute atoms, b is the Burgers vector, and E_L is the line energy for the dislocation. The grain size strengthening is generally expressed through the Hall-Petch {64,65} relation which can be expressed as

$$\sigma_{hp} = \Delta\sigma_i + \Delta\sigma_{gs} = \Delta\sigma_i + k_{hp}D^{-0.5} \quad (30)$$

where σ_{hp} is the Hall-Petch grain size strengthening for the Hall-Petch model, $\Delta\sigma_i$ is the intrinsic lattice strengthening representing the overall resistance of the crystal lattice to dislocation movement, k_{hp} represents the Hall-Petch constant which measures the relative hardening contribution of the grain boundaries, and D is the average grain size diameter. Thus, in the Hall-Petch relation given in equation {32}, the term $k_{hp}D^{-0.5}$ represents the actual grain size strengthening contribution ($\Delta\sigma_{gs}$).In addition, the grain size strengthening contribution is very small for order hardened alloys such as Ni-Al and Al-Li alloys.

Predicting the Strengthening Response for the Ni-Al Alloy System

A nickel-aluminum alloy was selected to demonstrate the method of predicting the precipitation strengthening response based on the microstructure and the appropriate strengthening mechanisms in terms of the heat treatment. Nickel-aluminum alloys are strengthened by the ordered face centered cubic γ' (Ni_3Al) precipitates, which have a superlattice crystal structure of the type $L1_2$ {31}, similar to the fcc δ' (Al_3Li) phase in binary Al-Li alloys. Ni_3Al has the Cu_3Au structure in which the nickel atoms occupy the face centers and the aluminum atoms the corners. The γ' (Ni_3Al) precipitate particles in aged Ni-Al are not the same morphology as the spherical δ' (Al_3Li) phase in aged Al-Li, but rather appear to be cube shaped resembling orthogonally diced cubes. Although the γ' (Ni_3Al) particles are cubical in shape rather than spherical, the tendency toward a cubical shape increases only as the particle size increases. In addition, there is a small lattice mismatch of $\varepsilon \cong 0.35\%$ {11} of the cubical shaped Ni_3Al particles of edge length "a" with the face centered cubic γ (nickel solid solution) matrix. Only when the misfit is greater than 0.1%, as in the Ni-Al or in Ni-Ti or Ni-Si systems, does γ' appear as cubes. In Ni-Al alloys when the ordered precipitates are sheared during the deformation process, the dislocations responsible for plastic flow travel in pairs, with the trailing dislocation restoring the disorder on the {111} slip planes created by the leading one. Gliding dislocations are split in their easy {111} glide plane into two superpartials dislocations with ½<110> burgers vectors and an antiphase boundary in between them. But since the antiphase boundary on {100} is energetically more favorable than on {111}, dislocations may locally cross-slip to {100}. In Ni-Al alloys, screw dislocations are known to cross-slip from {111} onto {100} planes. In addition, the antiphase boundary is smaller on {100} than on

{111} planes. Huther and Reppich {47} have demonstrated that the dislocation particle interaction mechanism can change from pairwise particle cutting to Orowan looping of single dislocations, depending upon the particle shape.

In order to accurately predict the strength of a precipitation strengthened Ni-Al alloy, the various maicrostructural parameters must first be accurately determined. The microstructural variables include the antiphase boundary energy, the dislocation line tension, the matrix Burger's vector, the matrix shear modulus, the Taylor or texture factor, and the looping diameter. Some of these microstructural parameters are very difficult to determine but are necessary in predicting the variation in strengthening with artificial aging treatment. In addition, contributions such as the solid solution strength, grain size strength, and the intrinsic lattice strength need evaluated.

Determining the Antiphase Boundary Energy

Critical to determining the strength of a Ni-Al alloy is an accurate determination of the antiphase boundary energy on {111}, γ_{apb}, of the ordered precipitates. Various investigators have estimated the antiphase boundary energy on {111} of Ni-Al either by experimental techniques or by theoretical models. The antiphase boundary energy can be estimated from measurements of dislocation pair spacings {38}, or from phase diagram information {39}, or from theoretical approximations {36,37}. Thus, the antiphase boundary energy is determined indirectly, and its usefulness depends on the accuracy of the approach used to estimate it from the experimental data if bulk quantities of the ordered phases can be prepared. For example, Douin, Veyssiere, and Beauchamp {40} measured the distances separating superlattice dislocations using weak-beam transmission electron microscopy on isolated dislocations and included the effect of elastic anisotropy on the dislocation line energy and found γ_{apb} for Ni_3Al polycrystals to be 0.111 ± 0.015 Jm^{-2}. The previous calculations of Veyssiere, Douin, and Beauchamp {41} did not include the effects of anisotropy and resulted in a value of the antiphase boundary energy of 0.180 ± 0.030 Jm^{-2}. Other approaches to estimate γ_{apb} have been taken. For example, Khachaturyan and Morris {39} have shown that the antiphase boundary energy can be calculated from thermodynamic information. Beauchamp, Douin, and Veyssiere {42} have shown that the antiphase boundary energy depends upon orientation in the $L1_2$ alloys with different glide antiphase boundary energies on {001} and {111}. Glazer and Morris {36} presented a method for estimating γ_{apb} based on estimating the critical planar particle diameter of the precipitates, d_{loop}, at which the transition from particle shearing to Orowan looping occurs, and using the theoretical results of Bacon, Kocks, and Scattergood {43}. Glazer and Morris {36} predicted the antiphase boundary energy for Ni_3Al to be $\gamma_{apb} = 0.102\pm0.035$ Jm^{-2}. However, Glazer and Morris {36} used an overestimated predicted value for looping diameter in their calculations which resulted in a underestimated prediction for γ_{apb}. Ardell and Huang {37} theoretically predicted an average value of the antiphase boundary energy on the {111} slip planes to be equal to 0.161 ± 0.024 Jm^{-2} for the γ' (Ni_3Al) precipitate phase. Analysis of the strengthening by γ' precipitates of underaged binary Ni-Al single crystals by Ardell {11} yielded the result of $\gamma_{apb} = 0.140$ Jm^{-2}. In addition, Munjal and Ardell {25} supported a value of $\gamma_{apb} = 0.140$ Jm^{-2} for Ni-6wt.%Al. The theoretical calculations of Chen, Voter, and Srolovitz {44} for γ_{apb}

of Ni_3Al-γ' phase yielded a value of $\gamma_{apb} = 0.142$ Jm^{-2}. The data of Dimiduk, Thompson, and Williams {45} suggest that the antiphase boundary energy is only weakly dependent upon alloy composition and that this moderate dependence is of essentially no consequence in explaining the behavior of Ni_3Al Alloys {45}. Ardell, Munjal, and Chellman {31} determined the antiphase boundaries energies of Ni-7.8wt.%Al, Ni-7.92wt%Al, and Ni-8.12wt.%Al to be 0.156, 0.129, and 0.129 Jm^{-2} respectively. Phillips {33} determined an antiphase boundary energy of 0.182 Jm^{-2} for a Ni-6.27wt.%Al alloy. Nembach and Neite {59} evaluated the antiphase boundary energy of a Ni-6.05wt.%Al and Ni-6.59wt.%Al to be 0.182 and 0.188 Jm^{-2} respectively.

Determining the Orowan Looping diameter

As the precipitates grow and coarsen with aging time, eventually they will reach a critical particle size where they will be too large in size to be sheared by the dislocations. There is a definite average size below which particles are sheared by dislocations and above which the dislocations pass around the particles and form a loop {46,50}. The looping diameter can be determined from transmission electron microscopy of plastically deformed material or from theoretical calculations. Thus, Orowan looping occurs when the size of a particle has become greater than a critical size, referred to as either the looping diameter (d_{loop}) or the looping radius (r_{loop}). At this size the force required by the dislocations to shear a particle is greater than the force required to loop a particle. Experimental values for the Orowan looping radius can be measured from the size of the smallest Orowan loops observed by TEM on plastically deformed alloys or deduced from the average particle size of a the corresponding size distribution at the peak-aged condition. The looping diameter is not easy to determine experimentally, since Orowan loops can occur in underaged alloys. Even if the particles are strong enough to support bypassing by the Orowan mechanism, they may be unable to support an Orowan loop if the bypassing events have occurred {37}. Thus, the looping diameter is most likely to be smaller than the sizes observed experimentally. In determining appropriate values for d_{loop}, the reported values of the average particle size diameter in the peak-aged condition, d_{max}, can be multiplied by $\pi/4$, so therefore {37}

$$d_{loop} = \pi\, d_{max}\, /4 \tag{31}$$

This converts the measurements of the average particle size in the peak-aged condition to average planar diameters, which is what the looping diameters actually are. For a Ni-6.0wt.%Al alloy, Munjal and Ardell {63} determined the looping diameter to be 15.7 nm. Also Ardell, Munjal, and Chellman {31} determined looping diameters of 15.7 nm, 19.9 nm, and 19.9 nm for Ni-7.8wt.%Al, Al-7.92wt.%Al, and Ni-8.12wt.%Al alloys respectively. Phillips {33} determined a looping diameter of 11.8 nm for a Ni-6.27wt.%Al. Nembach and Neite {59} found a looping diameter of approximately 11.8 nm for both Ni-6.59wt.%Al and Ni-6.05wt.%Al alloys.

Determining the Texture Factor

The Taylor factor is frequently used to express the flow stress or yield stress measured in a polycrystal in terms of the single crystal critical resolved shear stress. It is an average orientation factor which depends on the texture of

the material and the crystallographic nature of the assumed slip systems. According to Taylor {49} the selection of five slip systems out of twelve in face centered cubic polycrystals is necessary to produce a total strain which corresponds to the least work of deformation. The Taylor Model {49} can be derived from the single crystal shear strength given by

$$\tau_{crss} = \sigma_y / M \tag{32}$$

where τ_{crss} is the total single crystal critical resolved shear stress, M is a grain orientation factor which is the inverse of the Schmidt factor, and σ_y is the total polycrystalline flow stress. Using the Von Mises compatibility condition and assuming the same deformation for all of the grains, Taylor calculated M = 3.06. An applied Taylor factor of 3.0 is in accordance with the work of Melander and Persson {48} who determined a value of 3.03 for the Taylor factor of nickel bar strengthened by γ' precipitates by direct measurements of the texture using Mo K_α radiation, to convert the single crystal to polycrystalline yield strength. Nembach and Chow {26} applied a Taylor factor of 3.06 for an aged nimonic alloy PE16. Since in a polycrystalline material slip occurs on several slip systems simultaneously, the stress-strain curves should be correlated with the highly symmetrical <111> and <100> orientations rather than the orientations of easy glide. Dislocation glide is more difficult with <111> and <100>.

Determining the Growth Rate Constant and Volume Fraction for γ' (Ni_3Al) Precipitates

The growth rate constant K_c and volume fraction of γ' in Ni-Al alloys has been determined by various investigators. Ardell {52} has indicated that the coarsening rate of the γ' particles is independent of the volume fraction. The growth rate constants K_c for γ' particle coarsening was determined by Ardell and Nicholson {53} for both Ni-6.71wt.%Al and Ni-6.35wt.%Al alloys which were in excellent quantitative agreement of the Lifshitz-Slyozov-Wagner theory of Ostwald ripening. For Ni-6.35wt.%Al, Ardell and Nicholson {53} found a growth rate constant of $6.13(10^4)$ A^3/hr at 625°C, and $2.09(10^6)$ A^3/hr at 715°C. For Ni-6.71Al alloys, Ardell and Nicholson {53} found a growth rate constants of $5.77(10^4)$ A^3/hr at 625°C, $1.95(10^6)$ A^3/hr at 715°C, $4.60(10^6)$ A^3/hr at 750°C, and $1.04(10^7)$ A^3/hr at 775°C. The volume fraction of the γ' particles in aged Ni-Al has been found to be influenced by aging temperature. For example, for a Ni-6.35wt.%Al alloy, Ardell and Nicholson {53} found the volume fraction to be 0.145 at 625°C and 0.091 at 715°C. This is consistent with the results of Tsumuraya and Miyata {54} which reported a value of $f_{\gamma'} = 0.096$ for a Ni-6.31Al alloy for the value of the volume fraction for γ' aged at 650°C for 120 hours. In addition, Ardell and Nicholson found the volume fraction of γ' in Ni-6.71wt.%Al alloy to be 0.198 at 625°C, 0.148 at 715°C, and 0.102 at 775°C.

Determining the Solid Solution Strengthening Contribution and the Grain Size Effect

The solid solution strengthening for Ni-Al has been estimated by various investigators. The room temperature value of the single crystal solid solution strengthening contribution τ_{ss} of Ni-6wt.%Al is approximately 41 MN/m^2 (MPa) {25}. As a comparison, Nembach and Chow {26} estimated the total matrix yield strength i.e., the total combined

strengthening from solid solution hardening, grain size and Peierls strengthening, based on Labusch's {13} theory of solid solution hardening, to be 58.7 MN/m^2 (MPa) for nimonic alloy PE 16 with a volume fraction of 0.129, and 63.0 MN/m^2 (MPa) for a volume fraction of 0.082. The grain size effect on the strengthening in Ni-Al alloys is also rather small due to the small value of the Hall-Petch coefficient. For example, the grain size effect on the CRSS can be neglected in view of a Hall-Petch analysis of Nickel base alloy Nimonic PE 16, which yields a grain size effect on the CRSS of less than 1.7 N/mm^2 {51}.

Determining the shear Modulus G and Burgers Vector

Ardell {11} estimated the shear modulus G on the slip plane in the slip direction of the matrix ({111} glide planes) to be approximately 59.3 GN/m^2 for Ni-Al alloys, and a Poission's ratio of $\underline{\nu} = 0.374$. Pottebohn, Neite, and Nembach {55,66} derived for stiffness constants to estimate the shear modulus of a Ni-8at.%Al solid solution to be approximately 71.9 GPa, and also found the shear modulus for a Nimonic all PE16 to 65 GPa. Phillips {32} measured the shear modulus of Ni-12.7at.%Al at 77°K to be 79.2 GPa, and at 303°K to be 75.2 Gpa. Thomas and Brooks {60} used a average value of the shear modulus of 73.9 GPa. Glazer and Morris {36} determined from elastic constants for Ni_3Al at 650°C a value for G of 38.6 GPa and a value of $\underline{\nu} = 0.404$ for the Ni_3Al phase of Ni-Al. The Burgers vector in Ni-Al varies slightly with composition, but is approximately 0.255 nm {36}.

Determining the Average Particle Size and Particle Size Distribution

To develop a model for predicting the precipitation hardening response, the average particle size for each aging time and temperature along the age hardening curve must be accurately determined. The particle size can be determined from the Lifshitz-Slyozov-Wagner {56,57} cubic coarsening theory and can be expressed by the relation {56,57}

$$r^3 - r_o^3 = K_c(t-t_o) \tag{33}$$

where r_o is the initial particle size, at $t_o=0$, and can taken as approximately equal to zero for the Ni-Al alloy. The growth rate constant K_c is given by the expression {56,57}

$$K_c = 8\gamma V_m^2 C_{eq} D/9RT \tag{34}$$

where γ is the interfacial energy, V_m is the molar volume of the precipitate or dispersed phase, C_{eq} is the equilibrium concentration or solubility limit of the solute in the matrix, D is the diffusion coefficient or the diffusivity of the rate-controlling solute, R is the universal gas constant (8.314 J/mole °K), and T is the temperature. The cubic coarsening expression can thus be expressed by the expression {56,57}

$$r^3 - r_o^3 = (8\gamma V_m^2 C_{eq} Dt)/(9RT) \tag{35}$$

where t is the aging heat treatment time. The diffusion coefficient D can be expressed as {58}

$$D = D_o \exp \{-Q_A/RT\} \quad (36)$$

where Q_A is the activation energy for diffusion, and D_o is the diffusion coefficient temperature independent exponent. For Ni-Al, D_o is approximately 1.87 cm^2/sec {58}, and Q_A is approximately 63,900 Cal/mole (267.8 KJ/mole) {58}. This compares with the results of Ardell and Nicholson {35} who determined an activation energy of 62,600 cal/mole (262.2 KJ/mole) for γ' particle growth in a Ni-6.27wt.%Li. Ardell and Nicholson {53} also found for aged Ni-6.71wt.%Al the activation energy to be equal to 64,400 cal/mole (269.6 KJ/mole). In addition, the growth rate constant which is dependent of the value of D has been determined for some Ni-Al alloys. Ardell and Nicholson {35} determined the growth rate constant for Ni-6.35wt.%Al to be $2.09(10^6)$ A^3/hr at 715°C.

Predicting the Peak-Strength

The predicted strengthening for the Ni-6.27wt.%Al alloy can be can be compared with the experimental tensile data obtained from Phillips {32,33} for a Ni-6.27wt.%Al (Ni-12.7at.%Al) alloy aged at either 600°C or 700°C for a given aging time. The experimental peak yield strength of 363.7 MPa (52,500 psi) from Phillips {32,33} for the Ni-6.27wt.%Al alloy was for artificial aging at 700°C for 5.5 hours with an average particle cube edge length of 148 Angstroms. The as-quenched strength or zero hour aging time strength for the alloy is 175.9 MPa {32} which represents the contributions from solid solution, grain size, and intrinsic lattice strengthening.. In Ni-Al alloys the average particle size is taken as the half the average particle cube edge length. Thus the corresponding average particle size radius would be 74 Angstroms. Using equation (5a), the average interparticle spacing was determined to be 2.3 Angstroms. The constant dislocation line tension approximation, given by $\Gamma = Gb^2/2$ {51} for Ni-Al alloys {51}, can be used for the calculations. Thus, for the Ni-6.27wt.%Al alloy Γ is approximately 2.458E-9 Newtons. The Poissons ratio is $\nu = 0.374$ {11}, the shear modulus is G = 75.2 GPa {32}, the antiphase boundary is $\gamma_{apb} = 0.182$ J/m^2 {21,32,59}, and the Burgers vector is b = 0.255 {36}. The contribution to the strengthening in the peak-aged condition is predominately due to dislocation particle shearing by the order hardening mechanism and given by equation (24). In both the underaged and the peak-aged conditions, precipitates would be predominately sheared by the dislocations and thus the particle strengthening would be controlled by order strengthening. Therefore equation (24) can be used to evaluate the particle strengthening contribution in the underaged up to the peak-aged condition. Thus, using equation (24), the value for the critical resolved shear strength due to order strengthening was determined to be 62.2 MPa. The order strengthening contribution can then be multiplied by a Taylor texture factor of 3.1 {49} to obtain a value of 192.8 MPa for the polycrystalline age-hardening contribution. The as-quenched strength of 175.9 MPa can be added the value of 186.6 MPa to determine a total mechanical peak-strength of 368.7 MPa predicted for the Ni-6.27wt.%Al alloy, which is in agreement with the experimental yield strength

value of 363.7 MPa for the Ni-6.27wt.Al alloy. Thus, equations (27) and (28) along with equations (24) and (33) were used to determine the total mechanical peak strength of the Ni-6.27wt.Al alloy.

Extensions of Model to Other Alloy Systems

The current modeling effort will be extended to several different high temperature discontinuously reinforced aluminum (DRA) alloys that also achieve substantial strengthening through intermetallic strengthening precipitates such as γ' (Ni_3Al). These particular alloys also provide reduced cost and weight of aircraft fan and compressor components when compared with various titanium aerospace alloys often used for these components. Also when compared with polymer matrix composites, high temperature discontinuously reinforced aluminum alloys (for the F119 Joint Strike Fighter Aircraft Program) have reduced cost and improved reliability of fan and compressor components. High temperature discontinuously reinforced aluminum alloys consist basically of a three part microstructure. The microstructure includes an aluminum matrix, a precipitated second phase for mechanical strengthening and creep resistance, and a particulate reinforcement. The particulate reinforcement is most often silicon carbide (SiC) for stiffness/modulus improvement, creep resistance, and fatigue failure resistance. An important feature of this research effort is to study the effect of the precipitated second phase on the overall mechanical strength of the alloy, and to better understand the strength of alloy from the microstructure in terms of the precipitated second phase. However, the microstructure is slightly more complex in that the precipitated phase is a combination of a distribution of large precipitate particles and a distribution much smaller precipitate particles. Thus, the precipitated second phase is often a bimodal particle size distribution of the dispersoid. Bimodal particle size distributions can be produced by double aging, but with a difference in temperatures large enough so that a new population of particles nucleates, grows, and coarsens independently at the lower aging temperature. As a result of the bimodal particle size distribution, the particles of the smaller size distribution may be sheared by the dislocations rather than bypassed via. Orowan looping. However, the situation may be more complex in that some of the particles of the smaller size distribution may be sheared while others may be looped. However, if all of the particles of the smaller size distribution are larger than the Orowan looping radius, then they will not be sheared but rather bypassed by the dislocations. Thus, accurately knowing the Orowan looping radius size is critical for assessing the dislocation particle interactions mechanism(s) when modeling the strengthening of the alloy. An important related consideration is how does the strength of the matrix and the particles add together.

Summary and Conclusions

A Nickel-Aluminum alloy strengthened by γ' (Ni_3Al) intermetallic ordered coherent precipitates with a small misfit strain was used a demonstration material to develop a model to predict strengthening behavior during plastic deformation as a consequence of the γ' particles acting as obstacles to the dislocations and thus impeding their glide motion through the alloy. It was determined that the two most dominate strengthening mechanisms in the Ni-Al system were order hardening when the particles were smaller than the critical looping radius, and Orowan strengthening when the particles were larger than the looping radius. In the overaged condition when the particles

are large in size, the dislocations bypass and loop the particles by the Orowan mechanism. In the underaged to peak aged conditions where the particles are usually smaller than the looping radius, the dislocations shear the precipitates during deformation. The total polycrystalline yield strength included contributions from the intrinsic lattice strength, the solid solution strengthening, grain size strengthening, and particle strengthening which included the order hardening and Orowan strengthening contributions. The total mechanical yield strength for a Ni-6.27wt.%Al alloy was predicted for the peak-aged condition based on the theory for order strengthening by Brown and Ham {12,24}, and was found to be in good agreement with the experimental peak-strength data for Ni-6.27Al. The theoretical work described provides the basis for future modeling work with much more complex alloy systems such as high temperature discontinuously reinforced aluminum alloys (HTDRA). Thus, the current model provides a foundation for the development of a more detailed model which would incorporate the microstructural parameters and processing variables of alloys with very complex microstructures, such as various high temperature superalloys.

References

{1} E. Orowan, *Symposium on Internal Stresses in Metals and Alloys*, Session III Discussion, Institute of Metals, London, England, p. 451, 1948.

{2} J.F. Nie, B.C. Muddle, and I.J. Polmear, "The Effect of Precipitate Shape and Orientation in Dispersion Strengthening in High Strength Aluminum Alloys", *Materials Science Forum*, pp. 1257-1262, Vols. 217-222, 1996.

{3} U.F. Kocks, "A statistical Theory of Flow Stress and Work-Hardening", *Philosophical Magazine*, pp. 541-549, Vol. 13, 1965.

{4} A.J.E. Foreman, and M.J. Makin, "Dislocation Movement Through Random Arrays of Obstacles", *Canadian Journal of Physics*, pp. 511-517, Vol. 45, 1967.

{5} E.E. Underwood, *Quantitative Stereology*, Addison-Wesley Publishing Company, 1970.

{6} M.F. Ashby, and R. Ebeling, "On the Determination of the Number, Size, Spacing, and Volume Fraction of Spherical Second-Phase Particles from Extraction Replicas", *Transactions of the American Society of AIME*, Vol. 236, pp. 1396-1404, 1966.

{7} E.E. Underwood, "The Effect of a Dispersed Second Phase on the Recrystallization of Aluminum-Copper Alloys", *Journal Institute of Metals*, Vol. 92, pp. 124-125, 1964.

{8} C.W. Corti, P. Cotteril, and G.A. Fitzpatrick, "The Evaluation of the Interparticle Spacing in Dispersion Alloys", *International Metallurgical Reviews*, Vol. 19, pp. 77-88, 1974.

{9} J.E. Hillard, "The Counting and Sizing of Particles in Transmission Microscopy", *Transactions of the Metallurgical Society of AIME*, Vol. 224, pp. 906-917, 1962.

{10} P.P. Bansal, and A.J. Ardell, "Average Nearest-Neighbor Distances Between Uniformly Distributed Finite Particles", *Metallography*, Vol. 5, pp. 97-111, 1972.

{11} A.J. Ardell, "Precipitation Hardening", *Metallurgical Transactions A*, Vol. 16A, pp. 2131-2164, Dec.1985.

{12} L.M. Brown, and R.K. Ham, *Strengthening Methods in Crystals*, Edited by A. Kelly and R.B. Nicholson, John Wiley & Sons, Halsted Press Division, New York, NY, 1970.

{13} R. Labusch, "A Statistical Theory of Solid Solution Hardening", *Physical Status Solidus*, Vol. 41, pp. 659-669, 1970.

{14} G. Neite, M. Sieve, M. Mrotzek, and E. Nembach, *Deformation of Multi-Phase and Particle Containing Materials*, Proc. 4th Riso Int.Symposium on Metallurgy and Materials Science, J.B. Bilde-Sorensen, N. hansen, A. Horsewell, T. Leffers, and H. Liholt, eds., Riso National Laboratory, Roskilde, Denmark, p. 447, 1983.

{15} J.C. Huang, and A.J. Ardell, "Addition Rules and the Contribution of δ' Precipitates to Strengthening of Aged Al-Li-Cu Alloys", *Acta Metallurgica*, Vol. 36, No. 11, pp. 2995-3006, 1988.

{16} T.J. Koppenaal and D. Kuhlmann-Wilsdorf, "The Effect of Prestressing on the Strength of Neutron-Irradiated Copper Single Crystals", *Applied Physics Letters*, Vol. 4, pp. 59-61, 1964.

{17} A. Nitz and E. Nembach, "The Critical Resolved Shear Stress of a Superalloy as a Combination of Those of its γ Matrix and γ' Precipitates", *Metallurgical and Materials Transactions A*, Vol. 29A, pp. 799-807, March 1998.
{18} J. Glazer, and J.W. Morris, Jr., "The Effect of the Precipitate Size Distribution on the Aging Curve of Order Hardening Alloys", *Acta Metallurgicia*, Vol. 36, No. 4, pp. 907-915, 1988.
{19} J. Glazer, and J.W. Morris, Jr., "Strengthening Contributions of Strong Ordered Precipitates", *Philosophical Magazine A*, Vol. 56, No. 4, pp. 507-515, 1987.
{20} A.J.E., Forman, and M.J. Makin, "Dislocation Movement Through Random Arrays of Obstacles", *Philosophical Magazine A*, Vol. 14, pp. 911-924, 1966.
{21} A.J. Ardell, and J.C. Huang, "Antiphase Boundary Energies and the Transition from Shearing to Looping in Alloys Strengthened by Ordered Precipitates", *Philosophical Magazine Letters*, Vol. 58, No. 4, pp. 189-197, 1988.
{22} H. Gleiter, and E. Hornbogen, "Precipitation Hardening by Coherent Precipitates", Review Paper, *Materials Science and Engineering*, Vol. 2, pp. 285-302, 1967/68.
{23} J.W. Martin, "Micromechanisms in Particle-Hardened Alloys", Cambridge University Press, 1980.
{24} R.K. Ham, "Strengthening by Ordered Precipitates", *Transactions Journal Institute of Metals*, Vol. 9 Supplement, pp. 173-175, 1968.
{25} V. Munjal, and A.J. Ardell, "The Effect of Particle Size Distributions on the CRSS of Aged Ni-Al Alloys", *Acta Metallurgica*, Vol. 24, pp. 827-833, 1976.
{26} E. Nembach, and C. Chow, "Experimental Investigation into the Relation Between γ' Particles Dispersion and the Yield Stress of PE 16", *Materials Science and Engineering*, Vol. 36, pp. 271-279, 1978.
{27} D. Raynor and J.M. Silcock, "Strengthening Mechanisms in γ' Precipitating Alloys", *Metal Science Journal*, Vol. 4, pp. 121-130, 1970.
{28} P. Haasen, "Line Defects - Dislocations", *Physical Metallurgy*, Cambridge University Press, 2nd edition, p. 264, 1986.
{29} T.H. Sanders, Jr., "Al-Li Alloys - An Overview", *Aluminum-Lithium Alloys*, Proceedings of the 1st. International Aluminum-Lithium Conference, The Metallurgical Society of AIME, 1980.
{30} E.A Starke, Jr., "Aluminum Alloys of the 70's: Scientific Solutions to Engineering Problems", *Materials Science and Engineering*, Vol. 29, p. 99, 1977.
{31} A.J. Ardell, V. Munjal, and D.J. Chellman, "Precipitation Hardening of Ni-Al Alloys Containing Large Volume Fractions of γ', *Metallurgical transactions*, Vol. 7a, pp. 1263-1268, 1976.
{32} V.A. Phillips, "Hardening Mechanisms in a Precipitation Hardenable Nickel-12.71 at.% Aluminum Alloy", *Philosophical Magazine*, Vol. 16, pp. 103-117, 1967.
{33} V.A. Phillips, "A Metallographic Study of Precipitation in a Ni-12.7at.% Al Alloy", *Acta Metallurgica,* Vol. 14, pp. 1533-1547, 1966.
{34} J.C. Huang, "Microstructural Evolution and Strengthening Mechanisms in Two Ternary Aluminum-Lithium Copper Alloys", *Ph.D. Dissertation*, The University of California, p. 277, 1986.
{35} A.J. Ardell, and R.B. Nicholson, "On the Modulated Structure of Aged Ni-Al Alloys", *Acta Metallurgica,* Vol. 14, pp. 1295-1309, 1966.
{36} J. Glazer, and J.W. Morris, "Strengthening Contributions of Strong Ordered Precipitates", *Philosophical Magazine A*, Vol. 56, pp. 507-515, 1987.
{37} A.J. Ardell, and J.C. Huang, "Antiphase Boundary Energies and the Transition from Shearing to Looping in Alloys Strengthened by Ordered Precipitates", *Philosophical Magazine Letters*, Vol. 58, pp. 189-197, 1988.
{38} J.C. Huang, and A.J. Ardell, "Strengthening Mechanisms in Two Al-Li-Cu Alloys", *Aluminum Technology 86,* edited by T. Sheppard, The Institute of Metals, London, pp. 434-441, 1986.
{39} A.G. Khachaturyan and J.W. Morris, Jr., "The Interfacial Tension of a Sharp Antiphase Domain Boundary", *Philosophical Magazine A*, Vol. 56, No. 4, pp. 517-532, 1987.
{40} J. Douin, P. Veyssiere, and P. Beauchamp, "Dislocation Line Stability in Ni_3Al", *Philosophical Magazine A*, Vol. 54, No. 3, pp. 375-393, 1986.
{41} P. Veyssiere, J. Douin, and P. Beauchamp, "On the Presence of Superlattice Intrinsic Stacking Faults in Plastically Deformed Ni_3Al", *Philosophical Magazine A*, Vol. 51, No. 3, pp. 469-483, 1985.
{42} P. Beauchamp, J. Douin, and P. Veyssiere, "Dependence of the Antiphase Boundary Energy Upon

Orientation in the $L1_2$ Structure", *Philosophical Magazine A*, Vol. 55, No. 5, pp. 565-581, 1987.
{43} D.J. Bacon, U.F. Kocks, and R.O. Scattergood, "The Effect of Dislocation Self-Interaction on the Orowan Stress", *Philosophical Magazine A*, Vol. 28, pp. 1241-1262, 1973.
{44} S. P. Chen, A.F. Voter, and D.J. Srolovitz, "Computer Simulation of Grain Boundaries in Ni_3Al: The Effect of Grain Boundary Composition", *Scripta Metallurgica*, Vol. 20, pp. 1389-1394, 1986.
{45} D.M. Dimiduk, A.W. Thompson, and J.C. Williams, "The Compositional Dependence of Antiphase-Boundary Energies and the Mechanism of Anomalous Flow in Ni_3Al Alloys", *Philosophical Magazine A*, Vol. 67, No. 3, pp. 675-698, 1993.
{46} A. Pineau, F. Lecroisey, J.L. Castagne, and M. Sindzingre, "Yielding of an Alloy Hardened by Coherent and Ordered Precipitation", *Acta Metallurgica*, Vol. 17, p. 905, 1969.
{47} W. Huther, and B. Reppich, "Interactions of Dislocations With Coherent, Stress-Free, Ordered Particles", *Zeitschrift Fur Metallkunde*, Vol. 69, pp. 628-634, 1978.
{48} A. Melander, and P.A. Persson, "Strength of $\underline{\delta}$' Hardened Nickel-Base Alloy", *Metal Science*, pp. 391-398, Sept.1978.
{49} G.J. Taylor, "Plastic Strain in Metals", *Journal Institute of Metals*, Vol. 62, pp. 307-325, 1938.
{50} J.C. Fisher, E.W. Hart, and R.H. Pry, "The Hardening of Metal Crystals by Precipitate Particles", *Acta Metallurgicia*, Vol. 1, pp. 336-339, 1953.
{51} B. Reppich, P. Schepp, and G. Wehner, "Some New Aspects Concerning Particle Hardening Mechanisms in γ' Precipitating Nickel-Base Alloys-II Experiments", *Acta Metallurgica*, Vol. 30, pp. 95-104, 1982.
{52} A.J. Ardell, "Observations on the Effect of Volume Fraction on the Coarsening of γ' Precipitates in Binary Ni-Al Alloys", *Scripta Metallurgica et. Materialia*, Vol. 24, pp. 343-346, 1990.
{53} A.J. Ardell and R.B. Nicholson, "The Coarsening of γ' in Ni-Al Alloys", *The Journal of Physics and Chemistry of Solids*, Vol. 27, pp. 1793-1804, 1966.
{54} K. Tsumuraya, and Y. Miyata, "Coarsening Models Incorporating Both Diffusion Geometry and Volume Fraction of Particles", *Acta Metallurgicia*, Vol. 31(3), pp. 437-452, 1983.
{55} H. Pottebohm, G. Neite, and E. Nembach, "Elastic Properties (the Stiffness Constants, the Shear Modulus and the Dislocation Line Energy and Tension) of Ni-Al Solid Solutions and of the Nimonic Alloy PE16", *Materials Science and Engineering*, Vol. 60, pp. 189-194, 1983.
{56} I.M. Lifshitz, and V.V. Slyozov, "The Kinetics of Precipitation from Supersaturated Solid Solution", *Journal Physical Chemical Solids*, Vol. 19, pp. 35-50, 1961.
{57} C. Wagner, "Theories Associated With Age Hardening and Overaging During Ostwald Ripening", *Zeitschrift fur Electrochemie*, Vol. 37, pp. 581-591, 1931.
{58} R.A. Swalin, and A. Martin, "Solute Diffusion in Nickel-Base Substitutional Solid Solutions", *Transactions of the AIME Journal of Metals*, Vol. 206, pp. 567-572, 1956.
{59} E. Nembach, and G. Neite, "Precipitation Hardening of Superalloys by Ordered γ' Particles", *Progress in Materials Science*, Vol. 29, pp. 177-319, 1985.
{60} A.W. Thompson, and J.A. Brooks, "The Mechanism of Precipitation Strengthening in an Iron-Base Superalloy", *Acta Metallurgicia*, Vol. 30, pp. 2197-2203, 1982.
{61} A.W. Thompson, *Material Science and Engineering*, Vol. 14, pp. 253-264, 1974.
{62} R.L. Fleischer, "Substitutional Solution Hardening", *Acta Metallurgicia*, Vol. 11, pp. 203-209, 1963.
{63} V. Munjal, and A.J. Ardell, "Precipitation Hardening of Ni-12.19at.%Al Alloy Single Crystals", *Acta Metallurgica*, Vol. 23, pp. 513-519, 1975.
{64} E.O. Hall, "The Deformation and Ageing of Mild Steel: III Discussion and Results", *Proceedings of the Physical Society of London*, Vol. 64(B), pp. 747-753, 1951.
{65} N.J. Petch, "The Cleavage Strength of Polycrystals", *Journal of the Iron and Steel Institute*, Vol. 173, pp. 25-28, 1953.
{66} E.Z. Vintaikin, "X-Ray Determination of the Elastic Constants of Pure Nickel and its Alloy With Aluminum", *Soviet Physics-Doklady*, Vol. 11, No. 1, pp. 91-93, 1966.
{67} B.H. Kear, and H.G.F Wilsdorf., "Dislocation Configurations in Plastically Deformed Polycrystalline Cu_3Au Alloys", *Transactions of the Metallurgical Society of AIME*, Vol. 224, pp. 382-386, 1962.
{68} B. Reppich, "Some New Aspects Concerning Particle Hardening Mechanisms in γ' Precipitating Ni-Base Alloys-I. Theoretical Concept", *Acta Metallurgica*, Vol. 30, pp. 87-94, 1982.

SYNTHESIS AND CHARACTERIZATION OF METAL-THIOACID AND - DIHYDROGEN PHOSPHATE COMPLEXES USEFUL AS NONLINEAR (NLO) MATERIALS

Zewdu Gebeyehu
Associate Professor
Department of Chemistry

Tuskegee University
Tuskegee, AL 36088

Final Report for:
Summer Faculty Research Program
Wright Laboratory

Sponsored by:
Air Force Office of Scientific Research
Wright-Patterson Air Force Base, OH

and

Wright Laboratory

August 1998

SYNTHESIS AND CHARACTERIZATION OF METAL-THIOACID AND - DIHYDROGEN PHOSPHATE COMPLEXES USEFUL AS NONLINEAR (NLO) MATERIALS

Zewdu Gebeyehu
Associate Professor
Department of Chemistry

Abstract

Several complexes of chromium, copper, zinc, cadmium, mercury and lead were prepared by the reaction of aqueous solutions of the corresponding metal salts with xanthates, dithiocarbamate and dihydrogen phosphate as ligands, and their nonlinear optical properties studied. Attempts were also made to synthesize complexes with mixed xanthate-dithiocarbamate ligand around a single metal center. The synthesis of the complexes were carried out by mixing one equivalent of a metal salt with two equivalents of the ligands at room temperature or at around 80°C. The complexes isolated were characterized mainly by IR-spectroscopy and in few cases by elemental analysis. The nonlinear optical (NLO) properties of these complexes were screened by powder test using Kurtz's method. Complexes obtained from the reaction of mercury salt with xanthate ligands; mercury, cadmium and lead slats with mixed xanthate-dithiocarbamate ligands, and dihydrogen phosphate complexes of chromium, zinc and lead were found to show moderate to intense nonlinearities, although those of mercury and lead showed low damage thresholds by tending to burn under the heat of the laser beam. Most complexes obtained from the reactions of metal salts with dithiocarbamate ligand showed no NLO properties.

SYNTHESIS AND CHARACTERIZATION OF METAL-THIOACID AND - DIHYDROGEN PHOSPHATE COMPLEXES USEFUL AS NONLINEAR (NLO) MATERIALS

Zewdu Gebeyehu

Introduction

Nonlinear optical materials are of paramount importance for a variety of both commercial and military applications, among which are high resolution spectroscopy, remote sensing, telecommunications, data transmission and processing, and real time holography {1-3}. The key to turning device concepts into reality is the discovery and development of materials which have sufficient optical nonlinearities combined with other properties such as high transmission at the wavelengths of interest, physical, chemical, and thermal stability, and high damage thresholds. In addition, the materials must be reasonably easy to prepare and process and also have a reasonable cost.

The necessary criteria for a material to exhibit large second order optical non-linearity are that it should possess highly polarisable group, or contain conjugated π-system, or crystallize in noncentrosymmetric space group {4-7}. Based on these criteria, we have attempted to synthesize complexes which may have the desired properties of NLO activity. The synthesis involved the reaction of main group or transition metal salts with some common reagents including thioacid ligands such as xanthates and dithiocarbamates. The reason for using thioacid ligands is that sulfur containing ligands are very well known to complex with metals easily thereby resulting in higher molecular polarization. We have also investigated the reaction of mixed xanthate-thiocarbamate ligands, which are similar structurally and electronically, with the metal salts in order to synthesize complexes with the two ligands around a single metal center. The reaction of such mixed ligands with metal salts may result in compounds which can crystallize in noncentrosymmetric space group. Similarly we have investigated the reactivity of some metal salts with potassium dihydrogen phosphate (KDP), a well known nonlinear optical material {8}, in order to see how the NLO property of this compound is affected by replacing the potassium ion by some other metals. After isolating the products of each reaction, the NLO properties of the substances were studied using powder test.

Experimental

The preparation of all compounds were carried out in air using water as a solvent for all the synthesis. All metal salts MX_2 (M= Zn, Cd, Hg, Pb; Cu, X = Cl, CH_3COO^-) were purchased from Fluka or Aldrich. $CrCl_3.6H_2O$ and potassium dihydrogen phosphate were obtained from Fisher. The xanthic acid potassium salts were prepared by slow addition of CS_2 to a solution of KOH in the corresponding alcohol or purchased from Aldrich. Sodium dimethyldithiocarbamate was purchased from Fluka. IR-spectra were recorded from 4000 - 370 cm^{-1} on Perkin-Elmer FTIR spectrometer using KBr. Microanalysis was done in the analytical section of Wright Laboratory. Kurtz method was used to determine NLO activity in Wright Laboratory using ND: YAG laser (1.06 mm).

I. Reaction of Mercuric acetate with potassium xanthates

1. Reaction with ethylxanthate, potassium salt

a) 1.00 g (3.13 mmol) of $Hg(CH_3COO)_2$ was dissolved in about 30 mL of distilled water. To this solution, 1.05 g (6.55 mmol) of $KS_2C\text{-}O\text{-}C_2H_5$, dissolved in 20 mL of water, was added. This resulted in the formation of a white precipitate initially, which changed yellow after adding all of the xanthate salt. On heating the reaction mixture at 80°C for digestion, the color of the precipitate changed to dark gray. This was filtered, washed three times with distilled water and dried. The substance does not dissolve in any solvent. Yield = 0.878 g (63.2%). The nonlinear optical property of the product was investigated using Kurtz powder test and showed no second harmonic generation.

b) 0.50 g (1.57 mmol) of $Hg(CH_3COO)_2$ and 0.657 g (4.10 mmol) of $KS_2C\text{-}O\text{-}C_2H_5$ were reacted as above at room temperature. The mixture was not heated. The formed yellow precipitate was filtered, washed with water and dried. The substance was found to be insoluble in water and organic solvents. Yield = 0.515 g (74.10%). Unlike the product in a, the product showed a second harmonic generation, but tends to burn by the laser beam.

2. Reaction with isopropyl xanthate, potassium salt

0.25g (0.78 mmol) $Hg(CH_3COO)_2$ and 0.41 g (2.35 mmol) potassium isopropylxanthate, $KS_2C\text{-}O\text{-}C_3H_7$ were reacted after dissolving each in 15 mL of distilled water. A yellowish precipitate formed, which was filtered, washed with water and dried. Substance was insoluble in most organic solvents. Yield = 0.30 g (81.08%). The product isolated was tested for its nonlinear optical property. It showed a second harmonic property, but burns under the heat of the laser light.

II. Reaction of metal salts with sodium dimethyldithiocarbamate

1. Reaction of mercuric acetate with sodium dimethyldithiocarbamate

0.508 g (1.59 mmol) of $Hg(CH_3COO)_2$ was dissolved in 20 mL of distilled water. To this solution, 0.700 g (4.88 mmol) of sodium dimethyldithiocarbamate dissolved in 15 mL of water was added while stirring the solution. A white precipitate formed. The reaction mixture was stirred for one more hour and then filtered. The white precipitate was washed three times with water and dried. It was found to be insoluble in most solvents. Yield = 0.560 g (80.0%). The product isolated was tested for its nonlinear optical property. It showed no second harmonic property, but does not tend to burn under the heat of the laser light.

2. Reaction of cadmium chloride with sodium dimethyldithiocarbamate

0.25 g (1.36 mmol) $CdCl_2$ was dissolved in 20 mL of distilled water. To this solution, 0.45 g (3.14 mmol) of sodium dimethyldithiocarbamate dissolved in 30 mL of water, was added slowly while stirring. A white precipitate formed. The reaction mixture was heated at 80°C for one hour and filtered. The precipitate was washed with water and dried. The substance does not dissolve in any solvent. Yield = 0.212 g (44.25%). Second harmonic generation test showed negative result.

3. Reaction of lead acetate with sodium dimethyldithiocarbamate

0.250 g (0.659 mmol) $Pb(Ac)_2.3H_2O$ was dissolved in 20 mL of water and mixed with an aqueous solution of 0.230g (1.61 mmol) of sodium dimethyldithiocarbamate. A white precipitate formed immediately. This was filtered, washed and dried. The substance is not soluble in any solvent. Yield = 0.248g (84.35%). The product showed a second harmonic generation, but decomposed under the heat of the laser.

4. Reaction of zinc chloride with sodium dimethyldithiocarbamate

0.360 g (2.64 mmol) of $ZnCl_2$ was dissolved in 15 mL of water. To this solution, 0.850 g (5.94 mmol) of sodium dimethyldithiocarbamate dissolved in 15 mL of water was added. White precipitate formed immediately. This was filtered, washed with water three times and dried. The solid is little soluble in acetone and most organic solvents. Yield = 0.320 g (39.7%). The powder showed no frequency conversion.

III. Reactions of metal salts with mixed xanthate-thiocarbamate ligands

1. Reaction of $CdCl_2$ with potassium isopropylxanthate-sodium dimethyldithiocarbamate mixture.

0.250 g of $CdCl_2$ (1.36 mmol) was dissolved in 15 mL of water. To this solution, a solution of 0.238 g(1.36 mmol) potassium isopropylxanthate and 0.195 g (1.36 mmol) sodium dimethyldithiocarbamate mixture was added while stirring at room temperature. A yellowish precipitate was formed. The product was digested at 90°C for one hour and then filtered. Unlike the cadmium xanthate complex, this product is not or little soluble in most organic solvents. Yield = 0.485 g (97.19%). Product showed strong second harmonic generation

2. Reaction of $Hg(CH_3COO)_2$ with potassium isopropylxanthate-sodium dimethyldithiocarbamate mixture

0.250 g (0.784 mmol) $Hg(CH_3COO)_2$ was dissolve in 15 mL of water. A mixture of 0.139 g (0.784 mmol) potassium isopropylxanthate and 0.115 g (0.784 mmol) sodium dimethyldithiocarbamate dissolved in 20 mL of water was added to the mercuric salt solution drop by drop. A white precipitate formed initially which changed to light yellow, then yellowish gray and finally to dark solid. This was filtered, washed three times with water and acetone and dried. The product was found to be insoluble in most solvents. Yield = 0.220 g (61.62%). There is a frequency conversion, but it tends to burn.

3. Reaction of $ZnCl_2$ with potassium isopropylxanthate-sodium dimethyldithiocarbamate mixture

0.250 g (1.83 mmol) of $ZnCl_2$ dissolved in 20 mL of water was reacted with a mixture of aqueous solution of 0.320 g (1.83 mmol) potassium isopropylxanthate and 0.262 g (1.83 mmol) sodium dimethyldithiocarbamate. White precipitate formed, which was filtered, washed with water and dried. The precipitate is slightly soluble in acetone and other organic solvents. Yield = 0.320 g (54.6%). No second harmonic generation was observed.

4. Reaction of lead acetate with potassium isopropylxanthate-sodium dimethyldithiocarbamate mixture

A 20 mL of aqueous solution of 0.250 g (0.659 mmol) $Pb(Ac)_2.3H_2O$ was reacted with a mixture of aqueous solutions of 0.115 g (0.659 mmol) potassium isopropylxanthate and 0.095 g (0.659 mmol) sodium dimethyldithiocarbamate. The formed yellowish-white precipitate was filtered, washed with water three times and dried. The substance is little soluble in THF, acetone and CH_2Cl_2. Yield = 0.304 g (99.8%). The substance showed a frequency conversion, but tends to burn by the laser beam.

IV. Reaction of metal salts with potassium dihydrogen phosphate (KDP)

1. Reaction of cadmium chloride with potassium dihydrogen phosphate

0.270 g (1.47 mmol) of $CdCl_2$ was dissolved in 20 mL of water. To this solution 0.450g of (3.31 mmol) KDP, dissolved in 35 mL of water was added drop by drop while stirring. No precipitate formation was observed even after adding all the KDP solution. The reaction mixture was heated for one hour at 80°C. A white precipitate in small yield formed. This was filtered, washed with water and dried. Concentrating the filtrate by evaporation and cooling did not give more of the product. However, upon addition of acetone, more of the product was obtained. The isolated product was insoluble in any solvent. Yield = 0.110 g (24.44%). No frequency conversion was observed.

2. Reaction of zinc chloride with potassium dihydrogen phosphate

0.250 g (1.83 mmol) of $ZnCl_2$ was dissolved in 15 mL of water and reacted with 0.550 g (4.04 mmol) of KDP in 25 mL of water. No precipitate was detected. Upon heating the reaction mixture at 80°C for one hour, a white crystalline solid formed. This was filtered, washed with water and acetone and dried. More white solid was formed as acetone was added to the filtrate. The product does not dissolve in water or any other solvent. Yield = 0.260 g (54.85%). Powder test showed a weak frequency conversion.

3. Reaction of lead acetate with potassium dihydrogen phosphate

0.250 g (0.659 mmol) of $Pb(Ac)_2.3H_2O$ was dissolved in 20 mL of water and reacted with 20 mL of aqueous solution of 0.220 g (1.62 mmol) of KDP. A white precipitate was formed, which was filtered, washed with water and dried. The substance is insoluble in most solvents. Yield = 0.150 g (56.6%). Powder test showed, greenish color indicating a weak SHG.

4. Reaction of chromium chloride with potassium dihydrogen phosphate

0.270 g (1.01 mmol) of $CrCl_3.6H_2O$ dissolved in 10 mL of water was reacted with 0.630 g (4.63 mmol) of KDP in 20 mL. There was no precipitate formation when the reaction was stirred at room temperature. Heating the reaction mixture to 80°C, gave a greenish precipitate, which was filtered, washed with water and dried. The substance is insoluble in most solvents. Yield = 0.233 g (67.34%). Powder test showed significant NLO properties.

5. Reaction of copper chloride with potassium dihydrogen phosphate

0.250 g (1.47 mmol) of $CuCl_2$ was dissolved in 15 mL of water. To this solution, an aqueous solution of 0.50 g (3.68 mmol) of KDP was added. No precipitate formed initially, however, on heating the mixture, a bluish precipitate appeared. This was filtered, washed with water and acetone and dried. The substance does not dissolve in most solvents. Yield = 0.130 g (34.3%). No NLO property was seen when the powder was tested.

Results and Discussion

i) Reaction of mercuric acetate with xanthates

Both mercuric ethylxanthate, $Hg(S_2COC_2H_5)_2$, and mercuric isopropylxanthate, $Hg(S_2COC_3H_7)_2$, were prepared by the reaction of $Hg(Ac)_2$ and the corresponding potassium xanthate in aqueous solution in 1:2 molar ratios as shown in equations 1 and 2.

$$Hg(Ac)_2(aq) + 2KS_2COC_2H_5(aq) \text{ -----> } Hg(S_2COC_2H_5)_2\,(s) + 2KAc \qquad 1$$

$$Hg(Ac)_2(aq) + 2\,KS_2COC_3H_7\,(aq) \text{ -----> } Hg(S_2COC_3H_7)_2\,(s) + 2KAc \qquad 2$$

In the first reaction a dark-yellow precipitate and in the second a red-yellow precipitate were isolated when the reactions were carried out at room temperature. However, on heating the reaction mixture, the color of the precipitate changed gray and then dark. This may be a result of the decomposition of xanthates to give most likely a metal sulfide. The isolated compounds are insoluble in water and most organic solvents.

The compounds isolated were characterized by IR-spectroscopy. The IR-spectra of both $Hg(S_2COC_2H_5)_2$ and $Hg(S_2COC_3H_7)_2$ are rich in bands due to various modes of

vibrations of the xanthate ligands. The spectra are shown in Fig. 1 and 2. The IR-spectrum of $Hg(S_2COC_2H_5)_2$ shows three very strong bands between 1200-1000 cm^{-1} which may be due to mono- or bi-dentate xanthate group. The band at 1195 cm^{-1} is due to C-OEt stretching and that at 1112 cm^{-1} is mainly from ν(C-O). The band at 1024 cm^{-1} is entirely ν(C=S) {9}. Similarly the spectrum of $Hg(S_2COC_3H_7)_2$ shows very strong bands at 1213, 1092, and 1008 cm^{-1} also due to ν(C-O^{i}Pr), ν(C-O) and ν(C=S) respectively.

The nonlinear optical (NLO) properties of both $Hg(S_2COC_2H_5)_2$ and $Hg(S_2COC_3H_7)_2$ were studied by Kurtz method. Both are found to be NLO active, but tend to burn by the laser light indicating that they have low damage threshold.

ii) Reaction of metal salts with sodium dimethyldithiocarbamate

The reaction of sodium dimethyldithiocarbamate with various salts, such as $ZnCl_2$, $CdCl_2$, $Hg(Ac)_2$, and $Pb(Ac)_2$ was investigated according to equation 3.

$$MX_2 + 2NaS_2CNMe_2 \text{ ------> } M(S_2CNMe_2)_2 + 2NaX \qquad 3$$

In all cases, a one to two molar ratio of the salt to ligand was taken in aqueous solution and reacted initially at room temperature and then heated to 80°C for digestion. All the reactions gave white precipitate in quantitative yield. The products isolated from all reactions were found to be insoluble in most organic solvents.

The compounds isolated were characterized by IR-spectroscopy. The IR-spectra of all the compounds depicted six strong bands in the region of 1500-955 cm^{-1}. A comparison of the spectrum of sodium dimethyldithiocarbamate with the spectra of the isolated compounds clearly confirms that the six strong bands observed in the complexes are due to various vibrational modes of the dimethyldithiocarbamate group. The IR-spectra of the compounds are shown in Fig. 3-6.

The nonlinear optical properties of these complexes were also tested. The complexes formed from the reaction of $ZnCl_2$, $CdCl_2$ and $Hg(Ac)_2$ showed no second harmonic generation. However, the product formed from the reaction of $Pb(Ac)_2$ and sodium dimethyldithiocarbamate showed a nonlinear activity, but burned under the heat of the laser light.

iii) Reactions of metal salts with a mixture of potassium isopropylxanthate and sodium dimethyldithiocarbamate

The reaction of some metal salts, MX_2 (M = Zn, Cd, Hg, Pb; X = Cl, CH_3COO^-) with a mixture of potassium isopropylxanthate and sodium dimethyldithiocarbamate was investigated hoping that the reaction will result in the compounds shown by the following equation.

$$MX_2 + KS_2COC_3H_7 + NaS_2CNMe_2 \text{ ---> } M(S_2COC_3H_7)(S_2CNMe_2) + KX + NaX \quad 4$$

In all reactions, equimolar quantities of the two ligands were dissolved in water and then added to an equivalent amount of the corresponding metal salt in aqueous solution. In adding the solution of the ligands in to the solution of the metal salts slowly, one observes formation of precipitate immediately which are white or yellowish-white in color. Unlike the xanthate complexes, some of the products, such as the one formed from Cd and Hg were found to be difficult to dissolve in most organic solvents. However, the complexes formed from Zn and Pb were found to be little soluble in THF, acetone and methylene chloride.

The IR-spectra of all the compounds are very similar and are reach in bands due to both the isopropylxanthate and dimethyldithiocarbamate ligands. In all the spectra about eight strong bands are observed. Those in the region of 1200, 1090 and 1010 cm^{-1} belong to the isopropylxanthate and the others in the region of 1500, 1375, 1140, 1020 and 960 cm^{-1} belong to the dimethyldithiocarbamate group. The IR-spectra of the complexes are shown in Fig. 7-10.

The nonlinear optical property studies of these complexes showed that the complex formed from Cd showed strong second harmonic generation and is stable under the laser heat, while the complexes formed from Hg and Pb, although showed SHG, tended to burn under the laser heat. The product isolated from the reaction of zinc and the ligands failed to show SHG.

iv) Reaction of metal salts with potassium dihydrogen phosphate (KDP)

Potassium dihydrogen phosphate (KDP) is a well known NLO material {8}. Its reaction with some metal salts was investigated in order to see how its NLO property is affected when the potassium is replaced by other divalent or trivalent metal ions. To perform

these reactions, an aqueous solution of the metal salt was reacted with excess of KDP as shown by equations 5 and 6.

$$MX_2 + 2KDP \text{ ------> } M(DP)_2 + 2KX \quad (5)$$

$$M'X_3 + 3KDP \text{ ------> } M(DP)_3 + 3KX \quad (6)$$

(M = Zn, Cd, Pb, Cu; M' = Cr; X = Cl, CH_3COO^-, DP = $H_2PO_4^-$)

The reactions of $ZnCl_2$, $CdCl_2$ and $Pb(Ac)_2$ with KDP in 1:2 molar ratio gave no precipitate initially but white precipitates on heating for one hour at 75°C. Similarly, the reaction of $CuCl_2$ and $CrCl_3$ with KDP gave light blue and green precipitates respectively after heating the reaction mixture. All the isolated products were difficult to dissolve in any solvent. The compounds isolated were characterized by IR-spectroscopy. The IR-spectra of all the products show strong bands at around 1300 cm^{-1} to 900 cm^{-1} and medium bands varying in number in the region of 620-500 cm^{-1}. Similar bands in these regions have been identified for pure KDP. Fig. 11 and 12 show the IR-spectra of two of the compounds.

The isolated compounds were screened for their NLO activities. Both products isolated from the reaction of $CdCl_2$ and $CuCl_2$ with KDP showed no second harmonic generation. The other products, namely the one isolated from the reaction of $ZnCl_2$, $Pb(Ac)_2$ and $CrCl_3$ showed weak second harmonic generation. Unlike other reaction products of Pb, discussed earlier, the complex formed from reaction with KDP is very stable under the influence of the laser heat.

Conclusions

About fifteen complex, synthesized by the reaction of transition and main group metal salts with thio ligands and dihydrogen phosphate were characterized by spectroscopic method, and screened for their NLO activities. Some of the compounds were found to be promising NLO materials and worth further characterization and thorough investigation. Further more a full scale powder test using Kurtz method has to be conducted for the materials found to be NLO active.

References

1. S. P. Velsko, A. Ruggiero, and Mark Herman, "Frequency Agile OPO based transmitters for multiwavelength DIAL", in Applications of LIDAR to Current Atmospheric Topics, Arthur J. Sedlacek III, Ed., SPIE v2833, p144, 1996

2. G. I. Stegeman, Pavel Mamyshev, and William Torruellas, "Photonic Applications of Spatial Soliton Switches", in Photonic Device Engineering for Dual Use Applications, Andrew Pirich, Ed., SPIE v2481, p270, 1995

3. B. Fleck and L. Wenke, "Application of Photoanisotropic Polymer Films in Optical Information Processing and as Nonlinear Storage Media", in International Conference On Holography and Optical Information Processing (ICHOIP 96), Jiabi Chen, Anzhi He, and Dahsiung Hsu, Eds., SPIE v2866, p384, 1996.

4. Hann, R. A., Bloor, D. Eds. Organic Materials for Nonlinear Optics: Royal Society of Chemistry: London, 1989.

5 Chemla. D. S. Zyss, J. Eds. Nonlinear Optical Properties of Organic Molecules and Crystals; Academic Press: Orlando, Vols. 1 and 2, 1987

6. Williams, D. J., Ed. Nonlinear Optical Properties of Organic and Polymeric Materials; ACS Symposium Series 233; American Chemical Society: Washington, DC, 1983.

7. Polagie, T. P.; Stoner, T. C.; Dallinger, R. F.; Gilbert, T. M.;Hopkins, M. D. *J. Am. Chem. Soc.*, 113, 703-704, 1991

8. Eimer, D. "Electro-optic, linear and nonlinear optical properties of KDP and its isomorphs," *Ferroelectrics*, Vol. 72, P.95, 1987.

9. Pandey, O. P.; Sengupta, S. K.; Tripathy, S. C. *Polyhedron* 3, 659, 1984

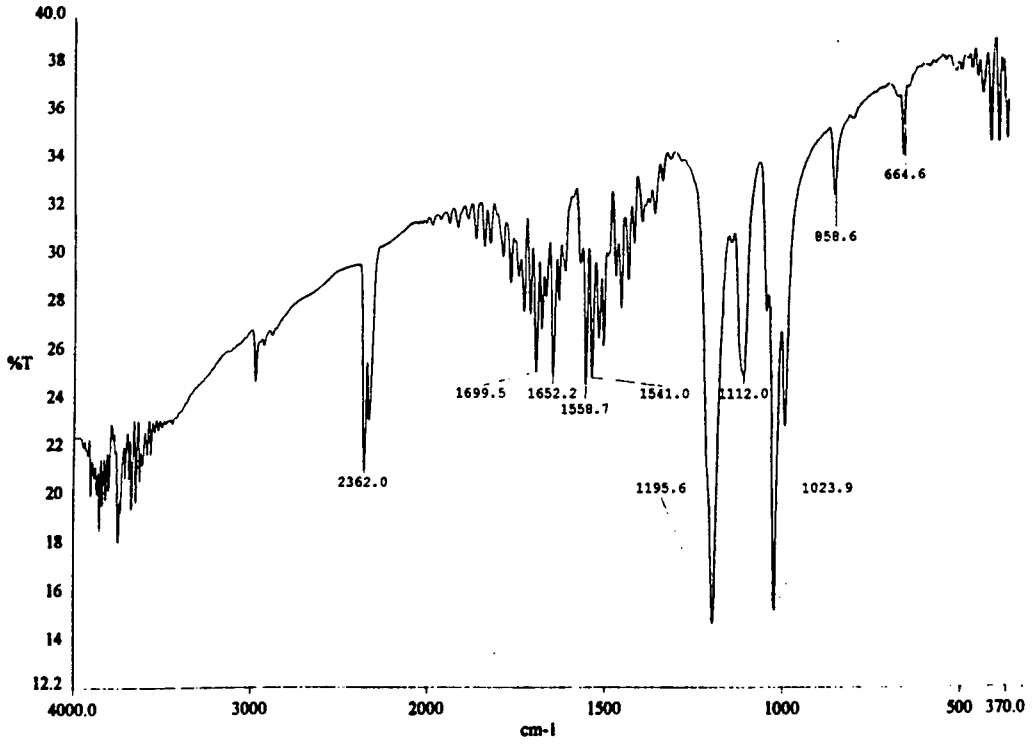


Fig. 1. IR-spectrum of $Hg(S_2COC_2H_5)_2$

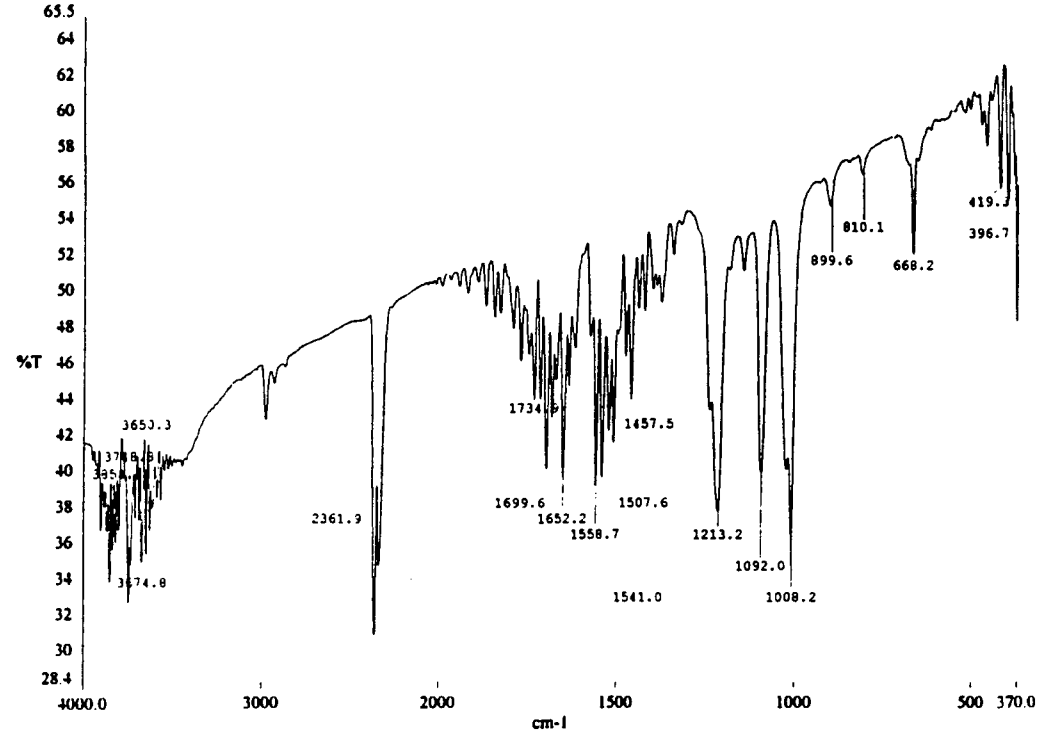


Fig. 2. IR-spectrum of $Hg(S_2COC_3H_7)_2$

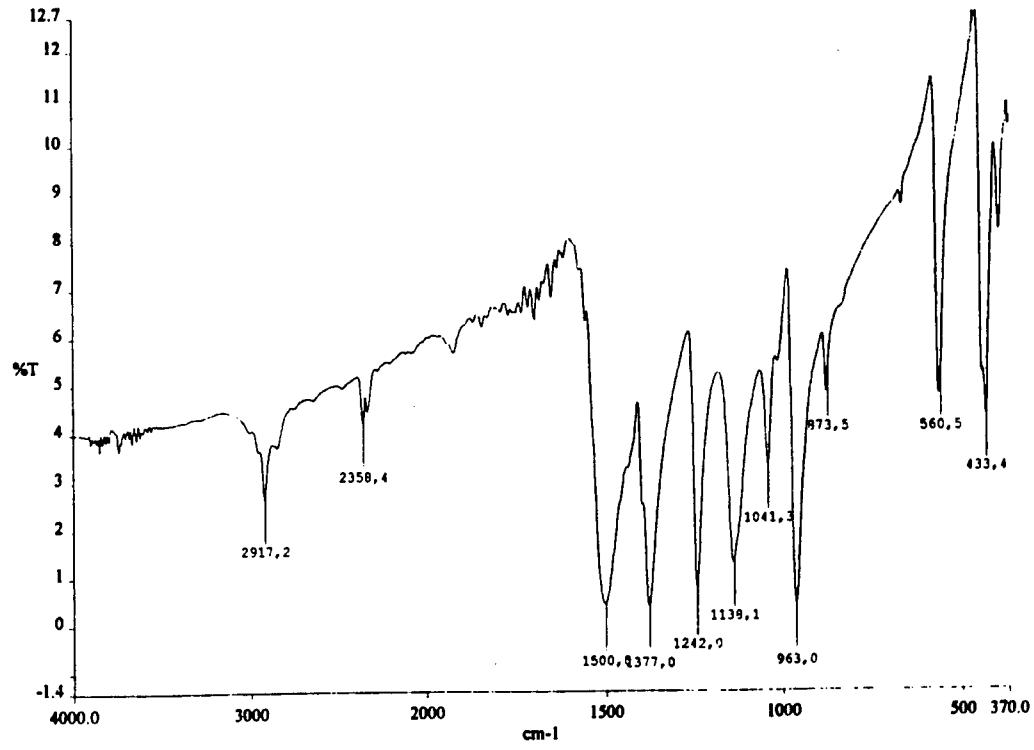


Fig. 3. IR-spectrum of $Hg(S_2CNMe_2)_2$

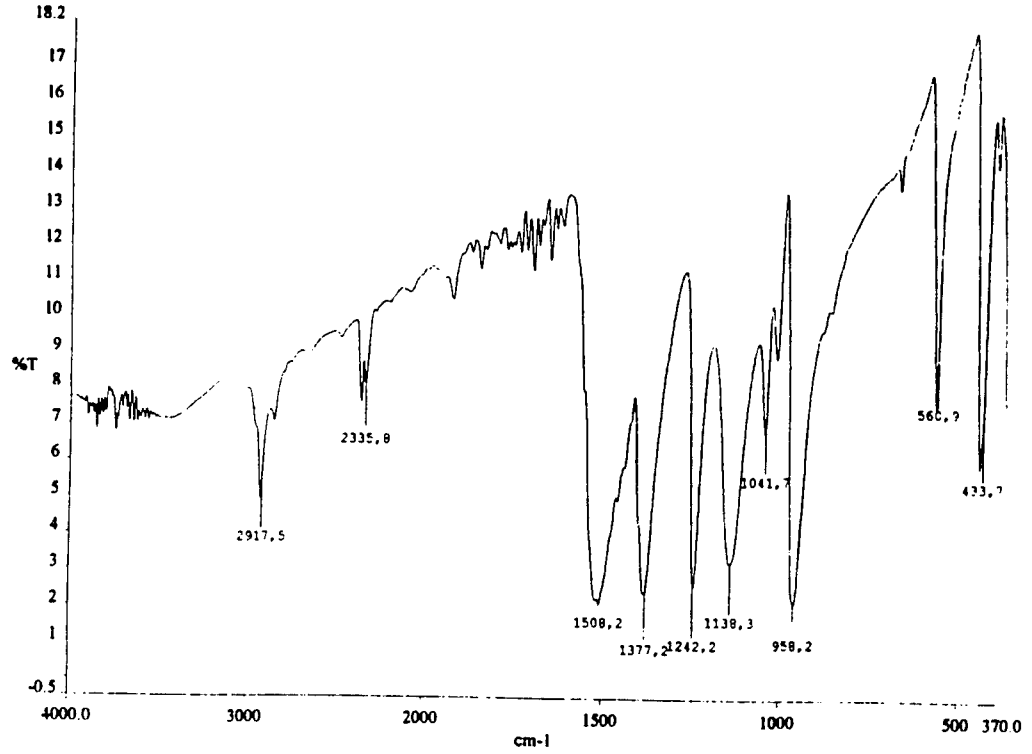


Fig. 4. IR-spectrum of $Cd(S_2CNMe_2)_2$

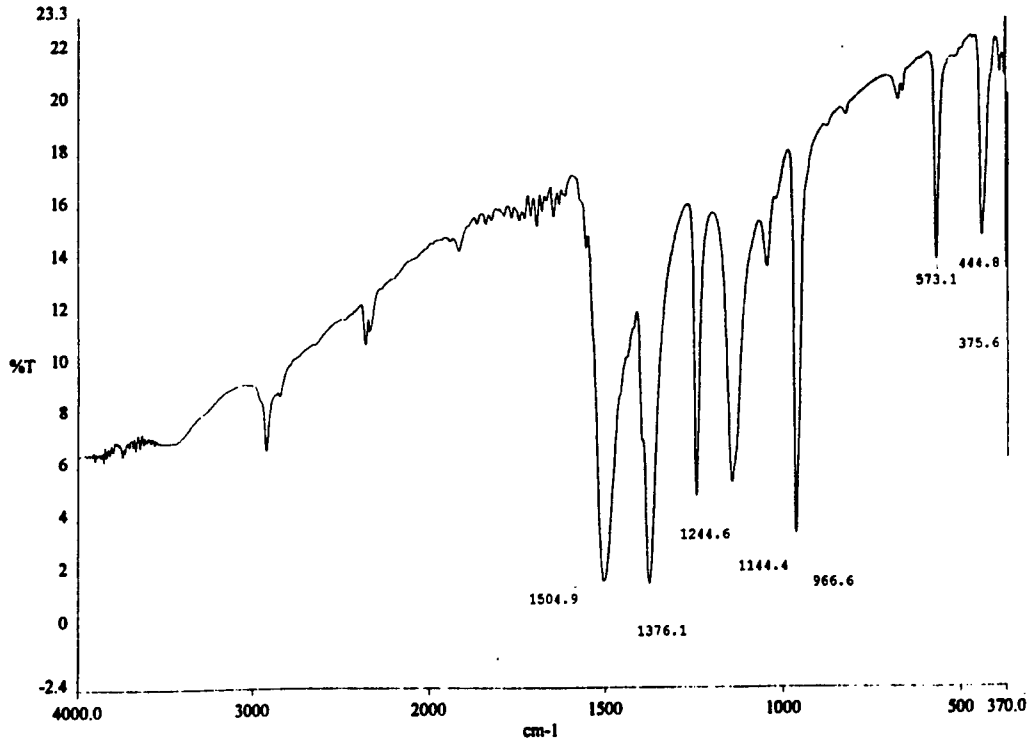


Fig. 5. IR-spectrum of $Pb(S_2CNMe_2)_2$

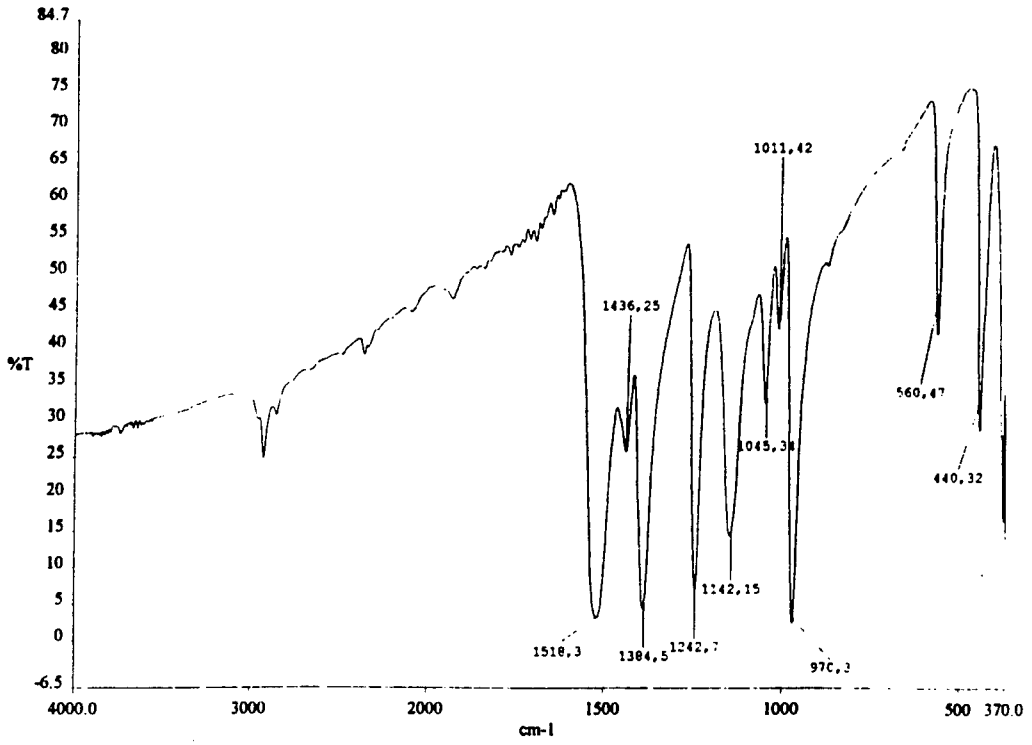


Fig. 6. IR-spectrum of $Zn(S_2CNMe_2)_2$

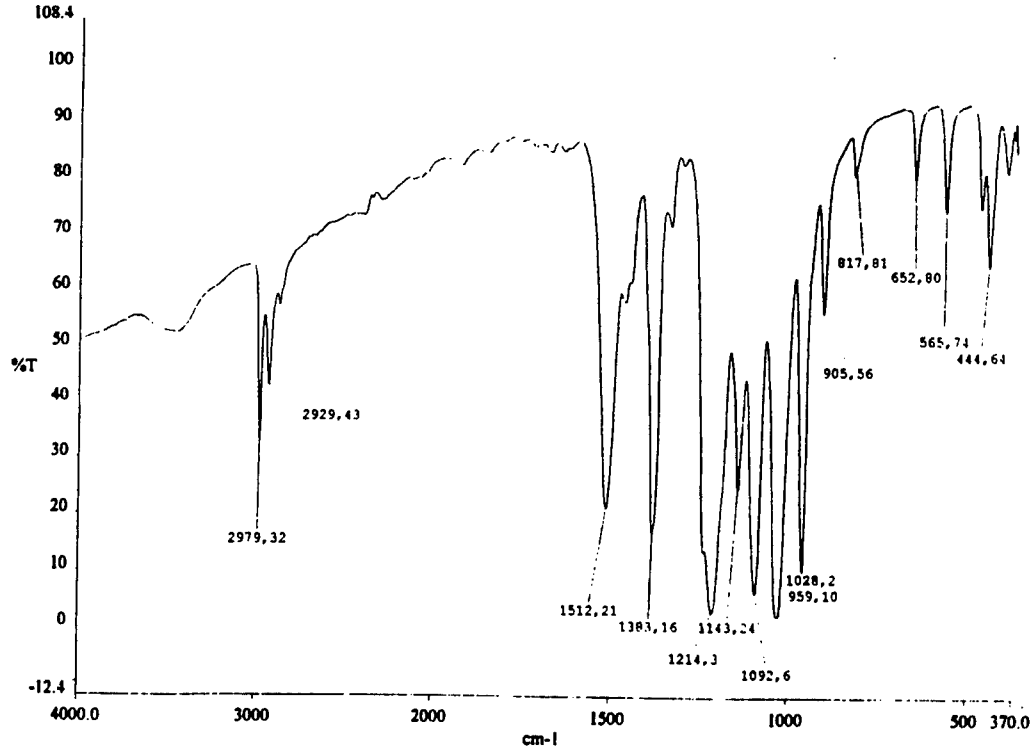


Fig. 7. IR-spectrum of $Cd(S_2COC_3H_7)(S_2CNMe_2)$

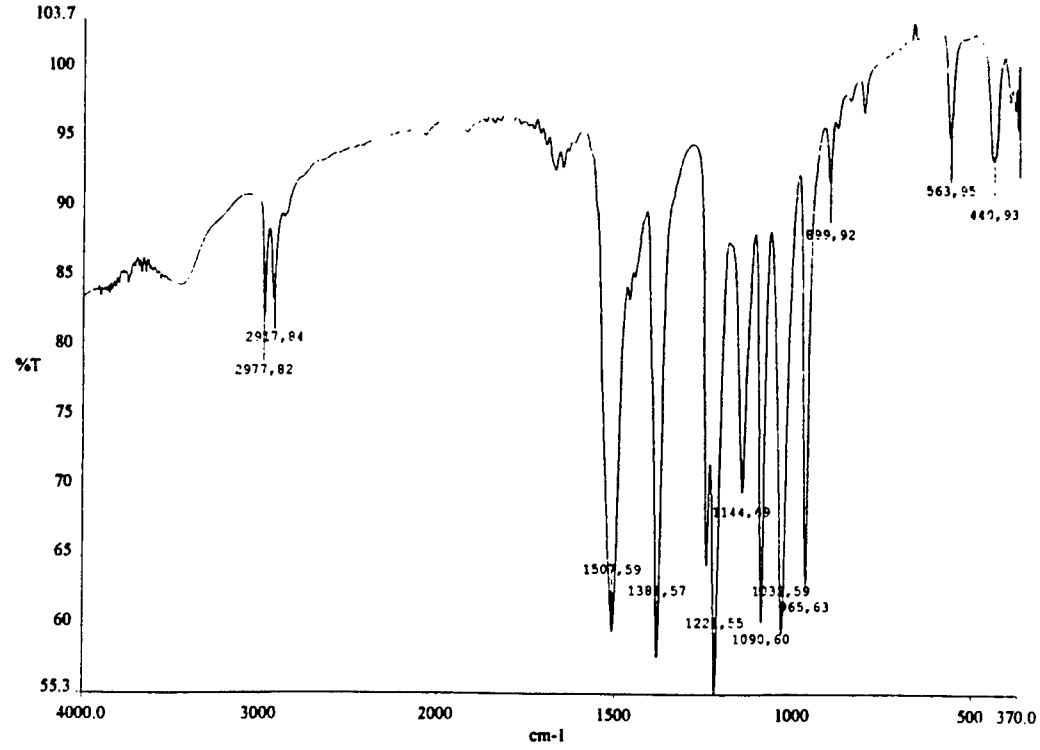


Fig. 8. IR-spectrum of $Hg(S_2COC_3H_7)(S_2CNMe_2)$

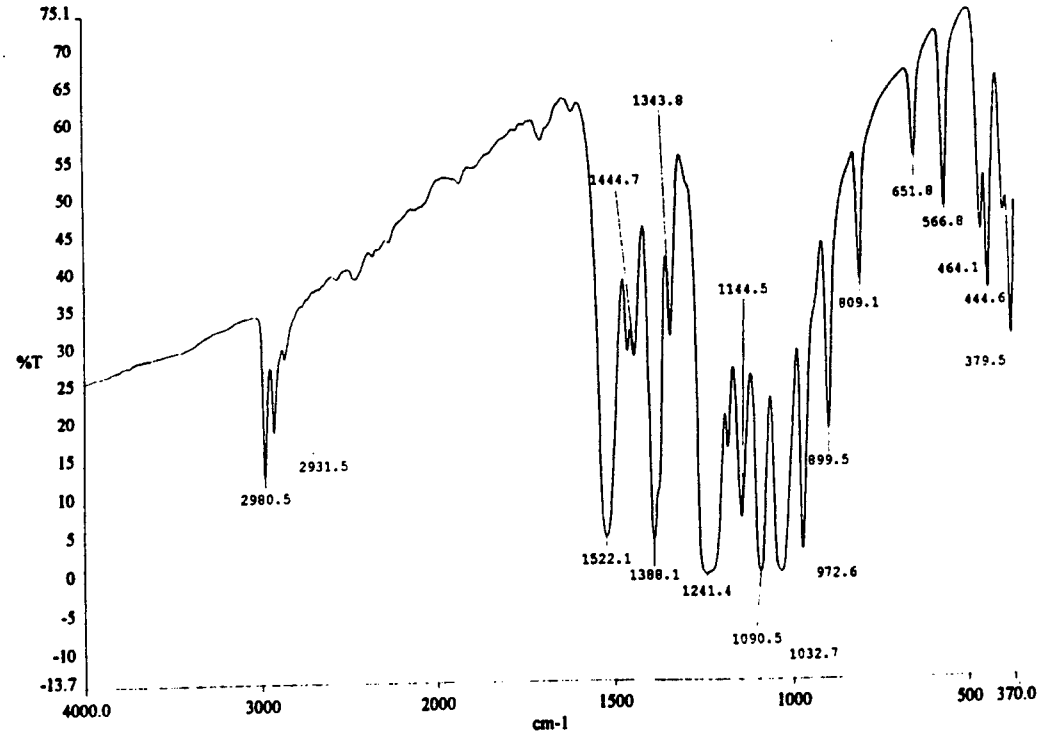


Fig. 9. IR-spectrum of $Zn(S_2COC_3H_7)(S_2CNMe_2)$

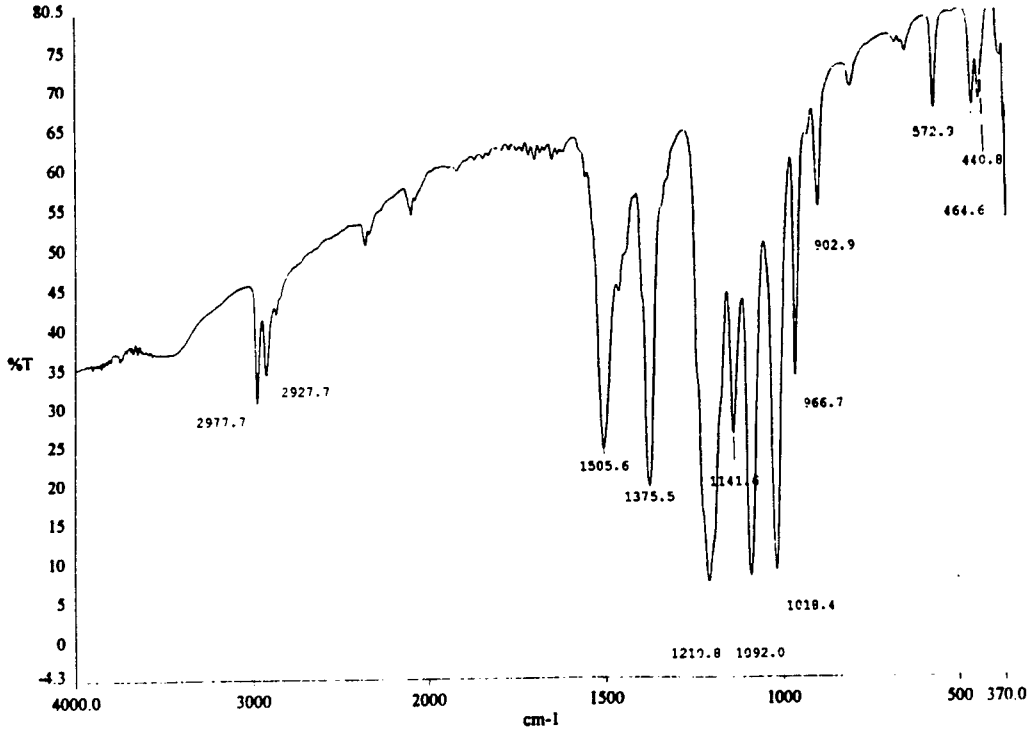


Fig. 10. IR-spectrum of $Pb(S_2COC_3H_7)(S_2CNMe_2)$

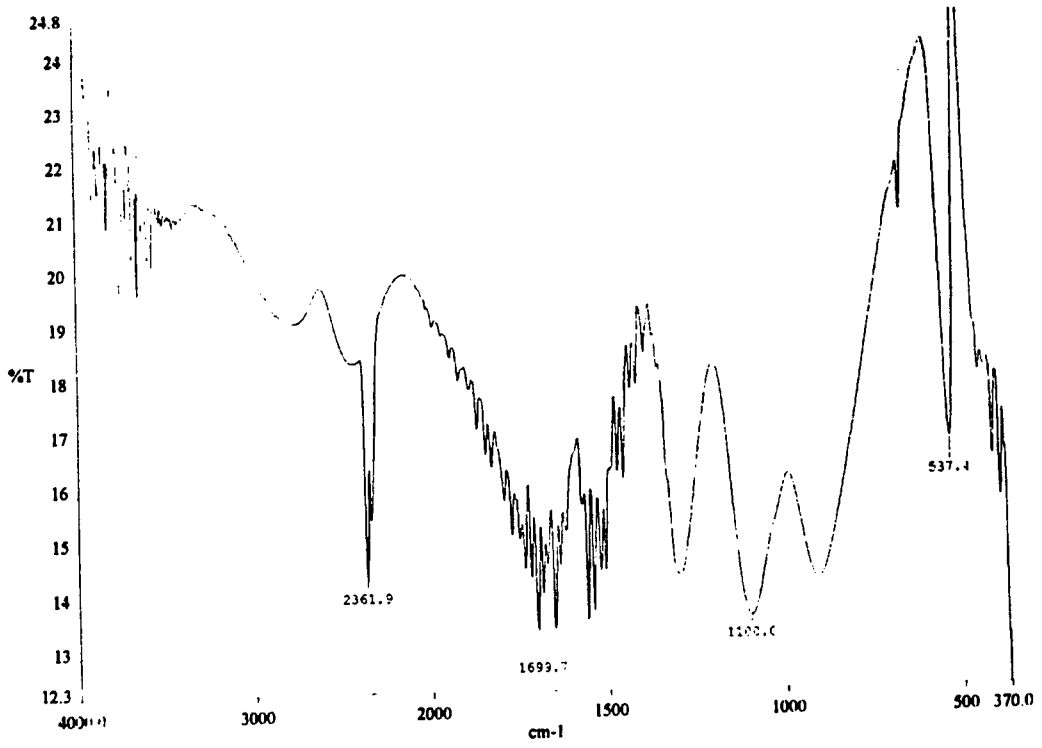


Fig. 11. IR-spectrum of $Zn(H_2PO_4)_2$

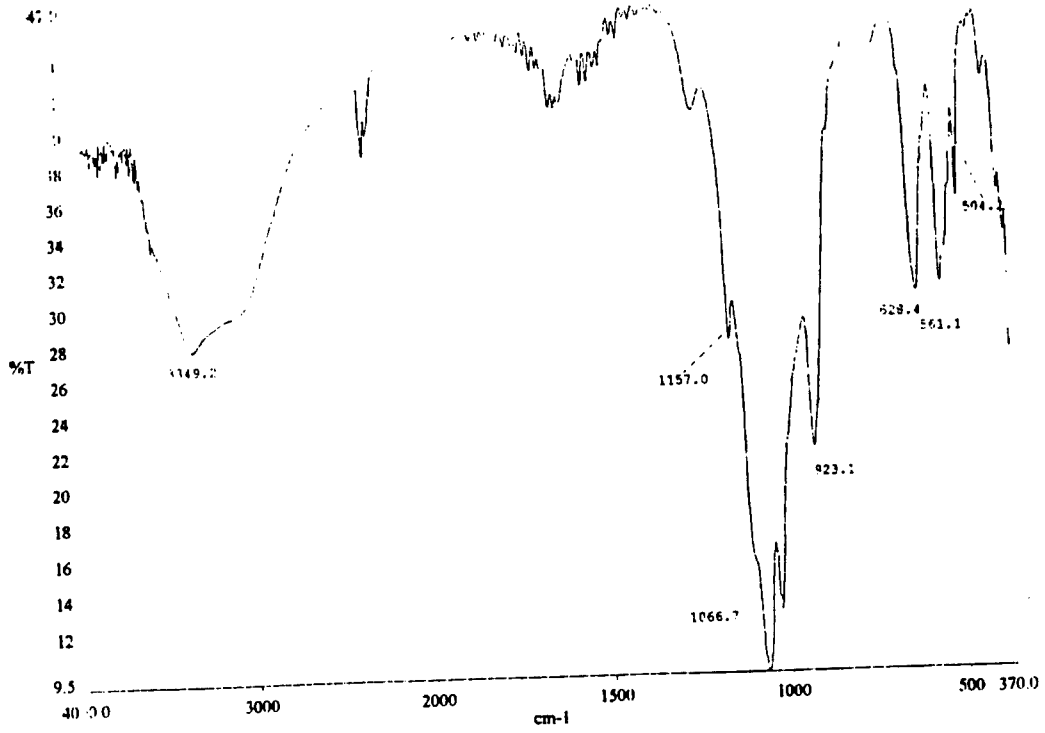


Fig. 12. IR-spectrum of $Cu(H_2PO_4)_2$

BIOCATALYSIS OF BIPHENYL AND DIPHENYLACETYLENE TO SYNTHESIZE POLYMER PRECURSORS

Patrick C. Gilcrease
Assistant Professor
Department of Chemical and Petroleum Engineering
University of Wyoming
P.O. Box 3295
Laramie, WY 82071-3295

Final Report for:
Summer Faculty Research Program
Tyndall Research Site
Tyndall AFB, FL 38403

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Tyndall Research Site

September 1998

BIOCATALYSIS OF BIPHENYL AND DIPHENYLACETYLENE TO SYNTHESIZE POLYMER PRECURSORS

Patrick C. Gilcrease
Assistant Professor, Department of Chemical and Petroleum Engineering
University of Wyoming, Laramie, WY 82071-3295

ABSTRACT

The U.S. Air Force Material Science directorate has identified the phenolic compounds hydroxybiphenyl (HBP) and hydroxydiphenylacetylene (HDPA) as critical precursors in the production of high performance polymers. Epoxy resins synthesized using HBP or HDPA have exceptional thermal resistance and could be used in a number of important aerospace applications; however, properties vary depending on which phenolic isomer is used as well as the purity of that particular isomer. While *m*-HBP and *m*-HDPA are difficult to produce by normal chemical synthesis, biocatalysis may offer an alternative method. This research demonstrates that two toluene monooxygenase enzymes (Tb2m and Tb4m) are capable of producing either *m*-HBP from biphenyl or *m*-HDPA from diphenylacetylene (DPA). For the enzymatic hydroxylation of biphenyl, specific activities varied between 2.1×10^{-4} and 8.1×10^{-4} μmol/mg protein/min, with the HBP product consisting of 76 to 91% *meta* isomer (remainder *para*). With diphenylacetylene as the substrate, specific hydroxylation activities varied between 0.11×10^{-4} and 2.0×10^{-4} μmol/mg protein/min, and 24 to 30% of the HDPA product was the *para* isomer (remainder most likely *meta*). The addition of glucose to the resting cell systems inhibited the enzyme activity initially, but prevented a drop in activity after 2 h. Tb2m activity was insensitive to medium pH in the range of 5 to 9, while Tb4m activity was maximized at pH 9. While the *meta* regiospecificity of these enzymes was promising, specific activities were only 1 to 3% of the values observed with toluene as the substrate (3 mM). Since the low activities may in part be due to the low aqueous solubility of biphenyl (50 μM) and diphenylacetylene (6 μM), four different surfactants (Triton X-100, Tween 80, Brij 58, and SDS) were evaluated for their effect on biphenyl hydroxylation rates. The specific hydroxylation activity of Tb4m was enhanced by 45% with 500 μg/mL SDS; this surfactant may have increased the effective biphenyl concentration in solution. Future work should further investigate methods for increasing the rate of biocatalysis.

INTRODUCTION

Biocatalysis can be defined as the use of biological cells or their enzymes in the production of chemicals for commercial uses. Bacteria are the most common organisms in biocatalysis, but yeasts, molds, algae, plant cells and even animal cells are also used. Biocatalysis is not a new concept; it has been used for centuries in the production of cheese, wine, and beer, and more recently to produce antibiotics and other pharmaceuticals. Scientists are now discovering that microorganisms can also transform a wide range of unnatural or man-made compounds; thus, they can also be used in the production of specialty organic chemicals. There are several advantages to using biocatalysis for organic synthesis: (1) mild temperatures and pressures are used, making the process energy efficient (2) specific compounds can be produced, often with very few byproducts being formed, (3) biocatalysis can promote reactions that are difficult or impossible to produce by conventional means, and (4) reactions are normally conducted in aqueous solutions, reducing the need for toxic organic solvents [1].

An organic synthesis of interest to polymer scientists is the hydroxylation of biphenyl or diphenylacetylene (DPA). Hydroxybiphenyl (HBP) and hydroxydiphenylacetylene (HDPA) form the backbone of specific phenolic resins used to synthesize epoxy thermosets and aromatic polyesters. These high performance polymers can have exceptional thermal resistance, and the U.S. Air Force Material Science directorate and NASA have recognized the potential aerospace applications for such polymers. The mechanical properties of these important polymers are dependent upon the purity of the phenolic isomer used, but regiospecific hydroxylation by purely chemical methods is notoriously difficult [2]. Thus, extensive purification is required to obtain the pure isomer.

An alternative method of producing hydroxylated aromatic compounds is biocatalysis. Oxygenase enzymes from various bacterial strains can hydroxylate a number of aromatic substrates [3]. Unlike conventional catalysts, enzymes exhibit high regiospecificity, favoring the production of one isomer over others. Therefore, biocatalysis can provide the means of producing specific

isomers that are difficult to formulate or separate using conventional methods. Previous work described the characterization of two multicomponent toluene monooxygenases from *Burkholderia* sp. strain JS150 that hydroxylate a broad range of aromatic compounds [4, 5]. The goal of this study was to characterize the specific activity of these enzymes in the hydroxylation of biphenyl and diphenylacetylene. To establish the stoichiometry of these reactions, substrate and hydroxylated product concentrations were quantified. Since higher reaction rates are directly correlated with improved process economics, the specific biphenyl activity of both enzymes was optimized with respect to induction time, added glucose, and pH. Finally, four different surfactants were evaluated for their effect on biphenyl hydroxylation rates; these agents can potentially improve the biological uptake/conversion of sparingly soluble substrates.

MATERIALS AND METHODS

Microorganism

Biocatalysis experiments used recombinant *Pseudomonas aeruginosa* strains - PAO4032 (pRO2037, pRO2369) and PAO4032 (pRO2038, pRO2369). These strains carry genes for the individual toluene monooxygenase enzymes (Tb2m and Tb4m, respectively) and regulatory proteins which were isolated from *Burkholderia* sp. strain JS150 [4, 5].

Culture media

Growth medium consisted of Stanier's Mineral Salts Broth (MSB) [6] plus the following compounds (g/L): glucose 5.0; caseamino acids 2.5; ticarcillin 0.25; trimethoprim 0.3. The antibiotics ticarcillin and trimethoprim were added after autoclaving and just prior to inoculation. Reaction medium was 20 mM KH_2PO_4 adjusted to pH 7.0 (Tb2m experiments) or 8.0 (Tb4m experiments). Biphenyl, diphenylacetylene, *m*-hydroxybiphenyl and *p*-hydroxybiphenyl were added to the reaction medium as noted.

Culture methods

Isolated colonies were spread on agar plates consisting of the growth medium plus 15 g/L agar. After ~24 h of incubation at 37 °C, the bacterial lawn was used to create a cell suspension in 20 mM KH_2PO_4 (A_{540} ~ 0.6). Three separate 500 mL shake flasks containing 75 mL of growth medium were inoculated with 0.1, 0.25, and 0.5 mL of the cell suspension. These flasks were then placed on a 37 °C, 250 rpm shaker; after ~12 h, the flask with A_{540} between 0.25 and 0.4 was selected for induction. Toluene was added to the liquid phase (40 µL = 3mM) and the vapor phase (150 µL in a vapor tube). After 6 h of induction at 37 °C, 250 rpm, cells were harvested via centrifugation (5000 rpm for 5 min). Cells were washed twice in 20 mM KH_2PO_4 buffer, then suspended in the reaction medium. The resting cell biocatalysis experiments were then conducted at 30 °C.

Biocatalysis of biphenyl and diphenylacetylene

Stoichiometry experiments were conducted in sealed serum bottles to minimize substrate volatilization losses. To prevent biphenyl/DPA adsorption, the bottles were baked in a 450 °C annealing oven for at least 2 h; Teflon-faced silicone septa were used to seal the bottles. The reaction medium was presaturated with either biphenyl or DPA, then filtered to remove any excess solid substrate (1.0 µM glass fiber). Liquid full bottles were inoculated at time zero, sealed, and placed on a 30 °C, 250 rpm shaker. Samples were drawn with a glass 1.5 mL gas-tight syringe; for each 0.5 mL liquid sample withdrawn, an equal volume of oxygen gas was added. Samples were immediately diluted with an equal volume of acetonitrile and microfuged for 1 min at 14,000 rpm to remove cellular material. The supernatant was then transferred to a glass 2.0 mL HPLC vial and analyzed within 2 h. For protein analysis, 90 µL samples were diluted with 10 µL of 1.0 N NaOH and then frozen.

Activity experiments were conducted in Erlenmeyer shake flasks or test tubes (pH and surfactant screening) sealed with Morton closures. Excess solid substrate (5.0 g/l biphenyl or 1.0 g/L diphenylacetylene) was present to maintain the aqueous concentration at saturation values. Reaction slurries were shaken at 30 °C, 250 rpm overnight to presaturate the aqueous phase prior

to inoculation. Washed cells were added at time zero, and 0.5 mL samples were taken with a 1.0 mL micropipeter. Sample prep was as before; since the aqueous samples contained solid substrate which dissolved in the acetonitrile, only the product concentrations were quantified in these experiments. Experiments with *m*-HBP or *p*-HBP as the starting substrate were also conducted in shake flasks with Morton closures, but no excess solids were present.

Analysis

Biphenyl and hydroxylated metabolites were detected and quantified by high-pressure liquid chromatography (HPLC). The reversed-phase column was an Alltech Octyl (C8) 5U (length 250 mm, I.D. 4.6 mm). The isocratic method used 40 vol% trifluoroacetic acid (1 mL TFA per liter of water) and 60 vol% acetonitrile at a total flow rate of 1.0 mL/min. UV A_{254} was used for detection; compound identification was via comparison with standards. Sample injection volume was 100 µL. The following gradient method was used for analysis of DPA cultures: 0-10.5 min, 50% TFA-50% CH_3CN; 11.5-15 min, 20% TFA-80% CH_3CN; 16-18 min, 50% TFA-50% CH_3CN. Separation of *meta* and *para*-HBP isomers required a C18 column (Alltech Adsorbosphere C18 H5 5U) with a gradient program as follows: 0-4 min, 90% TFA-10% CH_3CN; 5-16.5 min, 50% TFA-50% CH_3CN; 17.5-21.5 min, 15% TFA-85% CH_3CN; 22.5-25 min, 90% TFA-10% CH_3CN.

Cell concentrations were quantified using the BCA Protein assay. Protein samples were thawed, heated to 96°C for 10 min, and then vortexed. Protein concentrations were then obtained using the standard microtiter method [7]. Protein standards were prepared from bovine serum albumin.

HPLC standards

Biphenyl, diphenylacetylene, 4-phenylphenol, and 4,4'-biphenol were obtained from Aldrich. 2-biphenylol, 3-biphenylol, 3,4-biphenol, and 2,2'-biphenol were obtained from Ultra Scientific. 4-hydroxydiphenylacetylene was obtained from NASA.

RESULTS

Induction of monooxygenase activity. In order to determine the maximum specific activity as a function of induction time, Tb2m and Tb4m cells were induced with toluene at mid-log phase, and monooxygenase activity was quantified with a toluene assay 3, 6, 9, and 11 h after toluene was first added. For Tb2m, a maximum specific activity of 2.3×10^{-2} μmol/mg protein/min was observed after 6 h. Tb4m showed a maximum specific activity of 1.1×10^{-2} μmol/mg protein/min after only 3 h of induction, and 0.76×10^{-2} μmol/mg protein/min at 6 h. A six hour toluene induction period was used for both cultures in all subsequent experiments. In separate experiments, growing cultures of Tb2m and Tb4m were transferred to biphenyl coated shake flasks (0.069 g/150 mL medium) at mid-log phase; cell activity was assayed with toluene at 3,6, 9, and 12.5 h. No cresol was observed in the toluene assays, and no HBP products were detected in the original cultures.

Biphenyl hydroxylation by Tb2m and Tb4m. Initial biphenyl hydroxylation experiments were conducted with resting cells and excess biphenyl solids (5.0 g/L compared to 6.5 mg/L at aqueous saturation). Excess biphenyl was added to maintain the solution concentration at or near the saturation level; thus, the measured activities should be maximized with respect to aqueous phase concentration. Experiments were conducted with and without glucose to determine if resting cell activity could be enhanced or extended with an energy source present. For Tb2m, initial hydroxylation activity was actually 43% lower when 5.0 g/L glucose was added to the media (2.1×10^{-4} vs. 3.7×10^{-4} μmol/mg protein/min for the no glucose control). Glucose also affected the isomeric distribution of HBP; after six hours, the HBP product was 91% *meta* isomer (9% *para*) compared to 76% *meta* product without glucose. After six hours, specific activity had declined for the no glucose flasks (3.3×10^{-4}), but increased in the presence of glucose (2.4×10^{-4}). Dihydroxybiphenyl (DHBP) compounds were also observed with Tb2m (confirmed with GC-MS); these will be quantified as soon as appropriate standards are obtained.

Results for Tb4m cells in the presence of 5.0 g/L biphenyl solids are shown in *Figure 1*. Experiments were again conducted with and without glucose in the medium, but a lower glucose concentration (2.5 g/L) was used. As with Tb2m, the initial specific activity was lower (by 44%) for the flasks containing glucose (2.5×10^{-4} versus 4.5×10^{-4} µmol/mg protein/min without glucose). However, after two hours, specific activity without glucose declined dramatically, while activity in the glucose flasks only showed a modest decrease. At 8.5 h, specific activity without glucose had dropped to 0.23×10^{-4}, while specific activity with glucose was 1.4×10^{-4} µmol/mg protein/min. Glucose did not have a significant effect on the final HBP isomer distribution in this case (77% *meta* for the control vs. 81% *meta* with glucose). The Tb4m enzyme has a much lower activity for HBP as compared to Tb2m; while DHBP was observed in concentrated extracts analyzed via GC-MS, no significant peaks were observed in the HPLC analysis.

To establish the stoichiometry of biphenyl conversion by Tb2m and Tb4m, resting cell experiments were conducted in sealed serum bottles. The serum bottles prevented volatilization and sorption losses of the biphenyl substrate; in abiotic controls without cells, 97% of the initial biphenyl in solution was still present after 7 h. The conversion of biphenyl to HBP by Tb4m is depicted in *Figure 2*. In this experiment, 98 to 101% of the initial biphenyl concentration was accounted for as biphenyl or HBP product over the 2.5 h incubation. To minimize potential biphenyl losses after sampling, the shorter HPLC method was used for analysis (40/60 isocratic, C8 column); this method did not provide separation between the *meta* and *para* isomers of HBP. An isomeric distribution of 77% *meta*, 23% *para* was assumed in quantifying the single HBP peak. Specific activity for biphenyl conversion (as determined by the slope of the biphenyl curve from 0 to 100 min) was 7.6×10^{-4} µmol/mg protein/min in this experiment.

Biphenyl serum bottle experiments were also conducted with Tb2m for stoichiometry purposes; assumed DHBP peaks at 3.6, 4.4, and 4.8 minutes were quantified using a standard curve for *p,p'*-DHBP. While 97% of the initial biphenyl concentration was accounted for after 40 min, the

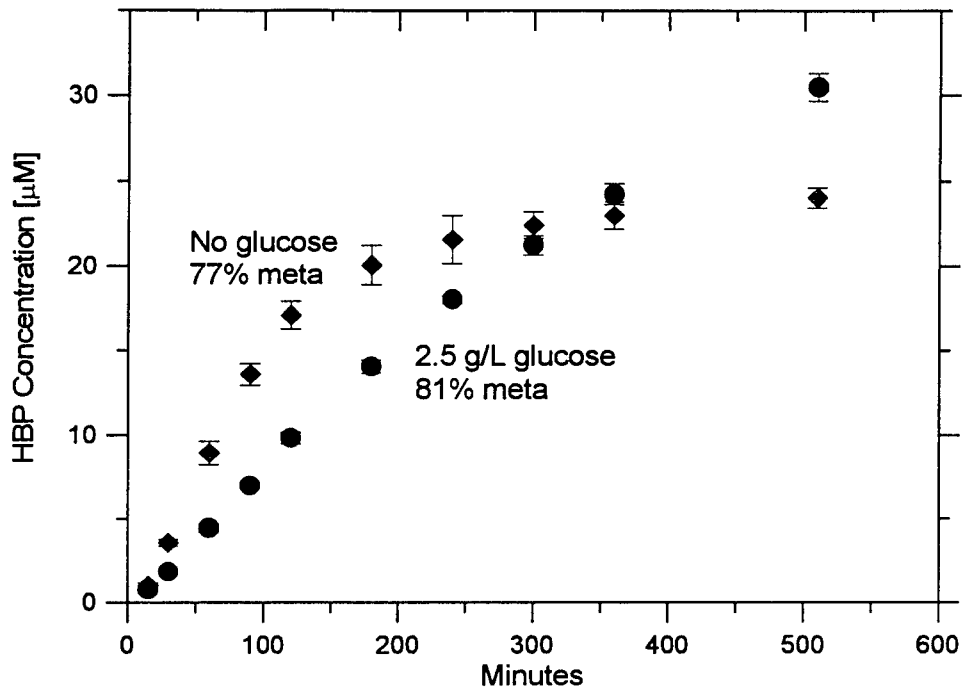


Figure 1. Total hydroxybiphenyl products formed by resting Tb4m cells in the presence of 5.0 g/L biphenyl solids. 20 mM KH_2PO_4 adjusted to pH 8.0. Results shown with and without 2.5 g/L glucose. 30 °C, 250 rpm. Error bars are 1 standard deviation for duplicate experiments. Avg protein concentration = 318 mg BSA/L.

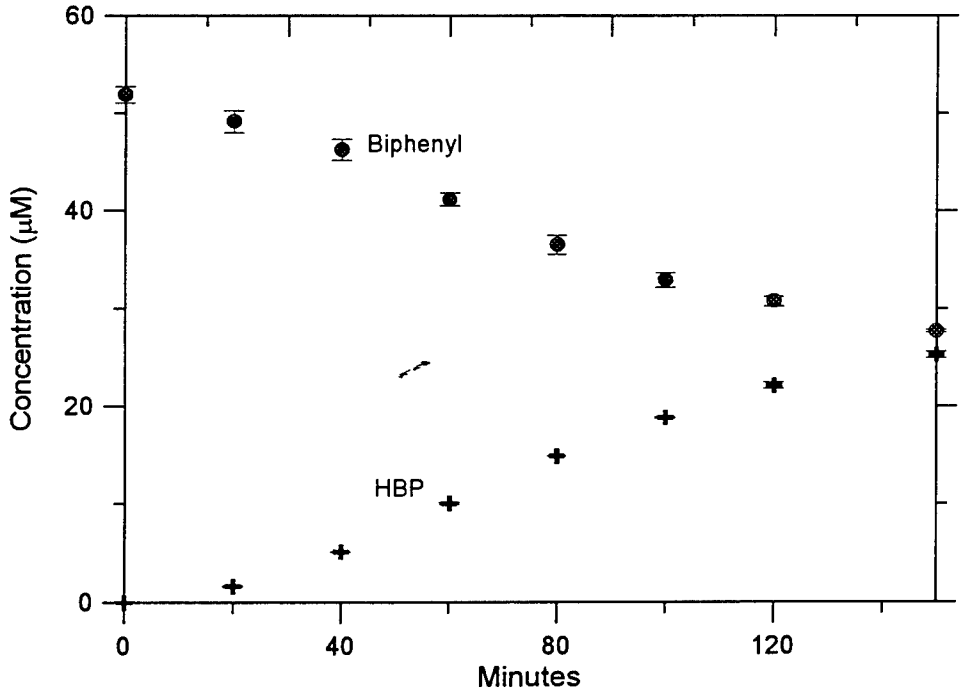


Figure 2. Serum bottle experiment for biphenyl hydroxylation by Tb4m. 52μM biphenyl in solution initially, no biphenyl solids. 20 mM KH_2PO_4 adjusted to pH 8.0. 30 °C, 250 rpm. Error bars are 1 standard deviation for duplicate experiments. Avg protein concentration = 259 mg BSA/L.

final material balance at 3 h was only 92% of the initial biphenyl present. Specific activity for biphenyl conversion was 7.2 x 10^{-4} µmol/mg protein/min in this case. Biphenyl disappearance with respect to time was linear from 0 to 100 min; the slope of the biphenyl curve increased afterwards.

To confirm the hydroxylation of HBP by Tb2m to form dihydroxybiphenyl (DHBP), experiments were conducted with either *meta* or *para*-HBP as the starting substrate. Disappearance of *meta*-HBP coincided with the formation of a compound which eluted at 4.8' on the HPLC. This unknown did not match HPLC standards for *p,p'*-HBP, *o,o'*-HBP, or *m,p*-HBP; however, it was identical to the 4.8' unknown observed in Tb2m experiments with biphenyl as the starting substrate. The initial concentration of *m*-HBP was 193 µM, and the initial specific activity for *m*-HBP conversion was 15.8 x 10^{-4} µmol/mg protein/min. The *m*-HBP concentration declined by 37% over 3 h; at this point, 81% of the initial *m*-HBP was still accounted for (assuming the 4.8' peak was a DHBP compound with an extinction coefficient equal to *p,p'*-HBP). When *m*-HBP solutions were inoculated with Tb4m cells, 97% of the substrate remained after 3 hours.

When resting cells of Tb2m were added to solutions containing *p*-HBP, substrate disappearance coincided with the formation of a 4.4' unknown on the HPLC chromatogram. Again, this unknown did not match any available DHBP standards, but did match 4.4' peaks observed in Tb2m-biphenyl experiments. The *p*-HBP concentration declined from 156 µM at time zero to 64 µM after 3 h (a decrease of 59%). Assuming the 4.4' metabolite to be DHBP with an extinction coefficient equal to that of *p,p'*-DHBP, the overall material balance was 70% after 3 h. The specific activity for *p*-HBP conversion was 29.8 x 10^{-4} µmol/mg protein/min initially. As with *m*-HBP, no significant *p*-HBP conversion was observed with Tb4m cells (97% of the initial *p*-HBP remained after 3 h).

Test tube experiments were conducted with 5.0 g/L biphenyl solids to determine pH effects on specific hydroxylation activity. 20 mM KH_2PO_4 buffer was modified with 1.0 N NaOH to obtain the desired pH between 5.0 and 9.0. Results are shown in *Figure 3* for both Tb2m and Tb4m.

For Tb2m, no significant trend was observed for specific activity as a function of pH between 5 to 9; average values varied between 7.0×10^{-4} and 8.1×10^{-4} μmol/mg protein/min. Conversely, pH did affect the specific biphenyl activity for Tb4m; the average activity increased an order of magnitude from 0.45×10^{-4} at pH 5 to 5.1×10^{-4} μmol/mg protein/min at pH 9. pH 9 was beyond the effective buffering range of KH_2PO_4; the final pH value was measured as 8.1 after 3 h.

Hydroxylation of diphenylacetylene. Shake flask experiments were also conducted with excess diphenylacetylene solids, since HDPA also has important polymer applications. In this case 1.0 g/L DPA solids were added to the phosphate medium (aqueous DPA concentration at saturation ~1 mg/L). Results for Tb2m and Tb4m are shown in *Figure 4*. For Tb2m, the initial specific hydroxylation activity was 2.0×10^{-4} μmol HDPA/mg protein/min, compared with 0.11×10^{-4} μmol HDPA/mg protein/min for Tb4m cells. DPA metabolites were observed at 12.3' and 12.6' in the HPLC analysis; the 12.3' peak matched the *para*-HDPA standard. The 12.6' peak was speculated to be *meta*-HDPA (due to similarities in residence time and DAD spectra, as well as the predominance of *meta*-HBP in the biphenyl experiments), but an HPLC standard has yet to be obtained. The data in *Figure 4* assumes the 12.6' metabolite is *meta*-HDPA with an extinction coefficient (A_{254}) identical to *para*-HDPA. The Tb2m HDPA product was 70% *meta*/30% *para* isomer after 6 h, while the Tb4m HDPA product was 76% *meta*/24% *para* after 6 h.

To establish reaction stoichiometry, DPA experiments were also conducted in sealed serum bottles to minimize any sorption/volatilization losses. Phosphate buffer was saturated with DPA overnight, then filtered (1 μM glass fiber) to remove all DPA solids. Reactions were conducted in duplicate with Tb2m cells only (due to the low DPA activity observed in *Figure 4* for Tb4m cells). The DPA concentration in solution decreased from 6.4 μM at time zero to 3.5 μM after 3 h. The HDPA concentration increased from zero initially to a plateau concentration of 0.7 μM after 1 h. After 5.4 h, only 48% of the initial DPA was accounted for as either DPA or HDPA; unknown peaks at 2.7' and 2.9' steadily increased in area from time zero, and were 138% of the initial DPA peak area after 5.4 h. It was speculated that HDPA further hydroxylated to form dihydroxydiphenylacetylene (DHDPA) compounds observed at 2.7 and 2.9 min; DHDPA

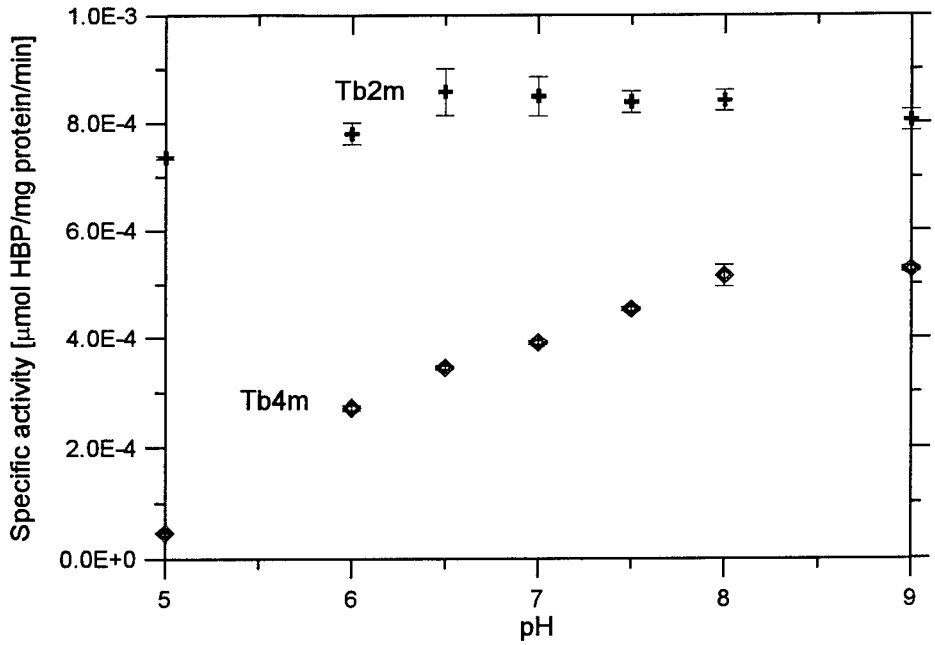


Figure 3 Specific hydroxylation activity for resting cells in the presence of 5.0 g/L biphenyl solids. Results are for Tb2m and Tb4m cells as a function of pH. 20 mM KH_2PO_4 adjusted to pH shown. 30 °C, 250 rpm. Error bars are 1 standard deviation for duplicate experiments. Activities based on measured HBP product after 60 min.

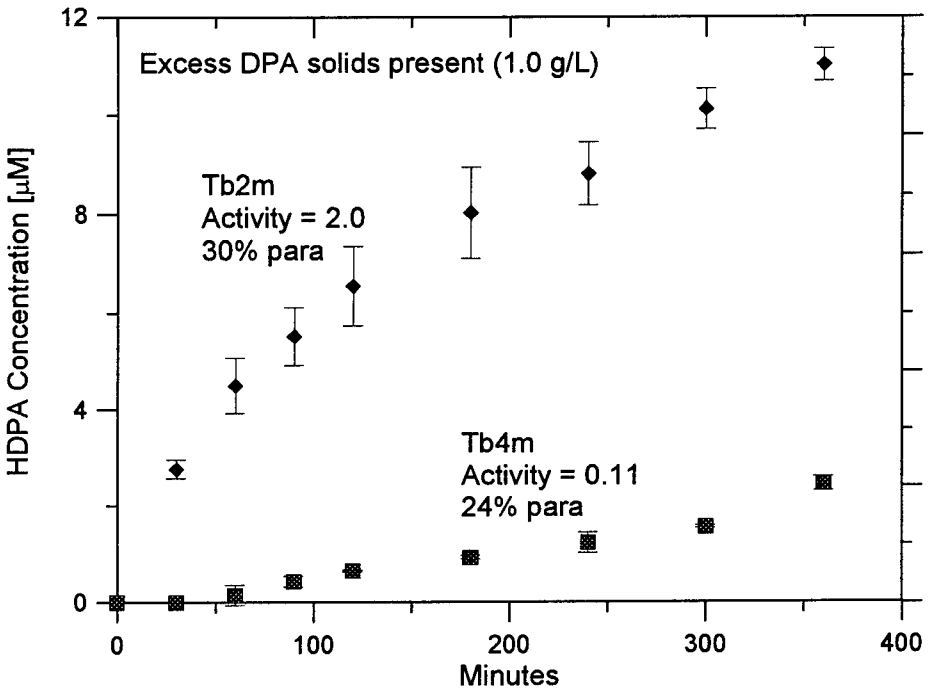


Figure 4 Total hydroxydiphenylacetylene (HDPA) products formed by resting cells in the presence of 1.0 g/L DPA solids. 20 mM KH_2PO_4 adjusted to pH 7.0 (Tb2m), 8.0(Tb4m). 30 °C, 250 rpm. Error bars are 1 standard deviation for duplicate experiments. Avg. protein concentration = 339 (Tb2m), 333 (Tb4m) mg BSA/L.

compounds were detected in culture extracts analyzed via GC-MS. Specific activity for Tb2m in the serum bottle experiments was lower than observed in *Figure 4* (0.93×10^{-4} μmol HDPA/mg protein/min). No DPA losses were observed in an abiotic serum bottle control without cells; 97% of the initial DPA in solution remained after 5.4 h.

Effect of surfactants. Four different surfactants were added to resting cell experiments with 5.0 g/L biphenyl solids and Tb4m cells; specific activity values (as determined by the amount of HBP produced after 30 min) for the various surfactant treatments and the no surfactant control are shown in *Figure 5*. Specific activities for Triton X-100, Tween 80, and Brij 58 were 6 to 31% lower than the no surfactant control. These three surfactants were evaluated at concentrations of 100 and 300 μg/mL. In contrast, experiments with sodium dodecyl sulfate (SDS) showed activities 30 and 45% greater than the control value (at SDS concentrations of 200 and 500 μg/mL, respectively).

DISCUSSION

The preferential formation of *meta*-HBP is quite fortuitous; *meta* hydroxylation of aromatic compounds is uncommon by either chemical or biochemical means [2]. Phenolic resins synthesized using *meta*-substituted phenols provide polymer chains with a narrow molecular weight distribution; this results in an increased glass transition temperature for the final polymers [8]. It appears that Tb2m can further catalyze HBP to form dihydroxybiphenyl; this compound is also an important polymer intermediate [9]. Hydroxylation of DPA resulted in mixtures of 24 to 30% *para*-HDPA, with the remainder of the product presumably the *meta* isomer; both isomers are potentially valuable precursors. Observed and speculated reactions are summarized in *Figure 6*.

Specific activities for the hydroxylation of biphenyl/DPA seem relatively low (0.11 to 8.1×10^{-4} μmol/mg protein/min), particularly when compared with specific activities for 3mM toluene

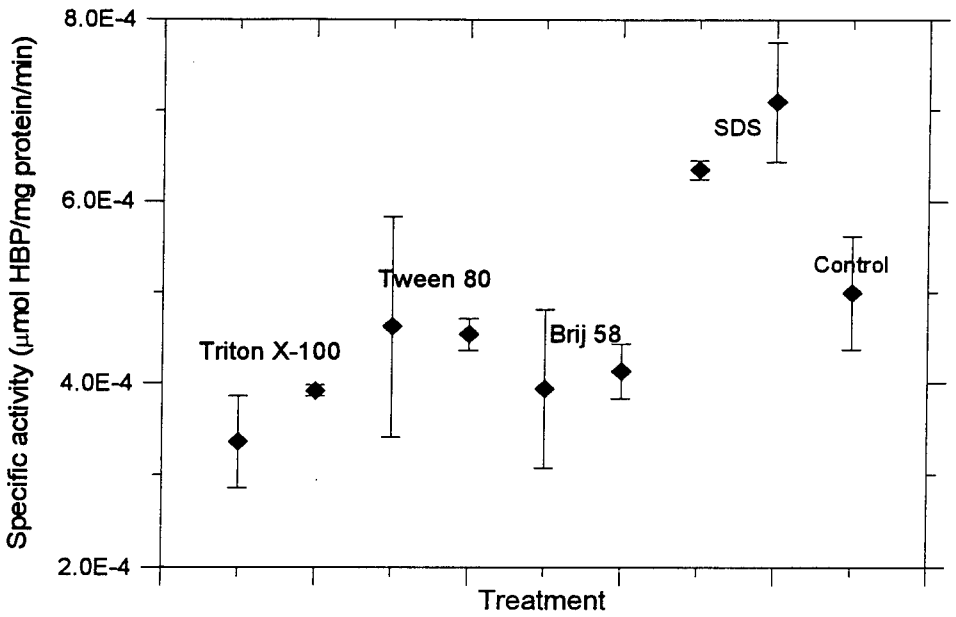


Figure 5 Specific hydroxylation activity for resting Tb4m cells in the presence of 5.0 g/L biphenyl solids and various surfactants. Surfactant concentrations (increasing from left to right) were 100 and 300 µg/mL except for SDS, which was 200 and 500 µg/mL. 20 mM KH_2PO_4 adjusted to pH 8.0. 30 °C, 250 rpm. Error bars are 1 standard deviation for duplicate experiments. Activities based on measured HBP product after 30 min.

(110 to 230 x 10^{-4} µmol/mg protein/min). At this point, it is difficult to state whether biphenyl/DPA are relatively poor substrates for Tb2m/Tb4m, or if the low activities are simply due to low substrate concentrations in aqueous solution. Lower activities for DPA compared to biphenyl are consistent with the lower solubility of DPA, but the slightly different structure of the DPA molecule could also decrease enzyme activity. Tb2m activities for *m/p*-HBP were 2 to 3.7 times the maximum biphenyl activity; in this case the molecular structure was slightly different, but the aqueous solubility was approximately 3 to 4 times that of biphenyl. While not conclusive, this suggests that agents which increase the aqueous concentration of biphenyl/DPA (surfactants and organic solvents) might also increase the specific hydroxylation activity.

Abramowicz et al. used a fine suspension of 4-bromobiphenyl particles (~5µm diameter) to synthesize 4-bromo-4'-hydroxybiphenyl using *Aspergillus parasiticus* cultures [9]. As the cell

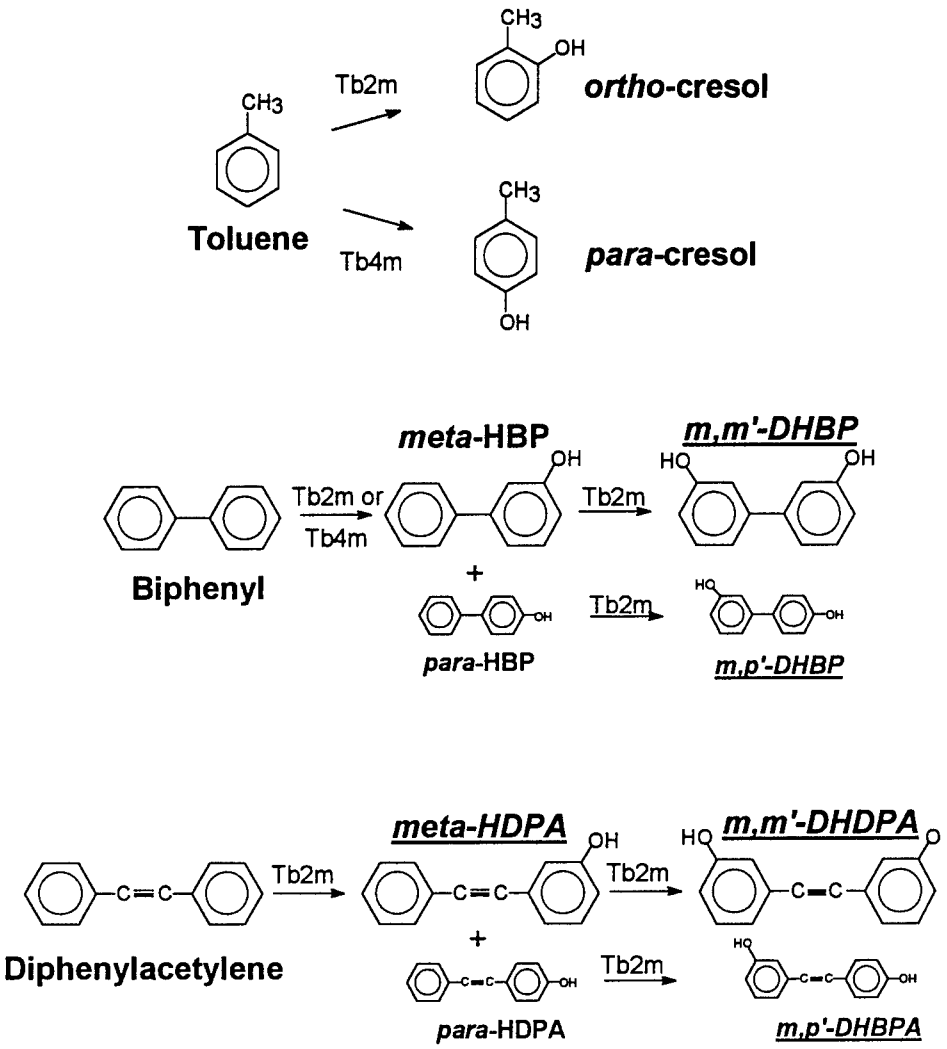


Figure 6. Summary of reactions observed with Tb2m and Tb4m in this study. Underlined isomers have yet to be confirmed with HPLC standards. Smaller structures denote less significant pathways.

mass was decreased from 10 to 0.6 g/L, specific activity increased from 7 x 10^{-4} to 62 x 10^{-4} mol/g cell/day, respectively. This increase in specific activity suggests the hydroxylation may have been substrate (mass transfer) limited at the higher cell mass concentration. In comparison, the specific biphenyl activity was estimated at 5 x 10^{-4} mol/g cell/day for our *Pseudomonas aeruginosa* cultures (approximately 0.7 g cell mass/L and 1000 μm biphenyl particle diameter). Since this value is comparable to the lower *A. parasiticus* value, higher activities may be possible with improved agitation/mass transfer.

In comparing the Tb2m enzyme with Tb4m, higher specific activities were observed with Tb2m when toluene, HBP, and DPA were the substrates. Specific activities for biphenyl varied; *Figure 3* shows higher biphenyl activities for Tb2m, while Tb4m activities for the experiment shown in *Figure 1* were higher than observed with analogous Tb2m experiments. Cultures were maintained by continuously streaking isolated plate colonies onto fresh agar; while this agar contained antibiotics to select for cells containing (pRO2037, pRO2369) or (pRO2038, pRO2369) plasmids, it is possible that Tb2m mutations may have led to decreased activity in later experiments. If a monohydroxylated biphenyl is desired as the final product, the Tb4m enzyme would be advantageous since it does not react with HBP to form DHBP. Similarly, Tb2m would be the enzyme of choice if HDPA or DHBP are the desired phenolic products.

Experiments with Tb2m and *m*-HBP as the starting substrate yielded a 4.8' metabolite in the HPLC analysis. Since any *m*-HBP metabolite must have one hydroxyl group at the *meta* position, and the 4.8' compound did not match the *m,p*-DHBP standard, this compound is most likely *m,m'*-DHBP. A supplier of *m,m'*-DHBP has been identified, and a follow-up analysis will be made when this standard is obtained. Similarly, the 4.4' metabolite of *p*-HBP must have at least one hydroxyl group at the *para* position; since *p,p'*-HBP and *m,p*-HBP standards did not match, this metabolite is most likely *m,p'*-HBP. Again, a follow up analysis will be made when a *m,p'*-HBP standard is obtained.

It is not readily apparent why glucose would inhibit the activity of toluene monooxygenase enzymes. Since the molecular structure of glucose is significantly different from toluene, it is unlikely that glucose would act as a competitive inhibitor for the active site of Tb2m or Tb4m. In malolactic fermentations with *Leuconostoc oenos,* glucose-induced inhibition was due to an accumulation of cellular NADH during glucose catabolism [10]; a similar, indirect effect may explain Tb2m/Tb4m inhibition with glucose. Sustained monooxygenase activity in the presence of glucose was expected (*Figure 1*, 2.5 g/L glucose curve), since reducing equivalents (NADH or NAD(P)H) are generated during glucose metabolism and required for the hydroxylation reaction.

It also seems unusual that pH would affect Tb4m but not Tb2m, given that both intracellular enzymes are expressed in the same host organism. Most microorganisms maintain an internal pH close to neutrality, despite wide variations in the external pH [11]; thus, the external pH must have some indirect effect that is specific to the Tb4m enzyme.

In the serum bottle experiments, disappearance of biphenyl from solution was linear with respect to time for the first 100 minutes. This suggests that the reaction kinetics were zero-order with respect to biphenyl concentration and that the Michaelis constant (K_m) for this hydroxylation is significantly less than 50 μM. This is unusual; one would expect to see first-order (with respect to biphenyl) reaction kinetics at such low aqueous concentrations. It may be that the absolute change in biphenyl concentration (~20 μM) was so small that changes in the reaction rate (slope of biphenyl curve in *Figure 2*) were not apparent. A decrease in the biphenyl conversion rate was observed after 100 min; this may be concentration related in the case of Tb2m, but *Figure 1* shows a decline in specific enzyme activity for Tb4m (no glucose) after 2 h. Thus, it becomes difficult to distinguish whether decreased rates are due to lower concentration or lower enzyme activity. Extended kinetic studies at different initial substrate concentrations are needed to truly establish hydroxylation rates as a function of biphenyl concentration.

The addition of surfactants can increase the apparent solubility of organic substrates in aqueous solution; however, mixed results have been observed when surfactants are added to cultures utilizing solid substrates [12]. While surfactants can increase solubility and substrate transport, they may also be toxic or act as a competitive substrate. The partitioning of a hydrophobic substrate in the micelle interior can sometimes prevent cellular uptake. Since surfactant effects are highly dependent on the specific system, screening experiments like the one shown in *Figure 5* are needed to find the best surfactant and the optimal concentration. Results shown in *Figure 5* are in contrast to other results in the literature; SDS inhibited the biodegradation of phenanthrene solids by a *Mycobacterium* species [13], and Triton X-100 enhanced the biphenyl biodegradation rate for a soil isolate [14]. The mechanism by which surfactants enhance biotransformation rates is uncertain; while surfactants can enhance substrate solubility, positive effects may also result

from an influence on cellular permeability [12]. A wide variety of surfactants are commercially available; further screening studies with other surfactants are warranted for this system.

CONCLUSIONS

This research demonstrated that toluene monooxygenase enzymes can be used to formulate important polymer precursors (phenolic compounds) via biocatalysis. Since the properties of these thermal-resistant polymers depend on the particular isomer as well as the isomeric purity of these phenolics, biocatalysis may provide a unique advantage over conventional organic synthesis. The enzymatic hydroxylation of biphenyl and diphenylacetylene was predominantly at the *meta* position (70 to 91% isomeric purity); this isomer is difficult to formulate, and may provide polymers with enhanced properties. Unfortunately, observed biocatalysis rates in this study were relatively low, possibly due to the low aqueous solubility of biphenyl/DPA. A preliminary study with four surfactants indicates it may be possible to enhance these rates. Future work should investigate additional methods for increasing biocatalysis rates with these enzymes.

REFERENCES

1. Roberts, S.M., *et al.*, *Introduction to biocatalysis using enzymes and micro-organisms.* 1995, Cambridge: Cambridge University Press.

2. Faber, K., *Biotransformations in organic chemistry.* 3rd ed. 1997, Berlin: Springer-Verlag.

3. Gibson, D.T., *Microbial metabolism of aromatic hydrocarbons and the carbon cycle,* in *Microbial metabolism and the carbon cycle,* S.R. Hagedorn, S. Hanson, and D.A. Kunz, Editor. 1988, Harwood Academic Publishers: Chur, Switzerland. p. 43-52.

4. Johnson, G.R. and R.H. Olsen, *Nucleotide sequence analysis of genes encoding a toluene/benzene-2-monooxygenase from Pseudomonas sp. Strain JS150.* Applied and Environmental Microbiology, 1995. **61**(9): p. 3336-3346.

5. Johnson, G.R. and R.H. Olsen, *Multiple pathways for toluene degradation in Burkholderia sp. strain JS150.* Applied and Environmental Microbiology, 1997. **63**: p. 4047-4052.

6. Stanier, R.Y., N.J. Palleroni, and M. Doudoroff, *The aerobic pseudomonads: a taxonomic study.* Journal of General Microbiology, 1966. **43**: p. 159-271.

7. Smith, P.K., *et al.*, *Measurement of protein using bicinchoninic acid.* Analytical Biochemistry, 1985. **150**: p. 76-85.

8. Zampini, A., et al. Esterification product of aromatic novolak resin with quinone diazide sulfonyl group. 1995. US Patent 5589553.

9. Abramowicz, D.A., C.R. Keese, and S.H. Lockwood, *Regiospecific hydroxylation of biphenyl and analogs by Aspergillus parasiticus,* in *Biocatalysis,* D.A. Abramowicz, Editor. 1990, Van Nostrand Reinhold: New York. p. 63-92.

10. Miranda, M., *et al.*, *Biochemical basis for glucose-induced inhibition of malolactic fermentation in Leuconostoc oenos.* Journal of Bacteriology, 1997. **179**(17): p. 5347-5354.

11. Prescott, L.M., J.P. Harley, and D.A. Klein, *Microbiology.* 1990, Dubuque, IA: Wm. C. Brown.

12. Mihelcic, J.R., *et al.*, *Bioavailability of sorbed- and separate-phase chemicals.* Biodegradation, 1993. **4**: p. 141-153.

13. Tiehm, A., *Degradation of polycyclic aromatic hydrocarbons in the presence of synthetic surfactants.* Applied and Environmental Microbiology, 1994. **60**(1): p. 258-263.

14. Roch, F. and M. Alexander, *Biodegradation of hydrophobic compounds in the presence of surfactants.* Environmental Toxicology and Chemistry, 1995. **14**(7): p. 1151-1158.

PERFORMANCE MODELING AND SCALABILITY ANALYSIS OF THE NAVIER-STOKES SOLVER FDL3DI ACROSS MULTIPLE PLATFORMS

David E. Hudak
Associate Professor
Department of Mathematics and Computer Science

Ohio Northern University
Ada, OH 45810

Final Report for:
Summer Faculty Research Program
Wright Laboratory
Aeromechanics Division

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Laboratory

August 1998

PERFORMANCE MODELING AND SCALABILITY ANALYSIS OF THE NAVIER-STOKES SOLVER FDL3DI ACROSS MULTIPLE PLATFORMS

David E. Hudak
Associate Professor
Department of Mathematics and Computer Science

Ohio Northern University

A preliminary parallel version of the FDL3DI application from Wright Laboratory was developed using MPI. This report describes the overall design of the parallel code and estimates the performance of the code. For this data-parallel implementation, a single grid is broken into subgrids and each subgrid is assigned to a separate processor. The boundary points are exchanged between processors as necessary.

Scalability analysis was performed to indicate the performance of the code for large data sets and large numbers of processors. Experimental results were obtained on the Cray T3E, IBM SP and SGI Origin 2000. These results demonstrate that parallel systems can provide significant performance improvements for this application.

PERFORMANCE MODELING AND SCALABILITY ANALYSIS OF THE NAVIER-STOKES SOLVER FDL3DI ACROSS MULTIPLE PLATFORMS

David E. Hudak

1 Introduction

The potential of massively parallel computer systems (MPPs) for the solution of large computational aeroscience problems has yet to be realized. In order for an application to attain high performance on a parallel computer it is essential that the problem is partitioned into subproblems such that the work is shared among the processors. This decomposition of a single problem into subproblems exploits *data parallelism*, and is achieved by splitting the global data set among the processors in such a way that the work load is distributed evenly.

In this project, I performed performance modeling, scalability analysis and obtained experimental timings for a data-parallel version of the FDL3DI application currently under development at the the Aeromechanics Division of Wright Laboratory. FDL3DI solves the three-dimensional Navier-Stokes equations using the approximate-factorization algorithm of Beam and Warming in conjunction with a Newton subiteration procedure to enhance the accuracy for rapid fluid motion. The parallel version of the code was written using the MPI message-passing library on an IBM SP2. The code was ported to an SGI Origin 2000 and Cray T3E. The sample problem was a model of uniform flow over a flat plate.

This report is organized as follows. The parallel version of the FDL3DI application is examined and its performance is modeled in Section 2. Scalability of the code is discussed in Section 3. Experimental runs on the target platforms are described in Section 4.

2 Parallelization Strategy and Performance Model

Parallelization of an application for a message-passing environment involves two decisions. First, a partition must be defined which allocates portions of the global data sets to the local memories of individual processors. Second, a format for messages must be established to ensure that processors are sending information according to the correct protocol, i.e., the right data to the right processor.

2.1 Partitioning Strategy

Valid partitions are required to satisfy the following conditions similar to [7] [6].

1. Tessellation. The partition must tile the data set so that each point in the iteration space is assigned to exactly one processor.

2. Identical by translation. Each partition should be the image of any other one by a translation, except for parts located near boundaries of the iteration space. This property promotes efficient

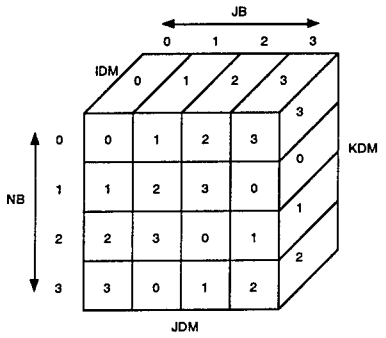


Figure 1: 2D Multipartitioning Strategy

code generation.

3. Straight Line Perimeters. In order to ensure tiling and simplify code generation, the perimeter of each part is composed of planes.

4. Symmetry. This requirement is satisfied by most shapes that tessellate and are simple to code.

The partitioning strategy chosen for this application was an extension of the multipartioning strategy devised by Smith et. al. [9] and examined by Jun [8]. For this report, I will consider a $N \times N \times N$ grid is being partitioned for P processors. Note that this is a restriction of the general case presented in Fig. 1 where $IDM = JDM = KDM = N$. In this case, the algorithm creates P^2 different partitions, each of size

$$\left\lceil \frac{P}{N} \right\rceil \times \left\lceil \frac{P}{N} \right\rceil \times N \tag{1}$$

as illustrated for $P = 4$ in Fig. 1. Note that this strategy satisfies the above four criteria.

In addition, it is important that a partitioning strategy evenly divide the computational load across the processors. For the FDL3DI application, the overall computation required is proportional to the size of the data set, i.e., $O(N^3)$. In order to achieve good load balance, each processor should be responsible for $O(N^3/P)$ computes. In the two-dimensional multipartitioning scheme, each processor must compute P partitions, each with volume equal to $\frac{N}{P} \times \frac{N}{P} \times N$. Therefore, the total number of points per processor is

$$\left(\frac{N}{P} \times \frac{N}{P} \times N\right) \times P = \frac{N^3}{P^2} \times P = \frac{N^3}{P} \tag{2}$$

and the computation per processor is $O(N^3/P)$, which is the desired result. Therefore, the two-dimensional multipartitioning strategy will achieve good load balancing. Each of the major three-dimensional arrays, X, Y, Z, U, V, W, P, RHO and EDDY, are partitioned in this manner.

2.2 Potential 3D Partitions

The major computation in this problem centers around dimensional sweeps. For example, a sweep along the I dimension would compute U(1,1,1), U(2,1,1), U(3,1,1), etc. This partition allows for all processors

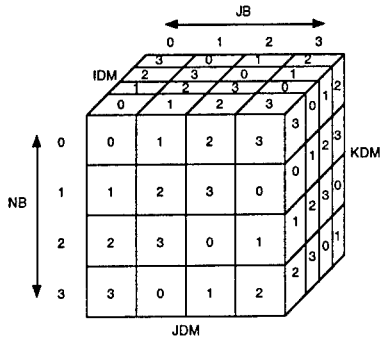


Figure 2: 3D Multipartitioning Strategy

to simultaneously work during the sweeps along the J and K dimensions. The sweep along the I dimension remains sequentialized for this implementation of the code.

Although not implemented, it is possible to extend the partitioning strategy to three dimensions. The partition is illustrated in Fig. 2. This would allow for more computation to be parallelized and hence should generate higher levels of parallel efficiency. However, this benefit must be weighed against the amount of communication and the programming difficulty.

When evaluating computation and communication requirements of such partitions, it is useful to compare the volume of the partitions (which is proportional to the amount of computational load) and the surface area of the partitions (which is proportional to the amount of interprocessor communication) [4]. For the two-dimensional partitioning in Fig. 1 involving an $N \times N \times N$ array split among P processors, the volume per part is $\frac{N}{P} \times \frac{N}{P} \times N$ and the total volume for all P parts per processor is $\frac{N^3}{P}$. The surface area for an interior part is $4 \times \frac{N}{P} \times N$ and, at P parts per processor has a total of $4N^2$. For the three-dimensional partitioning in Fig. 2, the volume per part is $\frac{N}{P} \times \frac{N}{P} \times \frac{N}{P}$ and the total volume for all P^2 parts per processor is again $\frac{N^3}{P}$. The surface area for an interior part is $6 \times \frac{N}{P} \times \frac{N}{P}$ and, at P^2 parts per processor has a total of $6N^2$.

From the above analysis, it is clear that the three-dimensional partitioning will have the same load balance as the two-dimensional partitioning ($\frac{N^3}{P}$) and will generate considerably more communication (50% more). However, it will allow for sweeps along the I dimension to be parallelized, which will improve scalability. This tradeoff is discussed in Section 3.1.

2.3 Interprocessor Communication

Interprocessor communication occurs during the sweeps along the parallelized dimensions. In order to compute the correct answer, each processor must know whom it is sending data to, what data to send and how it should be packed into messages. Let us consider each question individually.

In the partition illustrated by Fig. 1, sweeps along either parallelized dimension result in cyclical communication patterns. For example, in a sweep along the K-dimension, processor 0 always sends to processor 1, processor 1 always sends to processor 2, and so on as illustrated in Fig. 3. Note that for dimensional sweeps in the opposite direction (e.g., sweeping "down" along the K dimension and then

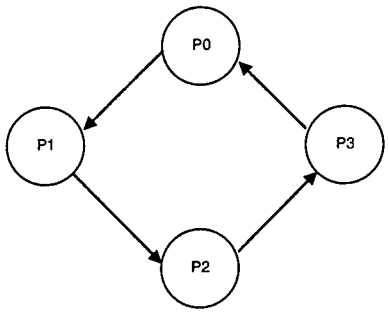


Figure 3: Communication Patterns for Parallel Sweeps

sweeping "up") the communication pattern is reversed.

The amount of data transmitted between processors may be different for the dimensional sweep along the K dimension as along the J dimension. In each case, the data transmitted (the border of the part) takes the shape of a two-dimensional plane. For communication along the K dimension, this plane has size $\lceil \frac{JDM}{P} \rceil \times IDM$ and has size $\lceil \frac{KDM}{P} \rceil \times IDM$ for sweeps along the J dimension. These borders must be transmitted for each of the nine three-dimensional arrays.

The current implementation of the code uses the MPI library for message passing [5]. In this implementation, one MPI message consists of a vector from each array involved in the computation. This means that the average message size is $9 \times IDM$ and approximately $\lceil \frac{N}{P} \rceil$ messages are sent for a general $N \times N \times N$ array in order to communicate across one border.

This model has two major implications. First, it may be more efficient to use MPI datatypes for communication. For example, the periodic boundary conditions in the flat-plate problem were coded using an MPI datatype for a two-dimensional border. Using this border, I was able to send 9 messages of size $\lceil \frac{N}{P} \rceil \times IDM$ as opposed to the strategy discussed in the preceding paragraph. This method greatly reduces the number of message sent and will reduce message start-up overhead. Second, it may be more efficient to use an MPI virtual topology to express the ring-structured communication (as seen in Fig. 3). Use of this topology would simplify programming and may be used as advice by the runtime system to effectively map the communication requirements of the application to the interconnection network of the computer.

3 Scalability Analysis of Parallel Code

The motivation for seeking a parallel solution for FDL3DI centered on increased performance and examination of larger grids. Therefore, it is important to consider the maximum number of processors which can be used for a grid of a certain size and to consider the maximum possible problem size for a given machine configuration, i.e., the number of processors and the amount of RAM per processor.

N	T_P	$S(\infty)$
41	92%	12
100	92.6%	13
200	94.1%	17

Table 1: Parallelizable Portion of Code for Various Grid Sizes.

3.1 Scaling the Number of Processors

When scaling a fixed-size problem to large numbers of processors, the two major considerations are parallel speedup and parallel efficiency. For an identical problem, let T_A be the time for running the sequential version of the application and $T_A(P)$ be the time for running the parallelized version of the application on P processors. The *speedup for P processors* is $S(P) = \frac{T_A}{T_A(P)}$. Note that, ideally, $S(P) = P$. The *parallel efficiency for P processors* is $E(P) = \frac{P}{S(P)}$. Note again that, ideally, $E(P) = 100\%$. Good performance is characterized by high efficiency.

Scalability of a parallel application is one application of **Amdahl's Law** [1]. The total execution time of an application can be divided into two components: T_P is the portion of the execution time which can be parallelized and T_S is the portion of the execution time which cannot be parallelized. Therefore, since $T_A = T_S + T_P$ the best possible execution time for P processors is

$$T_A(P) = T_S + \frac{T_P}{P} \tag{3}$$

and, as $P \rightarrow \infty$, $T_A(P) \rightarrow T_S$.

The implications of equation 3 are significant for parallel code. For example, suppose T_S is 25% of the total execution time. The maximum speedup is

$$S(\infty) = \frac{T_A}{T_A(\infty)} = \frac{T_A}{T_S} = \frac{T_A}{0.25T_A} = \frac{1}{0.25} = 4$$

Therefore, regardless of the number of processors devoted to this mythical application, the maximum attainable speedup is four.

Ultimately, the scalability of FDL3DI will be determined by the relative size of T_P with respect to T_A. In order to examine this, a code profiler (pixie) was used to determine the percentage of execution time required by routines which are parallelized in FDL3DI. The results are reported in Table 1. Note that for very large grids, the maximum attainable speedup is near twenty.

The major factor for the low scalability in the code is a major piece of computational work remains sequentialized, i.e., the sweep in the I dimension. As discussed previously in Section 2.2, a three-dimensional multipartition would allow for parallelization of the sweep along the I dimension and, as a result, would scale to higher speedups. It is an open question whether this increase in parallelism would be worth increasing the communication by fifty percent.

Machine	SGI Origin 2000	IBM SP	Cray T3E
CPU Speed	195 Mhz	135 Mhz	300 Mhz
Cache (L2)	4 MB	None	96 KB
RAM/node	512 MB	1 GB	128 MB
Link Bandwidth	800 MB/s	480 MB/s	600 MB/s

Table 2: Performance Statistics for Target Platforms

3.2 Scaling the Problem Size

Often in scientific codes such as FDL3DI, examination of interesting problems is restricted due to the ability of the target platforms to hold large data sets. For example, the size of the FDL3DI executable for a $100 \times 100 \times 100$ grid is approximately 150 MB on a Cray T3E. It is impossible to run this program on a single processor of the Cray T3E used for the experimental work since it had only 128 MB of RAM per processor. However, when parallelized for two processors, the size of the executable for each processor was 80 MB, and thus able to fit in RAM. Therefore, running larger data sets is possible by increasing the number of processors used on the problem.

The relationship between problem size, RAM per processor and a partition constraint will be used to give the minimum and maximum number of processors on a given machine which may be used for a given sized problem. The lower bound on P is simple. P must be large enough so that the data generated by $O(N^3/P)$ data points fit in the single processor RAM. So, for our $100 \times 100 \times 100$ example above, $P \geq 2$. The upper bound is also simple. Due to message-passing constraints, the minimum width of a part is restricted to 4. Since the width of a part is given by N/P (Fig. 1 and equation 1), we have $(N/P) \geq 4$. Therefore,

$$P \leq N/4 \tag{4}$$

The above results are useful. For example, consider the execution of of a $200 \times 200 \times 200$ grid on the T3E. From an examination of the size of the sequential executable, it was observed that P had to be larger than fifty. However, examination of equation 4 reveals that $P \leq 50$. Therefore, this particular problem size **cannot** be run on the T3E for any number of processors.

4 Performance of Parallel Code

4.1 Target Platforms

Three parallel platforms were used to examine the correctness and performance of the application. The SGI Origin 2000 and IBM SP are installed at ASC MSRC and the Cray T3E is installed at CEWES. Important performance statistics for the machines are summarized in Table 2.

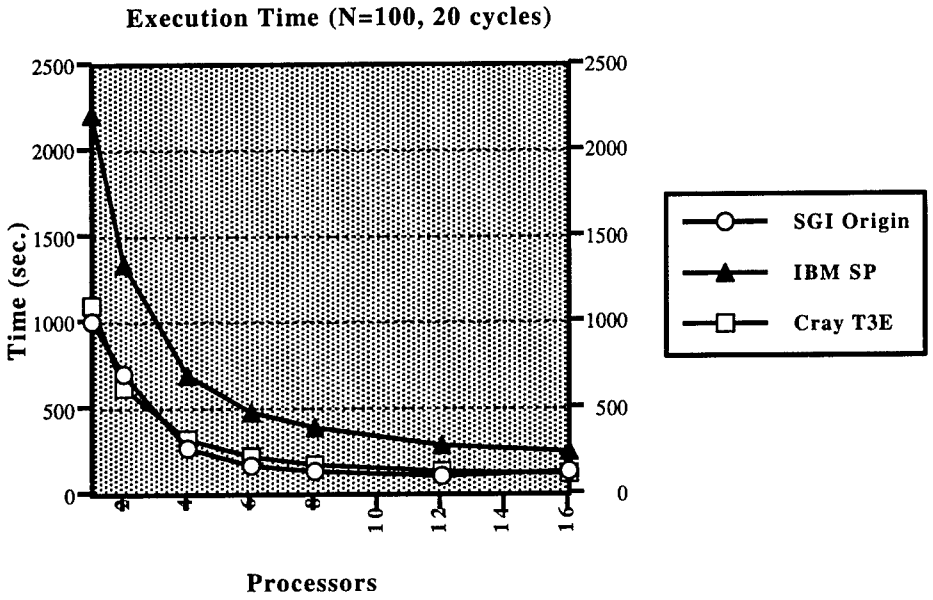


Figure 4: Execution Time for Multiple Platforms

4.2 Experimental Results

A problem was generated simulating uniform flow over a flat plate. The same source code and data were used for all three platforms. For all experiments, the timestep represented 0.005 sec. and the total time simulated in the experiment was 0.01 seconds. The execution times for a $100 \times 100 \times 100$ grid are given for each machine in Fig. 4. Due to the large per processor RAM requirements of this problem, it could not be run on a single processor of the IBM SP and Cray T3E. The sequential execution times for these machines reported in Fig. 4 are extrapolations of execution time relative to the performance of the SGI Origin 2000 on this problem. Note that the IBM SP is considerably slower than the other two machines. This effect is most likely due to the SP's slower processors (Table 2).

The parallel speedup is given in Fig. 5. The parallel efficiency is given in Fig. 6. Note that, for all machines, efficiency deteriorates for larger numbers of processors. Reasonable efficiency is maintained until about twelve processors.

In order to examine the impact of increased problem size on parallel performance, a grid with size $200 \times 200 \times 200$ was run on the Origin 2000 and compared with the results from the $100 \times 100 \times 100$ grid. The results are presented in Fig. 7. These results clearly demonstrate that large problem sizes allow the efficient use of greater numbers of processors. Note that it was impossible to run the sequential case on the Origin 2000 for the large grid. Therefore, the sequential times for smaller grids were obtained on the Origin 2000 and extrapolated via curve fitting, as shown in Fig. 8. Using this technique, I was able to generate an estimated execution time for the sequential code.

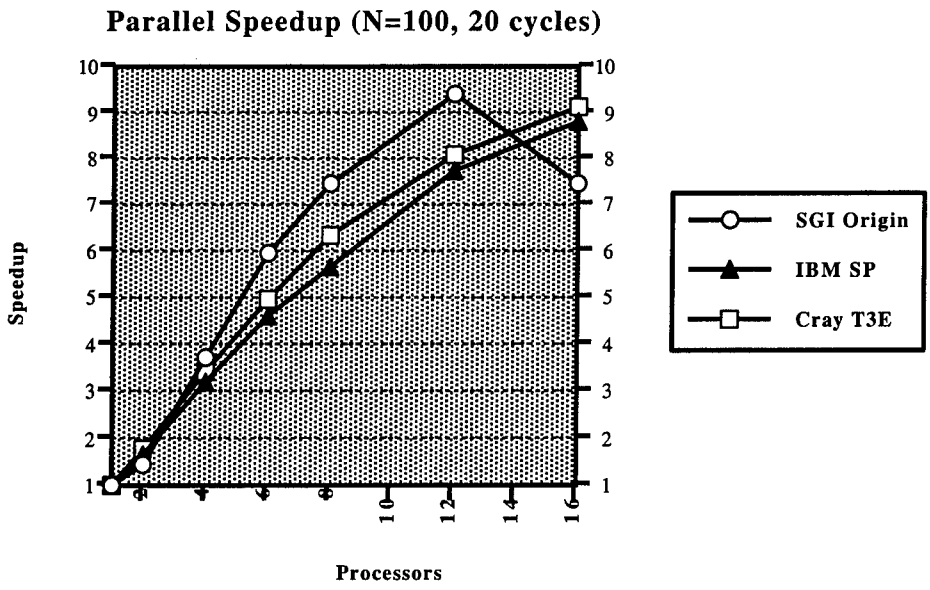


Figure 5: Parallel Speedup for Multiple Platforms

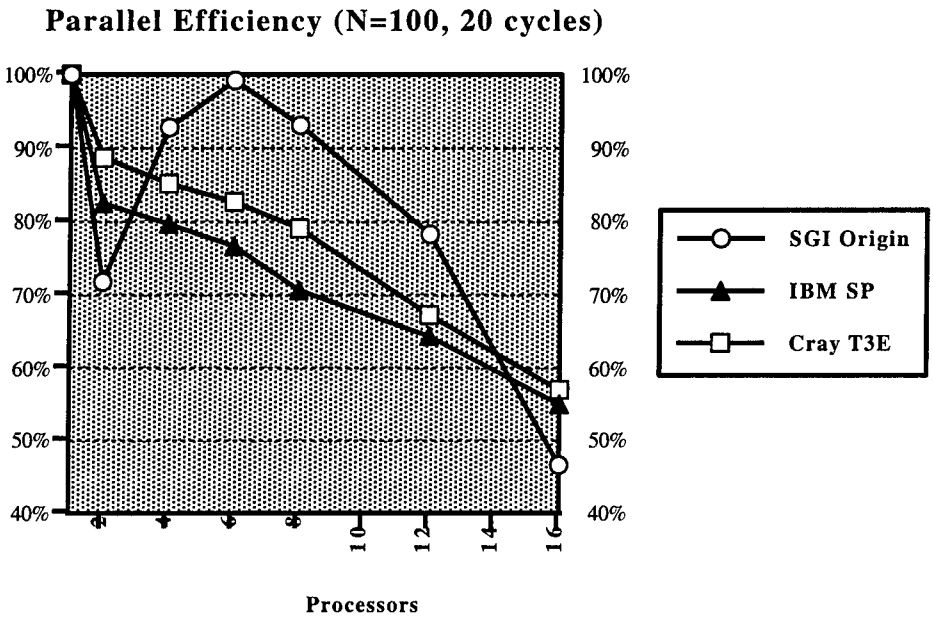


Figure 6: Parallel Efficiency for Multiple Platforms

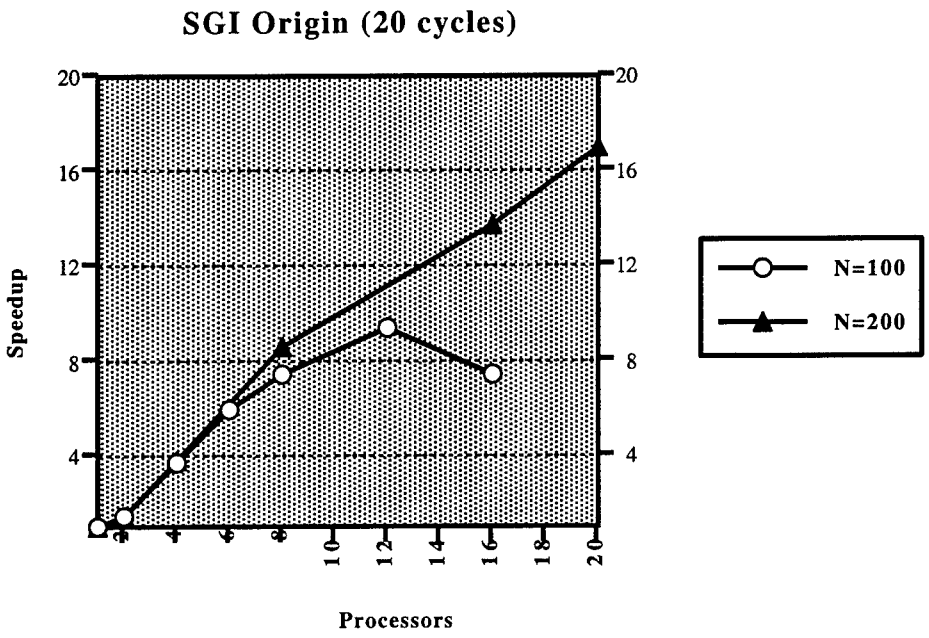


Figure 7: Parallel Speedup for Increasing Problem Size

5 Conclusion and Future Work

In this report I have demonstrated that the two-dimensional multipartition of FDL3DI can have good performance across a collection of parallel machines. This performance is maximized when the size of the grid and the number of processors are carefully balanced.

There are a collection of items which must be investigated to allow for maximum performance. These remain important areas for future work. The parallel efficiency of the code can be improved in multiple ways. First, the data alignment of the code can be improved. Current data alignment is consecutive by partition number. Alignment along the I dimension will improve performance. Second, the inclusion of MPI data types as described in Section 2.3 will improve parallel efficiency as well. Overlapping of computation with communication would amortize communication costs. Finally, if additional parallelism is required, three-dimensional partitioning could be used.

6 Acknowledgements

Thanks to Miguel Visbal and Joseph Shang for giving me the opportunity to work with the FIMC research group. Thanks to Drs. Sadayappan, Visbal, and Zha for their insight and advice.

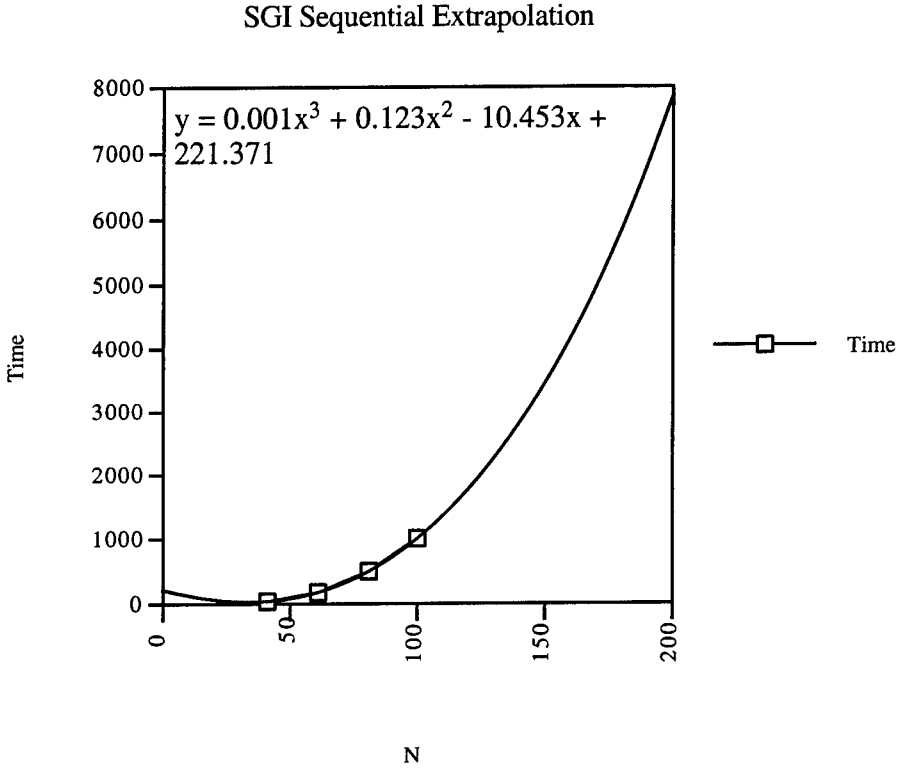


Figure 8: Curve Fitting to Extrapolate Sequential Execution Time

References

[1] G. M. Amdahl. Validity of single-processor approach to achieving large-scale computing capability. In *Proc. AFIPS Conf.*, pages 483–485, Reston, VA., 1967.

[2] Eric L. Boyd, Gheith A. Abandah, Hsien-Hsin Lee, and Edward S. Davidson. Modeling computation and communication performance of parallel scientific applications: A case study of the IBM SP2. Technical Report CSE-TR-236-95, University of Michigan, Department of Electrical Engineering and Computer Science, May 1995.

[3] T. Agerwala et al. SP2 system architecture. *IBM Systems Journal*, 34(2):152–184, 1995.

[4] G. Fox, M. Johnson, G. Lyzenga, S. Otto, J. Salmon, and D. Walker. *Solving Problems on Concurrent Computers*, volume I. Prentice-Hall, Englewood Cliffs, NJ, 1988.

[5] W. Gropp, E. Lusk, and A. Skjellum. *Using MPI: Portable Parallel Programming with the Message-Passing Interface.* MIT Press, Boston, MA, 1994.

[6] David E. Hudak and Santosh G. Abraham. *Compiling Parallel Loops for High Performance Computers: Partitioning, Data Assignment, and Remapping.* Kluwer Academic Publishers, 1993.

[7] F. Irigoin and R. Triolet. Supernode partitioning. In *Proceedings of the 15th Annual ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages*, pages 319–329, 1988.

[8] Yong-Du Jun. *Parallel Implementation of Factored Alternating Direction Implicit Method for Strongly Coupled Navier-Stokes and K-E Equations.* PhD thesis, University of Cincinatti, Cincinatti, OH, 1996.

[9] Merritt H. Smith, Rob F. Van der Wijngaart, and Maurice Yarrow. Improved multi-partition method for line-based iteration schemes. In *Computationl Aerosciences Workshop, NASA Ames Research Center*, March 1995.

[10] M. Snir, P. Hochschild, D. D. Frye, and K. J. Gildea. The communication software and parallel environment of the IBM SP2. *IBM Systems Journal*, 34(2):185–204, 1995.

SORPTION OF A NON-IONIC SURFACTANT VERSUS A DISSOLVED HUMIC SUBSTANCE TO A LOW ORGANIC CARBON SOIL

Dr. William P. Johnson
Assistant Professor
W. Wynn John
Graduate Research Assistant
Department of Geology and Geophysics

University of Utah
135 South 1460 East Rm. 719
Salt Lake City, UT 84112

Final Report for:
Summer Research Faculty Program
Air Force Research Laboratory
Materials & Manufacturing Directorate
Tyndall Air Force Base, FL

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

And

AFRL/MLQR Tyndall Air Force Base

September 1998

SORPTION OF A NON-IONIC SURFACTANT VERSUS A DISSOLVED HUMIC SUBSTANCE TO A LOW ORGANIC CARBON SOIL

Dr. William P. Johnson
Assistant Professor
W. Wynn John
Graduate Research Assistant
Department of Geology and Geophysics
University of Utah

Abstract

The limited ability to remove sorbed and non-aqueous phase contaminants from subsurface materials using simple pump-and-treat technology motivates research into the use of agents to solubilize these contaminants in order to facilitate their removal. Foremost among candidate agents being examined for this purpose are surfactants. In contrast, dissolved humic substances (DHS) have received little attention as potential remediation agents, although they strongly bind hydrophobic organic contaminants (HOC) and are readily available. In this study, a non-ionic octylphenol ethoxylate surfactant mixture was compared to Aldrich humic acid, a terrestrial DHS, with respect to sorption to a natural low organic carbon content sediment and depression of surface tension in high ionic strength and low ionic strength (artificial groundwater solution) conditions.

A comparison of aqueous concentrations of surfactant oligomers in the presence and absence of sediment indicated significant sorption of both small and large ethoxylate chain length oligomers. Sorption isotherms developed of individual surfactant oligomers indicated that surfactant oligomers with smaller ethoxylate (EO) chain lengths were preferentially adsorbed to the sediment. Sorption isotherms for total surfactant mass indicated that the maximum sorbed surfactant concentration 0.5 g/kg. Ionic strength did not appear to significantly affect the surfactant sorption characteristics under the conditions of the study. A comparison of aqueous concentrations of the DHS in the presence and absence of sediment indicated significant sorption of DHS to the sediment, with preferential sorption of the larger molecular weight DHS fraction. Sorption isotherms developed for the DHS indicate that the maximum sorbed DHS concentration equaled approximately 0.05 g/kg. Ionic strength did not appear to affect sorption of the DHS to the sediment, however, at low ionic strength (in the artificial groundwater solutions) a large molecular weight molecule (50,000-60,000 molecular weight), possibly an indication of DHS aggregation, was observed in the aqueous solutions. Further testing will be necessary to resolve this feature.

Surface tension measurements of the surfactant solutions indicated a critical micelle concentration (CMC) occurring at an aqueous surfactant concentration of approximately 0.5 g/L in the absence of sediment and approximately 1.0 g/L in the presence of sediment. The minimum surface tension achieved in the surfactant solution equaled approximately 37 dynes/cm regardless of the solution ionic strength. Only a small decrease in surface tension (~1-2 dynes/cm) was achieved in the presence of DHS even at the highest concentrations under both ionic strength conditions.

SORPTION OF A NON-IONIC SURFACTANT VERSUS A DISSOLVED HUMIC SUBSTANCE TO A LOW ORGANIC CARBON SOIL

Dr. William P. Johnson
W. Wynn John

Introduction

Coal tar and creosote form dense non-aqueous phase liquids (DNAPLs) comprised by highly hydrophobic organic compounds, specifically, polycyclic aromatic hydrocarbons (PAH). These compounds are highly insoluble (solubilities ranging from a few to hundreds of micrograms per liter), and are found at thousands of industrial sites in the United States (Peters and Luthy, 1993). Tetrachloroethene (PCE) and trichloroethene (TCE) are also DNAPL-forming compounds, and are among the most common contaminants detected in groundwater due to their higher solubility (few to hundreds of milligrams per liter) and relative recalcitrance to biodegradation (Pankow and Cherry, 1996). PCE and TCE comprise part of larger group of contaminants referred to as chlorinated solvents (CAH).

The relative insolubility of PAHs, and the insolubility of CAHs relative to the contaminant mass available in the parent DNAPL, makes these compounds untreatable by simple pump-and-treat approaches. Pump-and-treat inefficiency motivates research in the use of various solubilizing agents for remediation of subsurface contamination involving PAHs, CAHs, and the larger set of non-aqueous organic solvents. Because these contaminants occur in both non-aqueous and sediment-sorbed forms, remedial agents must be capable of enhancing contaminant dissolution from non-aqueous phases as well as enhancing contaminant desorption from soils and sediments. Surfactants are among the most extensively studied agents considered as candidates for enhancing the viability of pump-and-treat technology for remediation of the contaminants described above (West and Harwell, 1992; Edwards et al., 1991; Pennell et al., 1993; Abriola et al., 1993; Brown et al, 1994).

Dissolved humic substances (DHS) derived from soil hypothesized to represent as advantageous alternative to surfactants in situations where: 1) inadvertent mobilization of DNAPL is of concern, 2) sorption loss of agent to the stationary phase need to be minimized 3) the use of low micellar (sub-micellar) agent concentrations are effective. DHS, like surfactants, show strong solubilization of highly hydrophobic PAH, but unlike surfactants, do not require a minimum concentration (a critical micelle concentration, CMC) to form micelles and induce solubilization. This may allow the use of relatively low DHS concentrations in remediation (i.e. hundreds of mg/L for DHS as opposed to tens of g/L for surfactants).

For less hydrophobic aliphatic compounds such as PCE, soil-derived DHS may or may not solubilize these contaminants as well as surfactants. However, there may be an advantage of DHS use by virtue of their apparent relative small effect on DNAPL-water interfacial tension (IFT) (Johnson and John, 1998). The use of surfactants to enhance the dissolution of DNAPLs becomes potentially problematic

when mobilization of DNAPL is to be avoided. The known depression of NAPL-water IFT in the presence of surfactants may potentially cause the forces driving the DNAPL downward into the subsurface (DNAPL mass) to exceed the forces arresting DNAPL migration (DNAPL-water IFT) (West and Harwell, 1992; Pennell et al., 1996). DHS are hypothesized to result in lesser depression of the NAPL-water IFT for a given agent concentration due to the relatively complicated molecular structure of DHS (Schnitzer and Kahn, 1972) which prevents efficient accumulation of DHS at the NAPL-water interface. This property will be advantageous in situations where NAPL mobilization is detrimental.

An additional consideration in choice of solubilizing agent is the cost of the agent as affected by loss of agent to the stationary phase by sorption. DHS are hypothesized to undergo lesser sorption to the stationary phase relative to surfactants. Significant data exists regarding sorption of surfactants (Jafvert and Heath, 1991; Edwards et al.,1992; Liu et al., 1992) and sorption of DHS (Jardine et al., 1992; Dunnivant et al., 1992; Gu et al.,1994; McCarthy et al.,1996; Kibbey and Hayes, 1997) to porous media. However, despite the fact that both processes have been characterized by non-linear isotherms, there has been no direct comparison of the sorption behavior of the two agents. It is hypothesized that the relatively complicated structure of DHS, involving a hydrogen-bonded mixture of hydrophobic and hydrophilic substituents (Schnitzer and Kahn, 1972), results in lesser tendency to sorb to surfaces of the subsurface stationary phase relative to surfactants. The relatively simple structure of surfactant molecules, with hydrophobic and hydrophilic regions dominating opposite ends of the molecules, promotes relatively efficient packing at interfaces, and potential formation of bi-layers on stationary phase surfaces. Lesser sorption results in more efficient use of the mass of added agent.

This study examined the sorption of a non-ionic surfactant mixture versus a DHS to a natural sediment material. Surface tension measurements of the surfactant and DHS solutions were also performed as a proxy for DNAPL-water IFT measurement. The non-ionic surfactants utilized in this study were octylphenol ethoxylates commercially available as Igepal CA-720, with an average of 12 ethoxylate (EO) units per molecule, and Igepal CA-887, a 70% active (30% water) mixture with and average of 30 EO units per molecule. The DHS utilized in this study was Aldrich humic acid. Although this humic acid is not representative of natural aquatic and soil-derived humic substances (Malcolm and MacCarthy, 1986), it was chosen as the representative DHS due to its relative availability, an important consideration for a potential remediation agent. High performance liquid chromatography (HPLC) methods were used to assess sorption of surfactant to sediment in batch experiments. High pressure size exclusion chromatography (HPSEC), total organic carbon (TOC) analysis, and UV spectrophotometry methods were used to assess sorption of DHS to sediment in batch experiments equivalent to the surfactant experiments.

Methods

Materials

The octylphenol ethoxylate non-ionic surfactants Igepal CA-620, 720, and 887 were donated by Rhodia Incorporated. Pure octylphenol and octylphenol ethoxylate standards with 1 and 3 EO units per

molecule were donated by Dr. Jennifer Field, Oregon State University. Humic Acid was purchased from Aldrich Chemical Company. HPSEC molecular weight standards (1,800, 4,600, 8,000, 18,000 molecular weight) were purchased from Polysciences Incorporated. HPLC grade solvents (acetonitrile, tetrahydrofuran) were purchased from Fisher Scientific. All other chemicals utilized for aqueous solutions (artificial groundwater, NaN_3 solution, HPSEC mobile phase) were purchased from Aldrich Chemical Company or Sigma Corporation and were reagent grade or better. All chemicals (excluding Aldrich Humic Acid) were used as received from the manufacturer without further purification. Milli-Q ultra pure water adjusted to approximately pH=7 was used in all experiments. Natural aquifer sediments were obtained from Columbus Air Force Base, MS. The >2mm fraction of the sediment was used for these studies. This low f_{oc} (0.065%) sediment has been characterized extensively (MacIntyre et al., 1998).

Batch Experiments

A 1:1 surfactant solution mixture of Igepal CA-720 and CA-887 (with EO unit per molecule averages of 12 and 30, respectively) was used with total surfactant concentrations ranging from 0.2 to 5.0 g/L. Aldrich humic acid solution was used at concentrations ranging from 0.094 to 1.89 g/L. The Aldrich humic acid was treated using a method modified from the standard treatment method of the International Humic Substances Society. Briefly, this involved treatment of the humic material in concentrated HCl, NaOH, and HCl/HF solutions with centrifugation (at 2,500 rpm for 1 hour) to remove ash or to concentrate the humic acid. The humic acid was freeze-dried following these procedures.

Batch experiments were conducted using Fisherbrand borosilicate glass liquid scintillation vials (20 mL) with foil-lined screwtop caps. Each vial contained 10 grams of sediment and 20 mL of surfactant or humic acid solution. Samples were then allowed to mix on a laboratory rotator for 24 to 72 hours. Replicates were performed at all concentrations for both the sediment-equilibrated and blank (no sediment) samples. Two experiments examining surfactant sorption were run using a high ionic strength solution of 0.1 M (NaN_3) in one case, and artificial groundwater with a low ionic strength of 2.8 mM (McCaulou et al., 1995) in the other case. Two experiments examining humic acid sorption were run, one experiment using a phosphate buffered (1 mM K_2HPO_4, 1 mM KH_2PO_4) 0.1 M NaCl solution, and the other experiment using artificial groundwater. Surface tension measurements were made of both the blank and sediment-equilibrated solutions using the drop-weight method (Adamson, 1982).

Surfactant Analysis

Surfactant samples were analyzed using a Hewlett-Packard Series 1050 HPLC with a 21 vial autosampler, UV detector, and Chemstation software. The analytical method utilized was modified from a method developed for nonylphenol ethoxylates (Kibbey and Hayes, 1996). This reverse-phase method involves the use of a silica column and a C18 pre-column which allows for underivatized aqueous injections. A flow gradient from 0.3 to 2.0 mL/min over a period of 37 minutes (held at 2.0 mL/min for 5 minutes) and a solvent gradient from 100% acetonitrile to 30% water over a period of 43 minutes was

utilized. A 0.5 min 100% tetrahydrofuran flush was used to maintain the columns. A Supelco Supelcosil LC-Si (5μm, 2.5cm x 46cm ID) and an Alltech Econosphere C18 (4mm) guard column were used in these experiments.

After equilibration, batch samples were centrifuged for 30 minutes at 1,500 rpm to speed settling of fines. Supernatant (1.8 mL) was then collected from each batch sample and mixed with 0.2 mL of MeOH to prevent sorption of the surfactant to the sample vials (2 mL). Sample injections (50 μL) were then analyzed with UV detection at 220 nm. An example chromatogram showing separate peaks for each surfactant oligomer (differing EO chain lengths) is shown in *Figure 1*.

Chromatogram peak areas were related to surfactant concentration using a regression developed from analysis of pure octylphenol and octylphenol ethoxylates with 1 and 3 EO units per molecule, as well as Igepal CA-620, CA-720, and CA-887 (with average EO chain lengths of 7, 12, and 30, respectively). The surfactant solutions were mixed in borosilicate glass vials with 20% MeOH to prevent sorption to the vials. Applying Beer's Law, the observed peak area on the chromatogram was related to the known concentration using a response factor (RF) as shown in the following equation:

$$A = RF \cdot C$$

where A is the chromatogram peak area, and C is the molar concentration. A non-linear relationship between RF and EO group was observed due to flow and solvent gradients used in the method, and also possibly due to variation in the molar absorptivity of the chromophore with increased EO chain length. The decrease in RF with chain length was described using the following double exponential equation (*Figure 2*):

$$RF = RF_{OP}[(a)(e^{-k_1 n_{EO}}) + (1-a)(e^{-k_2 n_{EO}})]$$

where RF_{OP} is the response factor for pure octylphenol, and a is a coefficient which shifts control of the exponent to the two decay constant, k_1 and k_2, and n_{EO} represents the number of EO groups in the surfactant oligomer. The values of a, k_1, and k_2 were optimized by minimizing the sum of the squares of error between known surfactant concentrations and calculated surfactant concentrations using the RF value. Concentrations of each surfactant oligomer in the samples were calculated using the peak area observed divided by the appropriate RF value after optimization of a, k_1, and k_2.

DHS Analysis

DHS analysis was accomplished using high pressure size exclusion chromatography (HPSEC) on a Hewlett-Packard Series 1050 HPLC. The HPSEC technique utilized was developed for aquatic humic acid analysis (Chin et al., 1994). A Waters Protein-Pak 125 size exclusion chromatography column (10μm, 7.8mm x 300mm, molecular weight range: 2,000-80,000) was used in this analysis with a phosphate buffered (1 mM K_2HPO_4, 1 mM KH_2PO_4) 0.1 M NaCl mobile phase (Chin and Geschwend, 1991).

Polystyrene sulfonate (PSS) standards (1,800, 4,600, 8,000, and 18,000 molecular weights) were used to relate retention time to molecular weight.

Batch procedures involving humic acid samples were equivalent to those previously described for surfactant samples, excluding the addition of MeOH. Sample injections (50μL) were analyzed isocratically (at 1.0 mL/min) for 20 minutes using UV detection at 280 nm. Further quantification of sample DHS concentrations was accomplished by total organic carbon (TOC) analysis (Shimadzu TOC-5000A with ASI-5000A autosampler) and UV spectrophotometric analysis (Cary 3E UV-VIS Spectrophotometer). Both TOC and UV analysis required dilution of the samples to within the linear range of the instruments.

Results and Discussion

Surfactant Sorption

Aqueous concentrations of surfactant oligomers in the presence and absence (blank) of sediment are shown in *Figure 3* for the experiment which used 0.1 M NaN_3. Equivalent results were obtained for the artificial groundwater solution samples (not shown) indicating that addition of NaN_3 was not necessary to inhibit biodegredation of the surfactant during the duration of the experiments. An equilibration time of 48 hours was found to be sufficient for complete equilibration of surfactant (and humic acid) with the sediment (data not shown). Significant sorption of surfactant monomers with both small and large EO chain lengths was observed (*Figures 3*). Isotherms developed for specific surfactant oligomers (*Figures 4* and *5*) indicate that surfactant oligomers with small EO chains were preferentially adsorbed to the sediment used in this study. This is indicated by steeper isotherm slopes for oligomers with smaller EO chain lengths. The data also indicates that oligomers with smaller EO chain lengths out-competed other oligomers for the sediment surface, since these oligomers tended to show the highest maximum sorbed concentration (although this is not clear in the experiment with NaN_3). Oligomers with large EO chains did not achieve their maximum sorbed concentration presumably because relatively low aqueous concentrations of each oligomer were achieved due to spreading of the surfactant mass among a greater number of oligomers in the high EO range (larger EO distribution).

Sorption isotherms for total surfactant mass (*Figures 6* and *7*) also indicate that sorption equilibrium was reached within approximately 48 hours. The overall sorption isotherm indicates that the maximum sorbed surfactant concentration equals approximately 0.5 g/L of surfactant in solution and 0.5 g/kg of surfactant sorbed to sediment. Ionic strength did not appear to have strong affect on the surfactant sorption characteristics.

Surface tension measurements of samples with and without (blank) sediment show that the critical micelle concentration (CMC) of the mixture was approximately 0.5 g/L, whereas the apparent CMC of the solution in the presence of sediment was 1.0 g/L (*Figure 8*) in accordance with a maximum sorbed surfactant concentration achieved at an aqueous concentration of 0.5 g/L. The solution surface tension was reduced to approximately 37 dynes/cm in both the high and low ionic strength solutions.

DHS Sorption

Aqueous concentrations measured as UV-absorbance of Aldrich humic acid are shown in *Figure 9*. These results are qualitative since the UV absorbance has not yet been related to the mass concentration of each molecular weight fraction. Such a relation will require establishment of equivalent molar absorbtivites for the spectrum of molecular weights observed. The qualitative results show significant sorption of Aldrich humic acid to the sediment under the conditions of the experiment. The difference between humic acid concentrations in the absence (blank) versus the presence of sediment shows that larger molecular weight humic acids are preferentially sorbed. An unresolved feature of the HPSEC chromatograms is observed as a significant "hump" possibly indicating aggregation of humic acid to form an equivalent 50,000 to 60,000 molecular weight molecule. Further testing with appropriate molecular weight standards will be required to resolve this feature. This feature was amplified in the presence of sediment with artificial groundwater as the aqueous phase (data not shown). Qualitative isotherms describing humic acid sorption to the sediment were developed using TOC data (*Figures 6* and *7*). From these isotherms it is observed that a maximum sorbed humic acid concentration of about 0.05 g/kg was obtained under conditions of the experiment regardless of whether the phosphate/NaCl or a broader mixture of groundwater ions served as the solution electrolytes. Surface tension measurements of the humic acid solutions indicate that surface tension was only slightly decreased (~1-2 dynes/cm) by addition of up to almost 2 g/L of humic acid (*Figure 8*).

A comparison of the sorption isotherms for the surfactant solutions and the humic acid solutions indicates that both solubilizing agents sorb significantly to the low f_{oc} sediment used in these experiments, with non-ionic surfactant showing a maximum sorbed concentration that is an order of magnitude higher than that of Aldrich humic acid. The surfactant oligomers with smaller EO chains (lower molecular weight) exhibit preferential sorption to the sediment whereas in the case of humic acid, the larger molecular weight fraction sorbed preferentially to the sediment. In both cases, the component which underwent preferential sorption was the less polar component (the component with the highest percent aromaticity). Surface tension measurements indicated that similar concentrations of surfactants versus humic acids resulted in largely different effects on surface tension of the aqueous phase. These results suggest that humic acids may result in much smaller reductions in interfacial tension between non-aqueous phase liquids and water relative to non-ionic surfactants. These observed differences between non-ionic surfactants and humic acids in their extent of sorption and effects on surface tension of water have significant implications in determination of an appropriate potential remediation agent. Surfactant oligomers with short EO chains are known to have different solubilization characteristics relative to oligomers with long EO chains. Hence, preferential sorption of particular oligomers will affect distribution of contaminants between sediment and aqueous phases. The same is true of humic substance sorption, although these preliminary results indicate humic acids sorb to a much lesser extent than non-ionic surfactants.

Conclusion

The analytical methods developed during the course of this work will allow determination of the affects of preferential adsorption of particular surfactant oligomers or particular humic acid molecular weight fractions on the distribution of a variety of contaminants between sediment and aqueous phases. This determination will be made for a variety of surfactants and humic substance.

The preliminary results for surface tension reductions in the presence of surfactants and humic acids indicate that in cases where mobilization of dense non-aqueous phase liquids is not desired, humic acids may be advantageous. These comparisons will be made between additional surfactants and humic acids.

References

Abriola L.M., Dekker T.J., and Pennell K.D., 1993, Surfactant-Enhanced Solubilization of Residual Dodecane in Soil Columns. 2. Mathematical Modeling, ***Environ. Sci. Technol.***, Vol.27(10), 2341-2351.

Adamson, Arthur W. *Physical Chemistry of Surfaces.* John Wiley and Sons, 20-23p, 1982.

Brown C.L., Pope G.A., Abriola L.M., and Sepehrnoori K., 1994, Simulation of Surfactant-Enhanced Aquifer Remediation, ***Wat. Resour. Res.***, Vol. 30(11), 2959-2977.

Chin, Yu-Ping, Philip M. Geschwend, 1991, The abundance, distribution, and configuration of porewater organic colloids in recent sediments. ***Geochim. Cosmochim. Acta***. Vol. 55, 1309-1317.

Chin, Yu-Ping, George Aiken, Edward O'Loughlin, 1994, Molecular Weight, Polydispersivity, and Spectroscopic Properties in Aquatic Humic Substances. ***Environ. Sci. Technol.,*** Vol. 28, 1853-1858.

Dunnivant F.M., Jardine P.M., Taylor D.L., and McCarthy J.F., 1992, Transport of Naturally Occuring Dissolved Organic Carbon in Laboratory Columns: ***Soil Sci. Soc. of Am. J.***, Vol. 56, 437-444.

Edwards, D.A., Luthy R. G., and Liu Z. L., 1991, Solubilization fo polycyclic aromatic hydrocarbons in micellar nonionic surfactant solutions, **Environ. Sci. Technol.**, Vol 25, 127-133.

Edwards D. A., Liu Z. L., and Luthy R. G., 1992, Interactions between nonionic surfactant monomers, hydrophobic organic compounds, and soil, ***Water Sci. Tech.***, Vol. 26(1-2) 147-158.

Gu B., Schmitt J., Chen Z., Liang L., and McCarthy J.F., 1994, Adsorption and desorption of natural organic matter on iron oxide: mechansims and models. ***Environ. Sci. Technol.,*** Vol. 28, 38-46.

Jafvert C. T., and Heath J. K., 1991, Sediment and saturated-soil-associated reactions involving an anionic surfactant (dodecylsulfate). 1. Precipitation and micelle formation, ***Environ. Sci. Technol.***, Vol. 25(6), 1031-1038.

Jardine P.M., Dunnivant F.M., Selim H.M., and McCarthy J.F., 1992, Comparison of Models Describing the Transport of Dissolved Organic Carbon in laboratory columns containing aquifer material. ***Soil Sci. Soc. Am. J.*** Vol. 26, 393-401.

Johnson, W.P., and John, W.W., 1997, A Comparison of Humic Substances and Surfactants as Remediation Agents: PCE Solubilization and Mobilization by Commercial Humic Acid. ***Journ. Cont. Hydrology.***

Kibbey T. C. G., T.P. Yavaraski, and K.F. Hayes, 1996. HPLC Analysis of Polydisperse Ethoxylated Nonionic Surfactants in Aqueous Samples. ***J. Chrom.*** Vol. 752, 155.

Kibbey T. C. G., and K.F., Hayes 1997, A Multicomponent Analysis of the Sorption of Polydisperse Ethoxylated Nonionic Surfactants to Aquifer Materials: Equilibrium Sorption Behavior, ***Environ. Sci. Technol.,*** Vol. 31(4), 1171-1177.

Liu Z. L., Edwards D. A., and Luthy R. G., 1992, Sorption of non-ionic surfactants onto soil, ***Wat. Res.*** Vol. 26(10), 1337-1345.

Malcolm R.L., and McCarthy, P. 1986. Limitations on the Use of Commercial Humic Acids in Water and Soil Research. ***Environ. Sci. Technol.,*** Vol. 20(9), 904-911.

MacIntyre, William G, Christopher P. Antworth, Thomas B. Stauffer, Riki G. Young, 1998, Heterogeneity of sorption and transport properties in a sand-gravel aquifer at Columbus, Mississippi. ***J. Cont. Hydrol.*** Vol. 31 (3-4), 47-64.

McCarthy J.F., Gu B., Liang L., Mas-Pla J., Williams T.M., and Yeh T.C.J., 1996, Field Tracer Testes on the Mobility of Natural Organic Matter in a Sandy Aquifer: ***Water Resour. Res.***, Vol. 32 (5), p 1223-1238.

McCaulou, Douglas R., Roger C. Bales, Robert G. Arnold, 1995, Effect of temperature-controlled motility on transport of bacteria and microspheres through saturated sediment. ***Water Res. Res.***, Vol. 32, No. 2, 271-280.

Pankow,J.F. and Cherry,J.A. 1996. *Dense Chlorinated Solvents and Other DNAPLs in Groundwater.* Waterloo Press, Portland, OR, 522p.

Pennell K.D., Abriola L.M., and Weber W.J., 1993, Surfactant-Enhanced Solubilization of Residual Dodecane in Soil Columns. 1. Experimental Investigation, ***Environ. Sci. Technol.***, Vol. 27(12), 2332-2340.

Pennell K.D., Pope G.A., and Abriola L.M., 1996, Influence of Viscous and Buoyancy Forces on Mobilization of Residual Tetrachloroethylene During Surfactant Flushing, ***Environ. Sci. Technol.***, Vol. 30(4), 1328-1335.

Peters, C.A., and Luthy, R.G., 1993. Coal Tar Dissolution in Water-Miscible Solvents: Experimental Evaluation, ***Environ. Sci. Technol.***, Vol. 27(13), 2831-2843.

Schnitzer M., and Khan S.U., 1972, Humic Substances in the Environment, Marcel Dekker, Inc., New York, 201p.

West C.C., and Harwell J.H., 1992, Surfactants and Subsurface Remediation, ***Environ. Sci. Technol.***, Vol. 26(12), 2324-2330.

Figure 1: An example HPLC chromatogram of the Igepal CA-720, CA-887 1:1 mixture. Each peak represents a single oligomer with distinct ethoxylate (EO) chain length molecule (3 to 45 EO groups) with the left group of peaks representing the CA-720 (average EO chain length of 12) and the right group the CA-887 (average EO chain length of 30). The large peak at the beginning of the chromatogram (~9 minutes) represents NO_3^- in the artificial groundwater, whereas the large peak at the end of the chromatogram (~50 minutes) represents the THF column flush.

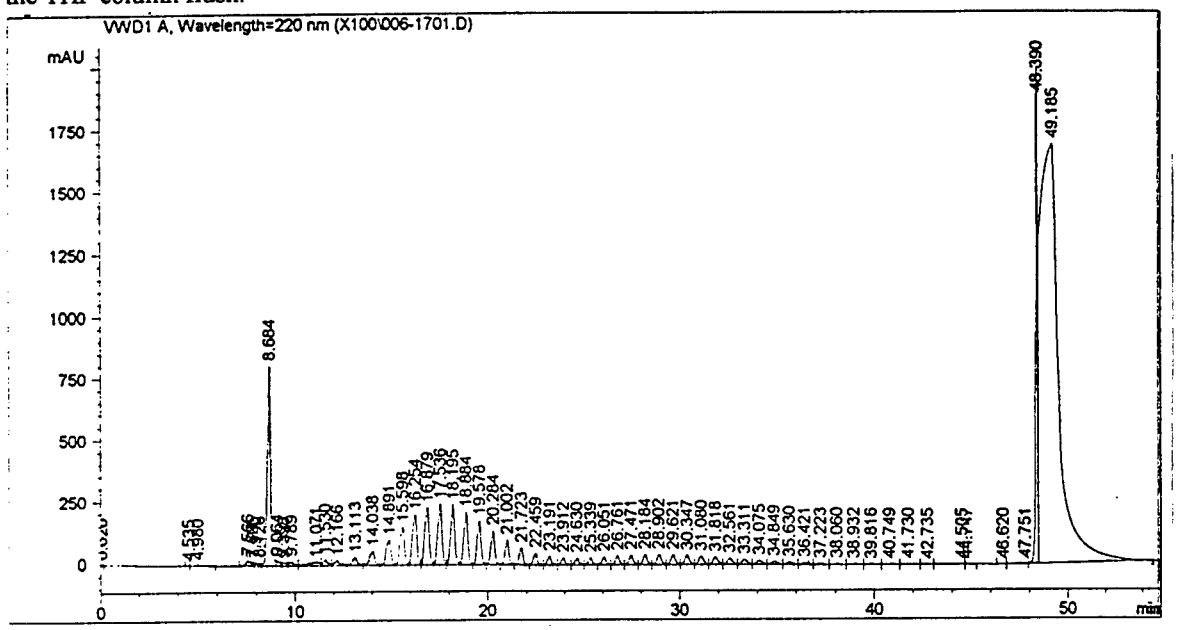


Figure 2: The Response Factor (RF) line, produced from a double exponential equation, developed to relate the chromatogram peak areas to EO oligomer concentrations. The values a, k_1, and k_2 were optimized by minimizing the sum of the squares of error between known surfactant concentrations and concentrations calculated for chromatogram peak areas using the RF values.

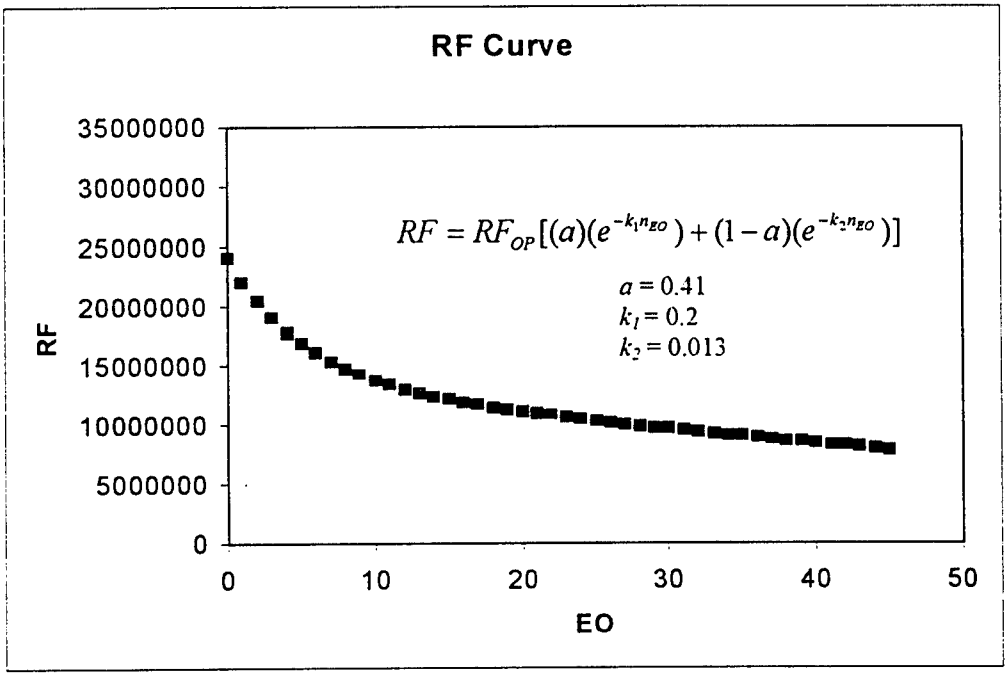


Figure 3: Equilibrium aqueous concentrations of surfactant oligomers in the presence and absence (blank) of sediment in high ionic strength sodium azide solutions (equivalent results were obtained in low ionic strength, artificial groundwater solutions).

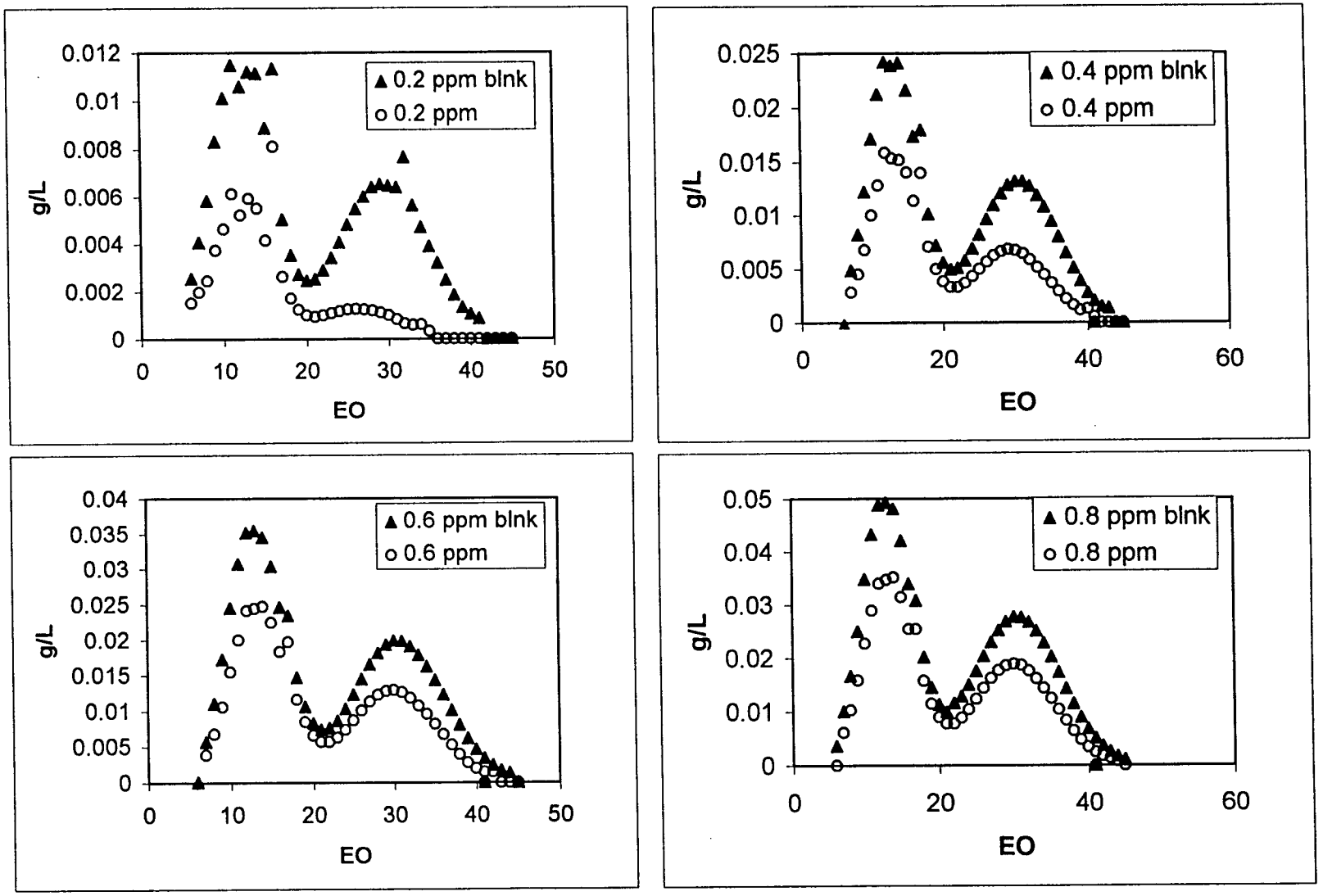
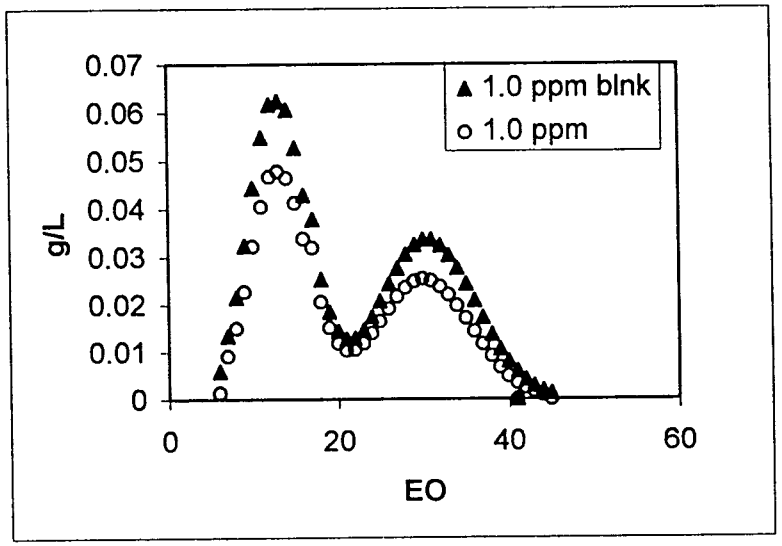
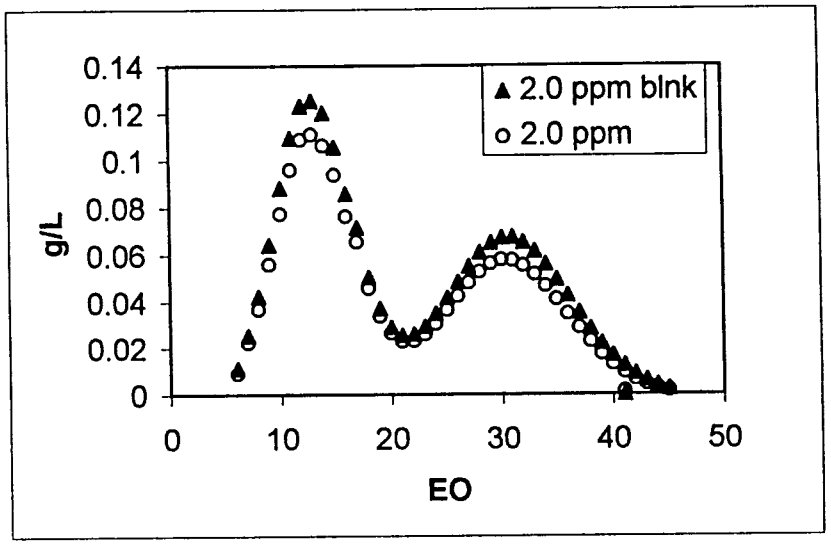
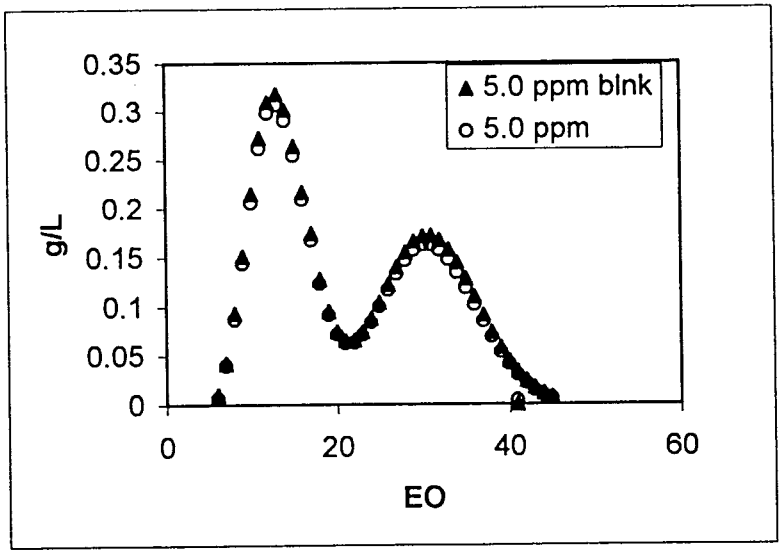


Figure 3: (continued)

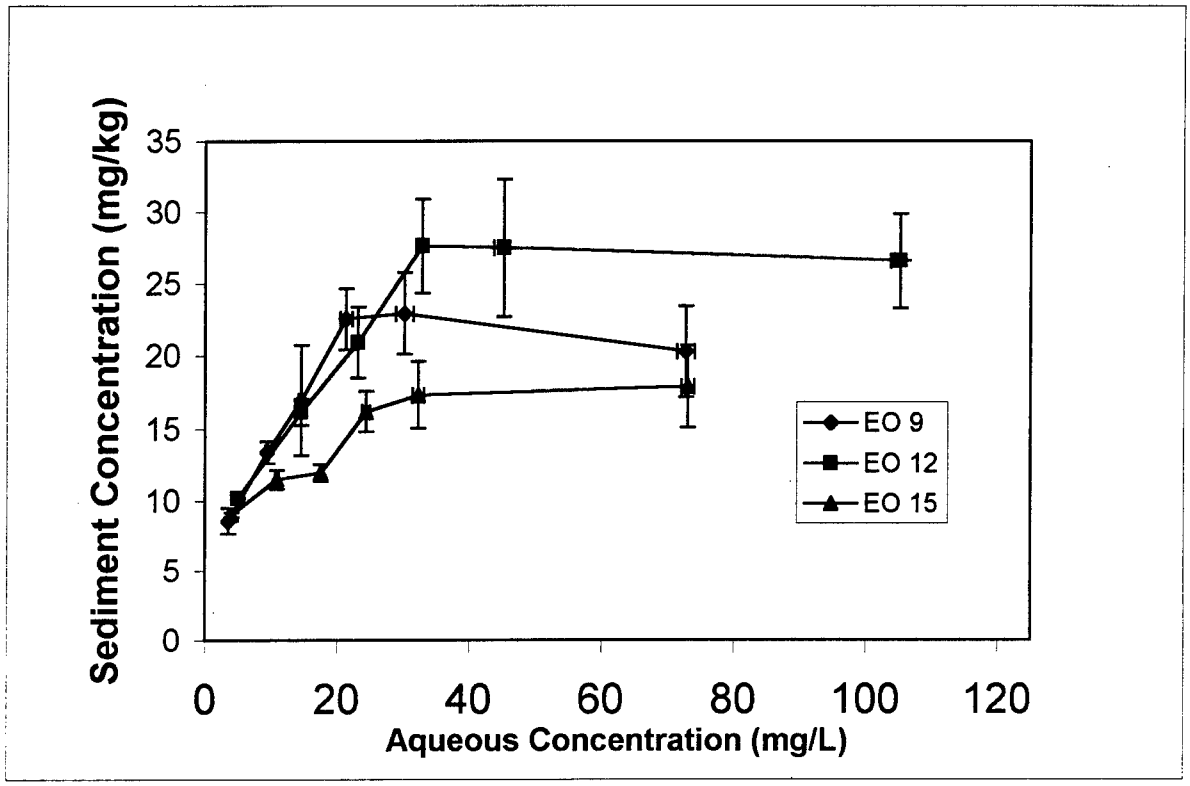






13-13

Figure 4: Equilibrium sorption isotherms developed for specific surfactant oligomers in high ionic strength sodium azide solutions.



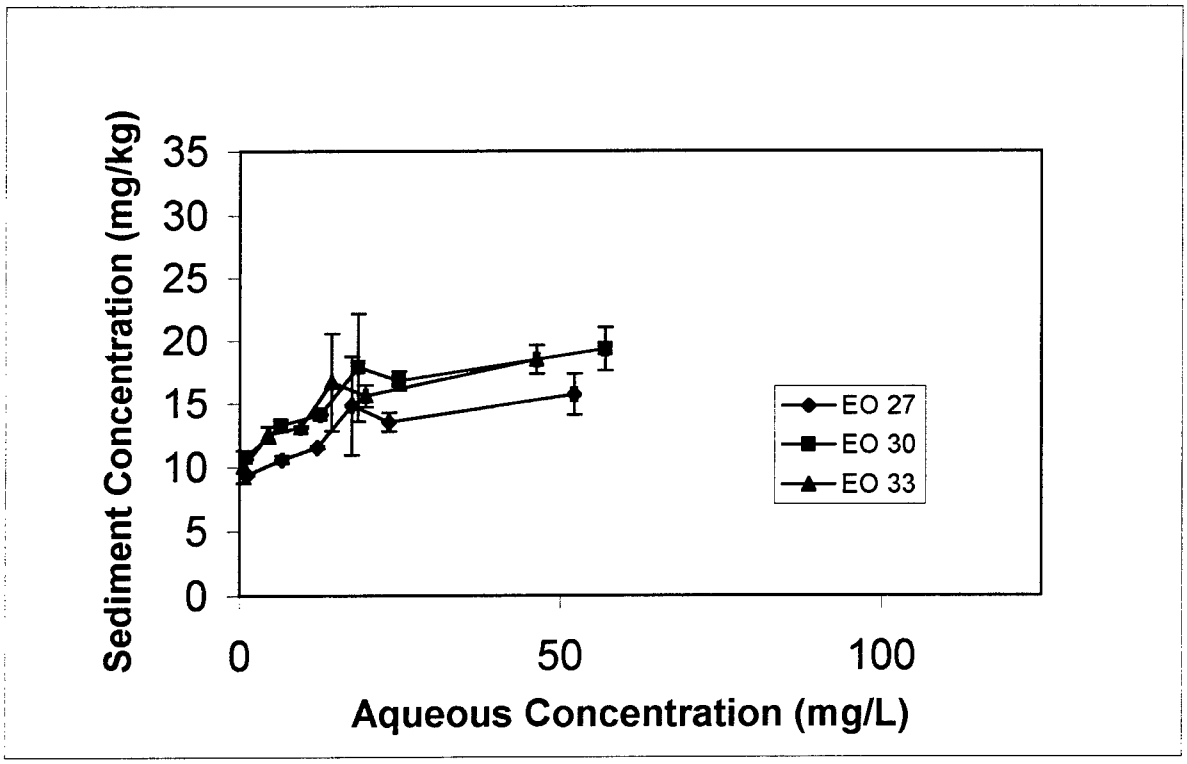
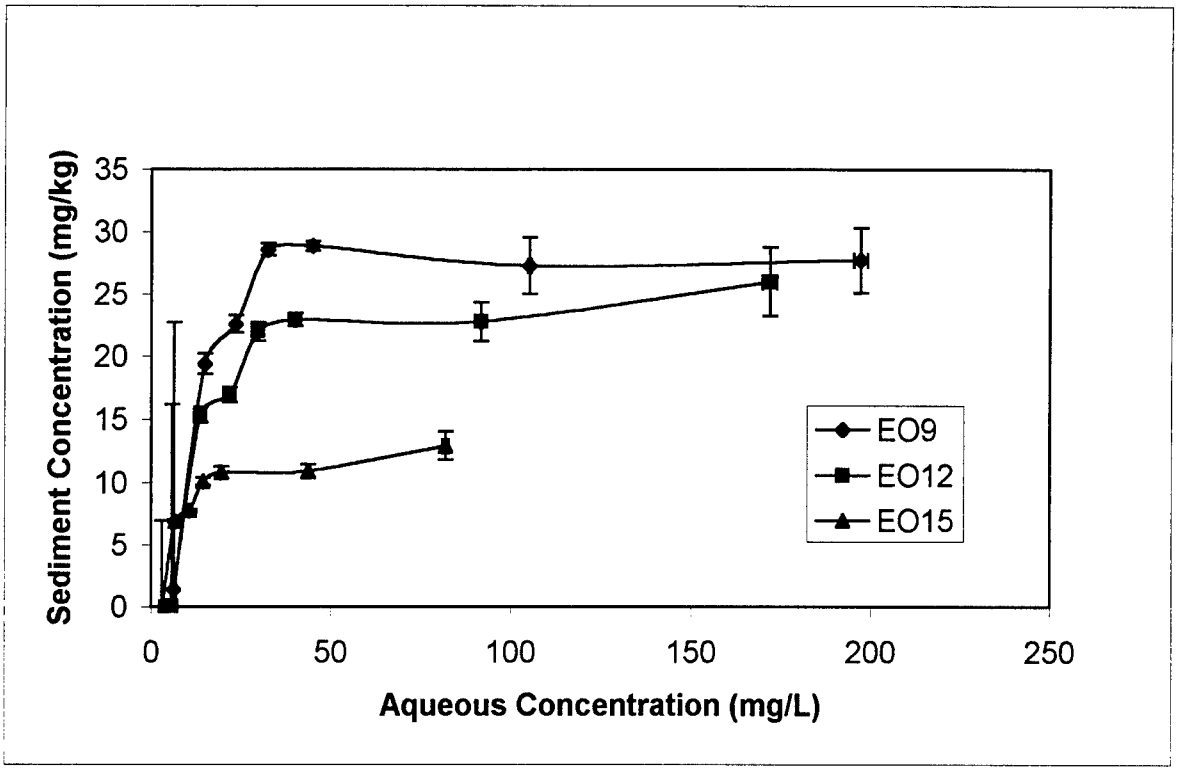


Figure 5: Equilibrium sorption isotherms for specific surfactant oligomers in low ionic strength artificial groundwater solutions.



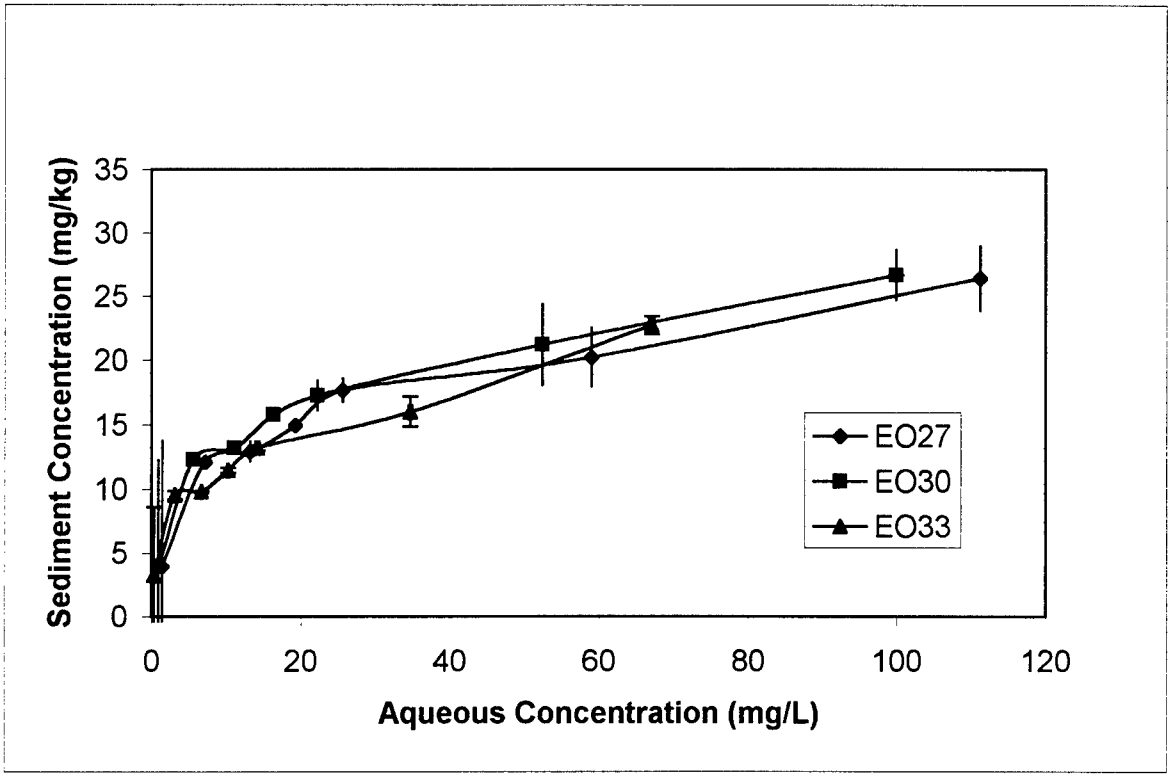


Figure 6: Sorption isotherms for total surfactant masses and DHS in high ionic strength solutions.

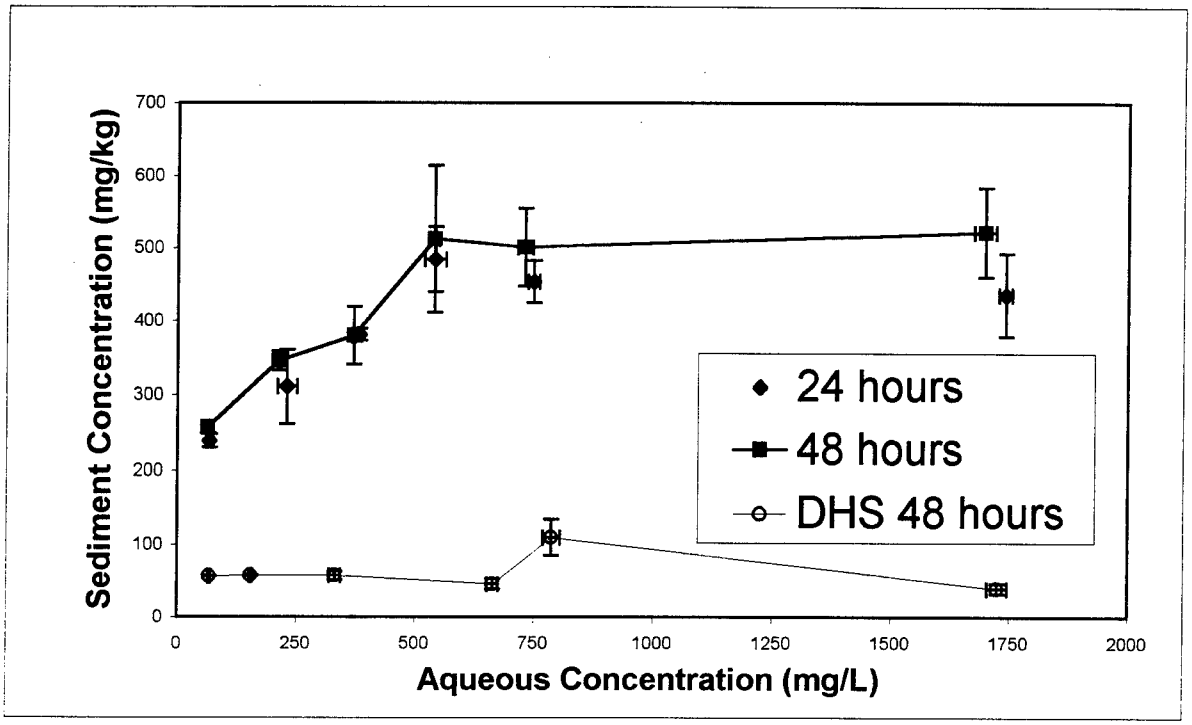


Figure 7: Sorption isotherms for total surfactant masses and DHS in low ionic strength, artificial groundwater solutions.

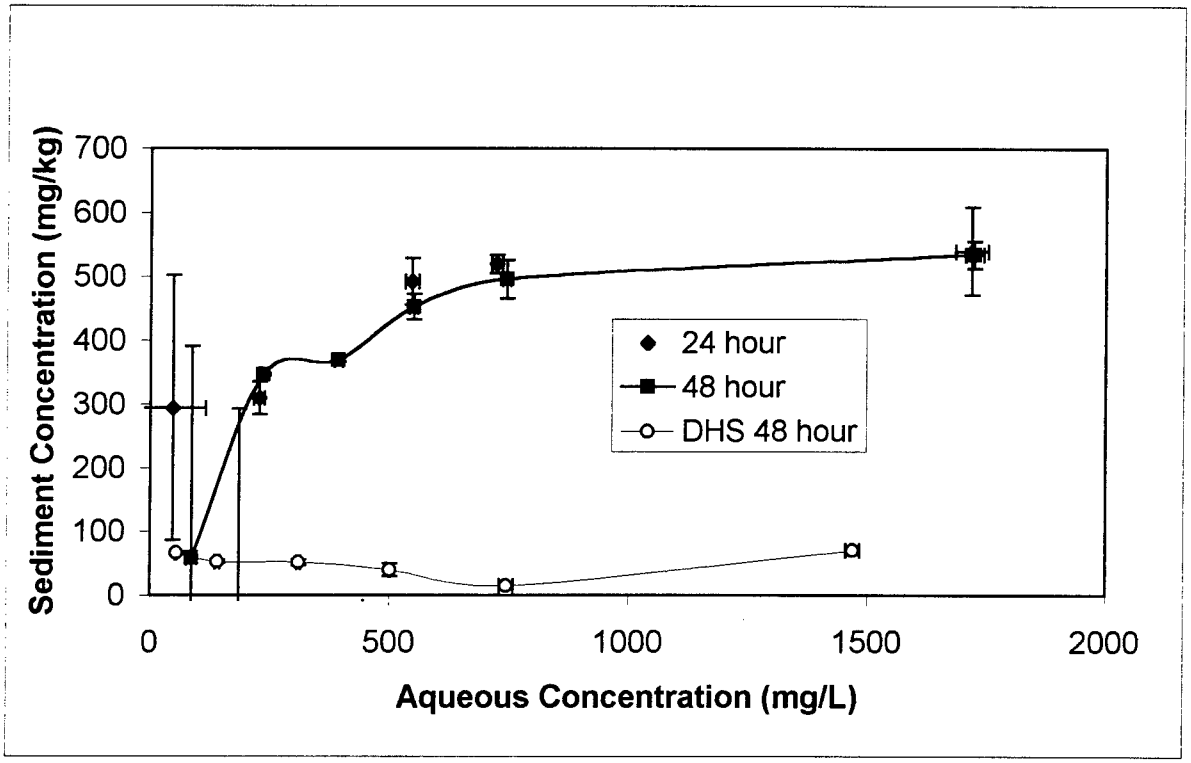


Figure 8: Solution surface tensions measured for surfactant mixtures with and without sediment, and DHS without sediment (DHS sorption to sediment results in a neglibible reduction of DHS solution surface tension). Results for all solutions were equivalent in the high and low ionic strength solutions.

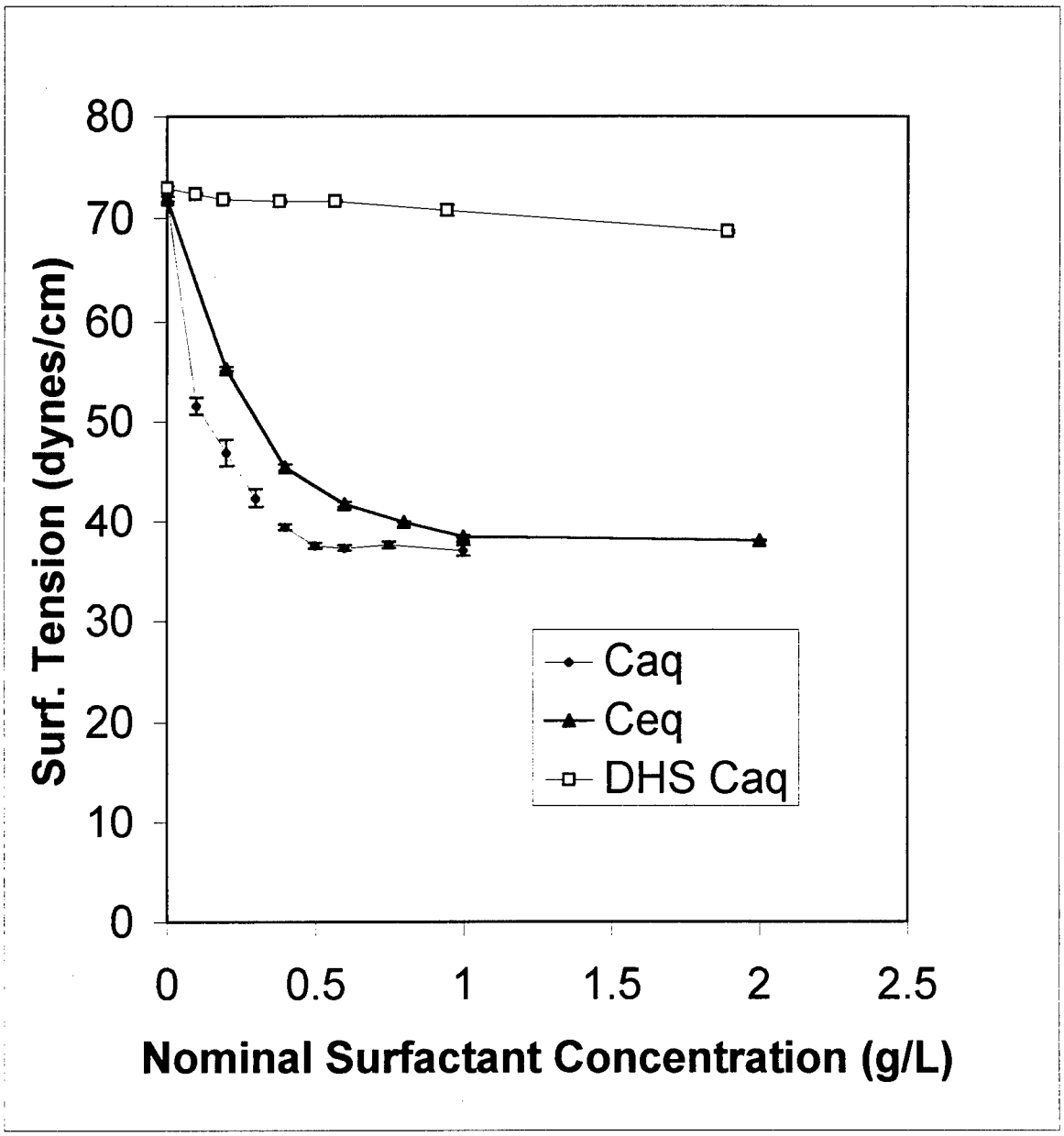


Figure 9: Equilibrium aqueous molecular weight DHS distributions in the presence and absence (blank) of sediment in high ionic strength, phosphate buffered solutions. Note the large molecular weight "hump" in both the sediment-equilibrated and blank samples.

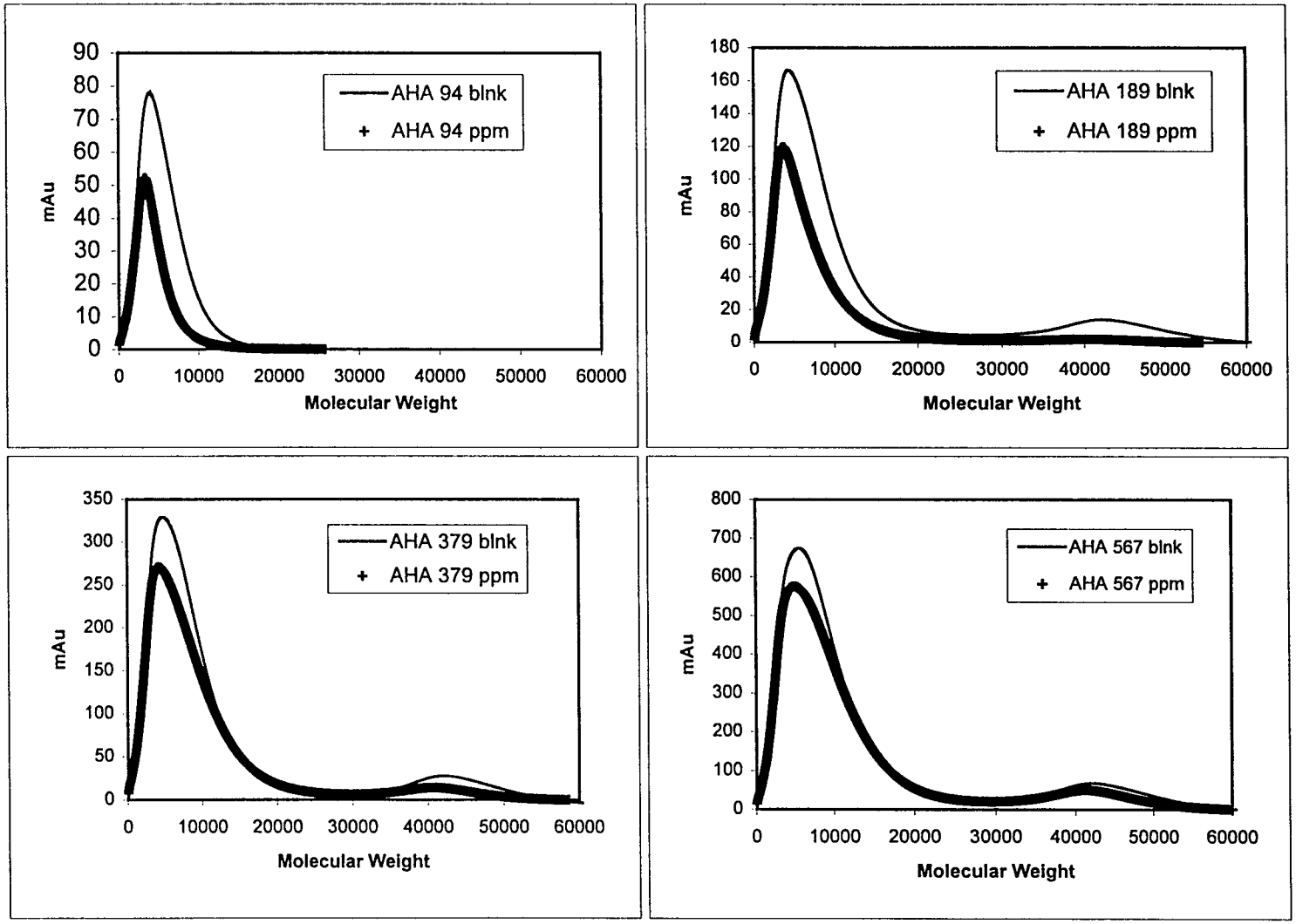
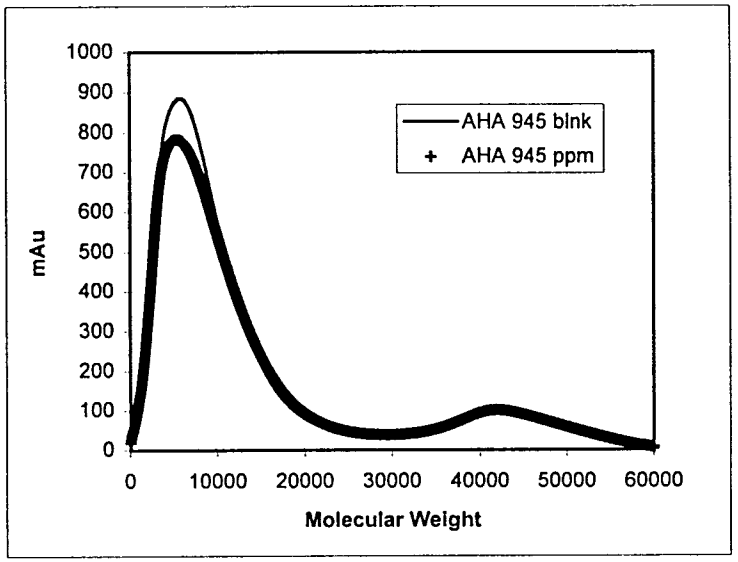
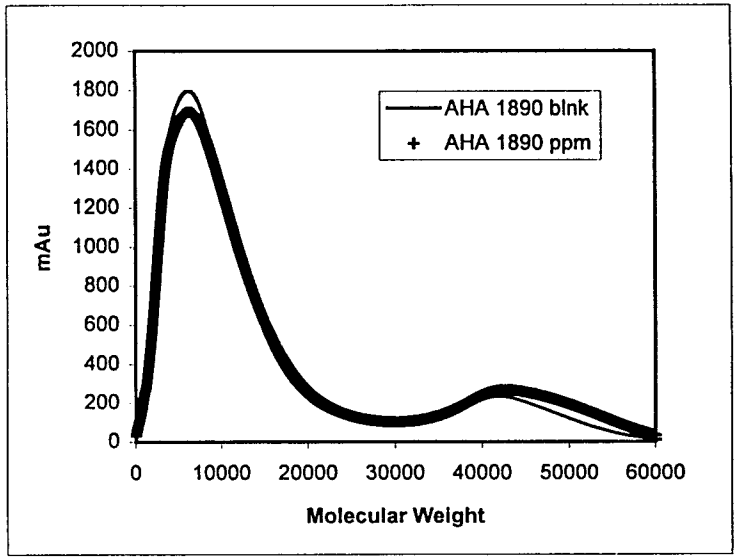


Figure 9: (continued)





13-19

Using Neural Networks to Control a Tailless Fighter Aircraft

Jeffrey D. Johnson
Assistant Professor
Department of Bioengineering

5031 Nitschke Hall
The University of Toledo
Toledo, Ohio 43606

Final Report for:
Summer Faculty Research Program
Wright Research Site

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Research Site

October 1998

Using Neural Networks to Control a Tailless Fighter Aircraft

Jeffrey D. Johnson
Assistant Professor
Department of Bioengineering
The University of Toledo

ABSTRACT

This paper presents the research conducted over a 12 week period during the Summer of 1998 while I was involved in the AFOSR Summer Research Faculty Program. The work was performed at Wright-Patterson Air Force Base in collaboration with Dr. Corey Schumacher of the Air Vehicles Directorate of the Air Force Research Laboratory (office symbol AFRL/VAAD). Dr. Siva Banda was the overall technical leader of the AFRL/VAAD effort and Dr. James Buffington also had a significant impact on the effort. I present an application of dynamic inversion control with adaptive neural networks to a tailless fighter aircraft. In the past, this technique has been studied with proportional-derivative desired dynamics, but not with proportional-integral desired dynamics. This paper presents an extension of the method to use with a dynamic inversion controller using proportional-integral desired dynamics. The performance of the PI adaptive dynamic inversion controller is compared to an adaptive dynamic inversion controller using PD desired dynamics. The controller was tested in a 6DOF simulation of the ICE tailless fighter aircraft. Simulation results are given for simultaneous longitudinal and lateral maneuvers performed under nominal conditions and also with the simulated loss of a major control surface. The PI dynamic inversion controller with neural network adaptation outperforms both the baseline PI controller without adaptation and also a PD dynamic inversion controller with neural network adaptation.

Using Neural Networks to Control a Tailless Fighter Aircraft

Jeffrey D. Johnson

Introduction

Next-generation air vehicles will have from 10-20 control surfaces. As the number of control surfaces increases, the number of control policies that accomplish any one maneuver increases exponentially. It becomes unrealistic to expect that a team of control engineers can create, a priori to the vehicle's maiden flight, the full range of control strategies necessary to control the vehicle in every circumstance. This point is reinforced by the unfortunate crash of the Darkstar UAV. The crash is attributed to a modification in the control system software that was expected to compensate for oscillations the UAV experienced during its first test flight. It was estimated that the accident cost the program $40 million and one year of development time [1]. In addition to these complications, designing control systems that can compensate for failures in control due to battle damage or component failure is becoming a priority. But, again, it is impossible to predict and account for all types of control failures that may arise. Adaptation in the control system removes the demand for completeness in the initial control design because the controller, through adaptation, will move toward a more optimal solution.

Reconfigurable Flight Control Systems

Reconfigurable control systems are becoming an important area of research [2]. With the cost to manufacture and maintain air vehicles rising exponentially, it has become important to find ways to 1) reduce the costs of maintaining the vehicle, 2) extending the lifetime of the vehicle, and 3) improving survivability. Self-reconfiguring control systems have the potential to reduce development costs and extend the lifetime of the air vehicle. The traditional approach to designing a reconfigurable flight control system consists of

1. Predicting the potential changes in the vehicle's dynamics due to failure,
2. Designing a fixed controller for each failure mode,
3. Using system identification methods to detect and identify failure modes, and
4. Scheduling the controller designed to compensate for the identified failure mode.

There have been many successes in parts or this approach [3-6], especially Step 2 [7]. Unfortunately, with the complexity of control in modern air vehicles, this method usually becomes impractical at Step 1. That is, with the growing number of control surfaces, and their interdependence, predicting and accounting for just a fraction of the possible failure modes would extend development time and costs indefinitely.

In this paper we explore the use dynamic inversion control augmented with an adaptive network to realize a self-reconfiguring flight control system. Specifically, we look to extend the work by [8, 9] by applying their method to a PI controller. Our test platform is a 6DOF model of a tailless fighter.

Dynamic Inversion Controller

In dynamic inversion control, we seek to linearize a nonlinear system [10-13]. The aircraft's dynamics can be put in the form

$$\dot{\mathbf{x}} = \mathrm{f}(\mathbf{x}) + \mathrm{g}(\mathbf{x})\mathbf{u}, \tag{1}$$

where $\mathbf{x}$ is a vector the controlled states and $\mathbf{u}$ is a vector of controls and $\mathrm{f}(\mathbf{x})$ and/or $\mathrm{g}(\mathbf{x})$ is nonlinear. The number of controlled states and controls need to be the same (a square system). The control can be calculated according to

$$\mathbf{u}_c = \mathrm{g}^{-1}(\mathbf{x})(\mathbf{u}_d - \mathrm{f}(\mathbf{x})), \tag{2}$$

where $\mathbf{u}_d$ is the desired response of $\dot{\mathbf{x}}$. Substituting (2) into (1), for $\mathbf{u}$, results in

$$\dot{\mathbf{x}}_d = \mathbf{u}_d \tag{3}$$

and any nonlinearities in f(**x**) and g(**x**) are cancelled. Figure 1 displays a diagram of the dynamic inversion process. In the figure, the pilot generates the commanded control variable, CV_{cmd}, it is shaped by command filters to produce the desired dynamics, CV'_{cmd} ($\mathbf{u}_d$, in (2)), the control system produces the necessary control, $\mathbf{u}_c$, with the desired result. Dynamic inversion control, if successful, results in a linear transform between desired and actual dynamics.

The task is to form $\mathbf{u}_d$. Depending on the control variable (roll, pitch, or yaw) the desired dynamics of the system may be of first- or second-order and the control law may be implemented as PD, PI, or PID. Two different sets of desired dynamics are used in this study. Control Law I employs a proportional-derivative set of desired dynamics, equivalent to a first-order response, such that

$$\mathbf{u}_d = K_p(\mathbf{x}_c - \mathbf{x}) + \dot{\mathbf{x}}_c, \tag{4}$$

where $\mathbf{u}_d, \mathbf{x}, \mathbf{x}_c \in \Re^3$, and K_p is a 3×3 diagonal matrix of control gains. Control Law II uses proportional-integral desired dynamics, with

$$\mathbf{u}_d = \tfrac{1}{2}K_b\mathbf{x}_c - K_b\mathbf{x} + \tfrac{1}{4}K_b^2\int(\mathbf{x}_c - \mathbf{x})dt\,. \tag{5}$$

Adaptive Artificial Neural Networks

One method by which adaptation can be added to a control system is through an adaptive artificial neural network [14, 15]. It is called an artificial neural network because it derives its structure and function from biological neural networks. We use the adjective adaptive to indicate that there is no long-term learning in this system, just short-term adaptation.

In this study we use a linear-in-parameters neural network (LIPNN). The LIPNN is a single-layer neural network but it can represent nonlinear functions because the inputs to the network

are nonlinear. Figure 2 is a diagram of a linear in parameters neural network with nonlinear inputs.

As an adaptive system, LIPNN plays two important roles in

1. It compensates for inaccuracies in the initial control law design. As mentioned previously, the complexity of modern control suites makes it difficult to completely specify the control law a priori to the vehicle's first flight. An adaptive control system has the potential to compensate for inaccuracies, due either to incomplete or approximated experimental data or minor errors in the control law synthesis.
2. It adapts the control system in response to actuator failures. While there have been successes in certain stages of the traditional approach to reconfigurable control, the inability to predict the myriad of possible failure modes makes this approach infeasible for next-generation aircraft. An adaptive system removes the requirement of designing explicit control laws for each potential failure mode.

Weight-Update Law

An important part of any neural network is the rule governing its adaptation, sometimes called the learning rule. We prefer, for this implementation, to call this the weight-update law because we reserve the word "learning" to indicate long-term changes in the response of the system. If the error dynamics are given by (5), then by taking the Laplace Transform and rearranging we arrive at the transfer function

$$\frac{p}{p_c}(s) = \frac{K_b}{2}\left(\frac{s+\frac{1}{2}K_b}{s^2+K_b s+\frac{1}{4}K_b{}^2}\right) = \frac{k_b}{2}\left(\frac{s+\frac{1}{2}K_b}{\left(s+\frac{1}{2}K_b\right)^2}\right) \tag{6}$$

which, by a stable pole-zero cancellation, reduces to,

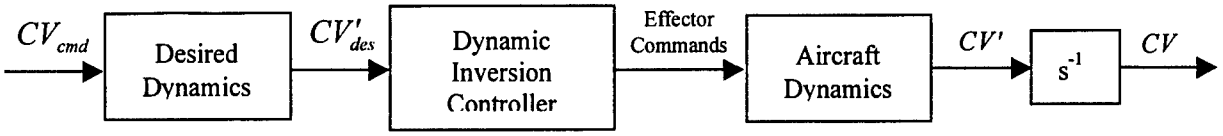


Figure 1. The Dynamic Inversion Control process.

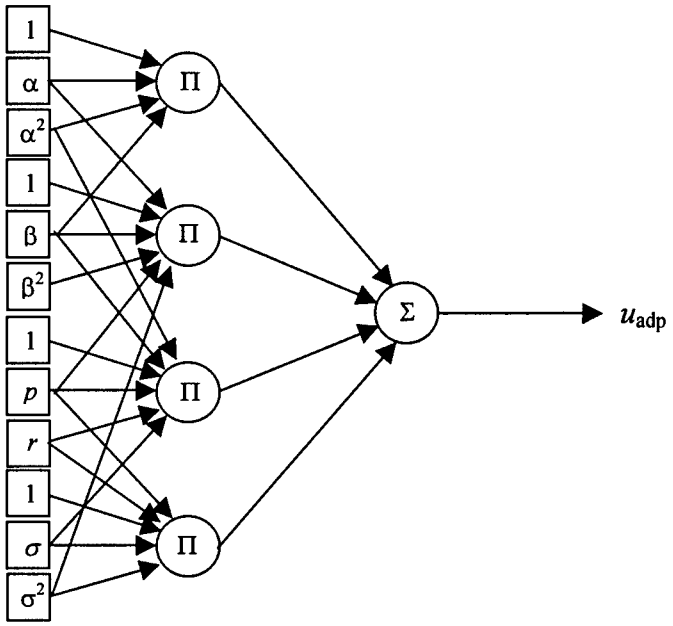


Figure 2. Representation of a Linear-in-Parameters Neural Network. For simplicity, all connections are between inputs and PI units are not shown.

$$\frac{p}{p_c}(s) = \frac{\frac{1}{2}K_b}{s + \frac{1}{2}K_b}. \tag{7}$$

Taking the inverse Laplace transform and rearrange the results, we arrive at a simplified differential equation for the desired dynamics:

$$\dot{p}_d = \tfrac{1}{2}K_b(p_c - p) \tag{8}$$

With the proper choice of gains, the PI dynamics will be that of a first-order system. We will take advantage of this fact, and use the derivation of the neural network weight-update law for first-order error dynamics found in [8, 9] to give us the weight update law for the PI error dynamics. The weight-update law for first order error-dynamics is corresponding to (4) is

$$\dot{w}_{pi} = \gamma_p\left((p_c - p)b_i + \eta|p_c - p|w_{pi}\right) \tag{9}$$

where w_{pi} is the i^{th} weight of the linear-in-parameters neural network, γ_p is the (roll channel) learning rate, b_i is the i^{th} basis function, and η is an *e-modification* factor to unsure that **w** is bounded [8, 9]. If we set $K_b = 2K_p$, then (8) is equivalent to (4) and we can use (9) as the update law for both controllers.

Nonlinear Adaptive Control

We designed a PD and a PI controller for each of the pitch, roll, and yaw channels. The same neural network structure was used for both controllers. In fact, the same weight-update law and desired dynamics command filter were used, due to the careful selection of K_p and K_b.

The pitch channel network used learning gains of $\gamma = 10$ and $\eta = 0.1$. The basis functions were polynomials in normalized α, α^2, q, $\sigma(u_i)$, and $\sigma^2(u_i)$ where

$$\sigma(u_1) = \frac{2}{1+e^{-0.1u_i}} - 1. \tag{10}$$

This squashing function has an output between 1 and –1 and guarantees a solution to the algebraic loop around the neural network [8, 9]. The state inputs are normalized to have magnitudes roughly between 1 and –1. There are a total of 18 basis functions and 18 network weights in the pitch network.

The roll and yaw channels use the same learning gains as the pitch network, that is, $\gamma = 10$ and $\eta = 0.1$. However, the network architecture for these channels is different from that of the architecture for the pitch channel. While the roll and yaw channels have separate neural networks, the networks are identical in structure. There are 72 basis functions formed as polynomials in normalized α, α^2, p, r, β, $\sigma(u_i)$, and $\sigma^2(u_i)$. The roll channel uses the roll control command as an input, and the yaw channel uses the yaw control command as an input.

Results

The simulation results presented are for a tailless fighter aircraft in a 6DOF simulation, flying at mach 0.35, 15,000 ft. The only control effectors used are those which are not bounded above or below by zero. This is done so that a simple ganging can be used and avoid confusing the effects of the neural network and a complex control allocation scheme. The fighter model has many other control surfaces that are not being used, limiting the size of maneuvers that can be performed. The control surfaces being used are pitch flap, left elevon, right elevon, and pitch and yaw thrust vectoring. Pitch flap, pitch thrust vectoring and the elevons provide pitching moment, yaw thrust vectoring provides yaw moment, and differential elevons provide roll moment.

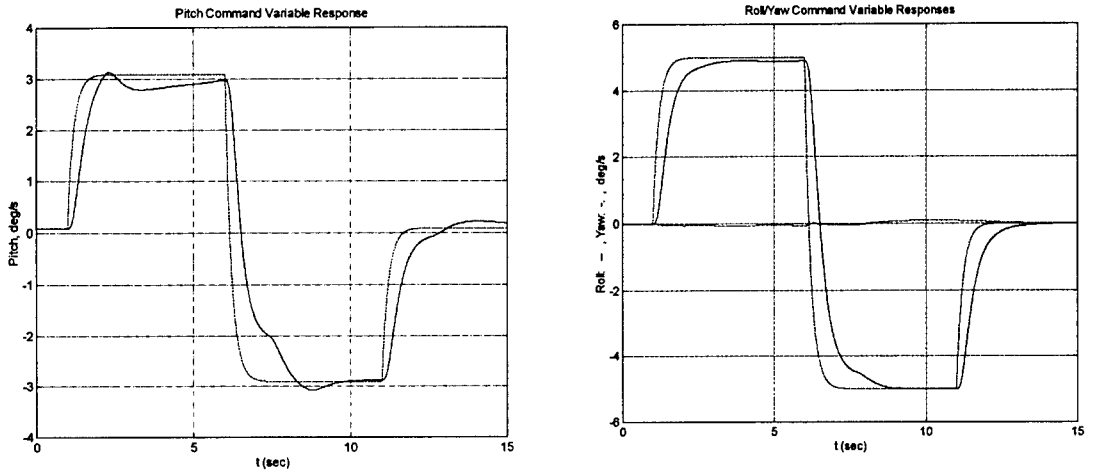
The controlled variables are blends of states. The pitch control variable is a blend of angle-of-attack and body-axis pitch rate. The weighting on alpha at this flight condition is small, giving an almost, but not quite, pure body-axis pitch rate command. The yaw command variable is a blend of sideslip angle and stability-axis yaw rate. The roll command variable is stability-axis roll rate. The dynamic inversion is being performed using linear models of the aircraft dynamics at each point, not the full nonlinear dynamics. The resulting inversion error is easily handled by the adaptive neural network or by integral desired dynamics, but without either method a tracking error exists.

Figure 3 shows the response of the system that uses a PI controller and no neural network to compensate for errors in the dynamic inversion. Figure 4 shows the same situation with a neural network augmenting the PI controller. Figure 5 shows the response of the system to a PD and neural network controller.

Reconfigurable Control in Response to Failure

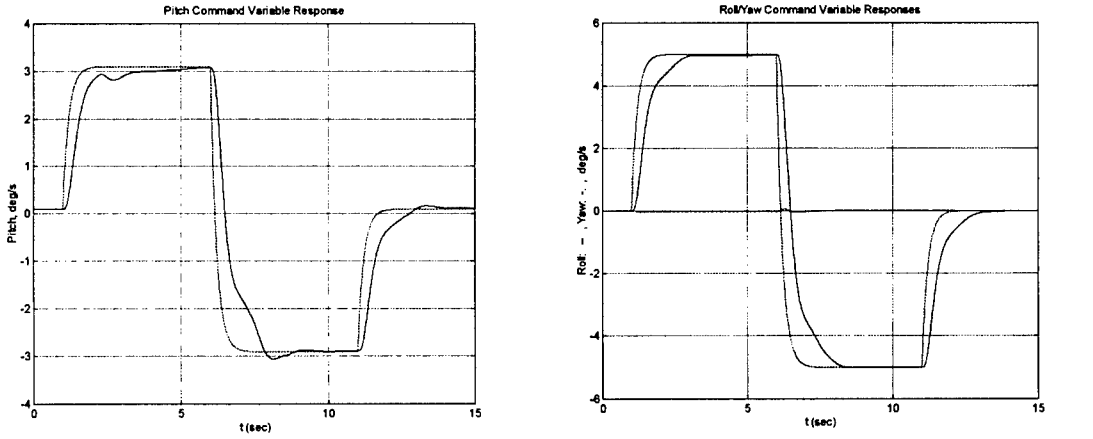
We then simulated a failure case in which the left elevon is lost at 0.5 seconds into a pitch doublet maneuver. The loss of this surface would be expected to cause large errors in both pitch and roll dynamics. In the case of the roll dynamics, this failure is a loss of 50% of control authority. The controller has no information about this failure, but the neural network adapts its weights to greatly reduce the resulting inversion errors.

Figure 6 displays the response of the system to the failure case for the PI controller. Figure 7 displays the response of the system to the failure case for the PI and neural network controller. Finally, Figure 8 displays the response of the system to the failure case with the PD and neural network controller. Results for the PD controller without neural network adaptation are not shown, as this controller performs poorly even without the failure.



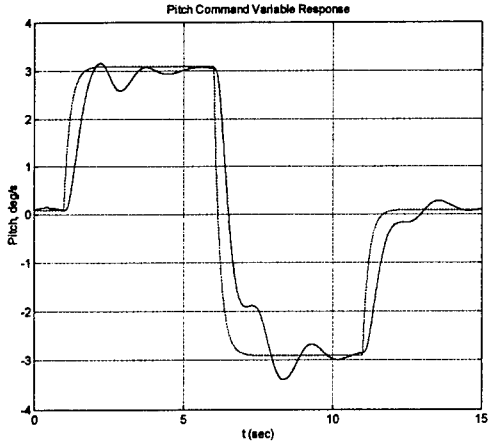
(a) (b)

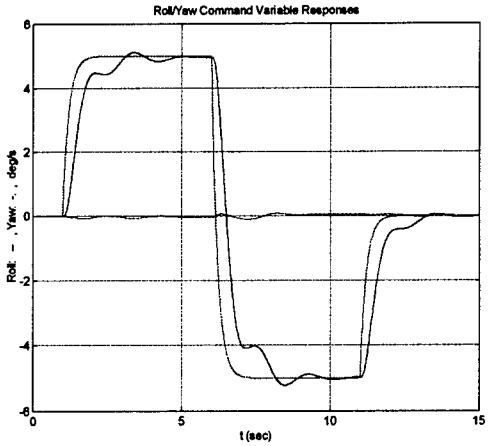
Figure 3. Response of the pitch CV (a) and roll and yaw CVs (b) to commanded pitch doublet. Controller: PI controller and no neural network, no control failure. Errors in the response are due to errors in the dynamic inversion process.



(a) (b)

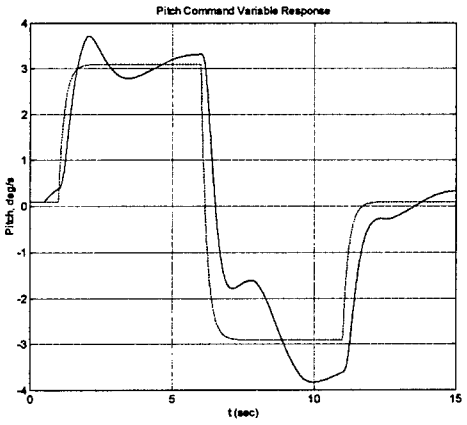
Figure 4. Response of the pitch CV (a) and roll and yaw CVs (b) to commanded pitch doublet. Controller: PI controller and neural network, no control failure. Neural network attempts to compensate for errors in the dynamic inversion process.

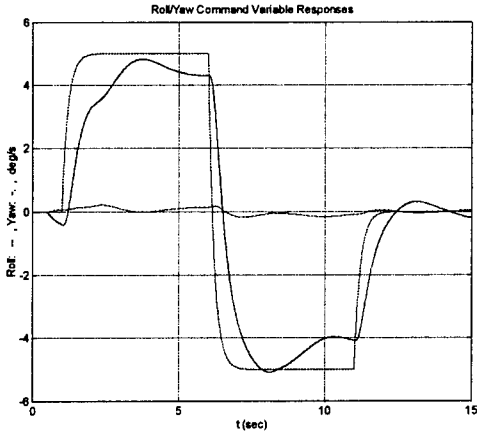




(a) (b)

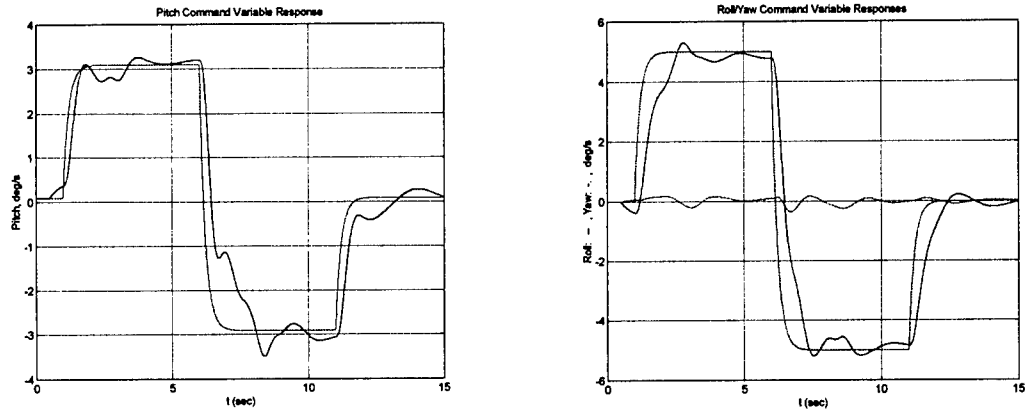
Figure 5. Response of the pitch CV (a) and roll and yaw CVs (b) to commanded pitch doublet. Controller: PD with neural network, no control failures. Errors in dynamic inversion have a greater effect on PD controller than on PI controller.





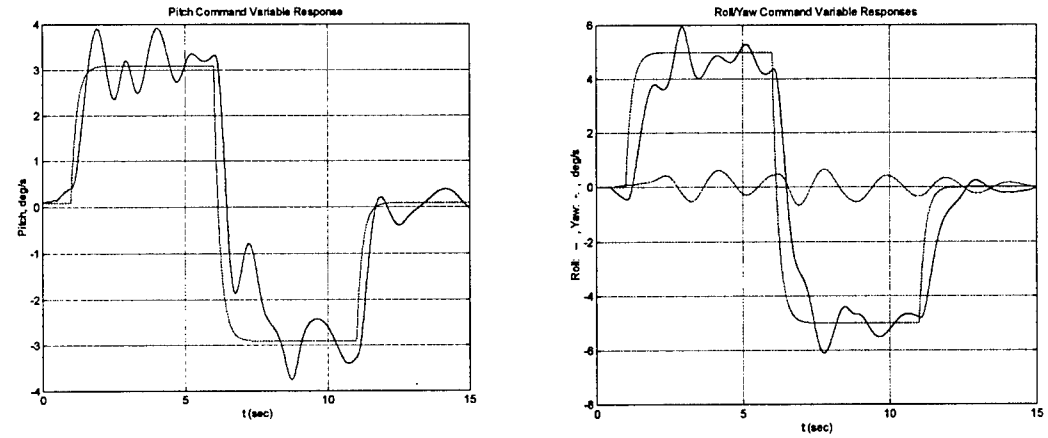
(a) (b)

Figure 6. Response of the pitch CV (a) and roll and yaw CVs (b) to commanded pitch doublet. Controller: PI controller and no neural network, loss of the left elevon at 0.5s.



(a) (b)

Figure 7. Response of the pitch CV (a) and roll and yaw CVs (b) to commanded pitch doublet. Controller: PI controller and neural network, loss of the left elevon at 0.5s.



(a) (b)

Figure 8. Response of the pitch CV (a) and roll and yaw CVs (b) to commanded pitch doublet. Controller: PD controller and neural network, loss of the left elevon at 0.5s.

Discussion

We have shown results of the effectiveness of using PI plus neural network controller to control a system linearized by dynamic inversion. The neural network compensates for errors in the dynamic inversion while PI control is less effected by remaining errors than PD control. We have also shown that this system is capable of compensating for the complete loss of a control surface. Furthermore, the use of PI plus neural network required no more complexity over that of the PD plus neural network as we used the same update laws and neural network architectures for both controllers.

Acknowledgements

I would like to thank AFOSR and RDL for their support of this program. I would also like to thank Drs. Siva Banda, Corey Schumacher, and James Buffington of AFRL/VAAD for their incorporating me into their group as if I was a permanent member.

References

1. Wilson, J.R., UAVs: a bird's eye view. Aerospace America, 1996: p. 38-43.
2. Rauch, H.E., Autonomous Control Reconfiguration. IEEE Control Systems Magzine, 1995. 15(6): p. 37-48.
3. Menke, T.E. and P.S. Maybeck, Sensor/actuator failure detection in the Vista F-16 by multiple model adaptive estimation. IEEE Transactions on Aerospace and Electronic Systems, 1995. 31(4): p. 1218-1228.
4. Maybeck, P. and P. Hanlon, Performance Enhancement of a multiple Model Adaptive Estimator. IEEE Transactions on Aerospace and Electronic Systems, 1995. 31(4): p. 1240-1254.
5. Narendra, K.S., Adaptive Control Using Multiple Models. IEEE transactions on Automatic Control, 1997. 42(2): p. 171-187.
6. Narendra, K. and K. Parthasarathy, Indentificiation and control of dynamical systems using neural networks. IEEE Transactions on Neural Networks, 1990. 1(1): p. 4-27.

7. Chandler, P.R., M. Pachter, and M. Mears, System identification for adaptive and reconfigurable control. Journal of Guidance, Control, and Dynamics, 1995. 18(3): p. 516-524.
8. McFarland, M.B., Adaptive Nonlinear Control of Missles using Neural Networks, in Aerospace Engineering. 1997, Georgia Institute of Technology: Atlanta, GA. p. 157.
9. Kim, B.S. and A.J. Calise, Nonlinear flight control using neural networks. Journal of Guidance, Control, and Dynamics, 1997. 20(1): p. 26-33.
10. Singh, S.N. and M. Steinberg, Adaptive control of feedback linearizable nonlinear systems with application to flight control. Journal of Guidance, Control, and Dynamics, 1996. 19(4): p. 871-877.
11. Reiner, J., G.J. Balas, and W.L. Garrard, Robust dynamic inversion for control of highly maneuverable aircraft. Journal of Guidance, Control, and Dynamics, 1995. 18(1): p. 18-24.
12. Brinker, J.S. and K.A. Wise, Stability and flying qualities robustness of a dynamic inversion aircraft control law. Journal of Guidance, Control, and Dynamics, 1996. 19(6): p. 1270-11277.
13. Adams, R.J. and S.S. Banda, Robust flight control design using dynamic inversion and structured singular value synthesis. IEEE Transactions on Control Systems Technology, 1993. 1(2): p. 80-92.
14. Agarwal, M., A systematic Classification of Neural-Network-Based Control. IEEE Control Systems Magzine, 1997. 17(2): p. 5-7.
15. Antsaklis, P.J., Intelligent Learning Control. IEEE Control Systems Magzine, 1995. 15(3): p. 5-7.

FUEL-AIR HEAT EXCHANGER FOR COOLED COOLING AIR SYSTEMS WITH FUEL-MIST AND AIR-JET IMPINGEMENT

Jayanta S. Kapat
Mechanical, Materials & Aerospace Engineering Department

University of Central Florida
Orlando, FL 32816-2450

Final Report for:
Summer Faculty Research Program
Air Force Research Laboratory, WPAFB

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Turbine Branch, Turbine Division
Air Force Research Laboratory, WPAFB

August 1998

FUEL-AIR HEAT EXCHANGER FOR COOLED COOLING AIR SYSTEMS WITH FUEL-MIST AND AIR-JET IMPINGEMENT

Jayanta S. Kapat
Assistant Professor
Department of Mechanical, Materials & Aerospace Engineering
University of Central Florida

Abstract

An innovative design of a compact fuel-air heat exchanger with fuel mist and air jet impingement is presented in this report. The new design not only enhances heat exchange through impingement and mist evaporation, but also is expected to reduce, or even eliminate, deposit formation, which is quite typical in a heat exchanger involving liquid fuel. This design should provide light, compact and maintenance-free fuel-air heat exchangers for cooled cooling air systems and other applications. The pre-vaporized fuel, when burnt in the combustor, also helps to eliminate fouling of combustor nozzles.

Different parameters affecting the overall design of such an exchanger are discussed in this report. An analysis of the fuel-side chemistry, flow and heat transfer is presented to compare the different characteristic times of the system. This analysis should help to identify the range of values for different operating parameters under which the heat exchanger is expected to perform satisfactorily.

FUEL-AIR HEAT EXCHANGER FOR COOLED COOLING AIR SYSTEMS WITH FUEL-MIST AND AIR-JET IMPINGEMENT

Jayanta S. Kapat

1. Introduction

Liquid jet fuel is the primary coolant for on-board heat sources in modern military aircraft (Edwards, 1993), and hence is the primary option for coolant in all on-board heat exchangers, including those for cooled cooling air systems that supply cooling air to turbine blades. However, at elevated temperatures, jet fuels undergo chemical reactions leading to formation of deposits, which can potentially cause fouling in a conventional shell-and-tube or cross-flow heat exchangers. Directly or indirectly, overall thermal management in aircraft and hence overall mission capability depend, to a large extent, on heat sink capabilities of liquid fuel.

To provide the necessary cooling, a part of the jet fuel is typically taken out of the fuel feed to the combustor. Fuel that remains liquid after heat exchange is recirculated back into the fuel tank, and fuel that gets vaporized in the process of cooling is introduced into the combustor. In alternate designs, remaining liquid fuel may be repressurized and introduced into the combustor. Hence, if properly used, the heat load on the fuel should not interfere with its primary mission, which is to provide energy in the combustor. However, there is a limit in heat load that can be imposed on liquid fuel. The heat absorbed by the fuel causes its temperature to increase. All hydrocarbon fuels have a temperature limit above which fuels go through irreversible decomposition reactions that lead to formations of gums and other deposits. These deposits may cause fouling of fuel nozzles and heat exchangers, and eventually, fuel and cooling system malfunction. There are two different mechanisms for deposit formation in hydrocarbon fuels: thermal-oxidative and pyrolytic (Edwards and Atria, 1997). Thermal-oxidative deposition occurs at lower temperatures, typically in a temperature range of 125 °C to 300 °C, whereas pyrolytic depositions typically occur above 480 °C. Typically, these two mechanisms require different strategies to mitigate them.

It has been observed that when the fuel is used in the form of droplets, rather than as bulk liquid, to cool hot objects, deposit formation is reduced or even completely eliminated (Hylton et al., 1996). It is conjectured that the fine droplets produced in a mist block the deposit formation process by one or both of two mechanisms. First, if the droplets in a mist are small enough, their temperature will rapidly increase as they

come in contact with hot walls. This short thermal transient time may not be enough for one or more of the deposit forming chemical reactions to be completed. Second, as droplets impinge on a hot surface, the chemistry may be completely altered and deposits may not be formed. Whatever be the actual mechanism in prevention or reduction of deposit formation, the results obtained in the study by Hylton et al. (1996) indicate the usefulness of fuel mist, as opposed to bulk liquid, in a fuel-air heat exchanger. Here an alternative design of a fuel-air heat exchanger is presented, which exploits this observed property of fuel mist in order to reduce problems of deposit formation. However, for proper design of such a heat exchanger, physical mechanisms involved in impingement of fuel mist on a hot surface must be understood.

2. Description of Proposed Heat Exchanger

In the proposed design, the fuel is introduced to the heat exchange surface in the form of mist jets (Figure 1). Fuel droplets in the mist get vaporized as they move out of the heat exchanger under a cross-flow of air that is taken from the bypass line. The cross-flow air is introduced with a circulation (Figure 1c) so that the mist droplets move radially out towards the hot heat exchanger surface under centrifugal action of circulating cross-flow. The hot side air is introduced to the heat exchanger surface also in the form of jets in order to enhance heat transfer on the hot side as well.

3. Technical Benefits over Conventional Fuel-Air Heat Exchangers

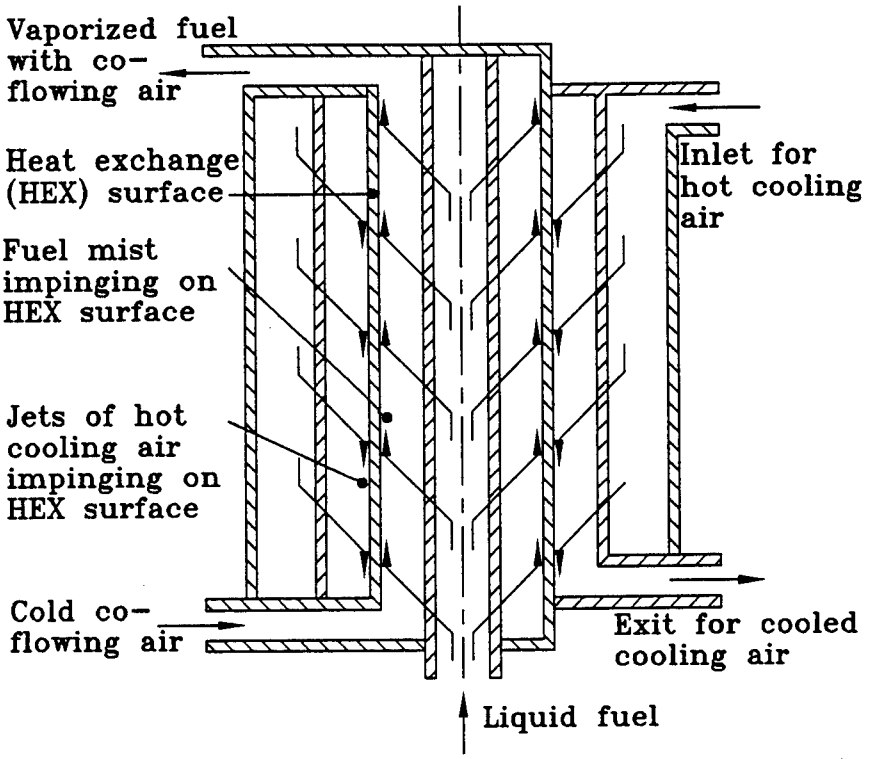
A. Heat transfer enhancement due to jet impingement on both cold (fuel) and hot (CD air) sides. Note: Jet impingement on only one side is not much useful because in that case heat transfer on the other side will become the limiting factor. Also, since cold side will have the extra benefit due to mist evaporation, the jet parameters on the hot side may have to be adjusted to provide similar heat transfer rate on the hot side.

B. Heat transfer enhancement due to mist evaporation on the cold (fuel) side.

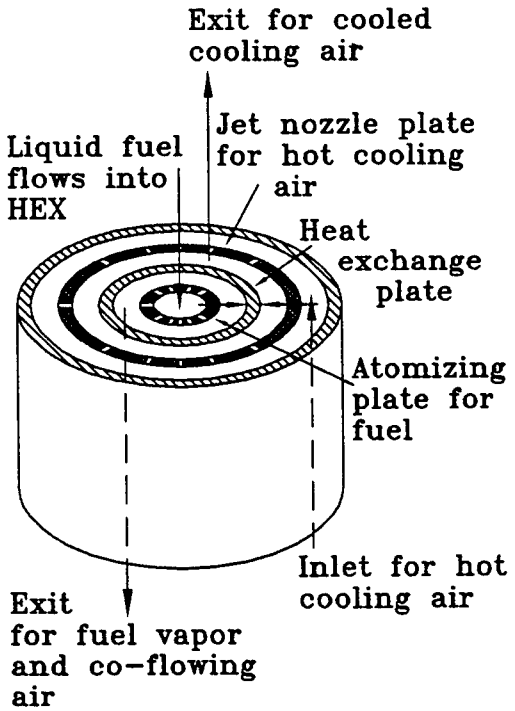
C. Small evaporating droplets may inhibit onset of chemical reactions that lead to oxidative and thermal deposits.

4. Payoff

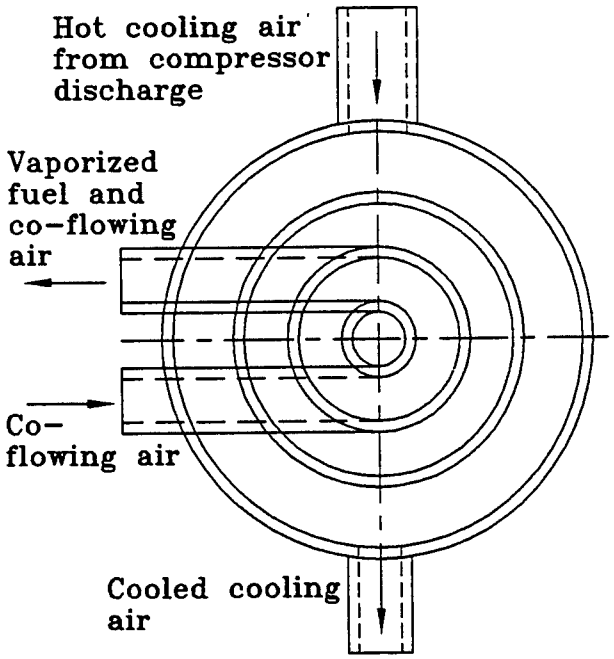
A. A light, compact, safe and maintenance-free fuel-air heat exchanger will be useful not just for Air Force, but for other branches of armed forces as well.



(a) Cross-sectional View (All inlets and outlets are not in the same plane, even though shown here to be so. (Correct orientations of inlets and outlets can be viewed in part c.)



(b) Overall Schematic



(c) Schematic showing circumferential locations of inlets and outlets

Figure 1. Schematic Diagram of the Proposed Fuel-Air HEX Based on Fuel Mist and Air Jet Impingement

B. The proposed heat exchanger, when used in a cooled cooling air system, can help in running gas turbine engines at lower blade temperature (thus increasing safety) and/or help in reducing cooling air consumption (thus increasing efficiency).

C. Pre-vaporized fuel, which is produced by this heat exchanger and which is to be burnt in combustor, will help combustion and reduce fouling of fuel nozzles, thus lowering maintenance frequency.

D. This design will complement and/or supplement anti-oxidant fuel additives that are currently being developed to suppress formation of thermal-oxidative deposits.

5. Issues Involved

The following questions need to be answered in order to determine the usefulness, feasibility and practicality of such a heat exchanger.

1. *Time Scale Comparison*: Can mist impingement heat transfer, as compared to bulk convective heat transfer, reduce deposit formation?

   It is conjectured that temperature of small droplets in a mist rapidly increases as the droplets come in contact with hot walls. This short thermal transient time may not allow one or more of the deposit forming chemical reactions to occur.

2. *Chemistry of Deposition*: Does reaction chemistry change in case of mist impingement on a hot surface, as compared to the case of bulk fluid flow over a hot surface?

   Answer to question 1 depends on the answer to question 2. However, by initially assuming that chemistry remains the same, question 1 can be answered independent of question 2 in order to obtain a rough idea of the problem involved.

3. *Behavior under Supercritical Conditions*: How do droplet behavior and chemistry (Question 1 and 2, respectively) change under supercritical conditions?

4. *Mist Generation*: Can enough atomization be produced without the use of any atomizing secondary flow? What should be the length and shape of the atomizing holes (which act like nozzles) in order to produce the optimum atomization?

5. *Amount of Co-flowing Air*: How much cold co-flowing air is needed to go with the atomized fuel in order to avoid any deposition? Can we completely avoid the need for co-flowing air where the pressure

difference between fuel inlet and vapor outlet would cause a strong enough flow that resists any deposition?

Eliminating the need for any co-flowing air will enhance inherent safety of the heat exchangers without additional safety measures. For example, in an actual flight, the liquid fuel flow rate may have to be changed because of changes in either cooling needs for turbine blades or fuel needs for combustor. If co-flowing air is used, then its flow rate also has to be appropriately adjusted to make sure that the resultant fuel-mixture is outside the flammability limit for all temperatures that may be encountered on the flow path. Thus, the optimum design will not use any co-flowing air, in which case the atomized mist jets will expand in an environment of fuel vapor.

6. *Generation of Air Jets*: How hot cooling air jets can be adjusted to provide optimum heat transfer on the hot side of the heat exchange plate? What should be length and shape of jet holes (or nozzles)?

   It should be noted that a low heat transfer coefficient on one side will cause a "bottle-neck" in the heat exchange path, in which case a high transfer rate on the other side will be a wasteful over-design.

7. *Nozzle Hole Patterns*: How should hole patterns on fuel and air sides be coordinated, so that highest possible heat exchange effectiveness be obtained? What should be optimum hole patterns for structural integrity?

8. *Safety*: What safety features must be adopted? Should new design codes be developed in order to ensure proper safety?

In this study, we focus our effort only to the first topic. We analyze and compare the different characteristic times that determine the performance of this heat exchanger. However, chemistry for deposit formation, which is presented in the next section, needs to be discussed first before characteristic times for chemical kinetics can be calculated.

6. Chemistry for Deposit Formation in Fuel Handling Systems

A. Chemistry for Thermal-Oxidative Deposition

Past studies (e.g. Jones et. al, 1995) have shown that there is a strong correlation between deposit formation in the lower temperature ranges and oxygen depletion in the fuel. This observation has prompted

suggestion of different reaction chemistry models in order to predict deposit formation, where the kinetics constants are determined empirically.

All of these models (e.g. Krazinski et. al, 1992) typically contain the following reaction or a variation of it.

$$F + O_2 \rightarrow P \tag{1}$$

Original model of Krazinski et. al (1992) assumes that rate of the above reaction is of first order in oxygen and zeroth order in fuel. However, after careful comparison with experimental results of Jones et. al (1995) and other researchers, Katta et. al (1993) suggested the following kinetics.

$$-\frac{dF}{dt} = -\frac{dO_2}{dt} = A_1 \exp\left(-E_1/RT\right); A_1 = 2.53 \mathrm{X} 10^{13}\ mole/m^3/s; E_1 = 32\ kcal/mole \tag{2}$$

The above reaction rate is of zeroth order in both fuel and oxygen. However, they found that when oxygen concentration drops below 10 ppm, the reaction mechanism switches to first order reaction in oxygen with the following constants:

$$A_{11} = 1.65 \mathrm{X} 10^{14}\ s^{-1}; \quad E_{11} = 35.5\ kcal/mole \tag{3}$$

Katta et. al (1993) have also proposed a more detailed chemistry model with 6 bulk fluid reactions and 3 wall reactions. These reactions and their chemical kinetic constants are listed in Table 1, where rate constants for all reactions are assumed to follow Arrhenius' law.

Table 1. 9-Step Chemistry Model for Thermal-Oxidative Deposition (Katta et. al, 1993)

No.	Reaction	O1	O2	E (kcal/mol)	A
B1	$F + O_2 \rightarrow ROO^*$	0	0	32.0	$2.5 \mathrm{X} 10^{13}$

B2	$ROO^* + F \rightarrow S$	1	0	10.0	1.0×10^4
B3	$ROO^* + F_S \rightarrow P$	1	0	15.0	8.0×10^9
B4	$ROO^* + F \rightarrow D_B$	1	0	10.0	2.0×10^2
B5	$P + F \rightarrow S$	1	0	30.0	3.2×10^{12}
B6	$D_B + F \rightarrow 2D_B$	1	0	0	1.0×10^{-3}
W1	$O_2 + F \rightarrow P$	1	0	12.0	5.2×10^{-3}
W2	$P \rightarrow D_W$	1	---	17.0	260
W3	$D_B \rightarrow D_W$	1	---	10.0	0.80

Concentration of dissolved oxygen is quite small compared to that of fuel, the concentrations of all intermediate or final products will also be negligible. Hence, fuel concentration will not have any appreciable changes during these reactions.

B. Chemistry for Pyrolytic Deposition

There have been recent efforts in modeling chemistry for pyrolytic deposition in jet fuels. One such model has been presented by Sheu et. al (1998). The kinetic parameters in this models have been obtained from curve-fitting of experimental data. The chemistry model is presented in Table 8.

Table 8. Chemistry Model for Pyrolytic Deposition (Sheu et. al, 1998)

Reactions	Order	Activation Energy (*kcal/mole*)	Pre-Exponential Factor
$F \rightarrow \beta$ *products*	1	63	3.64E+15
$F \rightarrow$ *deposits*		60	2.00E+16

7. Calculation of System Characteristic Times

A. Characteristic Time for Auto-Oxidation or Chain Propagation Reaction (B1):

Initial concentration of dissolved oxygen in typical jet fuels is around 60 to 80 ppm (Edwards, 1998). With a specific gravity of 0.8 for a typical jet fuel at room temperature, the average molar concentration of dissolved oxygen in jet fuels can be calculated to be 1.8 moles/m^3 (corresponding to 70 ppm of dissolved oxygen).

Since for reaction B1, rate of reaction, which is also oxygen depletion rate, is given by

$$-\frac{d[O_2]}{dt} = k_{B1} = A_{B1} \exp\left(-\frac{E_{B1}}{RT}\right), \tag{4}$$

characteristic time for this reaction can be calculated as

$$\tau_{B1} = {[O_2]_{initial}}/{k_{B1}}. \tag{5}$$

Table 2 shows value of this characteristic time for different temperatures and for an initial concentration of 1.8 moles/m^3.

Table 2. Characteristic Time for Reaction (B1) at Different Temperatures

$T\ (^oF)$	200	300	400	500	600
$\tau_{B1}\ (s)$	8.94×10^5	2.7×10^3	32.3	0.96	0.055

B. Characteristic Times for Chain Termination Reactions (B2, B3, B4):

Since these reactions are of first order (overall order as well as order in peroxy radical, ROO^*), corresponding reaction rates can be represented as

$$-\frac{d[ROO^*]}{dt} = [ROO^*]k, \text{where } k = A\exp\left(-\frac{E}{RT}\right) \tag{6}$$

In order to estimate characteristic times for these reactions, we can solve the above ordinary differential equation for each chain termination reaction as

$$[ROO^*] = C_1 \exp\left(-{t}/{\tau}\right), \text{ where } \tau = {1}/{k}. \tag{8}$$

Here, τ is the reaction time constant or characteristic time for each individual reaction.

Since (1) B1, B2 and B3 provide three parallel reaction pathways for chain termination, (2) there is unlimited supply of the other reactants of these reactions, and (3) each of these reactions are taken to be irreversible in the chemistry model presented in Table 1, total rate for consumption of peroxy radical is sum of the three individual rates, and the corresponding overall time constant (τ_{CTR}) will be reciprocal of that

overall rate. Individual characteristic times for these reactions as well as the overall one are listed in Table 3 for different temperatures.

Table 3. Characteristic Times for Chain Termination Reactions

$T\ (^{O}F)$	200	300	400	500	600
$\tau_{B2}\ (s)$	92.68	15.19	3.790863	1.263637	0.51827
$\tau_{B3}\ (s)$	0.1115	0.007398	0.000923	0.000178	4.66E-05
$\tau_{B4}\ (s)$	4633	759.3383	189.5431	63.18183	25.91351
$\tau_{CTR}\ (s)$	0.1114	0.007394	0.000922	0.000178	4.66E-05

It can be seen that reaction B3 is quite dominating over the other two chain termination reactions (B2 and B4), especially at higher temperatures. At 600°F, characteristic time for this reaction is at least 4 orders of magnitude smaller, corresponding to a reaction rate that is at least four orders of magnitude larger, compared to the same quantities for the other two reactions. These numbers indicate that when only a limited amount of peroxy radical is available, almost all of this radical will be consumed in reaction B3, producing precursor for deposit formation. This further indicates that among the following two pathways for formation of wall deposits:

$$O_2(dissolved) \rightarrow ROO^{*} \rightarrow P \rightarrow D_W \text{ , and} \tag{9}$$

$$O_2(dissolved) \rightarrow ROO^{*} \rightarrow D_B \rightarrow D_W \text{ ,} \tag{10}$$

the first pathway completely dominates over the second one. Moreover, formation of bulk particles (insolubles) is not significant. However, if the first pathway is somehow blocked, the second pathway will become important and/or bulk particles may become the only significant reason for fuel-system fouling.

C. Characteristic Time for Precursor Removing Reaction (B5):

Reaction B6 behaves similar to the last three reactions, as it is of first order in the first reactant and of zeroth order in the second. Hence, characteristic time for this reaction is also reciprocal of the reaction rate, and its values at different temperatures are presented in Table 4.

Table 4. Characteristic Time for Precursor Removal Reaction

$T\ (^{o}F)$	200	300	400	500	600
$\tau_{B5}\ (s)$	248677	1095	17.02	0.6305	0.04350

It should be noted that if reaction B5 can remove the precursor, P, at a fast enough rate from the bulk of the flow, then less P will be available for deposit formation at wall through reaction W2.

D. Characteristic Time for Agglomeration Reaction (B6):

Characteristic time for this reaction is also reciprocal of reaction rate. Since activation energy for this reaction is zero, the reaction rate does not depend on temperature. Hence, characteristic time does not depend on temperature either, and is presented in Table 5.

Table 5. Characteristic Time for Agglomeration Reaction

$\tau_{B5}\ (s)$	1000 (at all temperatures)

Value of characteristic time indicates that agglomeration reaction takes a long time to cause significant bulk agglomeration.

E. Characteristic Times for Thermal Transients in a Mist Flow:

As a relatively cold liquid drop enters a hot surrounding from an atomizer, almost all the heat supplied to the drop initially serves to raise its temperature. This initial period is called pre-heat period. As the liquid temperature rises, the fuel vapor formed at the drop surface has two effects. (1) A portion of the heat supplied is used to provide the heat of vaporization. (2) Outward flow of fuel vapor from the drop surface impedes the inward diffusion of heat towards the drop. Both of these effects reduce the rate of increase of surface temperature, and temperature distribution within the drop becomes more uniform. Finally, a stage is reached when drop temperature stabilizes at its wet-bulb temperature and all of the heat transferred to the drop is used for vaporization. This period is called the steady state evaporation period (Lefebvre, 1989).

In the beginning of initial heat-up, it may be expected that there will be a significant relative velocity between the drop and the surrounding fluid as the drop is forced out of the atomizer. As a result, heat

transfer at that time will be dominated by forced convection, and a lumped-body energy balance for a small drop will provide,

$$\frac{d}{dt}(\rho V c T) = -hA(T - T_\infty),$$

which can be solved to obtain

$$(T - T_\infty) = (T_i - T_\infty)\exp\left(-t/\tau_{FC}\right),$$

where T_i is the initial drop temperature and τ_{FC} is time constant for a drop in forced convection as given by

$$\tau_{FC} = \frac{\rho V c}{hA} = \frac{\rho c}{3} \cdot \frac{r}{h}.$$

The values of this time constant for different drop velocities and drop sizes are shown in Table 6 for JP-5 drops in air at 15 *atm.* and 550 *K.*

Table 6. Characteristic Times (in seconds) for Forced Convection for Different Drop Sizes and Velocities

	0.01 mm	0.1 mm	1 mm
1 m/s	9.15E-04	0.0544	2.24
10 m/s	5.44E-04	0.0224	0.717

In these calculations, heat transfer coefficients for spherical drops moving through a gas are obtained from correlations provided by Whitaker [1972]:

$$Nu_{D,ave} = 2 + \left(0.4\,Re_D^{0.5} + 0.06\,Re_D^{0.67}\right)Pr^{0.4}\left(\frac{\mu}{\mu_W}\right)^{0.25}.$$

As the droplets move through the surrounding hot gas and/or vapor, they will evaporate. The rate of evaporation depends on several factors (Lefebvre, 1989), including (1) thermo-physical properties of ambient gas: pressure, temperature, thermal conductivity, specific heat, viscosity, (2) slip velocity between droplets and ambient gas; (3) thermo-physical properties of liquid fuel and its vapor: density, vapor pressure, thermal conductivity, and specific heat, and (4) initial conditions of droplets: size and temperature. In the steady state of droplet vaporization, a balance is reached between rates of evaporation and outward diffusion and convection of the resulting fuel vapor. Under this condition and under the simplifying assumption of

$$\text{Lewis number, } Le \equiv \frac{\alpha_g}{\mathrm{D}_F} \equiv \frac{Sc(\text{Schmidt number})}{Pr(\text{Prandtl number})} \approx 1 ,$$

governing 1-dimensional mass transfer and heat transfer equations for a symmetrical. evaporating spherical droplet can be solved analytically (Kanury, 1984). The corresponding mass rate of vaporization per unit surface area of a liquid droplet is given by

$$\dot{m}_V = \frac{2\rho_g \alpha_g}{D} \ln(B+1)$$

where B is the heat transfer or mass transfer constant (which are equivalent because of unity Lewis number), and is defined as

$$B = \frac{c_g \left(T_\infty - T_{drop}\right)}{L} = \frac{Y_{drop\,surface} - Y_\infty}{1 - Y_{drop\,surface}}$$

with a lumped body assumption for the liquid drop. When $T_\infty >> T_{drop}$, $T_{drop} \approx T_{sat}(boiling)$. For JP-5 fuel droplets in air at 1273 K, the value of B is approximately 4.24. Here, the property values for JP-5 are obtained from CRC (1988). It should be noted that Hylton et. al (1996) used a gas temperature of 1922 K.

A mass balance of a single liquid droplet gives

$$-\rho_{F,liquid}\frac{\pi}{2}D^2\frac{dD}{dt}=\pi D^2\dot{m}_V=\pi D^2\frac{2\rho_g\alpha_g}{D}\ln(B+1)\Rightarrow D\frac{dD}{dt}=-\frac{4\rho_g\alpha_g}{\rho_{F,liquid}}\ln(B+1)$$

$$\Rightarrow\frac{d}{dt}(D^2)=-\lambda\quad or\quad D^2=D_0^2-\lambda t$$

where $\lambda=8\,(\rho_g/\rho_{F,liquid})\,\alpha_g\,\ln(B+1)$ is a constant and D_0 is the initial droplet diameter. Correspondingly, the characteristic time for droplet vaporization is given by

$$\tau_{vap}=(D_0^2/\lambda)$$

As expected, characteristic time for vaporization is shorter for a smaller droplet, higher value of thermal diffusivity for the surrounding gas or for larger heat/mass transfer number. Calculated values of this characteristic time for different initial sizes of JP-5 droplets in air at 1273 K are shown in the following table.

Table 7. Characteristic Times for Droplet Vaporization

(for JP-5 droplets in air at 1273 K)

D_0 (mm)	0.01	0.1	1
τ_{vap} (s)	1.06E-04	1.06E-02	1.06

The above analysis considers vaporization under low pressures in quiescent environments. Both higher pressure and slip velocity will affect the numbers calculated above. First, we discuss the effect of ambient pressure on characteristic vaporization time. At low pressures (compared to the critical pressure), the surface temperature of a droplet increases rapidly during the preheat period, and then levels off at a temperature slightly lower than the boiling temperature, condition which has been referred to as the steady state evaporation earlier. As the ambient pressure increases, the value of charactersitic vaporization time decreases only slightly. However, as the ambient pressure becomes more than the critical pressure, the surface temperature varies continuously and the vaporization process never reaches the steady state during the entire droplet lifetime (Yang et. al, 1995). As the ambient pressure becomes more than twice the critical

pressure and ambient temperature is also supercritical, droplet surface reaches the critical state within a short time. At that time, the distinct liquid/vapor interface vanishes and the enthalpy of vaporization drops down to zero, after which vaporization proceeds at a faster rate. Thus, in summary, the characteristic vaporization time decreases monotonically with pressure, and the rate of decrease increases slightly as the pressure exceeds the critical pressure.

When a droplet is vaporizing in quiescent environment, the primary mode of heat and mass transfer is diffusion. When a slip velocity develops between the droplets and the surrounding gas, rate of vaporization is enhanced due to convective effects. With the help of nonlinear data regression process, Yang et. al (1995) has proposed the following correlation.

$$\frac{\tau_{vap}}{\tau_{vap,\,Re=0}} = \left[1+0.09\,Re^{1.1}\left(\frac{p}{p_{ref}}\right)^{-0.88}\right]^{-1}$$

In this empirical relation, Yang et. al (1995) used twice the critical pressure as the value for p_{ref}.

F. Characteristic Time in Wall-Droplet Collision:

When a droplet collides with a surface that is at a temperature much higher than the saturation or boiling temperature, there is never a positive contact between the wall and the droplet. In this case, a thin vapor film is formed next to the wall because of flash vaporization (Figure 2), and separates the droplet from direct contact with the wall. The parameters that can possibly affect the collision time are surface tension of the liquid, and mass, volume and velocity of an individual liquid droplet. Since the droplets do not get in contact with the wall, the wall-droplet collision time does not affect the deposition process, and hence is not further considered, as the wall deposition reactions can not even get started.

8. Conclusion

From a comparison of the different characteristic times involved, it may be concluded that mist impingement impedes in deposit formation through several mechanisms.

(1) In the low temperature ranges, formation of bulk deposits is significantly impeded because of short vaporization times of fine droplets (less than 1 mm). However, under the same conditions, precursor molecules are still formed.

(2) Provided that the hot surface is heated above the Leidenfrot temperature, the mist droplets do not have any direct contact with the wall, and hence whatever precursor is formed inside the droplets do not transformed into wall deposits.

(3) Leidenfrost effect acts as an impediment also to pyrolytic deposition, which typically occurs at higher temperature ranges.

(4) Formation of pyrolytic deposit is further suppressed as the mist droplets get fully vaporized before they reach temperatures conducive for such depositions.

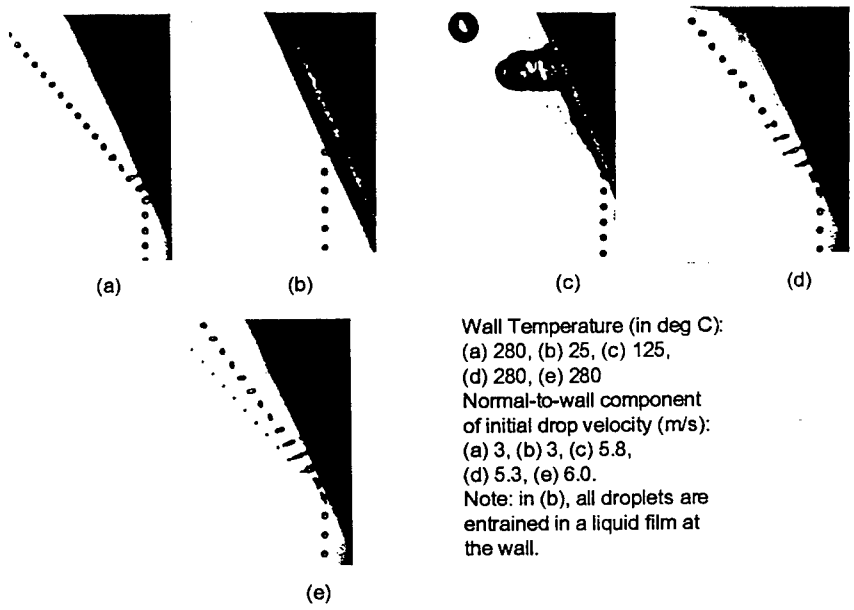


Figure 2. Wall collision of droplets with initial diameters of 140 microns. (Note: Here, approximate value for Leidenfrost temperature is 179 degrees C.)

9. References

Brandauer, M., A. Schulz, and S. Wittig, 1996, "Mechanism of Coke Formation in Gas Turbine Combustion Chambers," J. Eng. Gas Turbines and Power, vol. 118, pp. 265-270.

Daggett, D.L., A. Veninger, C. Lewis, S. Bullock, and R. Kamin, 1995, "The Development of an Aviation Fuel Thermal Stability Test Unit," J. Eng. Gas Turbines and Power, vol. 117, 468-474.

Edwards, T., 1993, "USAF Supercritical Hydrocarbon Fuels Interests," AIAA Paper No. 93-0807, 31st Aerospace Sciences Meeting & Exhibit, Jan 11-14, Reno, NV.

Edwards, T., 1998, personal communication.

Edwards, T., and J. V. Atria, 1997, "Thermal Stability of High Temperature Fuels," ASME Paper No. GT-143, ASME Turbo Expo, Jun 2-5, Orlando. FL.

Ervin, J. S., S. P. Heneghan, C. R. Martel. and T. F. Williams, 1996, "Surface Effects on Deposits from Jet Fuels," J. Eng. Gas Turbines and Power, vol. 118, pp. 278-285.

Ervin, J. S., and S. Zabarnick, 1997, "Numerical Simulation of Jet Fuel Oxidation and Fluid Dynamics," 6th International Conference on Stability and Handling of Liquid Fuels, Oct 13-17, Vancouver, BC.

Heneghan, S. P., S. Zabarnick, D. R. Ballal. and W. E. Harrison III, 1996, "JP-8+100: The Development of High Thermal Stability Jet Fuel," AIAA 96-0403, 34th Aerospace Sciences Meeting & Exhibit, Jan 15-18, Reno, NV.

Heneghan, S. P., C. R. Martel, T. F. Williams, and D. R. Ballal, 1993, "Studies of Jet Fuel Thermal Stability in a Flowing System," J. Eng. Gas Turbines and Power, vol. 115, pp. 480-485.

Hylton, L. D., N. V. Nirmalan, and P. C. Sweeney, 1996, "Advanced Cooling Concepts," NASA Contractor Report EDR 16666, March 1996.

Jones, E. G., W. J. Balster, and J. M. Pickard, 1996, "Surface Fouling in Aviation Fuels: An Isothermal Chemical Study," J. Eng. Gas Turbines and Power, vol. 118, pp. 286-291.

Jones, E. G., W. J. Balster, and M. E. Post, 1995, "Degradation of a Jet A Fuel in a Single-Pass Heat Exchanger," J. Eng. Gas Turbines and Power, vol. 117, pp. 125-130.

Katta, V. R., E. G. Jones, and W. M. Roquemore, 1993, "Development of Global Chemistry Model for Jet-Fuel Thermal Stability based on Observations from Static and Flowing Experiments," AGARD CP-536, Paper No. PEP-19, 81st Symposium on Fuels and Combustion Technology for Advanced Aircraft Engines, May 10-14, Colleferro, Italy.

Katta, V. R., and J. Blust, T. F. Williams, and C. R. Martel, 1995, "Role of Buoyancy in Fuel-Thermal-Stability Studies," J. Thermophysics and Heat Transfer, vol. 9, no. 1, pp. 159-168.

Katta, V. R., E. G. Jones, and W. M. Roquemore, 1997, "Modeling of Deposition Process in Liquid Fuels," AIAA Paper No. 97-3040, 33rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, July 6-9, Seattle, WA.

Katta, V. R., and W. M. Roquemore, 1993, "Numerical Methods for Simulating Fluid-Dynamic and Heat Transfer Changes in Jet-Engine Injector Feed-Arm due to Fouling," J. Thermophysics and Heat Transfer, vol. 7, no. 4, pp. 651-660.

Krazinski, J. L., S. P. Vanka, J. A. Pearce, and W. M. Roquemore, 1992, " A Computational Fluid Dynamics and Chemistry Model for Jet Fuel Thermal Stability," J. Eng. Gas Turbines and Power, vol. 114, pp. 104-110.

Sheu, J. C., N. Zhou, A. Krishnan, E. G. Jones, and V. R. Katta, 1998, "Modeling of Pyrolysis and Pyrolytic Deposition of Norpar-13 under Supercritical Conditions," Structure of Jet Fuels V, ACS National Meeting, Aug 23-27, Boston, MA.

Yang, V., G. C. Hsiao, J. S. Shuen, and K. C. Hsieh, 1995, "Droplet Behavior at Supercritical Conditions," in Recent Advances in Spray Combustion: Spray Atomization and Drop Burning Phenomena – Volume 1 (K. K. Kuo, ed.), Progress in Aeronautics and Astronautics Series, vol. 166, AIAA, Reston, VA.

Zabarnick, S., P. Zelesnik, and R. R. Grinstead, 1996, "Jet Fuel Deposition: Dilution, Materials, Oxygen, and Temperature Effects," J. Eng. Gas Turbines and Power, vol. 118, pp. 271-277.

Spacecraft Formation Flying: A Survey

Vikram Kapila
Assistant Professor
Department of Mechanical, Aerospace,
and Manufacturing Engineering

Polytechnic University
Six Metrotech Center
Brooklyn, NY 11201

Final Report for:
Summer Faculty Research Program
Wright Research Site

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Wright Research Site

September 1998

Spacecraft Formation Flying: A Survey

Vikram Kapila
Assistant Professor
Department of Mechanical, Aerospace,
and Manufacturing Engineering
Polytechnic University

Abstract

A novel concept of distributed array of small, low-cost, cooperative, and highly coordinated micro-satellites is vigorously being pursued for several future space missions. Implementation of the distributed coordinating satellite concept will require tight control of the relative distances and phases between the participating satellites. This research is intended to provide a unified treatment of relative satellite position modeling and linear quadratic control that is suitable for further advancement in autonomous multiple spacecraft formation flying technology. Specifically, it provides a complete development of the nonlinear relative satellite position dynamic equations as well as the linearized dynamics. In addition, a control relevant relative satellite position dynamic model is obtained and linear quadratic controllers are designed. Finally, some concluding remarks pointing to the open research problems in multiple spacecraft formation flying are given.

Spacecraft Formation Flying: A Survey

Vikram Kapila

1. Introduction

A novel concept of distributing the functionality of large satellites among smaller, less expensive, cooperative satellites is seriously being considered for numerous space missions [1,2,4]. A practical implementation of the concept relies on the control of relative distances and phases between the participating satellites. A ground-based command and control system for relative positioning of multiple satellites will be excessively burdened and complex and may not be able to provide sufficiently rapid corrective control commands for collision avoidance. Thus, the concept of autonomous formation flying of satellite clusters is vigorously being studied and has been identified as an enabling technology for the success of various future missions of NASA and the U.S. Air Force. For example, the Earth Orbiter-I (EO-I) and the New Millennium Interferometer (NMI), also known as the Deep Space-3 (DS3), are two current NASA projects that seek to demonstrate the feasibility of multiple spacecraft formation flight (MSFF) in upcoming years. Similarly, the recently proposed Air Force project TechSat-21 [1] seeks to push the frontier in micro-scale MSFF to enable global awareness and rapid access to space in the 21st century.

A number of space missions necessitate MSFF. For example, docking of space shuttle with space station requires spacecraft rendezvous where it is necessary to fly the two spacecraft in close formation in order to capture and dock at the specified time with zero relative velocity. Similarly, satellite recovery and servicing missions rely on MSFF. Most prior missions requiring MSFF have been carried using manual flight control [3] and have been limited to one-leader-one-follower configuration. In the case of formation flying of clusters of multiple satellites, collision avoidance becomes a significant issue for which ground/manual-control may not be reliable. Thus, the development of autonomous formation control strategies is critical to the success of MSFF. Even though the concept of autonomous MSFF has not been flight tested yet, several techniques have been reported in the literature for autonomous MSFF. Specifically, Vassar and Sherwood studied

the concept of formationkeeping of satellites for a ground-based terrestrial laser communication system [14] whereas Redding *et al.* considered stationkeeping for the space shuttle Orbiter [12]. Recent years have witnessed a resurgence of interest in autonomous MSFF. In particular, Lau *et al.* [10] have considered MSFF for NASA's NMI which relies on separated spacecraft interferometry. In addition, NASA's EO-I is scheduled for a flight demonstration in 1999 of a stereo imaging concept that relies on MSFF [8].

The increased interest in MSFF can be attributed to the criticality of the MSFF technology in future space missions that have been proposed by the U.S. Air Force and NASA. For example, in the next millennium the U.S. Air Force desires to maintain a global virtual presence. Rapid access to information and space-based imaging are key ingredients for the success of this mission objective. In this regard, a reconfigurable cluster of distributed satellites can be used to design instantaneous synthetic aperture radar. Using the long baseline of the space-based collaborative satellites accurate target detection can be achieved. Furthermore, sensor data fusion from multiple platforms can yield stereo images that are not available from single platform sensors. The distributed satellite clusters can also enable theater-wide surveillance with all weather operation and performance. The concept of space defense involving ground directed lasers reflected off the space-based mirrors can also be advanced by the autonomous formationkeeping technology. In addition, NASA has shown a keen interest in the development of a reliable autonomous formationkeeping strategy in order to deploy multiple satellites for deep space missions.

The principal objective of this summer research program was to conduct a state-of-the-art survey in the field of MSFF. This report is an outcome of the survey and it provides:

- A detailed modeling of satellite relative motion dynamics.
- Analysis of the nonlinear relative motion equations and linearization.
- Development of the linear quadratic control algorithms.

We note that most results provided in this report have been previously reported in technical publications by various authors. This report is intended to provide a unified treatment of relative

satellite position modeling and linear quadratic control that is suitable for further advancement in autonomous MSFF technology.

The report begins with the dynamic modeling of the relative position of satellites. Next, the nonlinear dynamic equations are linearized to obtain the Clohessy-Wiltshire equations, originally derived in the context of satellite rendezvous problem [6]. The linearized equations are critically evaluated *vis-a-vis* the nonlinear model. Next, we obtain a discrete-time control relevant model for impulsive control of the satellite relative position dynamics. Furthermore, we use a linear quadratic control approach to design full-state feedback regulator [14] and dynamic output feedback controllers. Detailed simulation results for open-loop and closed-loop relative satellite position dynamics are given. In addition, various competing control design techniques reported elsewhere in the literature are mentioned. Finally, some concluding remarks pointing to the open research problems in MSFF are given.

2. Relative Satellite Position Dynamic Modeling

In this section we begin by deriving the dynamic equations of motion for a single satellite. Thus, referring to Figure 1, we consider the motion of a mass m relative to mass M. Note that according to Newton's inverse square gravitation law force applied on the mass M by the mass m is given by

$$\overline{F}_{M/m} = \frac{GMm}{r^2}\left(\frac{\overline{r}}{r}\right), \tag{2.1}$$

where G is the universal gravitational constant [5] and $\overline{r}$ is the vector pointing from M to m. Furthermore, by Newton's third law of equal and opposite reaction, force applied on the mass m by the mass M is $\overline{F}_{m/M} = -\overline{F}_{M/m}$. Now using Newton's second law, the equations of motion for M and m in vector notation are given by

$$M\ddot{\overline{a}} = \frac{GMm}{r^2}\left(\frac{\overline{r}}{r}\right), \tag{2.2}$$

and

$$m\ddot{\overline{b}} = -\frac{GMm}{r^2}\left(\frac{\overline{r}}{r}\right), \tag{2.3}$$

respectively. Finally, the equation of motion of mass m relative to mass M is obtained by forming $\frac{(2.3)}{m} - \frac{(2.2)}{M}$ which yields

$$\ddot{\overline{b}} - \ddot{\overline{a}} = -\frac{G(M+m)}{r^2}\left(\frac{\overline{r}}{r}\right),$$

or

$$\ddot{\overline{r}} + \frac{G(M+m)}{r^2}\left(\frac{\overline{r}}{r}\right) = 0. \tag{2.4}$$

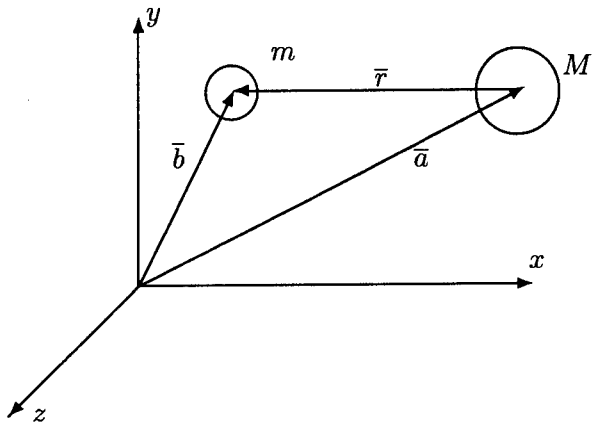


Figure 1: Motion of the Mass m Relative to the Mass M

In the case where we consider the motion of a satellite of mass m relative to the Earth of mass M, the mass M can be assumed to be fixed in the inertial reference. Furthermore, in this case $M >> m$. Hence, we assume that the motion of mass m does not influence the motion of mass M. In this case, (2.4) yields

$$\ddot{\overline{r}} + \frac{GM}{r^2}\left(\frac{\overline{r}}{r}\right) = 0, \tag{2.5}$$

or, equivalently,

$$\ddot{\overline{r}} + \frac{\mu}{r^2}\left(\frac{\overline{r}}{r}\right) = 0, \tag{2.6}$$

where $\mu \triangleq GM$. In the case of periodic motion, it can be shown that

$$\omega = \sqrt{\frac{\mu}{r^3}}, \tag{2.7}$$

where ω is the angular velocity of motion. Equation (2.7) is typically referred to as Keplar's equation [16]. Finally, (2.6) can be rewritten as

$$\ddot{\overline{r}} + w^2\overline{r} = 0. \tag{2.8}$$

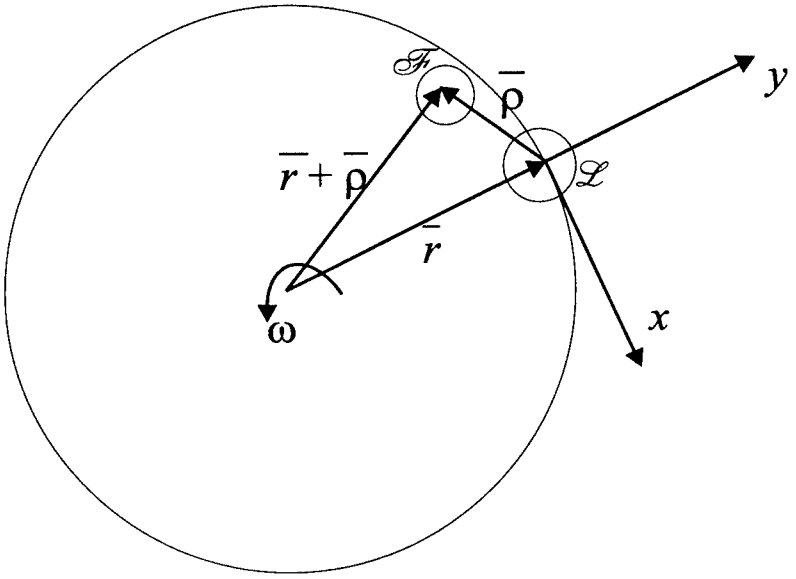


Figure 2: Motion of the Follower Satellite Relative to the Leader Satellite

Next, we develop the dynamic equations governing the relative motion of two satellites. Thus, consider Figure 2 where $\mathcal{L}$ denotes the leader satellite and $\mathcal{F}$ denotes the follower satellite. In the following development we assume that:

- Inertial reference is attached to the earth.
- The leader satellite is in a circular orbit around the Earth with an angular velocity ω.
- The instantaneous position vector from earth center to the leader satellite is $\overline{r}$.
- A rectangular moving coordinate frame is attached to the leader satellite with the x-axis pointing opposite to the instantaneous tangential velocity, the y-axis pointing along the instantaneous position vector $\overline{r}$, and the z-axis is mutually perpendicular to the x and y axis

and x-y-z form a right handed coordinate frame. Thus, in Figure 2 the z-axis is perpendicular to the plane of paper and points outwards.

Now we use (2.6) to develop the equations of relative motion. Note that the dynamics of leader and follower satellites are governed by

$$\ddot{\overline{r}} + \frac{\mu}{r^3}\overline{r} = \overline{F}_l, \tag{2.9}$$

and

$$\ddot{\overline{r}} + \ddot{\overline{\rho}} + \frac{\mu}{|\overline{r} + \overline{\rho}|^3}(\overline{r} + \overline{\rho}) = \overline{F}_f, \tag{2.10}$$

respectively, where $\overline{F}_l$ and $\overline{F}_f$ denote specific external disturbances and/or specific control forces (per unit mass) acting on the leader and follower satellites, respectively. The dynamic equations governing the motion of follower satellite relative to the leader satellite can be obtained by computing (2.10) − (2.9) which yields

$$\ddot{\overline{\rho}} + \frac{\mu}{|\overline{r} + \overline{\rho}|^3}(\overline{r} + \overline{\rho}) - \frac{\mu}{r^3}\overline{r} = \overline{F}, \tag{2.11}$$

where $\overline{F} \triangleq \overline{F}_f - \overline{F}_l$ is the resultant specific external disturbance and/or specific control force acting on the relative motion dynamics.

Next, the relative position vector $\overline{\rho}$ expressed in the moving reference x-y-z attached to the leader satellite is given by

$$\overline{\rho} = x\hat{\imath} + y\hat{\jmath} + z\hat{k}, \tag{2.12}$$

and the angular velocity vector $\overline{\omega}$ is given by

$$\overline{\omega} = \omega\hat{k}. \tag{2.13}$$

In this case, the relative acceleration vector is given by

$$\ddot{\overline{\rho}} = (\ddot{x} - 2\omega\dot{y} - \omega^2 x)\hat{\imath} + (\ddot{y} + 2\omega\dot{x} - \omega^2 y)\hat{\jmath} + \ddot{z}\hat{k}. \tag{2.14}$$

Furthermore, after some algebraic manipulations and using (2.7) it follows that

$$\begin{aligned}\frac{\mu}{|\overline{r}+\overline{\rho}|^3}(\overline{r}+\overline{\rho})-\frac{\mu}{r^3}\overline{r} &= \frac{\mu}{[(\overline{r}+\overline{\rho})\cdot(\overline{r}+\overline{\rho})]^{3/2}}(\overline{r}+\overline{\rho})-\frac{\mu}{r^3}\overline{r}\\ &= \frac{\mu}{r^3}\left[\left(1+\frac{2}{r^2}\overline{r}\cdot\overline{\rho}+\frac{1}{r^2}\overline{\rho}\cdot\overline{\rho}\right)^{-3/2}(\overline{r}+\overline{\rho})-\overline{r}\right]\\ &= \omega^2\left[\left(1+\frac{2y}{r}+\frac{1}{r^2}(x^2+y^2+z^2)\right)^{-3/2}[x\hat{i}+(y+r)\hat{j}+z\hat{k}]-r\hat{j}\right].\end{aligned} \tag{2.15}$$

Using (2.14) and (2.15) in (2.11), we now obtain the general nonlinear equations of relative satellite position dynamics

$$\ddot{x}-2\omega\dot{y}-\omega^2x+\omega^2g(x,y,z,r)x = F_x, \tag{2.16}$$

$$\ddot{y}+2\omega\dot{x}-\omega^2y+\omega^2\left[g(x,y,z,r)(y+r)-r\right] = F_y, \tag{2.17}$$

$$\ddot{z}+\omega^2g(x,y,z,r)z = F_z, \tag{2.18}$$

where F_x, F_y, and F_z are components of the resultant specific external disturbance and/or specific control force F along the x, y, and z axis, respectively, and

$$g(x,y,z,r) \triangleq \left[1+\frac{2y}{r}+\frac{1}{r^2}(x^2+y^2+z^2)\right]^{-3/2}.$$

Note that the relative satellite position dynamics (2.16)–(2.18) are nonlinear in general. However, in most MSFF missions, relative position components $x, y, z << r$. Thus, in this case, the relative satellite position dynamics can be simplified using a series expansion for $g(x,y,z,r)$ and truncating the higher-order terms. Specifically, a second-order and a linear approximation for (2.16)–(2.18) are given by

$$\ddot{x}-2\omega\dot{y}-\frac{3\omega^2}{r}xy = F_x, \tag{2.19}$$

$$\ddot{y}+2\omega\dot{x}-3\omega^2y+\frac{3\omega^2}{r}y^2-\frac{3\omega^2}{2r}(x^2+z^2) = F_y, \tag{2.20}$$

$$\ddot{z}+\omega^2z-\frac{3\omega^2}{r}yz = F_z, \tag{2.21}$$

and

$$\ddot{x}-2\omega\dot{y} = F_x, \tag{2.22}$$

$$\ddot{y} + 2\omega\dot{x} - 3\omega^2 y = F_y, \tag{2.23}$$

$$\ddot{z} + \omega^2 z = F_z, \tag{2.24}$$

respectively. Note that (2.22)–(2.24) are known as the Clohessy-Wiltshire equations and were originally derived in the context of satellite rendezvous problem [6].

We conclude this section by noting that even though the dynamic equations for relative attitude alignment are not discussed in this report the reader may consult [15] for an excellent treatment of the subject.

3. Analysis of the Relative Satellite Position Dynamics

In this section we briefly examine the behavior of the relative satellite position dynamics in the absence of any control action. Specifically, we analyze the open-loop response of the relative satellite position dynamics to non-zero initial relative position and velocity. In addition, we analyze the open-loop response of the relative satellite position dynamics in the face of a periodic solar pressure disturbance [14]. Note that for illustrative purposes we compute the open-loop system response using the fully nonlinear, the second-order, and the linear models.

Consider a leader-follower satellite pair to be in the geosynchronous orbit of radius $r = 42241$ km. Let the orbital period be 24 hours, thus $\omega = 7.2722 \times 10^{-5}$ radians/sec. Figure 3 illustrates the open-loop x and y follower position responses relative to the leader satellite with $x(0) = 1000$, $y(0) = -2000$, $z(0) = 0$ m, $\dot{x}(0) = 1$, and $\dot{y}(0) = \dot{z}(0) = 0$ m/sec. Although, in Figure 3 the responses for fully nonlinear (NL), second-order (QA), and linear (LN) models appear to be identical, it can be seen from Figure 4 that the linear model response is different from the fully nonlinear model response.

Next, we illustrate the response of the above leader-follower satellite pair to zero initial condition and a periodic solar pressure disturbance. The model of the specific solar pressure disturbance used is adopted from [14] and is given by

$$F_{f_r} = 4.66 \times 10^{-8} \cos\theta \quad m/s^2,$$

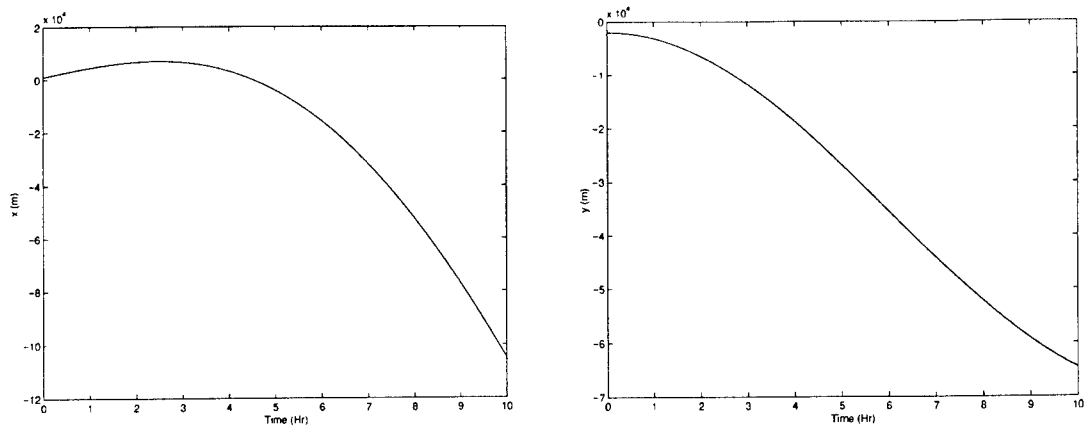


Figure 3: Relative Satellite Position Response to Nonzero Initial Condition

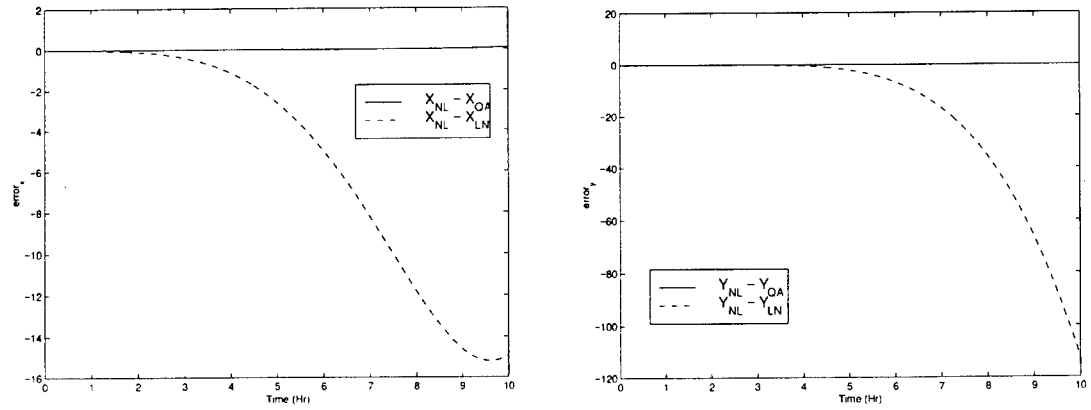


Figure 4: Error Analysis of the Quadratic and Linear Models *vis-a-vis* the Nonlinear Model

$$
\begin{aligned}
F_{f_y} &= -4.66 \times 10^{-8} \sin\theta \quad m/s^2, \\
F_{f_z} &= -3.7 \times 10^{-8} \quad m/s^2, \\
F_{l_x} &= 7.16 \times 10^{-8} \cos\theta - 2.79 \times 10^{-9} \cos 3\theta - 1.12 \times 10^{-9} \cos 5\theta \\
&\quad +4.43 \times 10^{-9} \sin 2\theta \quad m/s^2, \\
F_{l_y} &= 4.43 \times 10^{-9} - 4.43 \times 10^{-9} \cos 2\theta - 1.99 \times 10^{-7} \sin\theta \\
&\quad +2.79 \times 10^{-8} \sin 3\theta + 4.09 \times 10^{-9} \sin 5\theta \quad m/s^2, \\
F_{l_z} &= 2.19 \times 10^{-8} + 1.68 \times 10^{-9} \cos 2\theta + 7.41 \times 10^{-10} \cos 4\theta \\
&\quad -3.84 \times 10^{-9} \sin\theta \quad m/s^2,
\end{aligned}
$$

where $\theta = \omega t$. Note that the solar pressure disturbance models for the follower and the leader

satellite are different since these were obtained for the follower and the leader satellites weighing 410 and 1550 kgs, respectively. See [14] for further details. Figure 5 illustrates the response of the relative satellite position dynamics in the phase plane (x versus y plot) in addition to the system time responses. It is clear from the responses in Figure 5 that the open-loop relative satellite position dynamics is unstable. Thus, we conclude that the two satellites starting at an identical location in the orbit would drift with the passage of time due to differential solar pressure disturbances.

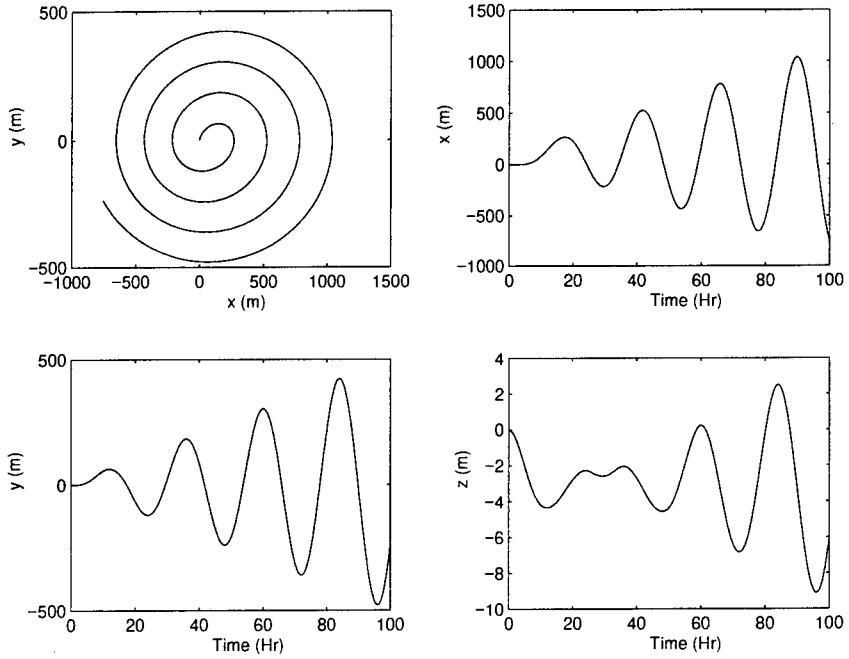


Figure 5: Relative Satellite Position Response to Periodic Differential Solar Pressure

4. Discrete-Time Relative Satellite Position Modeling

In MSFF the principal objective is to control the relative satellite position while minimizing the fuel consumption. Thus, in these applications, it is typical to sample the system at slow rates and apply control inputs only at the sampling instant, i.e., the control inputs are applied at the sampling instant impulsively and are not held constant over the entire sample duration. This requires sampled-data modeling of relative satellite position dynamics model (2.22)–(2.24). Thus, in this section we obtain a discrete-time state-space model of the linearized relative satellite

position dynamics model (2.22)–(2.24). In order to do so, we begin with a continuous-time state-space description of (2.22)–(2.24). Thus, define the state variables $x_1 \triangleq x$, $x_2 \triangleq \dot{x}$, $x_3 \triangleq y$, $x_4 \triangleq \dot{y}$, $x_5 \triangleq z$, and $x_6 \triangleq \dot{z}$. Using these state variables and observing the fact that the in-plane (x-y) dynamics are decoupled from the out-of-plane (z) dynamics, we obtain the state-space descriptions for the in-plane dynamics

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 3\omega^2 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_x(t) \\ u_y(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_x(t) \\ w_y(t) \end{bmatrix}, \tag{4.1}$$

and the out-of-plane dynamics

$$\begin{bmatrix} \dot{x}_5(t) \\ \dot{x}_6(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \begin{bmatrix} x_5(t) \\ x_6(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} u_z(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_z(t), \tag{4.2}$$

where u_i and w_i for $i = x, y, z$ denote the control and disturbance inputs, respectively.

Next, recall that in the typical sampled-data modeling a sample-and-hold device is used to hold the input variable constant over the entire sampling interval [9]. However, as discussed in the beginning of this section, in MSFF the control inputs are applied at the sampling instant impulsively and are not held constant over the entire sample duration. Thus, for the impulsive control design purpose we ignore the inputs u_i and w_i for $i = x, y, z$. Furthermore, we define the notation A_1 and A_2 to denote the in-plane and out-of-plane system dynamics, i.e.,

$$A_1 \triangleq \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 3\omega^2 & 0 \end{bmatrix}, \qquad A_2 \triangleq \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix}.$$

Next, let T_1 and T_2 denote the in-plane and out-of-plane sampling times and define

$$A_{1d} \triangleq e^{A_1 T_1}, \qquad A_{2d} \triangleq e^{A_2 T_2},$$

$$B_{1d} \triangleq \begin{bmatrix} A_{1d}(:,2) & A_{1d}(:,4) \end{bmatrix}, \qquad B_{2d} \triangleq A_{2d}(:,2),$$

where $Z(:,i)$ denotes the i^{th} column of Z. Furthermore, let ΔV_i, $i = x, y, z$, denote the impulsive velocity inputs applied at the sampling instants. With the above notation the combined in-plane

and out-of-plane discrete-time relative satellite position dynamics are given by

$$x(k+1) = \begin{bmatrix} A_{1d} & 0 \\ 0 & A_{2d} \end{bmatrix} x(k) + \begin{bmatrix} B_{1d} & 0 \\ 0 & B_{2d} \end{bmatrix} \begin{bmatrix} \Delta V_x \\ \Delta V_y \\ \Delta V_z \end{bmatrix}, \tag{4.3}$$

where $x(k) = \begin{bmatrix} x_1(k) & x_2(k) & x_3(k) & x_4(k) & x_5(k) & x_6(k) \end{bmatrix}^T$. Note that it is now clear from the definitions of B_{1d} and B_{2d} and (4.3) that the control input to the discretized system (4.3) consists of velocity impulses along the x-y-z directions.

5. Linear Quadratic Control Design

In this section we use the linear quadratic control design methodology to stabilize the in-plane and out-of-the plane relative satellite position dynamics. Specifically, we use the linear quadratic full-state feedback and output feedback techniques for the above control design problems. Hence, consider, the problem of designing

$$\begin{bmatrix} \Delta V_x(k) \\ \Delta V_y(k) \end{bmatrix} = K_{ip} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix}, \tag{5.1}$$

$$\Delta V_z(k) = K_{op} \begin{bmatrix} x_5(k) \\ x_6(k) \end{bmatrix}, \tag{5.2}$$

such that the system (4.3) is asymptotically stable and the following performance indices are minimized

$$J_1 = \sum_{k=0}^{\infty} x_1^2(k) + \frac{1}{\omega^2}[\Delta V_x^2(k) + \Delta V_y^2(k)],$$

$$J_2 = \sum_{k=0}^{\infty} x_5^2(k) + \frac{1}{\omega^2}\Delta V_z^2(k).$$

As in [14], in this research, the above linear quadratic full-state feedback control design problem was solved by sampling the in-plane and out-of-plane dynamics every one hour and four hours, respectively. The control objective is to regulate the relative x position component to 100 meters, relative y and z position components to zero, and relative x, y, and z velocities to zero. The system response to nonzero initial condition $x(0) = \begin{bmatrix} 200 & 0 & -100 & 0_{3\times 1} \end{bmatrix}^T$ and periodic solar pressure

differential described in Section 3 is given in Figure 6. The resultant solar pressure differential acting on the relative satellite position dynamics is also shown in Figure 6. Unlike the unstable responses in Figures 3 and 5, the closed-loop response in Figure 6 is bounded. However, the system response in Figure 6 is quite oscillatory. This is due to the fact that the system is subjected to periodic solar pressure differential and, under the slow sampling rate considered, the linear quadratic full-state feedback controller is unable to reject the periodic disturbance. We note that the controlled response was simulated for both the nonlinear model (2.16)–(2.18) and the linear model (2.22)–(2.24). Note that unlike the open-loop case (see Figure 4), the two models in the face of control action did not exhibit any significant differences in their responses.

Next, we design a full-state feedback controller by sampling the in-plane and out-of-plane dynamics every 15 minutes and 1 hour, respectively. A comparison of the regulated x position response for the two sampling schemes is given in Figure 7. Note that even though the fast sampling scheme improves the relative positioning accuracy the oscillatory behavior due to periodic solar pressure differential is still manifested, albeit, at a smaller scale. Furthermore, this improved accuracy is obtained at a cost of increased fuel consumption. Specifically, note that for a 1000 hour simulation the second sampling scheme led to a 20% increase in the ΔV requirement.

Finally, we consider the output feedback control synthesis problem for the restricted in-plane dynamics only. Specifically, we assume that only the tangential and radial relative position measurements are available for feedback. . Furthermore, we assume zero mean unit intensity uncorrelated plant and measurement noise. Now, we design a continuous time Kalman filter to obtain filtered estimates of the states $x_i(t)$, $i = 1, \ldots, 4$. The estimated states are used for feedback control purpose within a linear quadratic control framework described above. Next, let $x_1(0) = 200$, $x_2(0) = -100$, $\dot{x}_1(0) = \dot{x}_2(0) = 0$ and consider the periodic differential solar pressure disturbance model of Section 3. Figure 8 compares the regulation response of the tangential and radial relative positions under the full-state feedback regulator and the Kalman filter plus linear quadratic design described above. Clearly, the output feedback controller response is satisfactory. Note that the above design methodology can be extended to the case where sensor measurements consists of only relative velocities.

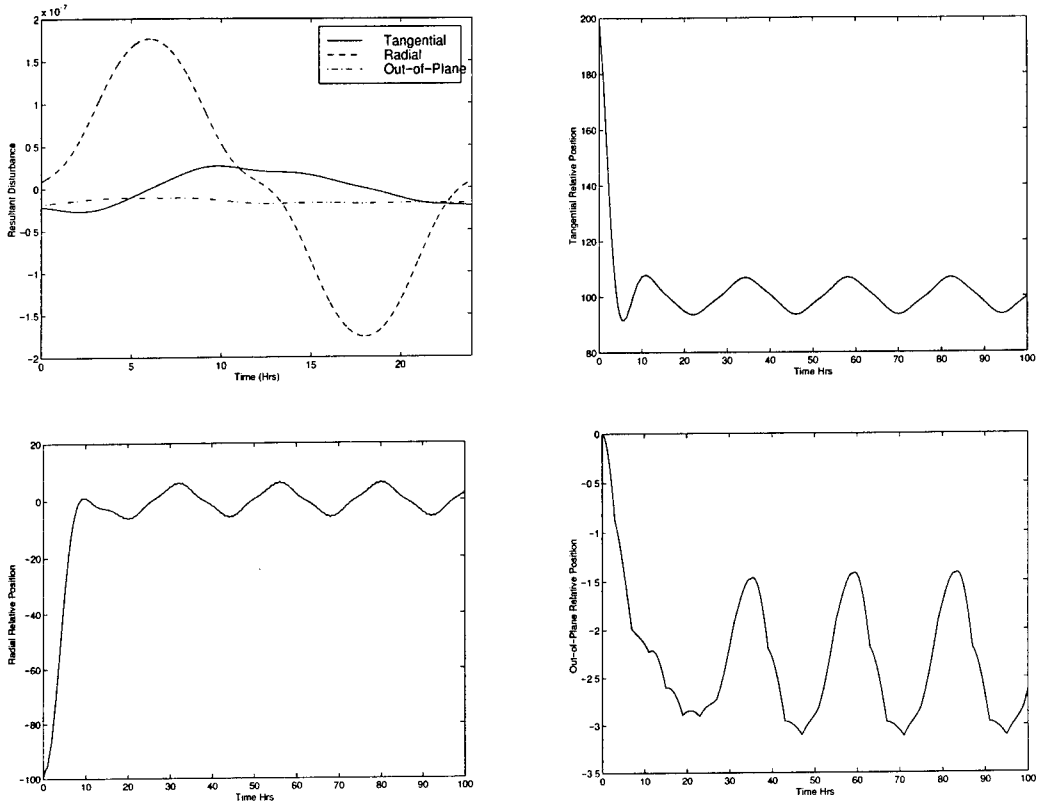


Figure 6: Differential Solar Pressure Disturbance and Closed-Loop Relative Position Response

Next, we point the reader to some competing control design methods for the relative satellite positioning problem. First, we note that the linear quadratic control design framework discussed above can be extended to non-zero set point tracking control using a combination of feedback and feedforward control techniques [12]. Furthermore, alternative control schemes using on-off phase-plane controller can provide ease of implementation. One such algorithm using differential drag elements for actuation is reported in [11]. Recent advances in the phase-plane based control design are reported in [13]. Finally, note that a Lyapunov-based MSFF control design framework has been developed in [15] which considers relative attitude alignment and translational motion control.

6. Concluding Remarks

In this research, we focused on the study of the relative satellite position dynamics and linear quadratic control design. Although linear control theory has been successfully applied to the MSFF

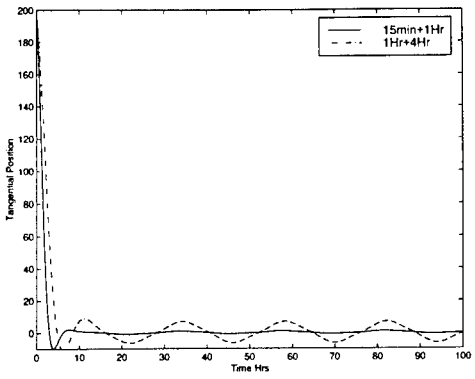


Figure 7: Closed-Loop Response Comparison

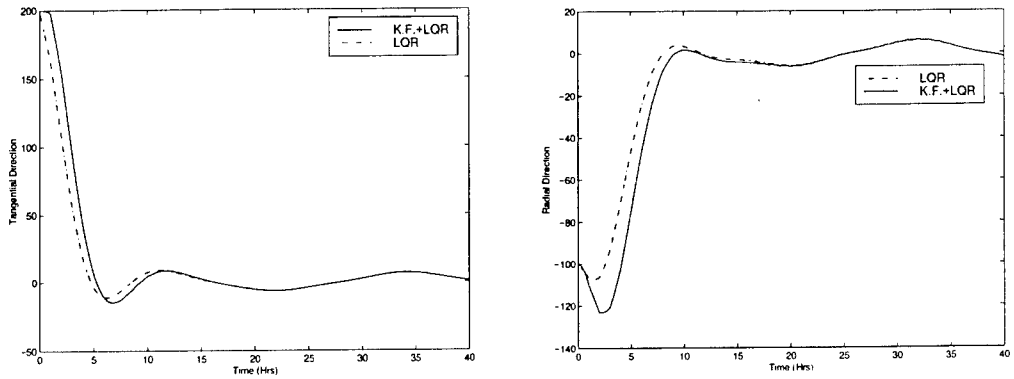


Figure 8: LQR and Kalman Filter plus LQ Control Comparison

problem [14], issues such as parametric variation, external disturbances, nonlinear effects, active collision avoidance, etc. have largely been neglected so far. One exception to this is [15] where a persistent disturbance is employed to account for model imperfection, measurement inaccuracies, etc. Another key issue in MSFF is the tradeoff between fuel consumption and achievable relative positioning accuracy. Furthermore, hardware issues such as reliable actuation and sensing technologies also remain less developed. Recent exception to this includes application of global positioning system (GPS) as an enabling sensor technology for MSFF [7, 13]. Further research is needed to address these critical issues in MSFF.

References

1. Available on the Web at **http://www.vs.afrl.af.mil/factsheets/TechSat21.html**.

2. Available on the Web at **http://origins.jpl.nasa.gov/missions/overview.html**.

3. "Shuttle 7 to Fly Approaches to Payload," *Aviation Week and Space Technology,* **118**-19, 51–57, 1983.

4. F. Bauer *et al.*, "Satellite Formation Flying using An Innovative Autonomous Control System (AUTOCON) Environment," in *AIAA/AAS Astrodynamics Specialists Conference*, (New Orleans, LA), Aug. 1997.

5. V. A. Chobotov (Ed.) *Orbital Mechanics*, AIAA, Washington, DC, 1996.

6. W. H. Clohessy and R. S. Wiltshire, "Terminal Guidance System for Satellite Rendezvous," *J. Aerospace Science*, **27**, 653–658, 1960.

7. T. Corazzini *et al.*, "GPS Sensing for Spacecraft Formation Flying," *Proc. ION GPS-97 Conf.*, Kansas City, MO, 1997.

8. J. R. Guinn, "Autonomous Navigation for the New Millenium Program Earth Orbiter 1 Mission," *Proc. AIAA GN&C Conf.*, New Orleans, LA, 612–617, 1997.

9. B. C. Kuo, *Digital Control Systems*, Saunders HBJ, New York, NY, 1992.

10. K. Lau *et al.*, "The New Millenium Formation Flying Optical Interferometer," *Proc. AIAA GN&C Conf.*, New Orleans, LA, 650–656, 1997.

11. C. L. Leonard, W. M. Hollister, and E. V. Bergmann, "Orbital Formationkeeping with Differential Drag," *AIAA J. Guid., Contr., and Dynm.*, **12**, 108–113, 1989.

12. D. C. Redding, N. J. Adams, and E. T. Kubiak, "Linear-Quadratic Stationkeeping for the STS Orbiter," *AIAA J. Guid., Contr., and Dynm.*, **12**, 248–255, 1989.

13. A. Robertson, T. Corazzini, and J. P. How, "Formation Sensing and Control Technologies for a Separated Spacecraft Interferometer," *Proc. ACC*, 1574–1579, 1998.

14. R. H. Vassar and R. B. Sherwood, "Formationkeeping for a Pair of Satellites in a Circular Orbit," *AIAA J. Guid., Contr., and Dynm.*, **8**, 235–242, 1985.

15. P. K. C. Wang and F. Y. Hadegh, "Coordination and Control of Multiple Microspacecraft Moving in Formation," *J. of Astronautical Sciences*, **44**, 315–355, 1996.

16. W. E. Wiesel, *Spacecraft Dynamics*, McGraw Hill, New York, NY, 1997.

MICRO-SCALE VISUALIZATION OF THIN MENISCUS AND CAPILLARY PORE FLOWS OF CAPILLARY-DRIVEN HEAT TRANSFER DEVICES

Kenneth D. Kihm (98-0044)
Associate Professor
Department of Mechanical Engineering
Texas A&M University
College Station, TX 77843-3123
Tel. (409) 845-2143 & Fax (409) 862-2418
e-mail: ken-kihm@tamu.edu

Final Report for:
Summer Faculty Research Program
Wright-Patterson Air Force Laboratory

Sponsored by:
Air Force Office of Scientific Research
Wright-Patterson Air Force Base, Dayton, OH

September 1998

MICRO-SCALE VISUALIZATION OF THIN MENISCUS AND CAPILLARY PORE FLOWS OF CAPILLARY-DRIVEN HEAT TRANSFER DEVICES

Kenneth D. Kihm
Associate Professor
Department of Mechanical Engineering
Texas A&M University

Abstract

Thin meniscus and thermocapillary stress-driven flows were examined to improve the performance of capillary-driven heat transfer devices. The goal of the present study was to develop optical diagnostic techniques to measure the thickness and movement of thin meniscus and to visualize the capillary-driven flows inside capillary pores. Fizeau interferometry was implemented to measure the location, thickness and slope of meniscus. The use of electrohydrodynamic (EHD) phoresis enables the extension of meniscus toward the electrode of higher field density. This unique method of microscale control of the thin film region was fully evidenced by the interferometric visualization. Microscopic particle image velocimetry (MPIV) visualizes the thermocapillary stress-driven flows inside capillary pores under heating. Depending on the relative heater location, with respect to the liquid-vapor interface, the circulation direction of the thermocapillary-driven flows was changed.

MICRO-SCALE VISUALIZATION OF THIN MENISCUS AND CAPILLARY PORE FLOWS OF CAPILLARY-DRIVEN HEAT TRANSFER DEVICES

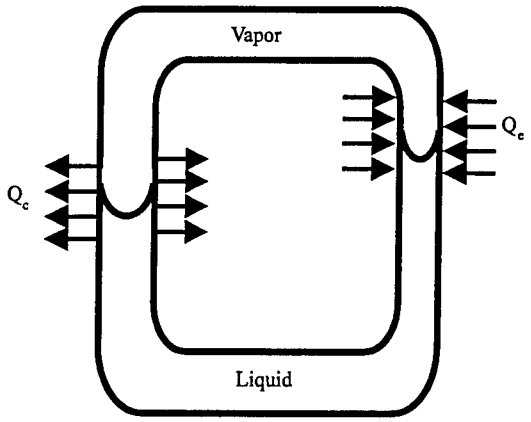
Kenneth D. Kihm

Introduction

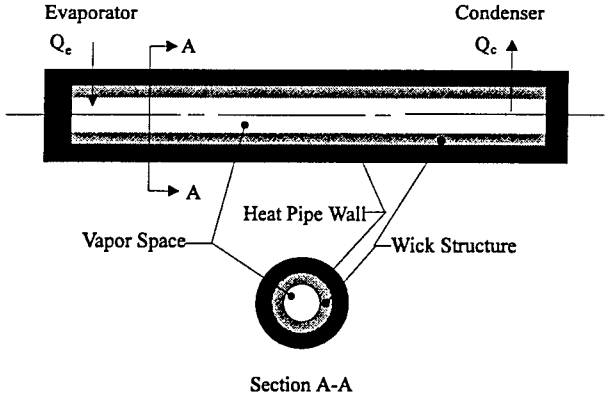
Some commercial and military applications in terrestrial and space thermal management systems require heat transport and dissipation capabilities in the hundreds of watts per square centimeter level. Of the heat transport devices presently under consideration in this regime, most utilize liquid-vapor phase change due to its superior heat transfer characteristics which can translate to a system with a reduced mass and at a near constant temperature. Additionally, if the phase change device is capillary-driven, the device will be passive, requiring no mechanical circulator for the working fluid.

The broad category of capillary-driven heat transfer devices includes capillary pumped loops (CPL) and heat pipes (Chang and Hager, 1990). Figure 1-a shows a simple CPL that is composed of a single pore evaporator and a single pore condenser. In this figure, Q_e and Q_c are the heat transferred into and out of the CPL, respectively. Figure 1-b depicts a traditional capillary-driven heat pipe, which is composed of a tube with its inner wall embedded with a porous substrate known as the wick.

A description of their operation follows. Heat addition to the evaporator results in evaporation from the liquid-saturated porous or grooved substrate. The vapor flows to the condenser where the removal of heat causes the vapor to condense. The liquid is driven back to the evaporator through the wick due to the capillary action. When evaporation exceeds the rate of liquid re-supply, the device is said to reach its capillary limit and fails. Other limitations on the heat carrying capacity of these devices include the sonic limit (choking of the vapor flow), the



(a)



(b)

Fig. 1 Capillary-driven heat transfer devices: (a) capillary pumped loop, and (b) heat pipe

entrainment limit (tearing of the liquid off the liquid-vapor interface by the high velocity flow of the vapor), and the boiling limit (disruption of the liquid flow by nucleate boiling in the wick) (Dunn and Reay, 1994). Of these limits, *the capillary limit* is of primary concern for the present study.

The basic mechanisms describing evaporating menisci in capillary tubes at the capillary limit are not well understood even though capillary-driven heat transfer devices have been studied widely. Most studies have concentrated on the macroscopic observables, such as the overall heat transfer rate or temperature distribution. The capillary limit has been typically used to define the maximum heat transfer capability of capillary-driven heat transfer devices. It presumes that the capillary potential driving the liquid flow from the condenser to the evaporator is maximum, i.e., the liquid floods the condenser and the meniscus curvature is equal to the pore radius in the evaporator. This implies that the liquid is perfectly wetting in the evaporator. In practice, the maximum heat transfer capability associated with the capillary limit is rarely achieved (Richter and Gottschlich, 1994).

The operation can be worse in application to fighter/attack aircraft. Fighter/attack aircraft requirements include up to 10 transient g while under maximum thermal load and the evaporator can be "dried-out" after the meniscus is detached from the pore surfaces. This occurs in rather critical moments, such as during attack maneuvers or "dog fighting." One solution to the problem of such an under-operation is to oversize the capillary driven systems, which opposes the minimum weight/volume design specification for fighter/attack aircraft. So the question arises how to minimize the weight or eliminate the need to oversize and compensate for the high g loads. The solution is conceptually simple, eliminate the deleterious effects of energy and mass transport on the capillary-driven heat transfer device and augment their capillary pumping.

Dynamics associated with fluid motion and heat transport in the vicinity of the evaporating meniscus can detrimentally affect the driving capillary potential by altering the wetting ability of the working fluid (Pratt and Hallinan, 1997). The change in wettability results from non-isothermal interface conditions due to both non-uniform substrate wall temperatures in the vicinity of the meniscus or contact line and non-uniform evaporation. Either or both of these conditions yield surface tension gradients on the liquid-vapor interface since the surface tension increases with decreasing temperature. These surface tension gradients result in thermocapillary stresses acting near the contact line, which degrade the wettability of the liquid and thus the ability of a heated meniscus to transfer heat via evaporation.

The present study focuses on the examination of microscopic observables, namely thin meniscus and microscale flows inside capillary pore, using advanced optical diagnostic techniques. In addition, it explores how to control the meniscus and capillary pore flows to prevent "dry-out" and maximize the operation efficiency of CPL and heat pipes. The ultimate goal of the study is to augment and maximize the macroscopic observable of the heat transfer rate from careful investigation of microscopic observables, such as the meniscus movement and thermocapillary stress driven flows.

Methodology

The methodology used for the study is two-fold: (1) Fizeau interferometry (Hecht, 1987) measures the movement and thickness of thin meniscus, and (2) microscopic particle image velocimetry (MPIV) visualizes the flow patterns inside capillary pore flows. Liquid meniscus extends to the molecular level of an order of sub-nanometer thickness along a clean solid surface (Fig. 2). The adsorbed film layer is too thin and its evaporative flux is often negligible. When the meniscus thickness exceeds a certain limit (bulk meniscus region), the evaporative flux is

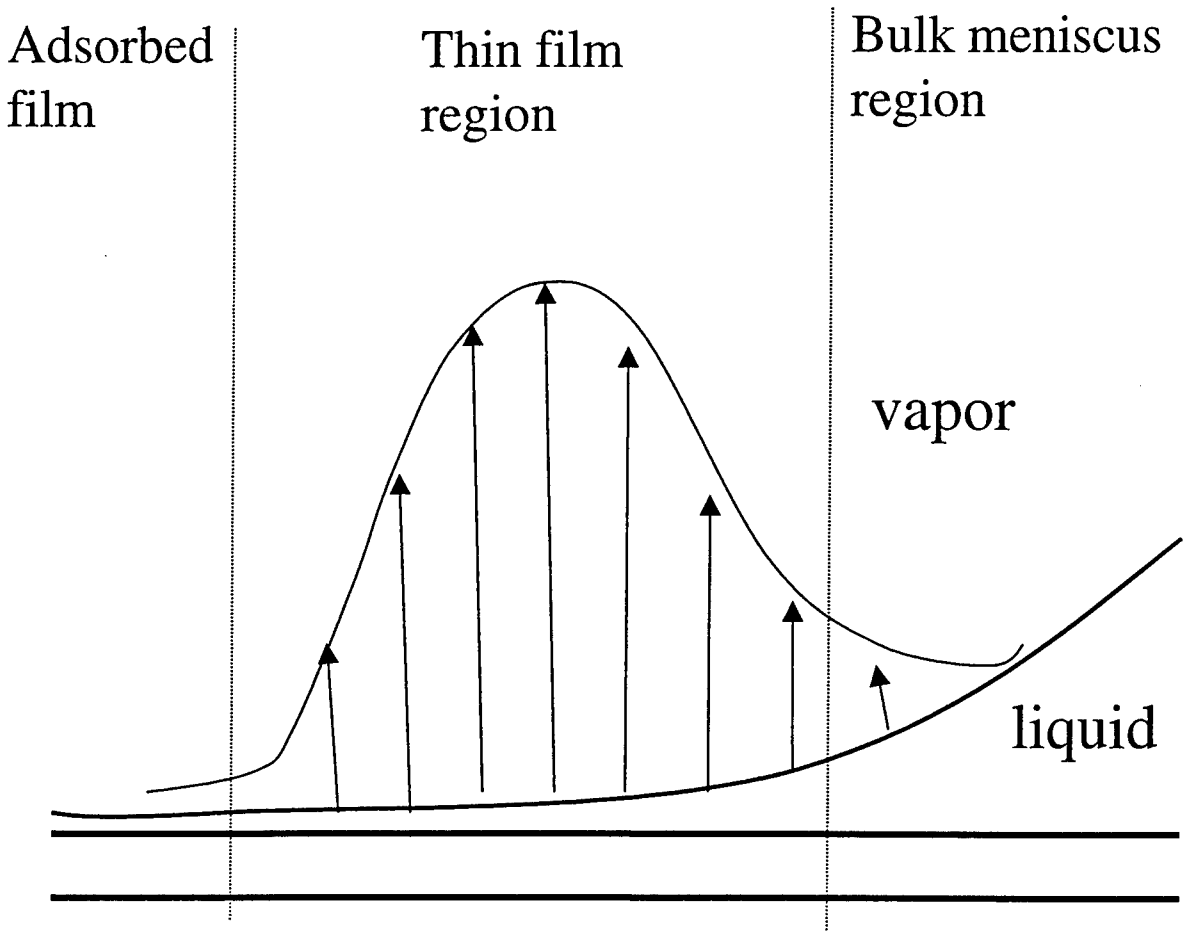


Fig. 2 Development of thin meniscus along a heated solid surface

again negligible since the meniscus surface temperature is not high enough for active evaporation. In between, called a thin film region, a maximum evaporative flux is expected and its thickness ranges from an order of one micron to one hundred microns. The key idea is how to identify the thin film region and possibly extend the region to augment the heat transfer rate of capillary-driven heat transfer devices.

Figure 3 illustrates the Fizeau interferometry that basically consists of a microscopic imaging system with a monochromatic light source. The rectangular test cell is made of two 105 mm square, 1 mm thick glass plates spaced 8 mm apart. The monochromatic light ray partially reflects from the inner surface of the top glass plate and partially reflects from the meniscus surface (see the inset drawing in Fig. 3). These two reflected rays create an optical path length differential that is equivalent to two times the meniscus thickness, $2t$. When the optical path length differential, $2t$ is equivalent to $m\lambda$, the two rays interfere constructively and result in the brightest intensity. When the optical path length differential is equivalent to $(m+1/2)\ \lambda$, the two rays interfere destructively and result in the darkest intensity. The resulting sinusoidal intensity pattern conforms to optical fringes. Thus, any two neighboring brightest (or darkest) locations on the fringes represent thickness differentials of one-half of the wavelength, $\lambda/2$.

The present study introduces the concept of electrohydrodynamic (EHD) phoresis to move and extend the thin film region of the meniscus. The net force that arises due to the action of the electric field can be shown to be (Singh, 1995)

$$F_e = qE + Const_1 \nabla E^2 + Const_2 E^2 \nabla \kappa \qquad (1)$$

where q is the electric field space charge density, E is the applied electric field strength, κ is the dielectric constant, and $Const_1$ and $Const_2$ are given by the electric properties of tested fluid. The first term on the right side of Eq. (1) presents Coulomb force for ionized molecules. In highly

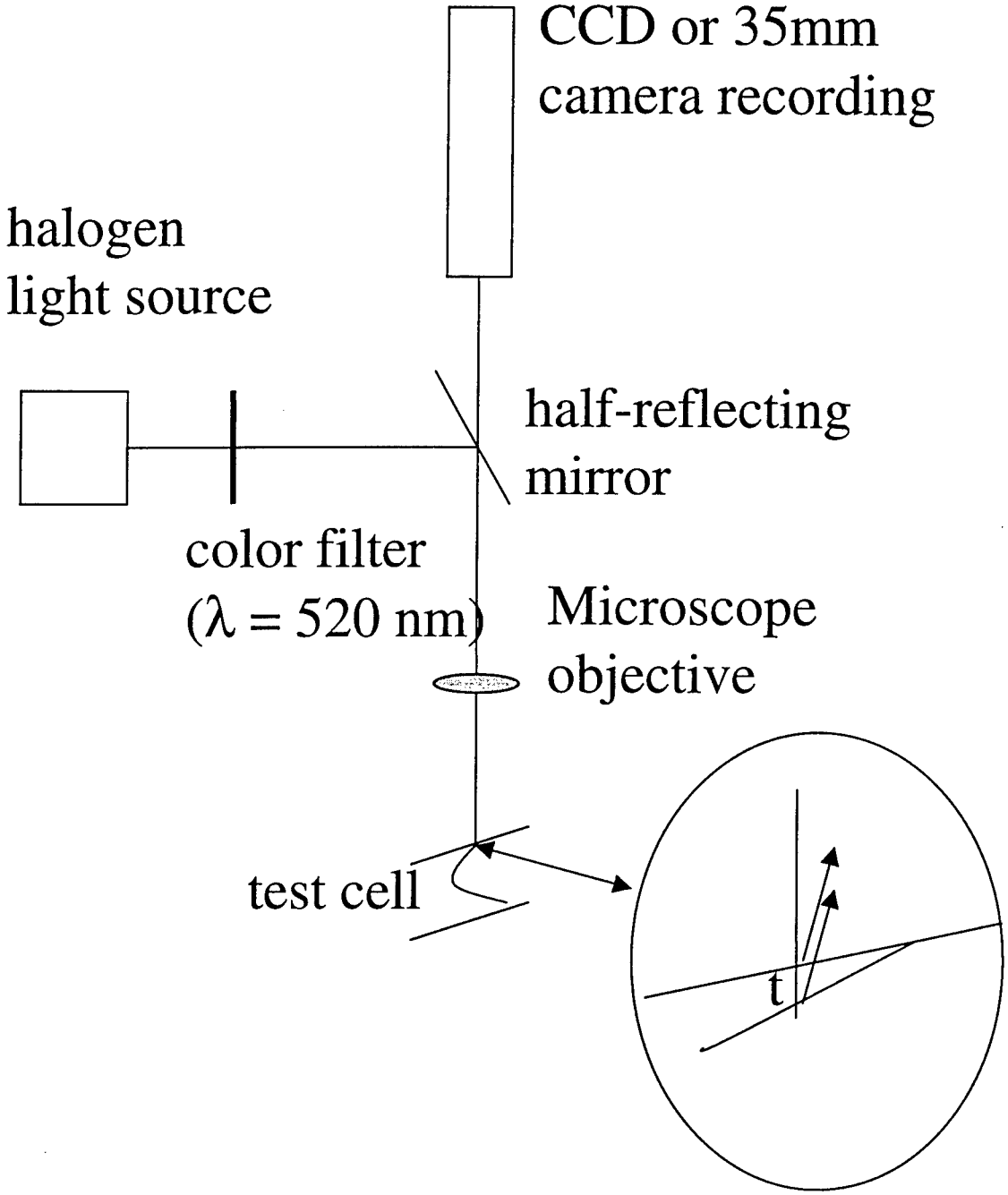


Fig. 3 Fizeau interferometry

dielectric fluids with low dielectric permittivity, the EHD current discharge is negligible and therefore the first term is insignificant. The second and third terms present forces resulting because of nonuniformity of the electric field and nonuniformity in the dielectric permittivity of the fluid, respectively. If there is no spatial variation in the dielectric constant ($\nabla\kappa = 0$), the EHD force will be solely due to nonuniformity in the electric field. Then, the fluid molecules will always move toward the region of the higher electric field. This phenomenon is called *dielectrophoresis*.

To ensure the phoresis toward the extension of meniscus, the electrode configuration has been devised to create a higher electric field away from the bulk of the fluid as shown in Fig. 4. Since the dielectrophoresis is proportional to the gradient of the scalar amount of E^2, it always directs toward an electrode of the higher E-field regardless of the polarity of the electrodes.

Also, a microscopic particle image velocimetry (MPIV) has been implemented to visualize the thermocapillery stress-driven flows near the liquid-vapor interface inside a capillary pore under heating. The MPIV system (Fig. 5) uses an Ar-ion cw laser sheet of about 100 μm thickness for illumination and the tested fluid of isopropyl (S.G.= 0.85) is seeded by homogeneous polystyrene latex particles (S.G. = 1.05) of 5 μm in diameter. A microscopic CCD camera at 30 fps makes the image recording with high magnifications. The tested capillary pore diameters are 0.5, 1.0 and 5.0 mm.

Results

Pentane (κ= 1.84) was used as the tested fluid for meniscus thickness measurements. Figures 6-a and 6-b show Fizeau fringes taken for meniscus with charging voltage, V_{ch} = 0 and 1.5 kV, respectively. The test cell is under heating at 0.1 watts onto the external glass wall above the liquid-vapor interface. The left-hand side leads to the linear electrode attached to the bulk fluid

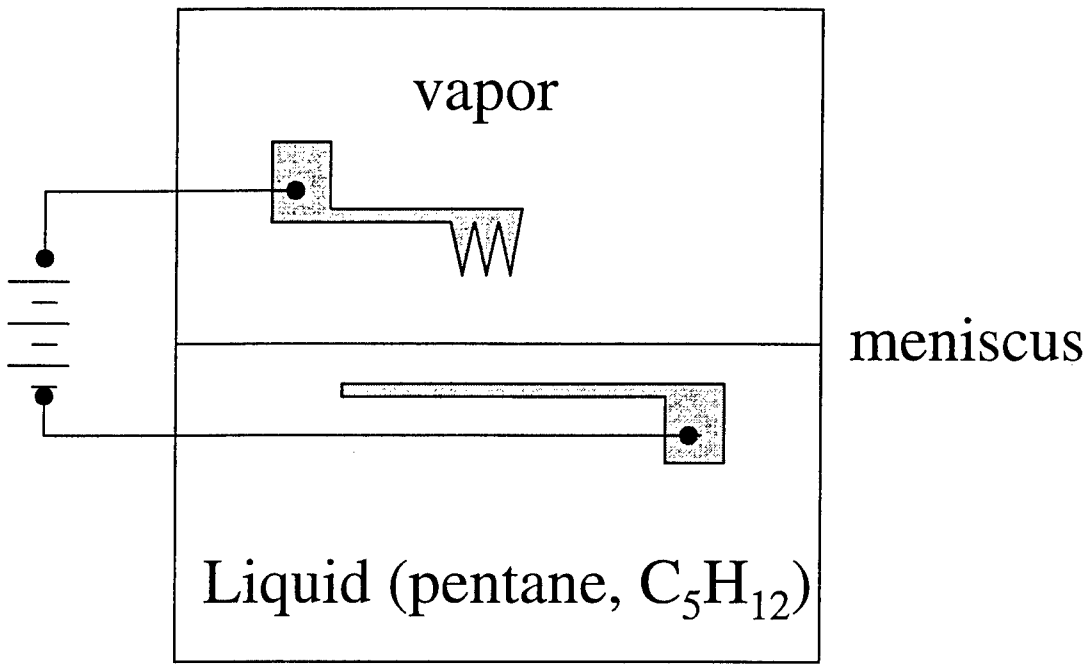
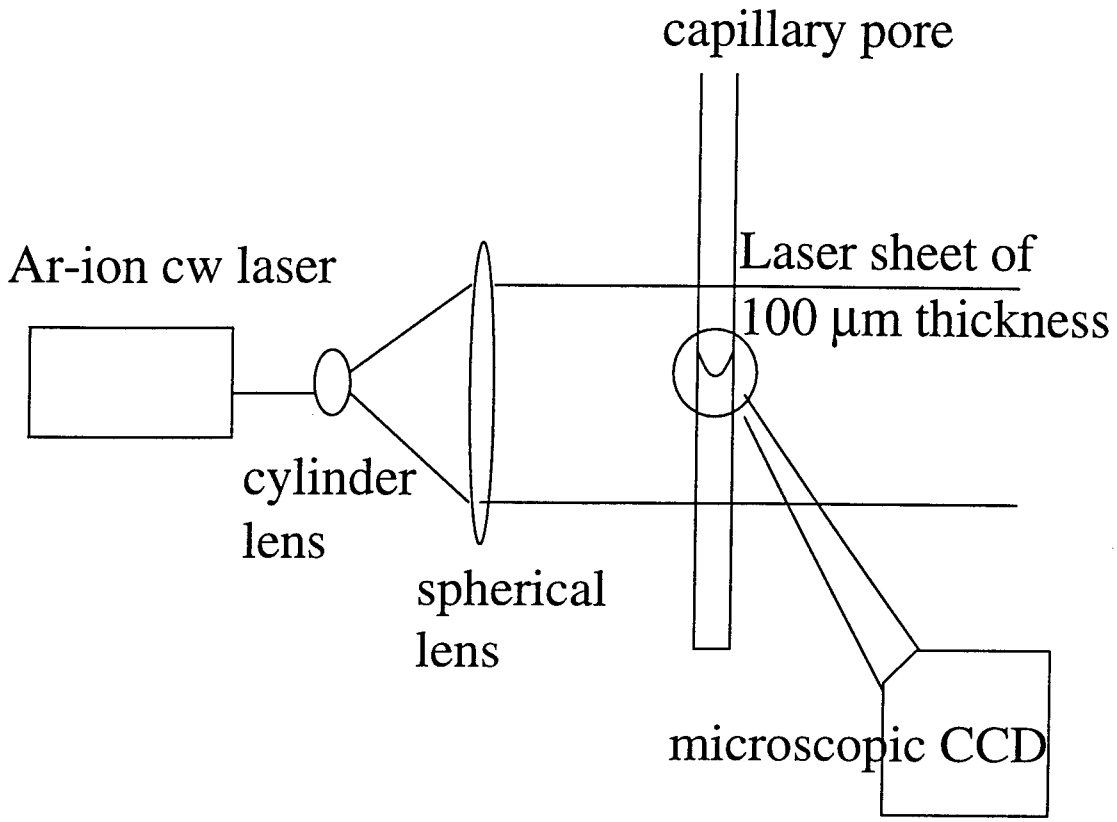
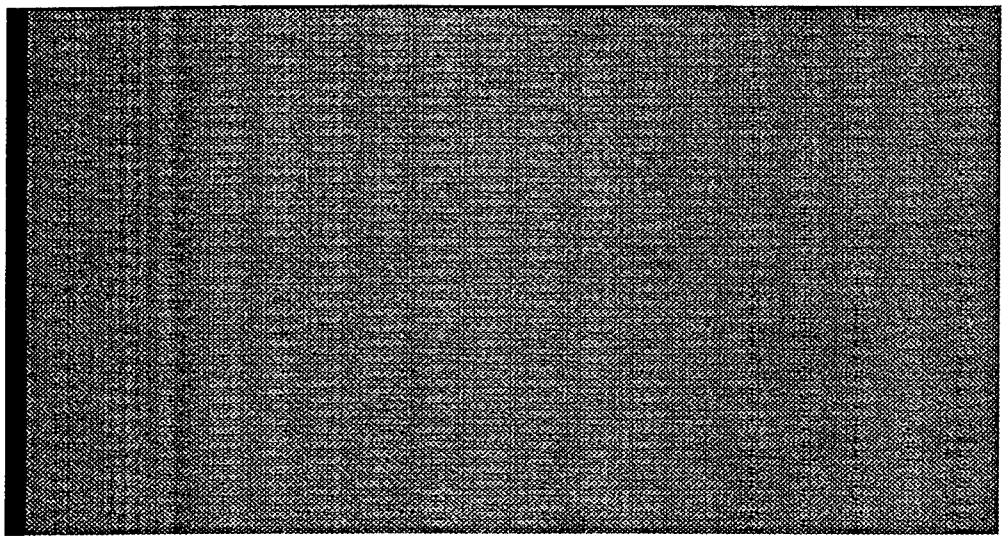


Fig. 4 Test cell with electrode configuration

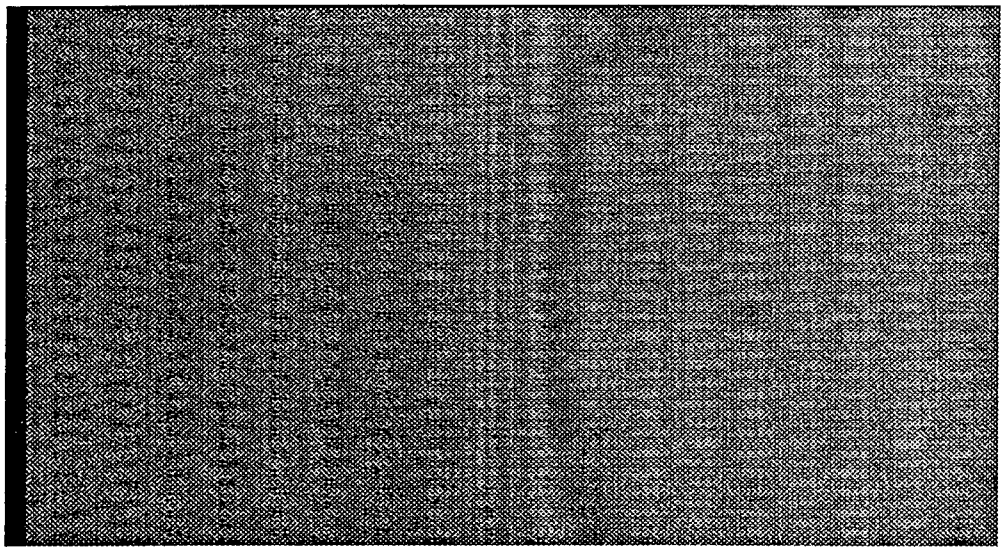


[seeded particles: d = 5μm, ρ = 1.05]
[tested fluid-isopropyl: ρ = 0.85]

Fig. 5 Schematic illustration of microscopic particle image velocimetry (MPIV) system



(a) $V_{ch} = 0$



(b) $V_{ch} = 1.5$ kV

Fig. 6 Fizeau fringes of thin meniscus (a); and their dielectrophoretic movement (b): a foil heater providing 0.1 watts is attached to the right-hand side surface.

and the right-hand side leads to the serrated teeth electrode resulting in a strongly non-uniform field. The first dark fringe indicates the film thickness, $t = \lambda/4$ or 130 nm, and the first bright fringe indicates $t = \lambda/2$ or 260 nm. The second bright band is equivalent to λ or $t = 520$ nm, and so on. Since the neighboring (dark or bright) fringes get narrower as we go to the left, the meniscus slope gradually increases toward the bulk fluid.

Measured film thickness profiles versus applied charging voltage are shown in Fig. 7. The meniscus front moves to the direction of the higher electric field and the movement increases with the amount of charging voltage. The dielectrophoresis is clearly evidenced and an active control of the meniscus is now possible by the use of EHD forces. The thickness measurement is available only up to about 3 μm because of the resolution limitation of the current imaging system. Note that the initial slope of meniscus remains nearly unchanged with increasing charging voltage.

The microscopic particle image velocimetry (MPIV) recording shows the thermocapillary-driven flow patterns for different L/d ratio, where L is the distance from the meniscus to the lower end of the heater and d is the capillary pore diameter (Fig. 8). When the heater is located below the free surface, the flow region surrounded by the heater shows radially inward circulations due to the natural convection heating along the inner surface of the pore.

The flow above the heater region is induced to counter-rotating circulations and the heated liquid rises along the pore center. The warm liquid first reaches the center of the interface and the liquid temperature at the center is higher than the thin film region near the pore wall. This temperature differential drives thermocapillary stress increasing from the center of the interface to the thin film meniscus as the liquid surface tension increases with decreasing liquid temperature. The higher thermocapillary stress near the film region tends to pull the liquid from the center and the resulting flow pattern will be blooming radially outward near the interface.

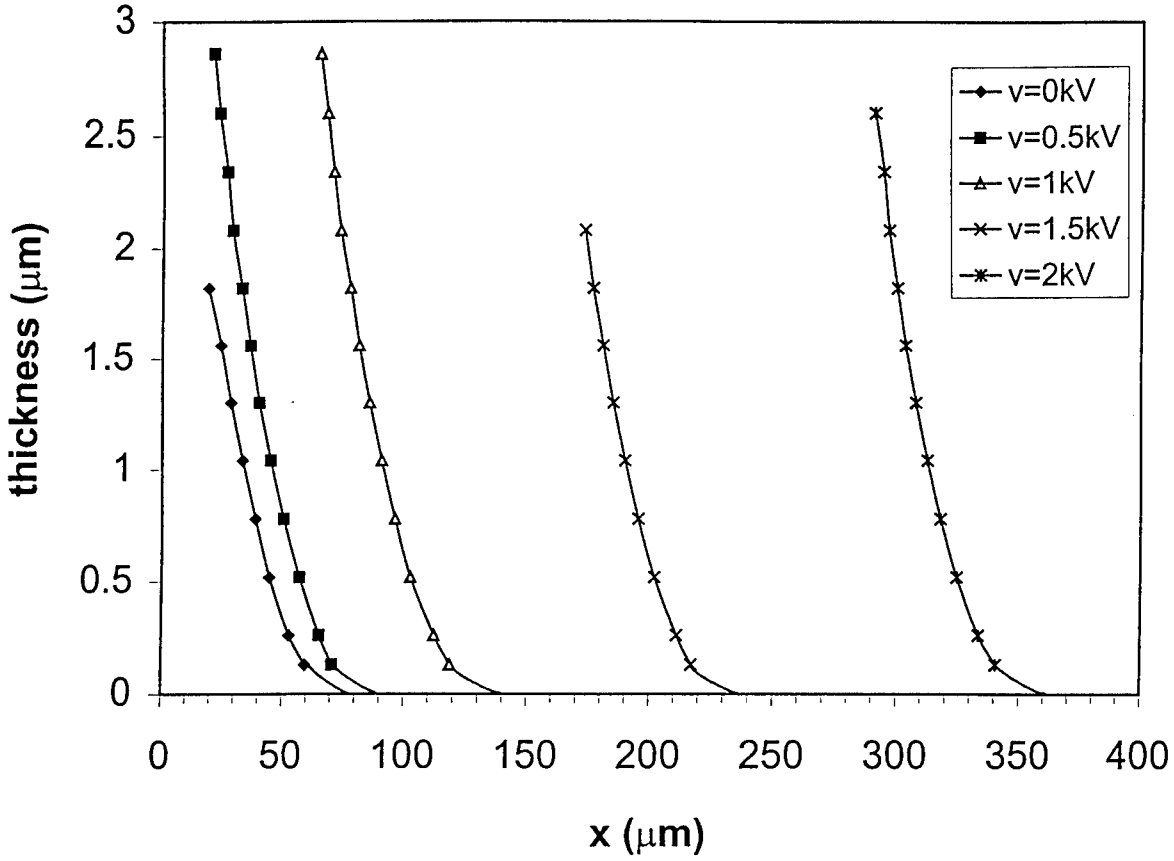


Fig. 7 Meniscus thickness versus displacement for different charging voltages.

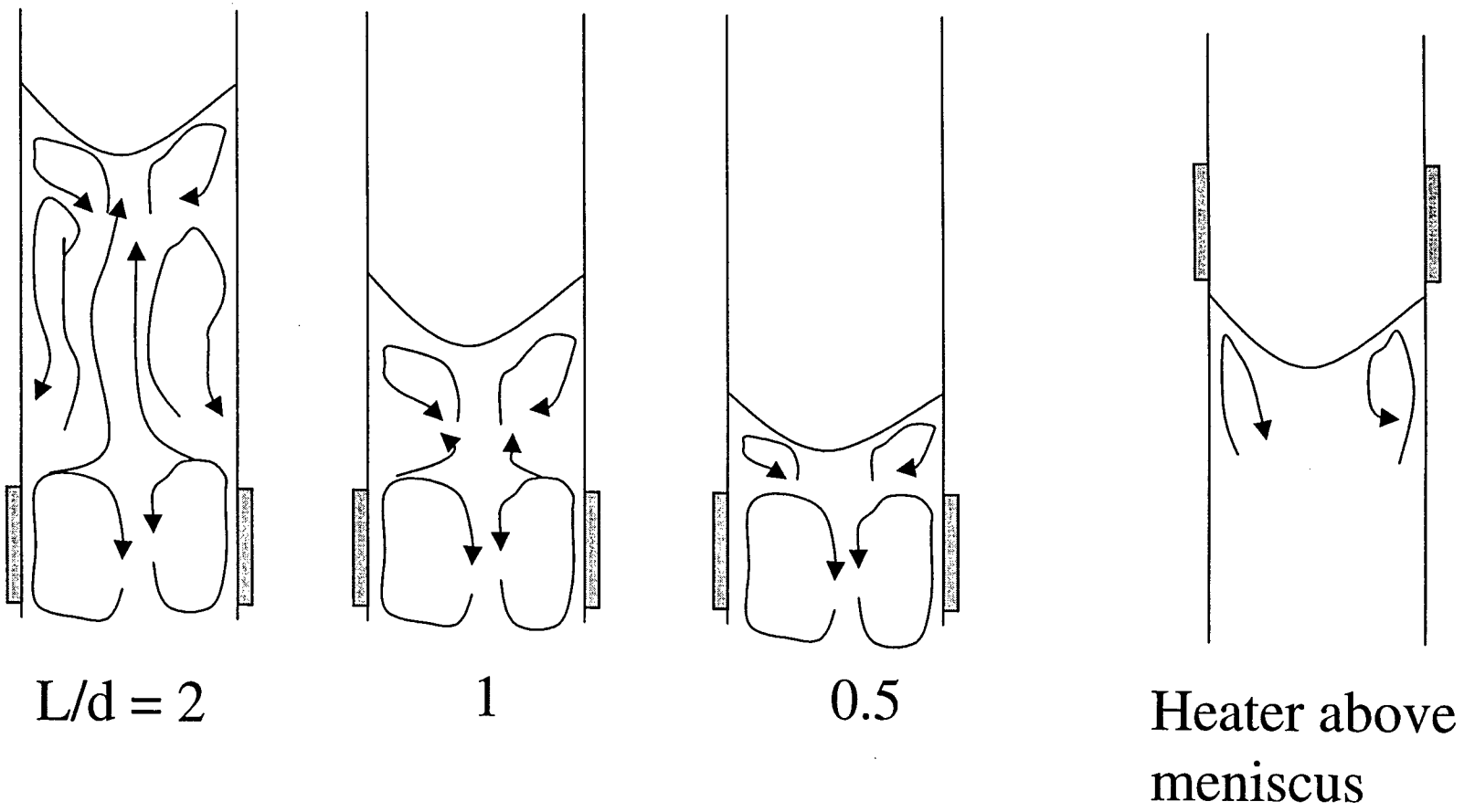


Fig. 8 Observation of thermocapillary-driven flow patterns from MPIV images.
(L: distance from the meniscus to the bottom of the heater; d: capillary pore diameter)

When the heater is located above the free surface (the last schematic in Fig. 8), however, the direction of thermocapillary stress is reversed since the liquid in the thin film region is warmer than the liquid at the center. Thus, the liquid is pulled to the center from the thin film region and the direction of flow circulation near the interface is reversed.

The direction of the flow circulation, associated with the heater location, is expected to have an important role in the heat transfer effectiveness of the thin film region of thermocapillary-driven heat exchanger devices. This also controls the wettability and the dimension of the thin film region where active evaporative flux exists. Further detailed study will be necessary including accurate heat transfer measurements associated with different capillary stress-driven flows.

Conclusion

Noble diagnostic methods have been developed and used to visualize thin meniscus and microscale capillary pore flows for capillary-driven heat transfer devices. The ultimate goal of the study is to seek a way to control the microscale phenomena and eventually augment the macroscopic heat transfer rate of capillary-driven heat transfer devices, namely, capillary pumped loops (CPL) or heat pipes.

Some fundamental conclusions reached from the present investigation are:

1. Fizeau interferometry system has been successfully developed to measure the thickness of thin liquid film to the resolution of $\lambda/4$.
2. Active control of the meniscus front and extension of the thin film region have been achieved by the use of electrohydrodynamic (EHD) phoresis with a specially designed electrode configuration.

3. Microscale particle image velocimetry (MPIV) system has been developed to visualize the thermocapillary stress-driven flows inside a capillary pore, whose diameter is as small as 0.5 mm.
4. The thermocapillary-driven flow reverses its rotational direction near the liquid-vapor interface depending on the heater location, with respect to the interface.

References

[1] Chang, W.S. and Hager, B.G., 1990, Advanced Two-Phase Thermal Management in Space, National Heat Transfer Conference, Minneapolis, MN.

[2] Dunn, R.D. and Reay, D.A., 1994, Heat Pipes, Pergamon.

[3] Hecht, E., 1987, Optics (2nd ed.), Addison-Wesley.

[4] Pratt, D.M. and Hallinan, K.P., 1997, Thermocapillary Effects on the Wetting Characteristics of a Heated Curved Meniscus, Journal of Thermophysics and Heat Transfer, Vol. 11, No. 4, pp. 519-525.

[5] Richter, R. and Gottschlich, J.M., 1994, Thermodynamic Aspects of Heat Pipe Operation, Journal of Thermophysics and Heat Transfer, Vol. 8, No. 2, pp. 334-340.

[6] Singh, A., 1995, Eletrohydrodynamic (EHD) Enhancement of In-Tube Boiling and Condensation of Alternate (Non-CFC) Refrigerants, Ph.D. Dissertation, University of Maryland.