



Design of an Optical Fabrication System for Nano-technology

Freddy Bruce, Christ Church

Jamieson Brynes, Somerville

James Chapman, Jesus

Lewis Fish, Lincoln

Adam Gibbons, St. Hugh's

Samuel Holt, Pembroke

Jin Jia, St. Hilda's

Xiaonan Lu, St. John's

Thomas Pretty, St. Edmund Hall

Shashwat Saxena, Oriel

Aditi Srivastava, Exeter

Mark Uckermann, Christ Church

May 11, 2016

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____

Date: _____

**FINAL HONOUR SCHOOL OF
ENG / EEM** (delete as appropriate)

DECLARATION OF AUTHORSHIP

You should complete this certificate. It should be bound into your third year project report, immediately after your title page. Three copies of the report should be submitted to the Chairman of examiners for your Honour School, c/o Clerk of the Schools, examination Schools, High Street, Oxford.

Name(in capitals): _____

College (in capitals): _____ **Supervisor:** _____

Title of Project (in capitals): _____

**Page count
(excluding risk and COSHH assessments):** _____

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*Essential Information for Students. The Proctors' and Assessor's Memorandum*, Section 9.6; also available at <http://www.admin.ox.ac.uk/proctors/info/pam/8examinations/>)

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at <http://www.admin.ox.ac.uk/edc/goodpractice/about/>

The project report I am submitting is entirely my own work except where otherwise indicated.

It has not been submitted, either partially or in full, for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution.

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs.

I have clearly indicated the presence of all paraphrased material with appropriate references.

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor.

I have not copied from the work of any other candidate.

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: <http://www.admin.ox.ac.uk/statutes/352-051a.shtml>)

The project report does not exceed 30 pages (including all diagrams, photographs, references and appendices).

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted.

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism.

Candidate's signature: _____ Date: _____

Abstract

We designed a general purpose optical nano-fabrication facility, which can prototype structures with nanometre features, and has a range of nanometre imaging systems, to evaluate sub-micron structures. The facility is designed and built for a variety of applications, which comes as a complete product, with detailed optical layout, all mechanical components appropriately housed, intelligent control, automatic automation and a bespoke software suite. The report will detail the design process, explaining the engineering theories, applying them to the system and the specifications of the designed system.

Contents

1	Preamble	0
1	Introduction	1
1.1	Overview	Freddy 1
1.2	Brief	Tom 1
2	Processes	2
2.1	Laser Micro-machining	Adam 2
2.2	Stereolithography	Jamie 2
2.3	Two-Photon Polymerisation	Sam 2
2.4	Confocal Microscopy	Xiaonan 2
2.5	Two-Photon Microscopy	Mark 2
2.6	Stimulated Emission Depletion (STED)	James 3
2.7	Optical Tweezing	Mark 3
3	Background Info	4
3.1	Magnification and Microscopes	Mark 4
3.2	Fourier Optics	Freddy 5
3.2.1	Diffraction	6
3.2.2	Fresnel and Fraunhofer Approximations	7
3.2.3	Fourier transforming properties of a lens	8
3.3	Point Spread Function	Freddy 11
3.3.1	Mathematical description of the Airy Disk	11
3.3.2	Modelling the PSF in MATLAB	12
3.3.3	The Airy Disk and Microscope Resolution	12
3.3.4	Summary	14
3.4	Gaussian Optics	Shashwat 15
3.5	Spatial Light Modulator - general introduction	Aditi 16

4 System Set-up		17
4.1 Mechanical		17
4.2 Optical Layout	Shashwat	18
II Fabrication		18
5 Fabrication Introduction		Adam 19
6 LASER Micro-Machining		Adam 20
6.1 Introduction		20
6.2 Method		21
6.2.1 Heat Affected Zone		21
6.2.2 Use of Femtosecond pulses		21
6.3 Motivation		22
6.4 How it works		22
6.4.1 Non-linear Absorption		22
6.4.2 Ionisation Modes		23
6.4.2.1 Photo-Ionisation		23
6.4.2.2 Avalanche Ionisation		23
6.4.2.3 Ionisation in Laser Micro-Machining		24
6.4.3 Inverse Bremsstrahlung and resonance Absorption		24
6.4.4 Strong (Hot) VS Gentle (cold) Ablation		25
6.4.5 Ablation Threshold		26
6.4.5.1 Analytical Techniques		26
6.4.5.2 Experimental Techniques		27
6.5 Set-up		28
6.5.1 Choice of Laser Crystal and Wavelength		28
6.5.2 Repetition rate		28
6.5.3 SLM		29
6.5.3.1 Parallelisation		29
6.5.3.2 Beam correction		29
6.5.3.3 Beam Shaping - Flat Top		29
6.5.3.4 Beam Shaping - Bessel Beams		30
6.5.4 Choice of Beam Targeting		31

6.5.4.1	Galvanometers	31
6.5.4.2	XYZ Stage	32
6.5.4.3	Comparison	32
6.5.5	Choice of Objective Lens	33
6.5.5.1	High NA	33
6.5.5.2	Low NA	34
6.5.5.3	Objectives Comparison	35
6.6	Capabilities	37
6.6.1	Resolution	37
6.6.2	Power	37
6.6.3	Speed	38
6.7	Summary	38
7	Stereolithography	Jin Jia 39
7.1	process overview	39
7.2	Optical Design	42
7.2.1	F-theta lens	42
7.2.1.1	why is F-theta lens needed?	42
7.2.1.2	F-theta lens selection	44
7.2.1.3	calculation of actual spot size	45
7.2.1.4	The scan area and distortion	46
7.2.2	Galvanometer for Stereolithography	47
7.2.2.1	Mechanism of galvanometers	47
7.2.2.2	Control design using PID tuner	49
7.2.2.3	Control design using state-space model	52
7.2.2.4	realisation of full state control - sensor	59
7.3	Mechanical design	60
8	Two-Photon Polymerisation	Sam 63
8.1	Literature Review	64
8.1.1	Theory	64
8.1.2	Resolution	65
8.2	TPP Optical System Design	67
8.3	Fabrication Process	70

8.4	TPP System Control	71
8.4.1	Spatial Control	71
8.4.2	Polymerization Control	74
8.4.2.1	Classification by Machine Learning	76
8.4.2.2	Classification by Threshold	78
8.4.2.3	Classification Conclusion	80
8.5	Summary	80
9	Optical Tweezing	Mark 81
9.1	Theory	81
9.1.1	General Principle	81
9.1.2	Ray Optics Explanation	81
9.1.3	Dielectric Explanation	82
9.2	HOT	84
9.3	What type of objects can be picked up	85
9.4	Simulation	85
9.5	Alternative Beam shapes	86
9.6	Use cases	87
9.6.1	Fabrication	87
9.6.2	Microfluids	87
9.6.3	Micromachines	87
9.6.4	Lab on a Chip	87
9.7	Setup	88
9.7.1	Conclusion	89
10	Excluded Techniques	90
10.1	Selective Laser Sintering	Jin Jia 90
III	Imaging	90
11	Imaging Introduction	James 91
11.1	Requirements and Specifications	91
11.2	Transmitted and Reflected Microscopy	Xiaonan 92
12	Live imaging modes	93

12.1	Technique: Bright-field Microscopy	93
12.1.1	Basic Principle of Bright-field imaging	93
12.1.2	Instrumentation of Bright-field imaging	94
12.1.2.1	Light source - Fibre coupled red LED	94
12.1.2.2	Detector - Scientific CMOS	95
12.1.3	Modelling of Bright-field imaging	96
12.1.3.1	Optical transfer function	97
12.1.3.2	MTF for bright field microscopy	97
12.2	Polarized Light	98
12.2.1	Basic Principle of polarized light imaging	98
12.2.2	Instrumentation of polarized light imaging	100
12.2.3	Modelling of polarized light imaging	100
12.3	Excluded Methods - James	101
12.3.1	Dark-field microscopy	101
12.3.1.1	Theory	101
12.3.1.2	Implementation	102
12.3.1.3	Reason for exclusion	102
12.3.2	Phase Contrast Microscopy	103
12.3.2.1	Theory and what it offers	103
12.3.3	Implementation	103
12.3.3.1	Reason for exclusion	104
13	Post-Process Imaging	Xiaonan 105
13.1	Confocal Laser Scanning Microscope	105
13.1.1	Basic Principle of confocal imaging	105
13.1.2	Fluorescence microscopy	106
13.1.3	Instrumentation of confocal imaging	107
13.1.4	laser	108
13.1.4.1	Single beam scanning mechanism	108
13.1.4.2	Detector - Photomultiplier tube	110
13.1.4.3	Detector - Pinhole size	110
13.1.5	Modelling of confocal imaging	112
13.1.5.1	Coherent and incoherent light source	112
13.1.5.2	Coherent system	112

13.1.5.3 Incoherent Confocal Imaging System	114
13.1.5.4 MATLAB simulation for confocal microscopy	115
13.2 Two-Photon Microscopy	120
13.2.1 Basic Principle of 2-Photon Imaging	Mark 120
13.2.2 Instrumentation of 2 photon excitation microscopy	Xiaonan 121
13.2.3 Modelling of 2 photon excitation microscopy	122
14 Superresolution	James 123
14.1 STED	123
14.1.1 Theory	123
14.1.2 Modelling Vortex Phase Plate in MATLAB	126
14.1.3 Resolution	129
14.1.3.1 Saturation Intensity	130
14.1.4 Continuous vs. Pulsed	131
14.1.5 Power Considerations and Photobleaching	133
14.1.6 Dyes	135
14.1.7 STED Fabrication	136
14.1.8 Beam Alignment and Drift	138
14.1.9 3D STED	140
14.1.10 Future Developments - Parallelisation	141
14.1.11 Summary	142
14.2 Excluded Methods	142
14.2.1 4-pi Illumination - James	142
14.2.1.1 Theory and what it offers	142
14.2.1.2 Resolution	143
14.2.1.3 Reason for exclusion	144
14.2.2 Structured illumination Microscopy - Xiaonan	Xiaonan 146
14.2.2.1 Reason for Exclusion	146
IV Optics	146
15 Spatial Light Modulators	Aditi 147
15.1 Motivation	147
15.2 Types of SLM	149

15.2.1 Optically and Electrically addressed SLM	149
15.2.2 MEMs	149
15.2.3 Multiple Quantum Well SLM	150
15.2.4 Liquid crystal on silicon SLM	150
15.3 Comparison of various SLM types	151
15.4 Liquid crystals	153
15.5 How it works	155
15.6 Key parameters that influence SLM choice	157
15.7 Choice of SLM	158
15.8 Power Considerations	161
15.9 Conclusion	162
16 Parallelisation	Lewis 163
16.1 Applications	163
16.1.1 Two-Photon Polymerisation and Laser Micro-machining	163
16.1.2 Holographic Optical Tweezing	164
16.2 Holography theory	165
16.2.1 Fourier Optics Revisited	166
16.2.2 Computer Generated Holography	167
16.3 Optical Layout of a holographic setup	167
16.3.1 Use of a Spatial Light Modulator (SLM)	169
16.3.2 Imaging system	169
16.4 Computer Generated Holograms	170
16.4.1 Phase Retrieval Algorithms	170
16.4.1.1 Gerchberg-Saxton	171
16.4.1.2 Weighted Gerchberg-Saxton	173
16.4.1.3 Direct Binary Search	173
16.4.1.4 Simulated Annealing	175
16.4.1.5 Extension to Three Dimensions	176
16.4.2 Modelling	177
16.4.2.1 Error metrics	177
16.4.2.2 Choice of algorithm	178
16.4.2.3 Results	179
16.4.2.4 Number of foci	180

16.4.2.5 Off-axis intensity	181
16.4.3 Speed	183
16.4.3.1 Considerations of GPUs, FPGAs	183
16.5 Parallelisation Conclusion	183
17 Microscope Objective Lenses	Mark 184
17.1 Classifications	184
17.1.1 Refractive Objective Lenses	184
17.1.1.1 Achromats	184
17.1.1.2 Fluorites	185
17.1.1.3 Apochromats	185
17.1.1.4 Flat Field - "Plan"	185
17.1.2 Reflective Objective Lenses	185
17.1.3 Finite Conjugate Objectives	185
17.1.4 Infinity Corrected Objectives	186
17.2 Specifications	186
17.2.1 Numerical Aperture	186
17.2.2 Magnification	187
17.2.2.1 Finite Conjugate Objectives	187
17.2.2.2 Infinite Conjugate Objectives	188
17.2.3 Parfocal Distance	188
17.2.4 Working Distance	188
17.2.5 Cover Slip Thickness	188
17.2.6 Design Wavelengths	189
17.3 Resolution	189
17.4 Magnification	190
17.4.1 Thin Lens Magnification	190
17.4.2 Empty Magnification	190
17.5 Rays, Pupils, Apertures	191
17.5.1 Aperture Stop	191
17.5.2 Marginal and Chief Rays	191
17.5.3 Pupils	191
17.5.4 Field Stop	192
17.6 Field of view	192

17.7 Rear Aperture and Back Focal Plane	193
17.7.1 Scanning and Epi-Illumination	194
17.8 Depth of Field	195
17.9 Light Gathering Ability	195
17.10 Choices and options	196
17.10.1 Confocal, STED, Brightfield	196
17.10.2 Two Photon Microscopy, Two Photon Polymerisation, Micro Machining . . .	196
17.10.3 Polarized Light	196
17.10.4 Non Oil Immersion	196
17.10.5 Manufacturer selection	197
17.10.6 Objective selection	197
17.11 Conclusion	199
18 4f	Freddy 200
18.1 Magnification and Demagnification	200
18.2 4f as a Spatial Filter	201
18.3 4f vs a Single Lens	201
18.4 4f Implementation	203
18.5 Summary	204
19 Aberrations	205
19.1 Achromatic Lens Design	Shashwat 205
19.1.1 Implementation and Supplier Choice	205
19.1.2 Lens Quality Control	207
19.1.3 Material Properties	207
19.1.4 Summary	209
19.2 Ultrashort Pulse Maintenance	209
19.2.1 Group Velocity Dispersion and Chirp	210
19.2.2 Spatio-Temporal Coupling	211
19.2.3 Pulse Control System Setup	214
19.2.3.1 Theory	214
19.2.3.2 System Implementation	215
19.2.4 Summary	217
19.3 Effects on the Point Spread Function	Freddy 218

19.3.1	Generalised Pupil Function	218
19.3.2	Matlab Model of Aberrated PSF	219
19.4	Adaptive Optics	Lewis 221
19.4.1	Modelling wavefront aberrations	221
19.4.2	Correcting wavefront aberrations	221
19.4.3	Shack-Hartmann Wavefront Sensor (SHWS)	223
19.4.4	Optical Layout	226
19.4.5	Adaptive Optics Conclusion	227
20	Beam Control	228
20.1	Generic Implementation	Shashwat 228
20.2	Choice of Lasers	229
20.2.1	Ti:sapphire Amplifier System	229
20.2.2	DPSS Laser Selections	231
20.3	Variable Attenuation	231
20.3.1	Alternatives Considered	232
20.3.2	Laser-Induced Damage	233
20.4	Shutter	234
20.4.1	Considerations and Specifications	234
20.4.2	Implementation	235
20.5	Beam Redirection	236
20.6	Fibre-coupled LED	236
20.7	Beam Expander/Condenser Design and Implementation	237
20.7.1	Design Alternatives	237
20.7.2	System Implementation	238
20.7.3	Summary	238
20.8	Pulse Shaping	239
20.8.1	Motivation	Freddy 239
20.8.2	Pulse Shaping Basics	Freddy 239
20.8.3	Diffraction Gratings	240
20.8.4	Implementation	241
20.8.5	Control	244
20.8.6	Comparison of Dispersive Components	Shashwat 245
20.9	Pulse shaping using an SLM	Aditi 246

20.9.1 Conclusion	Freddy	251
V Mechanical Design		251
21 Overall Mechanical Design	Tom	252
21.1 Introduction		252
21.2 Optical Table		252
21.2.1 Vibrational Theory		252
21.2.2 Resonance		254
21.2.3 Table Design		255
21.2.4 Vibrational Isolation		255
21.2.5 Table Choice		256
21.3 Mounting Optics		256
21.3.1 Adhesive Bonding		257
21.3.2 Clamping		258
21.3.3 Choice of Solution		260
21.4 Seeking a Clamping based Solution		260
21.4.1 Cage System		260
21.4.2 Custom Mounting		260
21.4.3 Choice of Solution		261
21.4.4 Custom Cage Components		261
21.5 Final Design		262
21.5.1 Final Cage System		262
21.5.2 SolidWorks Model		264
21.5.3 Design in Review		265
21.6 The Finite Element Method		266
21.6.1 Introduction		266
21.6.2 Theory		266
21.6.3 Meshing		269
21.6.4 Refinement		270
21.6.5 Examples in Design		271
21.7 Housing Design		273
22 Automated Stage and Galvanometers	Sam	276

22.1	Galvo scanning mirrors	276
22.1.1	Control	277
22.1.2	Scanning Area	277
22.2	Automated Stage	279
22.2.1	Linear Stages	279
22.2.1.1	Control	281
22.2.2	Transmissive Sample Illumination	281
22.2.3	Custom Cantilever Design	282
22.2.3.1	Cantilever Material Selection	283
22.2.3.2	Cantilever Geometry Optimization	283
22.2.3.3	Cantilever FEA	284
22.2.4	Overall Stage Specifications	286
22.2.5	Calibration	287
22.3	Summary	287
22.4	Further Development	287
22.4.1	Stage	287
22.4.2	Galvo scanning mirrors	287
23	Sample Logistics	Adam 288
23.1	Specification	288
23.2	Cassette	288
23.2.1	Build Area	288
23.3	Cassette I.D. Positioning and Calibration	289
23.3.1	Repeatable Positioning	289
23.3.1.1	Identifying the error	289
23.3.1.2	Encoders	290
23.3.1.3	Calibration	291
23.3.1.4	Clamping	291
23.3.2	Cassette Identification	293
23.3.2.1	Types of I.D.	293
23.3.2.2	Comparison	293
23.4	Storage	294
23.5	Transport	295
23.5.1	Specification	295

23.5.2 Result	295
23.6 System Summary	295
VI Control	295
24 Software	296
24.1 Overview	Jamieson 296
24.1.1 Capabilities	296
24.2 STL File Format	296
24.2.1 Data Structures for STL File	297
24.2.2 Parsing of STL File	298
24.3 Layer Slicing	298
24.3.1 Theory	298
24.3.2 Algorithm for Slicing	301
24.4 Contours	302
24.4.1 Linked List vs. Vector	303
24.4.2 Creating the Contour Objects	305
24.4.3 Floating Point Inaccuracies	305
24.4.4 The New Contour Creation Algorithm	306
24.4.5 Classification of Contours	306
24.4.6 Results	310
24.4.7 Performance	312
24.5 Resolution Separation	313
24.5.1 Purpose of Resolution Separation	313
24.5.2 Theory	314
24.5.3 Maximal Area Inscribed Rectangles	314
24.5.4 Polygonal Offsetting	316
24.5.5 Algorithm Implementation	316
24.5.6 Reorganising Data	317
24.5.7 Results and Performance	318
24.6 Tool Path Generation	318
24.6.1 Theory	319
24.6.2 Set of Intersecting Lines	319
24.6.3 Algorithm Implementation	320

24.6.4 Results and Performance	320
24.7 Conclusion	322
VII Conclusion	322
25 Applications	Aditi 323
26 Business Case	324
26.1 Unique selling points	324
26.2 Porter's Model	325
26.3 Cost	326
27 Conclusion	328
Appendices	329
Appendix A Optical Component List	Shashwat 330

Part I

Preamble

1 Introduction

1.1 Overview

This report will outline the design of a standalone optical nano-fabrication system that utilises light as a versatile engineering tool. It comes as an integrated product which incorporates a variety of fabrication and imaging techniques namely Laser Micro-Machining, Stereolithography, Two-Photon Polymerisation, Confocal Microscopy, Two-Photon Microscopy, STED and Optical Tweezing. These incorporated techniques allow users to fabricate, image and move full structures up to $8cm \times 8cm \times 8cm$ with feature sizes down to $55nm$. The system comes equipped with a cohesive and intelligent control system and a robust and bespoke software system. The facility is designed to work for a variety of materials such as: polymers, metals, glass and ceramics. This allows users to fabricate and image objects like: lab-on-a-chip, MEMS devices and other research activities. The originality of the system is its ability to incorporate the aforementioned techniques, thus producing remarkable versatility.

The report is divided into 7 main sections — preamble, fabrication, imaging, optics, mechanical design, control and conclusion. Section 1 provides an overview of the optical and mechanical layouts as well as an introduction to the optical theory used in the rest of the report. Sections 2-6 provide explanations of techniques or processes used within the system, their implementation in the system and finally discussions of further considered techniques. The report concludes by exploring the applications of the system and the business case.

1.2 Brief

A design brief was provided at the onset of the project so as to formalize the desired outcomes and design goals, it is provided here also, at the onset of the report so as to provide the reader with an appreciation for the catalyst that sparked into life the content of the following pages. The brief is as follows:

- Investigate several techniques for optical based fabrication
- Investigate and design appropriate light sources and modulation methods
- Develop an integrated optical system
- Develop a mechanical realization of the the optical system
- Develop a software suite for the system
- Develop the overall facility implementation

2 Processes

2.1 Laser Micro-machining

Laser Micromachining makes use of a high power laser to heat or ionise the material of the target sample ultimately resulting in a state change or the removal of that material. The use of an ultrashort femtosecond laser removes the issue of heat build up, and enables precise machining with steep wall edges and little change to the material surrounding the fabrication point. By controlling the laser intensity and location, holes can be cut, grooves machined, surfaces patterned and wave-guides created.

2.2 Stereolithography

The stereolithography fabrication technique utilizes a continuous wave, ultraviolet laser to selectively polymerize photo-sensitive resin to create a 3-D structure. The object is built layer by layer with a defined layer thickness until fabrication is complete. Our stereolithography process will be responsible for fabrication of the lower resolution portions of the build, to this end the resolution of the process is limited to 12 micrometres.

2.3 Two-Photon Polymerisation

High resolution sub-micron additive fabrication technique, analogous to the high resolution process of stereolithography. Included, to incorporate fabrication of sub-micron structures onto the larger stereolithography fabricated structures, and using the same resins. Two-photon polymerisation fabrication technique, uses the high intensities available from a femtosecond laser to induce polymerization in a photopolymer resin to create a free standing structure [1]. To give the highest resolution, and enable multiple polymerization beams we have designed a holographic, STED Two-photon polymerisation process. This allows X, Y features sizes down to 55nm to be fabricated, in parallel, which reduces the total fabrication time.

2.4 Confocal Microscopy

Confocal laser scanning microscopy is a simple setup with resolution up to 150nm. Compared with conventional bright field or polarized light microscopy, CLSM improves the resolution and contrast by placing a spatial pinhole at the focal plane of the lens to eliminate out of focus light. Its ability to image non-fluorescence material makes it an essential part of the imaging system.

2.5 Two-Photon Microscopy

Two-Photon Microscopy is a technique similar to confocal scanning microscopy but with an increased depth of imaging. Fluorophores are excited at and recorded from a precise point of focus which

enables the creation of three dimensional scans of fabricated materials as well as stained living cells. Its ability to image through IR transparent material makes this technique useful for observing hidden features in the samples.

2.6 Stimulated Emission Depletion (STED)

We use STED in our system to give users the ability to fabricate and image features well below 100nm lateral resolution. A method which allows us to engineer a narrower point spread function by depleting the area around the focal spot, STED is implemented in the system with the addition of the 532nm laser source, used as the excitation and depletion source in imaging and fabrication respectively.

2.7 Optical Tweezing

A highly focused laser beam can be used to exert forces on small particles and objects ranging from tens of nanometres to tens of micrometers. These forces can be used to trap, move, and manipulate objects as well as being used to measure displacements or to cause photochemical reactions. In combination with our ability to shape beams, we can use optical tweezers to assemble and power micromachines, giving users a high degree of flexibility.

3 Background Info

3.1 Magnification and Microscopes

The fundamental aspect of our system revolves around focusing light and controlling that point of focus on a small scale. This is analogous to imaging small objects which has been studied with vigour since the invention of the first microscope in the 16th century.

Microscopes are used to image and magnify objects onto the human eye or a digital sensor for further processing. The apparent size of an object is determined by the angle subtended between the rays emerging from the outermost points. The closer the object, the larger the apparent size. The human eye, however, has a limit in its ability to focus on small distances, the near point at distance n . The most basic system to overcome this limit is a simple magnifier which consists of a single lens, governed by the thin lens equation

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f} \quad (3.1)$$

where i is the image distance, o is the object distance and f is the focal length. This equation is derived in subsection 17.4.1 on page 190.

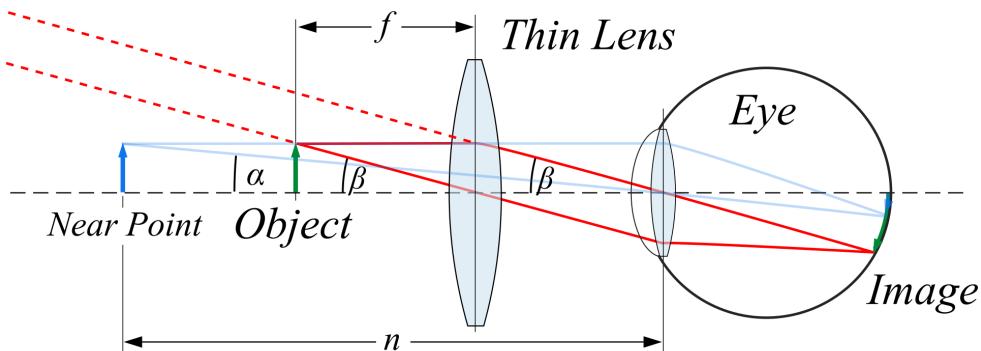


Figure 3.1: Ray Diagram of a Simple Magnifier

From Equation 3.1, if the object is placed at the focal length of the lens as in Figure 3.1, a virtual image will be formed behind the object at infinity. The **angular magnification** is defined as the increase in angle compared to when the object is at the near point without the lens, like the blue object in the figure. Using the small angle approximation $\tan(\alpha) \approx \alpha \approx \frac{h}{n}$ and $\tan(\beta) \approx \beta \approx \frac{h}{f}$, the magnification is determined by $m = \frac{\beta}{\alpha} = \frac{n}{f}$. Due to aberrations at small focal lengths, the maximum magnification which can be achieved with a simple magnifier is only about 20, even with multiple compound lenses. This is not nearly enough to image the nano structures required for this project.

To further increase magnification a typical compound microscope uses two lenses, both with short focal lengths. The objective lens, closest to the object, creates a real image inside the tube of the microscope as shown in Figure 3.2. The second lens, called the eye piece, works like the simple

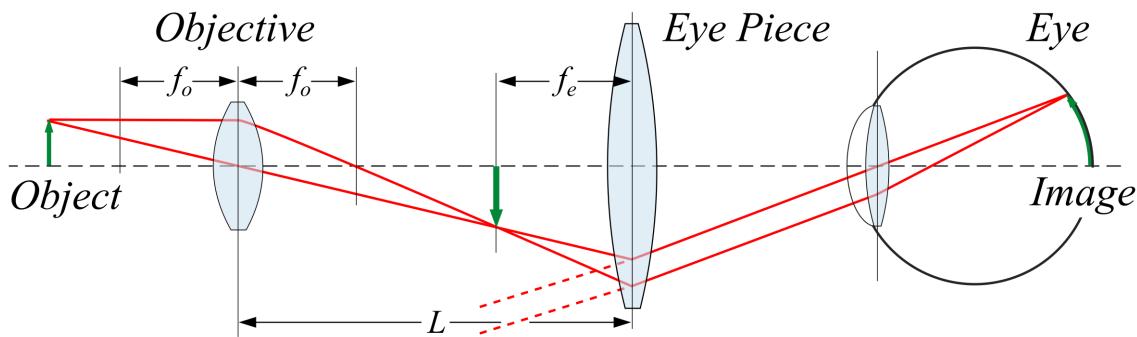


Figure 3.2: Schematic of a Compound Microscope

magnifier above but produces a virtual image from the already magnified real image, rather than the object itself. In this configuration, there are two magnifying steps and the total magnification is given by:

$$M_{\text{System}} = M_{\text{Objective}} \times M_{\text{Eyepiece}} \quad (3.2)$$

Real objective lenses consist of many elements in order to achieve high magnification with limited aberrations. The system we will use is more complicated than the simple microscope described but the principles of angular magnification are the same and form the basic theory of how we will optically interact with these small objects. More on this is provided in Part III (Imaging) and chapter 17 (Microscope Objective Lenses).

3.2 Fourier Optics

Fourier optics forms the theoretical backbone with which we can model and characterise how light propagates and diffracts. This is an essential tool for the appropriate design of a nanoscale fabrication and imaging system. In 1968, the publication of the book *Introduction to Fourier Optics*, by Joseph Goodman, provided a bedrock which explained and united the fundamental concepts of Fourier optics. This book is subsequently used to aid analysis.

Fourier optics provides the basis upon which diffraction, interference, imaging and holography can be explained. Geometrical optics, described in section 3.4, assumes a rectilinear propagation of the rays of light and ignores the wave nature of light and therefore, spreading due to diffraction. This means for small departures from perfection (diffraction limit) a wave description is required and therefore, an understanding of the Fourier description was essential to fully analyse and understand certain techniques and components within our system, such as imaging, the spatial light modulator, chapter 15, and the 4f system, chapter 18.

This section will start by describing the phenomena of diffraction, introducing the Fresnel-Kirchoff formula based on scalar diffraction theory and, subsequently providing two approximations of the formula, the Fresnel (near-field) and the Fraunhofer (far-field) — showing that the intensity pattern

at the far-field is actually the 2D Fourier transform of the aperture. Finally, this section will provide analysis of the Fourier transforming properties of lenses.

3.2.1 Diffraction

Diffraction refers to the various phenomena which occur when a wave encounters an obstacle or slit. According to Huygens' principle, Figure 3.3, diffraction can be described as the interference of waves — a wavefront can be thought of as many wavelets that interfere to form new wavefronts. In

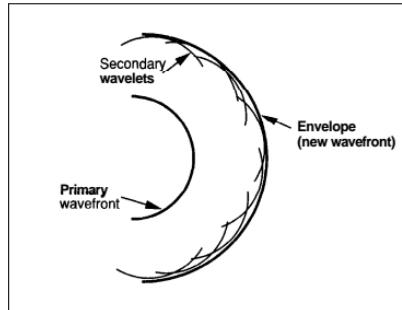


Figure 3.3: Illustration of Huygens' principle [2]

mathematical terms, this is outlined by the Fresnel-Kirchoff diffraction formula — which describes the optical field in a new plane (if we know the optical field in some other plane).

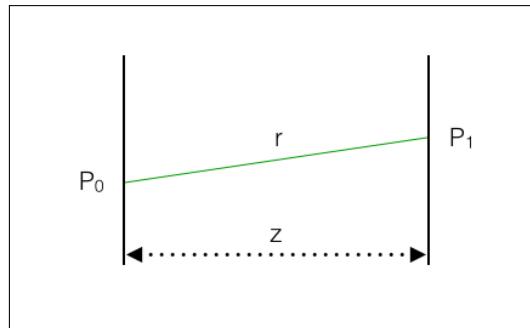


Figure 3.4: Illustration of an optical field that is propagating from one plane to another

The formula can be derived simply by considering the field at P_0 where the field strength is $U(x_0, y_0)$ and it is deemed to bring about a spherical wave. The strength of the field, therefore, decreases in proportion to the distance travelled and the phase is also related to the distance travelled. The field at another point P_1 a distance r from P_0 , as seen in Figure 3.4, can be written as:

$$U(x_0, y_0) \frac{\exp(-jkr)}{r} \quad (3.3)$$

where

$$r = \sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2 + z^2} \quad (3.4)$$

As Huygens principle states the field in a second plane is the summation of all the contributions from

the first plane. The field in the second plane can be written as:

$$U(x_1, y_1) = \frac{-j}{\lambda} \iint U(x_0, y_0) \frac{\exp(-jkr)}{r} \frac{1 + \cos(r, z)}{2} dx_0 dy_0 \quad (3.5)$$

In this equation there is a factor $\frac{-j}{\lambda}$ which ensures the field has the correct phase and strength. An obliquity factor has also been included in this because light cannot diffract backwards as easily as forwards. This was included by Kirchhoff to ensure Fresnel's previous work satisfied Maxwell's equations.[3]

3.2.2 Fresnel and Fraunhofer Approximations

This report will be using approximations to this formula for the Fresnel (near field) and Fraunhofer (far field) diffraction. These approximations make the integral in Equation 3.5 easier to evaluate in those regions of space.

The near field approximation says that our observation screen is sufficiently far away so $z = r$ on the denominator and $|x_0 - x_1|$ and $|y_0 - y_1|$ are sufficiently small so $r = z + \frac{(x_0 - x_1)^2}{2z} + \frac{(y_0 - y_1)^2}{2z}$ in the exponent. It is not reasonable to set $r = z$ in the exponent because small changes in r causes substantial changes in $\exp(jkr)$.[3] The diffraction integral becomes:

$$U(x_1, y_1, z) = \frac{-j}{\lambda} \frac{\exp(-jkr)}{z} \iint U(x_0, y_0) \exp\left(-\frac{jk}{2z}(x_0 - x_1)^2 + (y_0 - y_1)^2\right) dx_0 dy_0 \quad (3.6)$$

The far field approximation states that the distance to the observation plane is much greater than the aperture, therefore:

$$z \gg \frac{k}{2} (x_0^2 + y_0^2)_{\max} \quad (3.7)$$

$$\exp\left(\frac{-jk}{2z}(x_0^2 + y_0^2)_{\max}\right) \approx 1 \quad (3.8)$$

Thus the integral becomes:

$$U(x_1, y_1, z) \approx \frac{-j}{\lambda} \frac{\exp(-jkr)}{z} \exp\left(-\frac{jk}{2z}(x_1^2 + y_1^2)\right) \iint U(x_0, y_0) \left(\exp\frac{jk}{z}(x_0 x_1 + y_0 y_1)\right) dx_0 dy_0 \quad (3.9)$$

Ignoring the phase factor in front of the integral, it can be deduced that the Fraunhofer diffraction is proportional to the 2D spatial Fourier transform of the aperture function. It is important to note that in the Fraunhofer region the diffraction pattern does not differ with z but the pattern size is related to z .

Fourier optics has shown how light is diffracted through different apertures and how the optical field changes as the distance between the aperture and screen changes. The fabrication system employs Holography, section 16.2, which is the ability to define light intensity profiles, shapes and

patterns using diffractive optical elements. Fresnel and Fraunhofer formulas are particularly important for Holography as they allow us to engineer diffraction gratings to produce the desired holograms.

3.2.3 Fourier transforming properties of a lens

The most important components of an optical system is the lens and one of the most useful properties of a converging lens is its ability to perform 2D Fourier transformations. To understand how a lens performs a Fourier transformation, it must be first thought of as a phase transforming element rather than a ray bending element. A lens is composed of optically dense material in which the propagation velocity of an optical disturbance is less than the velocity in air. A lens is described as being thin if the co-ordinates (x, y) of entry of a ray are approximately the same as the coordinates of the exit; hence there is no translation of the ray. A thin lens, therefore, delays an incident wavefront by an amount proportional to the thickness of the lens at each point.[2] [3] Referring to Figure 3.5, the phase change, $\phi(x, y)$ of a wave can be expressed as

$$\phi(x, y) = kn\Delta(x, y) + k[\Delta_0 - \Delta(x, y)] \quad (3.10)$$

where n is the refractive index, $kn\Delta(x, y)$ is the phase delay induced by the lens, and $k[\Delta_0 - \Delta(x, y)]$ is the phase delay induced by the free space between the two planes.

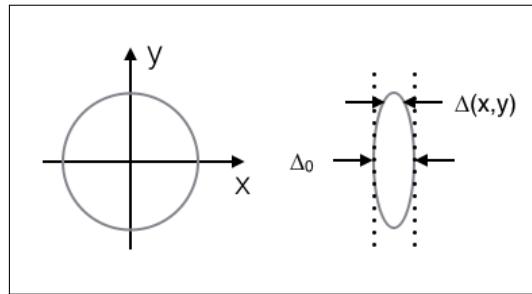


Figure 3.5: The thickness function. Front view (left), side view (right)

The lens may also be represented as a multiplicative phase transformation

$$t_l(x, y) = \exp[jk\Delta_0] \exp[jk(n - 1)\Delta(x, y)] \quad (3.11)$$

The complex field $U'_l(x, y)$ across a plane immediately behind the lens is then related to the incident optical field $U_l(x, y)$ by

$$U'_l(x, y) = t_l(x, y)U_l(x, y) \quad (3.12)$$

In order to fully understand the effects of the lens, a mathematical form of the thickness function $\Delta(x, y)$ must be defined. Using a paraxial approximation — that is the light rays are close to the optical axis — combined with the physical properties of the lens, an equation can be defined for the

thickness function. The lens phase transformation can then be re-written as:

$$U'_l(x, y) = \exp[jkn\Delta_0] \exp \left[-j \frac{k}{2f} (x^2 + y^2) \right] \quad (3.13)$$

The first term is a constant phase delay and the second term is a quadratic approximation to a spherical wave.

With this information the lens can be now considered as a Fourier transforming element. Referring to Figure 3.6, consider the case where the input plane is placed against the lens. Here the input field has a local amplitude transmittance $t_a(x, y)$, with a monochromatic plane wave of amplitude of A , to give a field at the lens of

$$U_l(x, y) = At_a(x, y) \quad (3.14)$$

A lens is a finite element, and therefore in order to account for this, a pupil function is defined.

$$U_l(x, y) = \begin{cases} 1 & \text{inside the aperture} \\ 0 & \text{otherwise} \end{cases} \quad (3.15)$$

The amplitude distribution behind the lens becomes

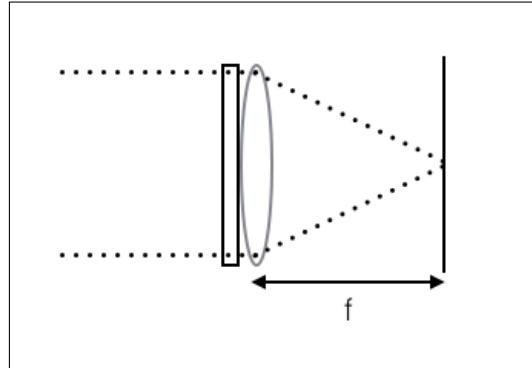


Figure 3.6: Geometry for performing the Fourier transform on a positive lens when the input is against the lens

$$U'_l(x, y) = U_l(x, y)P(x, y) \exp \left[-j \frac{k}{2f} (x^2 + y^2) \right] \quad (3.16)$$

The Fresnel diffraction is now applied to find the distribution $U_f(x, y)$ at the back focal plane. It should be noted that the constant phase delay arising from the lens transformation has been omitted.

Setting $z = f$.

$$U_f(x, y) = \exp \left[\frac{jk}{2f} (x_f^2 + y_f^2) \right] \iint_{-\infty}^{\infty} U'_l(x, y) \exp \left[\frac{jk}{2f} (x^2 + y^2) \right] \exp \left[\frac{-jk}{2f} (xx_f + yy_f) \right] dx dy \quad (3.17)$$

Substituting for $U'_l(x, y)$

$$U_f(x, y) = \exp\left[\frac{jk}{2f}(x_f^2 + y_f^2)\right] \iint_{-\infty}^{\infty} U_l(x, y) P(x, y) \exp\left[\frac{-jk}{2f}(xx_f + yy_f)\right] dx dy \quad (3.18)$$

The field distribution is thus proportional to the 2D Fourier transform of the incident field in the aperture.

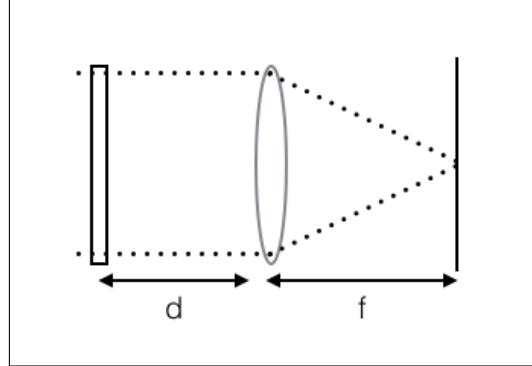


Figure 3.7: Geometry for performing the Fourier transform on a positive lens when the input is a distance d from the lens.

The final part of this section considers a more general geometry where the input is some distance from the lens. As per the previous example, the lens is illuminated with a normally incident plane wave of amplitude A and the input's amplitude transmittance is t_a again. Finally, let $F_0(f_x, f_y)$ represent the Fourier spectrum of the light at the input transparency and $F_l(f_x, f_y)$ represent the light incident on the lens. Using the paraxial approximation and an understanding of the angular spectrum of waves (see [2]) we can write:

$$F_l(f_x, f_y) = F_0(f_x, f_y) \exp[-j\pi\lambda d(f_x^2 + f_y^2)] \quad (3.19)$$

Using Equation 3.18 and ignoring the finite extent of the lens we can write:

$$U_f(u, v) = \frac{-j}{\lambda f} \exp\left[\frac{jk}{2f}(u^2 + v^2)\right] F_l\left(\frac{u}{\lambda f}, \frac{v}{\lambda f}\right) \quad (3.20)$$

Then substituting Equation 3.19 into Equation 3.20, we can write:

$$U_f(u, v) = \frac{-j}{\lambda f} \exp\left[\frac{jk}{2f}\left(1 - \frac{d}{f}\right)(u^2 + v^2)\right] F_0\left(\frac{u}{\lambda f}, \frac{v}{\lambda f}\right) \quad (3.21)$$

This shows that the phase and amplitude of the light at coordinates (u, v) are related to the phase and amplitude of the input spectrum at frequencies $(\frac{u}{\lambda f}, \frac{v}{\lambda f})$. It is important to note that when the input is placed in the front focal plane of the lens, $d = f$, the quadratic phase factor disappears, leaving an exact Fourier transform. The image at the back focal plane is the Fourier transform of the

light incident on the lens — multiplied by $\frac{-j}{\lambda f}$. This property is the fundamental reason why we use 4f imaging systems, discussed in chapter 18 on page 200.

3.3 Point Spread Function

The point spread function (referred to as the PSF onwards) provides a method to measure the imaging capabilities of an optical system. It is considered the fundamental unit of an image in theoretical models of image formation. The PSF is the 3 dimensional diffraction pattern seen when an infinitely small point source is emitted in a specimen and transmitted to the image plane through a lens. The lens collects a portion of the emitted light and focusses it to a point in the image plane. This point is not infinitely small due to light waves converging and interfering. A diffraction pattern of concentric rings of light surrounding a central, bright disk, is produced when viewed in the x-y plane, known as the Airy disk.[4] This section will focus on the theory of the PSF; the relationship between the PSF and the image of an object; model the PSF in MATLAB; and finally discuss how the Airy disk affects resolution.

3.3.1 Mathematical description of the Airy Disk

The coherent image of an object is given by the convolution between the object, $U_g(u, v)$ and the point spread function of the lens forming the object image, $h(u, v)$.[2]

$$U_i(u, v) = h(u, v) \star U_g(u, v) \quad (3.22)$$

Where the PSF is the 2D Fourier transform of the pupil function forming the image. The ideal PSF can, therefore, be calculated using an ideal lens, which has a circular pupil function. The pupil function can be written as:

$$circ(r) = \begin{cases} 1 & r < 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.23)$$

As stated the the pupil function is circularly symmetric, so we can take the Fourier Bessel transform or Hankel transform. We can, therefore, write the PSF as:

$$\mathcal{F}[circ(r)] = \frac{2J_1(r)}{r} \quad (3.24)$$

where J_1 is the Bessel function of the first kind of order one, which is more commonly known as the Airy disk (named after George Biddell Airy). From this equation, it can be seen that the resolving power of a lens can be evaluated by measuring the size of the Airy disk. It is worth noting that if the pupil function were equal to one (an infinite lens), then the PSF would be the Fourier transform of 1 -

which is a delta function. Thus Equation 3.22 reduces to $U_I = U_g$, as the convolution of a function with a delta function is just the original function and, therefore, the system has been reduced to geometrical optics. Having an understanding of how to manipulate the PSF is therefore important to improving the image quality of our system. This control of the PSF will be seen later in the imaging techniques: confocal and 2 photon microscopy, and STED (where the diffraction limit of light is removed).

3.3.2 Modelling the PSF in MATLAB

In order to demonstrate how changing the properties of a lens changes the PSF and thus the output image, a MATLAB simulation of a perfect lens was built. The function takes as inputs: the focal plane sampling; the wavelength of the light; the aperture diameter; the lens' focal length; and finally the PSF array size. The function then constructs a sampling plan — the centre to centre spacing of the pixels at the focal plane. The pixels are assumed to be square. The aperture is then imposed upon the system. Finally, it takes the Fourier transform of the system in order to produce the point spread function. The first null of the Airy disk has a radius of:

$$r = \frac{1.22\lambda f}{d} \quad (3.25)$$

where λ is the wavelength of the light, f the focal length of the lens and the d the aperture diameter. This information was used in order to check that the function was producing the correct output.

In Figure 3.12, relative intensity is plotted as a function of spatial position for PSFs from objectives having numerical apertures of 0.3 and 1.3. This shows graphically how the PSF depends upon the numerical aperture of the lens. The numerical aperture, which will be referred to as NA onwards, is a dimensionless number that describes the angles over which an imaging system can accept light, for a more detailed explanation see subsection 17.2.1. It can be seen that increasing the NA produces a narrower PSF - hence a better imaging performance. Similarly, Figure 3.17 shows that, for the same numerical aperture lens, decreasing the wavelength of the light results in a narrower PSF.

3.3.3 The Airy Disk and Microscope Resolution

The limit up to which two small objects are seen as separate entities is used to measure the resolving power of an imaging system. The distance where this limit is reached is known as the effective resolution of the system and is usually denoted by d_0 . This minimum distance between two point objects in order to distinguish them can be defined using the Abbe or Rayleigh criterion. These are closely related by (but differ on) how they define what is meant by two objects being resolvable from each other.[5] The Rayleigh criterion is defined as:

$$d = \frac{0.61\lambda}{NA} = \frac{0.61\lambda}{n \sin(\theta)} \quad (3.26)$$

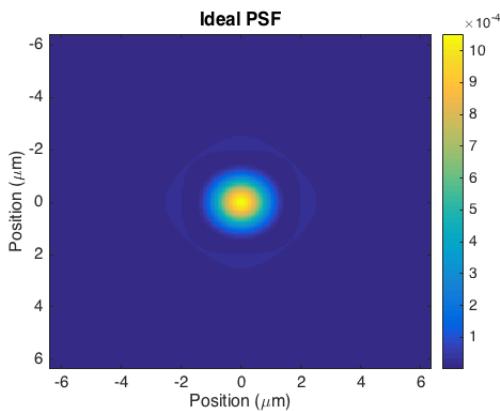


Figure 3.8: PSF plot NA = 0.3

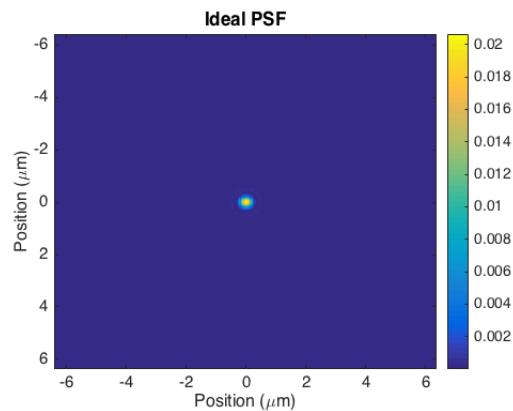


Figure 3.10: PSF plot for NA = 1.3

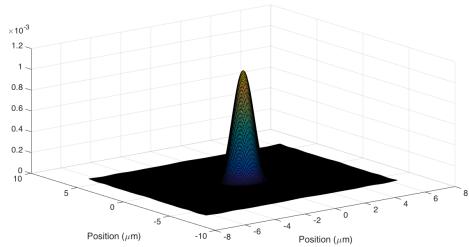


Figure 3.9: Surface plot for NA = 0.3

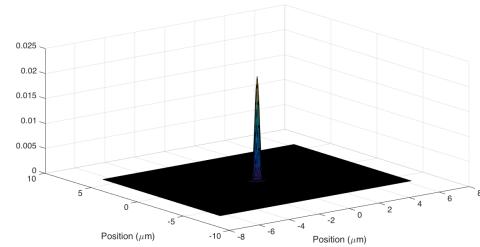


Figure 3.11: Surface Plot for NA = 1.3

Figure 3.12: Relative intensity plotted as a function of spatial position for PSFs from objectives having numerical apertures of 0.3 and 1.3. Plots are produced using the MATLAB simulation.

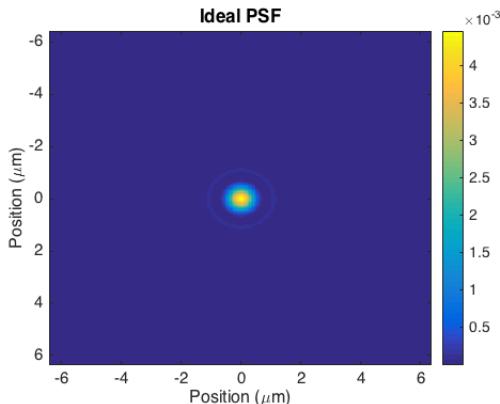
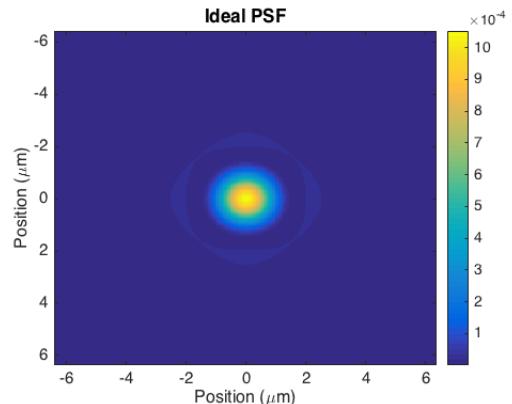
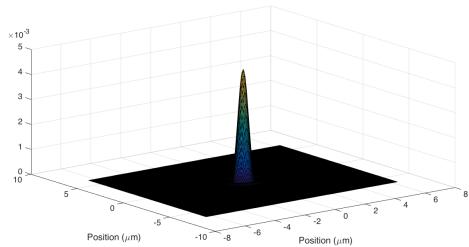
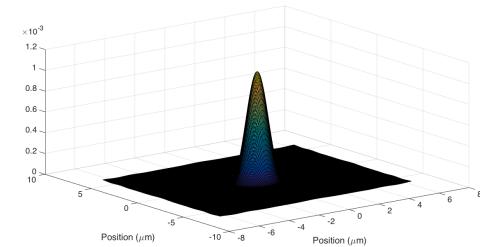
Figure 3.13: PSF plot for $\lambda = 400\text{nm}$ Figure 3.15: PSF plot for $\lambda = 800\text{nm}$ Figure 3.14: Surface plot for $\lambda = 400\text{nm}$ Figure 3.16: Surface plot for $\lambda = 800\text{nm}$

Figure 3.17: Relative intensity plotted as a function of spatial position for PSFs from objectives having $\lambda = 400\text{nm}$ and 800nm . Plots are produced using the MATLAB simulation.

while the Abbe criterion is defined mathematically as:

$$d = \frac{0.5\lambda}{NA} = \frac{0.5\lambda}{nsin(\theta)} \quad (3.27)$$

where d is the distance between the two objects; λ is the wavelength of light; NA is the numerical aperture of the imaging lens; n is the refractive index of the medium between the object and the lens; and finally θ is the angle at which light enters the lens. Using Abbe's criterion the theoretical limit of resolution using a 1.4 NA lens and a light wavelength of 510nm is 182nm whereas for Rayleigh's criterion the limit is 222nm . This difference is because, according to Abbe's theory, at least two orders of the refractive pattern must be captured to reconstruct an image; whereas Rayleigh assumes the limit to be where the peak of one airy disk meets the first null of the next one, as illustrated in Figure 3.18. The superimposed profiles display two brightness maxima that are separated by a valley. The intensity in the valley is reduced by approximately 20 percent compared with the two maxima.[5] This is just sufficient for the human eye to see two separate points.

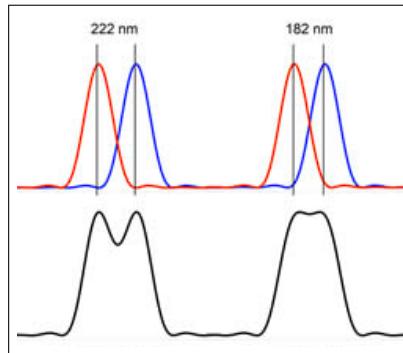


Figure 3.18: Illustrations of Rayleigh and Abbe's criterion [5]

3.3.4 Summary

In summary, the ideal (aberration free) PSF describes the response of a system to a point source input and is therefore the fundamental unit of image in theoretical models of image formation. The PSF is thus valuable for determining the resolving power of lenses and imaging systems. A greater understanding of the PSF function allows us to use techniques, such as confocal microscopy and STED, in order to manipulate the PSF to improve our systems resolution and image quality.

3.4 Gaussian Optics

Throughout this report, all laser beams have been modelled and considered as Gaussian beams. A Gaussian beam is a beam with an intensity distribution $I(r, z, t)$ that varies both spatially and temporally as a Gaussian function [6]. The parameters that define the spatial intensity distribution include the wavefront front radius r , and the distance of the beam from its minimum waist point z , as shown in Figure 3.19.

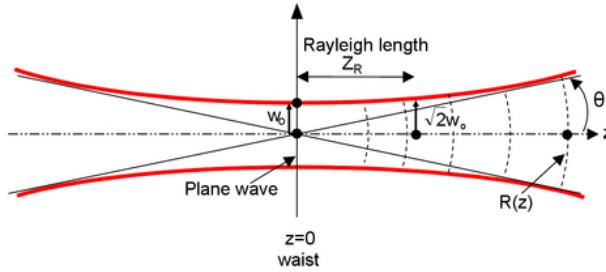


Figure 3.19: Gaussian Beam Model [6]

Given the maximum power at the cross-section of the beam P , and beam waist (radius) $w(z)$, the intensity distribution will follow the power distribution, with intensity $I = P/A$, where A is the spot size of the beam. The Gaussian beam was modelled using the expression in (3.28).

$$I(r, z, t) = \frac{P}{\pi w^2(z)} \exp \left(-2 \frac{r^2}{w^2(z)} + \frac{t^2}{\tau^2} \right) \quad (3.28)$$

Most beam profiles are either Gaussian or sech profiles, as they ensure that the power loss through free space propagation of the beam is minimal. The Rayleigh length z_R is defined as the approximate length of the beam for which it can be assumed to be collimated. This approximation is important, as unlike ideal rays, real beams are rarely collimated. The range is given by the following relation, where n is the refractive index of the medium, and λ is the centre wavelength of the beam:

$$z_R = \frac{\pi w_0^2 n}{\lambda} \quad (3.29)$$

It must be noted that the Rayleigh length is not valid after the beam has passed through a lens, and hence the Rayleigh length is valid only for the beam until it hits a lens.

This definition of a Gaussian beam can be applied to all the lasers used in this optical system, and further discussion will be conducted regarding the intensity distributions of such beams and the propagation constraints through the system.

3.5 Spatial Light Modulator - general introduction

A device used repeatedly by the fabrication and imaging techniques is the spatial light modulator (referred to as SLM onwards). SLMs are systems that modulate light according to a fixed spatial (pixel) pattern.[7] They are controlled by a computer software provided with the SLM and provide great advantage through electronic control, in terms of accuracy and ease of integration in a system. They are a basic optical control system used widely in optics. SLMs are used to provide both phase and amplitude modulation for an incoming beam, dependent on the angle of polarization. A liquid crystal on silicon based SLM is the most common choice because of the anisotropic properties of liquid crystals. On application of an electric field on a pixel by pixel basis, the molecules realign accordingly leading to a change in the refractive index and hence the optical path. This leads to a phase difference. A representation of this can be seen in Figure 3.20.

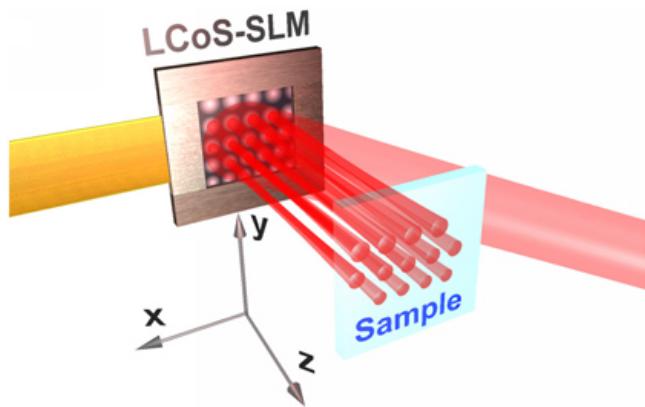


Figure 3.20: Pixel by pixel electronic control of phase[8]

Using the SLM, the phase and intensity of the beam can be modified which is used by the imaging and fabrication techniques. This allows complete beam control.

4 System Set-up

Having defined the design specifications of the system and identified the key processes incorporated into the system, the system design was then put together as presented in this section.

4.1 Mechanical

Full automation of fabrication processes providing feature resolutions spanning several orders of magnitude necessitated the use of highly specialized custom components in the system. Two bespoke staging areas are provided, one for Stereolithography, and the other for Two Photon Polymerization and Laser Micro-Machining, both facilitating their respective manufacturing techniques whilst providing compatibility with the others. Structures large enough for the eye to see be produced with nanoscopic features. To allow for all of this without the need for human intervention a fully integrated robotic arm is provided, along with a custom built cassette system. This not only allows for the transportation of samples between staging areas, but also to a storage unit, allowing fabrication in small batches.

4.2 Optical Layout

The optical layout of the system is produced in Figure 4.1. The colours of the beam paths indicate the optical paths through the system, where red indicates the 800 nm Ti:sapphire laser, green/navy paths indicate the 532 nm and the 355 nm lasers, respectively. The orange paths indicate the Red 632 nm fibre-coupled LED light that passes through the objective for transmissive live imaging, while the sky-blue path indicates the reflected 400 nm light from the sample for two-photon microscopy. The component list, containing the full description of the abbreviations used in this layout, has been presented in Appendix A. While each process has been incorporated into this layout separately

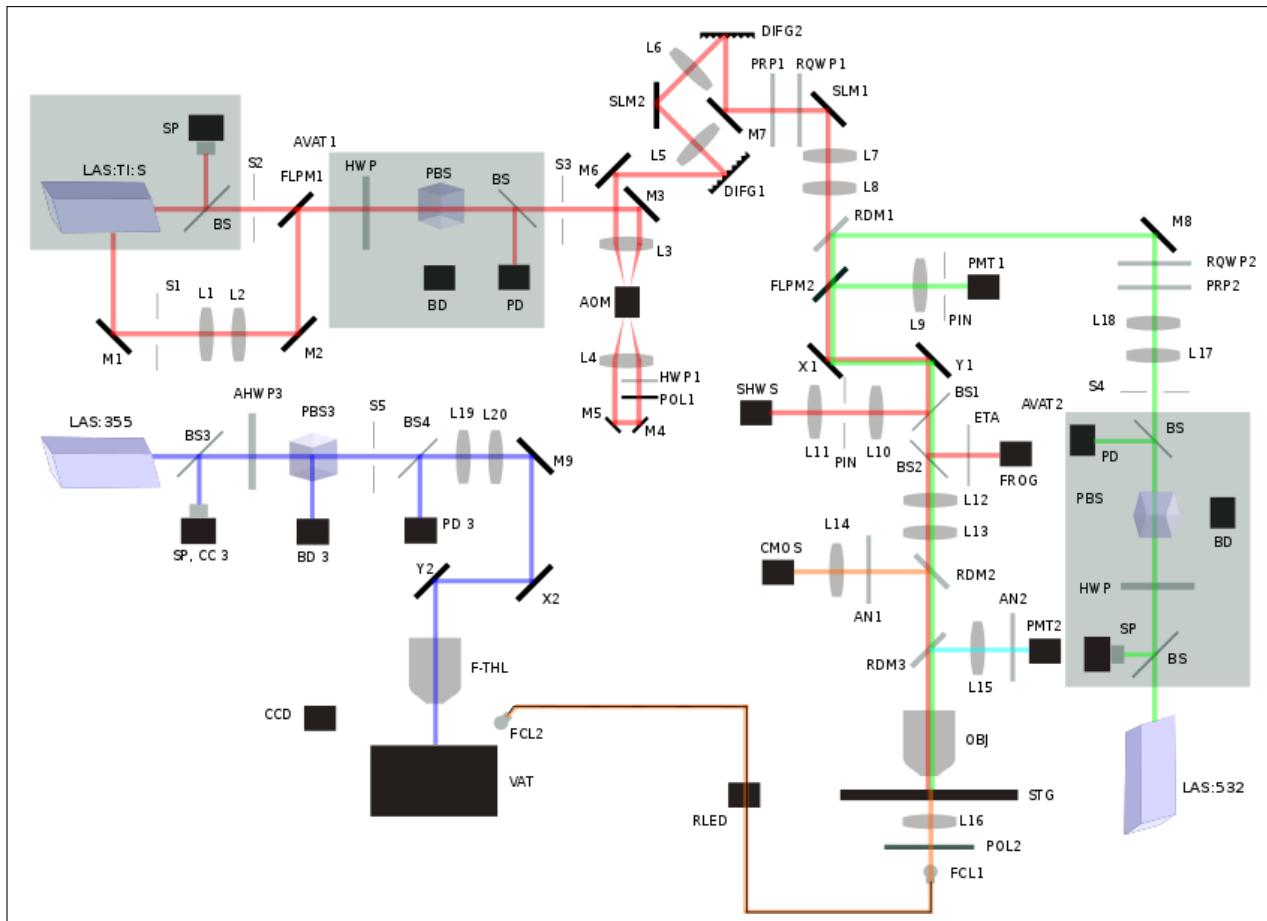


Figure 4.1: System Optical Layout

as illustrated in later chapters, the integration aspects of the optical layout, such as laser power control, high-speed shutter design, pulse shaper design, lens design, and beam redirection have been included in Chapters 19 - 23.

Part II

Fabrication

5 Fabrication Introduction

This section of the report covers in detail the 3 primary fabrication methods implemented in the system: 'Laser Micro-Machining', 'Stereolithography', 'Two Photon Polymerisation'. Also discussed is 'Optical Tweezing' and the methods considered before being excluded from the design. These process have each been included due to their unique combination of assets that they bring to the system, and hence the customer.

Laser Micro-Machining provides a sub-micrometer reductive manufacturing technique that can also utilise its thermal effects to induce property changes in materials such as the refractive index of glass allowing the fabrication of optical waveguides with resolution lower than 700nm .

Stereolithography provides a method by which large bulky structures can be rapidly fabricated. It is the key part of one of our unique selling points. Many 3D fabricated objects have relatively large volumes of material with no fine detail. Consequently, we use this like a paintbrush (a $12\mu\text{m}$ brush to be precise) in comparison to the pencils of Laser Micro-Machining and Two Photon Polymerisation. Enabling both fine detail and fast fabrication.

Two Photon Polymerisation builds on the Stereolithography process by reducing the resolution by 2 photon absorption and a novel STED implementation resulting in resolution of 100nm and 60nm respectively.

Also implemented is (Holographic) Optical Tweezing. This ties the systems together by proving a means by which particles can not only be moved around but also rotated. This Allows micro-mechanical mechanisms to be powered for 'Lab-on-a-chip' applications.

6 LASER Micro-Machining

6.1 Introduction

Laser Micro-Machining is the reductive manufacturing technique included in our system. It can ablate a broad range of materials from metals such as aluminium and titanium, to dielectrics like silicon and polymers - a driving reason behind including it. Additionally, it can be used to induce thermal effects to change the properties of materials such as the refractive index of borosilicate glass, a unique feature that adds to the reasons to include Laser Micro Machining in the system and for customers to invest.

Figure 6.1 shows the optical path of this system, there are some key parts to note in this path. The Ti:Sapphire laser that drives this process has 2 individually shuttered outputs, a seed output ($80MHz525mW$) and an amplified output ($10kHz0.15mJ$). These are directed onto the same path with a mirror assembly. The lasers have their power moderated by a closed loop control system consisting of power control (AVAT1) and detection (PD1). A high speed shutter assembly, from M4 through the AOM twice and to M7, provides a shutter action fast enough to be able to pick and choose pulses from the laser source. A pair of diffraction gratings (DIFG1/2) either side of an SLM (SLM2) also provide temporal pulse shaping allowing stretching and compression of the pulses as well as the creation of arbitrary pulse shapes. Another SLM (SLM1) then provides spatial beam shaping effects. To enable techniques such as multiple foci generation to provide efficiency boosts, and Bessel beams enabling high aspect ratio hole drilling and cutting through thick materials, among countless other benefits. The beam then passes through the galvanometers before reaching the high numerical aperture objective to focus the laser to a $700nm$ (see part 6.5.4.2) spot on a work area held in an interchangeable cassette positioned atop a translatable stage.

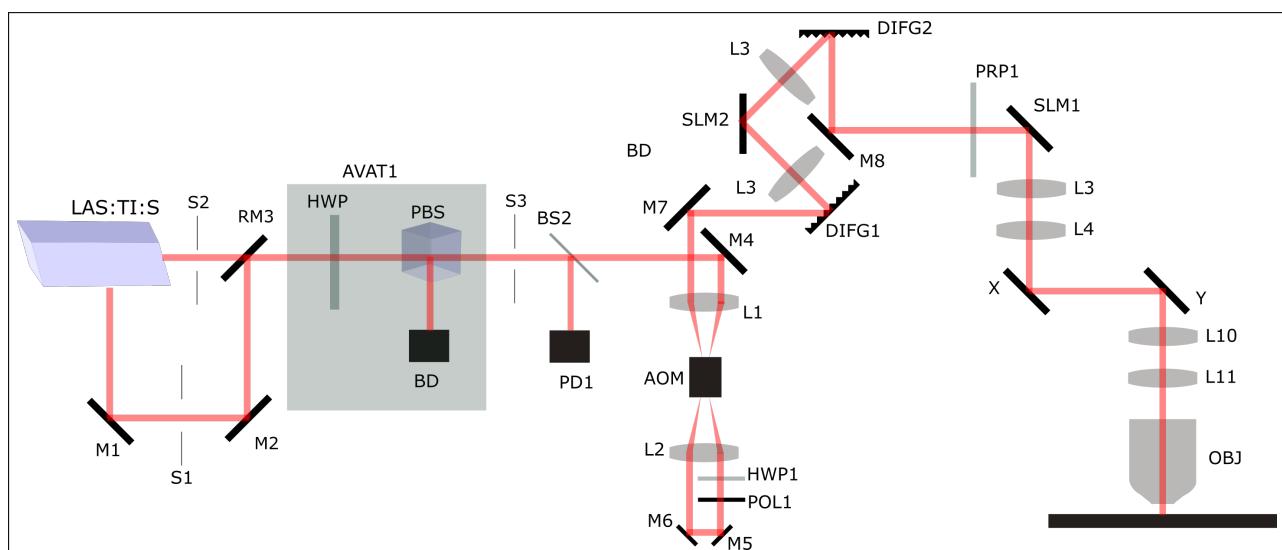


Figure 6.1: Laser Micro-Machining Optical Layout

6.2 Method

Laser Micro-Machining (also known as laser ablation) is where a laser with sufficient power is targeted at the sample to be machined, and thermal effects boil or combust away material. This has the benefits of being a contact-less machining method that is relatively fast. However, 'Heat Affected Zone' or HAZ causes limitation when one attempts to implement it at the nano-scale.

6.2.1 Heat Affected Zone

HAZ is the area around the focus where the material is not machined away, but has still been affected by the laser either directly or indirectly. The issue with this is that the radius of the HAZ is often much larger than the radius of the machined hole. In the HAZ there can be material deformities and stresses applied that change the properties, shape and structure of the material. This is rarely desired. The issues of HAZ are avoided by using a pulsed laser at higher intensity as opposed to a lower powered continuous wave laser.

6.2.2 Use of Femtosecond pulses

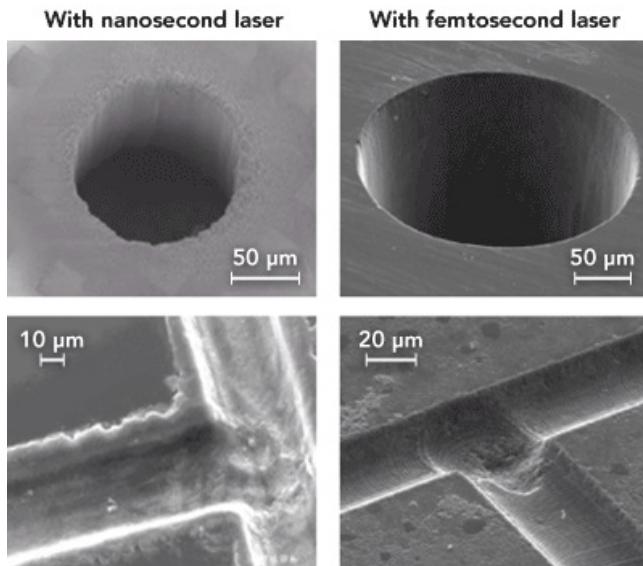


Figure 6.3: Comparison of ns machining to fs machining [9]

gated by the pulses.

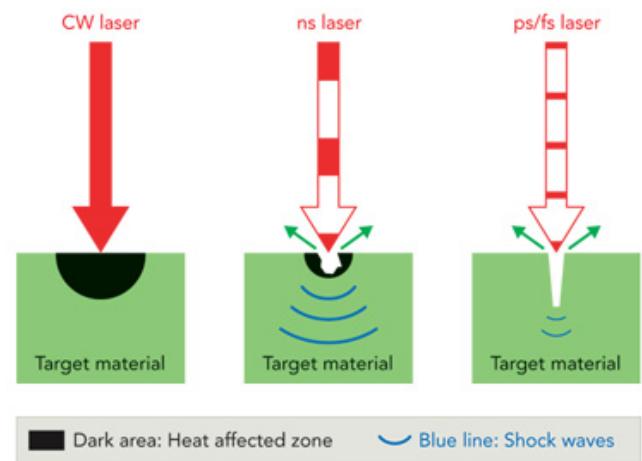


Figure 6.2: Comparison of material effects between pulse durations [9]

Once the pulses are down to femtosecond (fs) scale, the instantaneous power must be higher in order to carry the same energy. This means that the pulses require less energy to reach the required power for non-linear absorption, so the little energy that is deposited as heat is also minimised. Furthermore, once operating in a sub-pico-second regime, the absorption of laser energy takes less time than that required to transfer that energy to the lattice, hence decoupling lattice heating from absorption. Figure 6.2 shows the reduction in HAZ as shorter pulse lengths are introduced as well as the shock-waves ini-

Femtosecond laser micro-machining has the benefits over other laser machining methods of: finer resolution, minimal HAZ, no change to sample properties, sub-surface machining in transparent materials, and due to the non linear absorption being independent of wavelength most materials can be machined with the same laser source. Figure 6.3 provides an example of the difference between ns machining (in this case with a UV laser) and fs machining (100 fs pulses at 780nm).

6.3 Motivation

A key motivation for the inclusion of laser micro machining is the desire for a reductive manufacturing method that can operate at the nano scale. This is critical to the unique 'fast-nano' selling point that creates large structures quickly with Stereolithography before adding sub-micrometer detail with Laser Micro-Machining and Two Photon Polymerisation. Additionally, it brings flexibility in terms of materials and applications. For example, the additive methods are tightly restricted to polymers that can be photo-polymerised, or doped to behave as such, they cannot for example create a glass structure for photonics experiments. However, given a block of glass Laser Micro-Machining can change the refractive index and ablate cavities with great versatility in shape and positioning, even within the bulk of the material. While the laser settings required for ablation and thermal modification will vary between materials, a large number of different materials can still be ablated including metals and dielectrics.

6.4 How it works

6.4.1 Non-linear Absorption

Non-linear Absorption is what lies at the heart of femtosecond laser micro-machining, the non-linear absorption reduces HAZ, removes the wavelength dependence of the process and makes it more deterministic. The classical or linear interpretation of light can be understood by thinking of the material as a system and light as an input to that system which results in some output. This is a linear system and describes aspects that we may expect from light, such as light pulses not interacting with each other, and the material properties being independent of the field intensity. However the interaction of light with a material is described by Maxwell's equations the constitutive relations of which are non-linear with respect to the electronic and magnetic fields. This leads to the conventional ideas just mentioned to be refuted and in practice, exploited. Light pulses can interact with each other, and material properties are dependent upon the incident field intensity.

6.4.2 Ionisation Modes

6.4.2.1 Photo-Ionisation

Photo-Ionisation refers to the direct excitation of electrons in the target material by the laser field, and it comes in two forms: 'Tunnelling' and 'Multi-Photon'.

Tunnelling is a process where the electrical field of the laser suppresses the coulomb well in which the electron resides, until it can tunnel out of said well. This process dominates at lower laser frequencies and is visualised in figure 6.4b.

Multi-Photon Photo-ionisation however dominates at higher energies (although not so high as to provide ionisation by the energy of just one photon). It operates by an electron absorbing several photons; the combined energy of which is sufficient to exceed the band gap of the material and promote an electron from the valence band into the conduction band. A visual interpretation of this can be seen in figure 6.4a.

The transition between these two schemes was investigated by L.V. Keldysh and he defined the Keldysh parameter γ as in equation 6.1.

$$\gamma = \frac{\omega}{e} \left[\frac{mcn\epsilon_0 E_g}{I} \right]^{\frac{1}{2}} \quad (6.1)$$

Where ω is the laser frequency, I is the intensity at the focus, m is the reduced mass of an electron in the sample lattice, e is electron charge, c is the speed of light, n the refractive index of the sample, E_g the band gap and ϵ_0 the permittivity.

Using this parameter he defined 2 regions where each of the ionisation schemes are dominant, for $\gamma > 1.5$ the multi-photon process dominates however for $\gamma < 1.5$ Tunnelling Dominates.

6.4.2.2 Avalanche Ionisation

Unlike Photo-Ionisation, Avalanche Ionisation requires some electrons to have already moved to the conduction band from the valence band. This may be achieved by thermally excited impurity or defects, or alternatively can be provided by Photo-Ionisation. Once in the conduction band, the sequential absorption of photons (free carrier absorption) elevate the electron's energy level until it has the energy required to be at the level of the conduction band summed with the band gap energy. $E_c + E_g$ At this point the electron can collide with another electron and impart its excess energy into it, raising the new electron from the top of the valence band into the bottom of the conduction band; leaving two conduction band electrons available for the process to repeat visualised in figure 6.4c.

The growth of conduction band electrons obeys equation 6.2

$$\frac{dN}{dt} = \eta N \quad (6.2)$$

where η is the avalanche ionisation rate. Some have proposed that $\eta = \alpha \times I(t)$ [10] where $I(t)$ is the beam intensity and α is known as the avalanche rate and is determined by the scattering and band gap of the target sample. However Stuart's work makes a number of assumption in the creation of his model and it's accuracy has been questioned. Thornber proposes an alternative relationship as $\eta \propto \sqrt{I(t)}$ [11]

6.4.2.3 Ionisation in Laser Micro-Machining

For short pulses with high intensities these processes make the ionisation deterministic, in contrast to the stochastic properties when longer pulses are used. This is because Avalanche ionisation dominates with longer pulses and takes its seed electrons from impurities in the lattice that are distributed in a statistical manner. While for short pulses, photo-ionisation dominates and occurs early in the time-scale of the pulse, the rest of the pulse is absorbed by the free electron plasma produced by the ionisation of the material. The absorption is governed by the 'inverse Bremsstrahlung' and 'resonance absorption' effects which are independent of the initial state of the material, allowing highly repeatable machining due to the consistency of energy absorption. These deterministic results are critical to the process as they enable repeatable and predictable outcome in the nature of the ablation, allowing dimensions and surface finish to be controlled accurately, a key property that customers will appreciate.

6.4.3 Inverse Bremsstrahlung and resonance Absorption

Once an electron cloud (plasma) has been formed in the first few fs of the pulse, the remainder of the laser energy is absorbed efficiently by a combination of 'Inverse Bremsstrahlung effect' and 'Resonance Absorption' causing heating of the plasma [13].

The Inverse Bremsstrahlung effect operates on the opposite principle to the Bremsstrahlung effect. Laser light is absorbed near a 'critical surface' where the oscillating electric field causes oscillations in the electrons. These oscillations are then converted to thermal energy by electron ion collisions, the nature of which can be described by the cross-section for collisions (6.4) which is

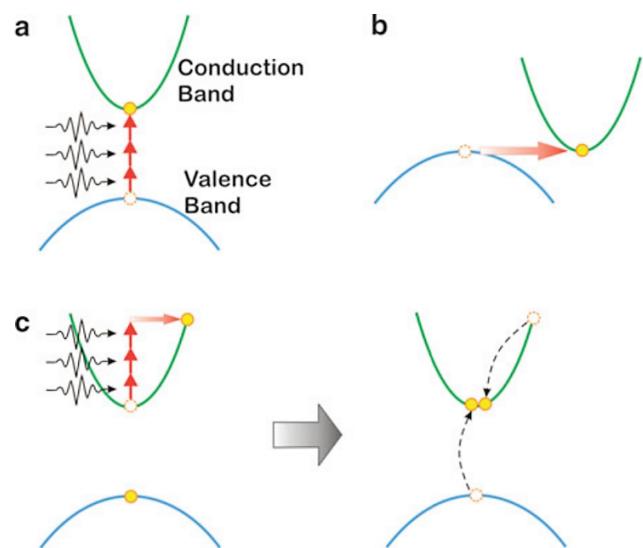


Figure 6.4: Potential Ionisation modes in Laser Micro-machining, (a) multi-photon, (b) tunnelling, (c) avalanche [12]

derived from Rutherford's formula (6.3) [14].

$$\frac{d\sigma_{ei}}{d\Omega} = \frac{1}{4} \left(\frac{Ze^2}{m_e v^2} \right)^2 \frac{1}{\sin^2(\theta/2)} \quad (6.3)$$

$$\sigma_{ei} = \frac{\pi}{2} \left(\frac{Ze^2}{m_e v^2} \right)^2 \int_{Ze^2/K_B T_e}^{\lambda_D} \frac{\sin \theta}{\sin^4(\theta/2)} d\theta \quad (6.4)$$

This can then be turned into a collision frequency before combining with Maxwell's equations and solving to find that an absorption coefficient exists (6.5) for which the differential equation (6.6) dictates the absorption of the laser energy. [14]

$$k_{ib} \approx \frac{Z_i}{T_e^{3/2}} n_e^2 \left(1 - \frac{n_e}{n_c} \right)^{-1/2} \quad (6.5)$$

$$\frac{dI}{dz} = -k_{ib} I \quad (6.6)$$

However this cannot be solved easily as the plasma is inhomogeneous, some have made approximations to linearity for long pulse durations however at short pulse durations this does not hold.

It is this inhomogeneity, that enables the second absorption mechanism, 'Resonance Absorption'. Electrostatic waves are excited in the plasma when a component of electric field for the laser coincides with the density gradient of the plasma. The electric field becomes large near critical surfaces and the waves resonate in the plasma. These waves are damped and hence the oscillatory energy is eventually converted into thermal energy. Collision based damping is what is seen in Inverse Bremsstrahlung effect, but there are other damping mechanisms in the plasma. However it can be shown [14] that the absorbed power is independent of the damping mechanism. Hence resonance absorption will dominate over Inverse Bremsstrahlung effect at the high intensities for which the system operates.

6.4.4 Strong (Hot) VS Gentle (cold) Ablation

Gentle ablation occurs at short pulse lengths (order of picoseconds and less) and operates on ionisation and coulomb explosion avoiding the unwanted thermal effects. Strong ablation occurs at longer pulse periods and inherently operates on thermal effects as the material is removed by explosive vaporisation. With the versatility of our system we can exploit both, however, cold ablation is the most highly sought after and hence is focussed on here.

6.4.5 Ablation Threshold

When working on a material the laser parameters need to be set. If the material had been tested or used in the fabrication system already then the details may be stored in a database as part of the user interface of the system. However, if the material has not been used before or unique settings are required for it then methods may be required to ascertain the settings. These come in two varieties: analytical techniques where theoretical calculations are followed, and experimental techniques where different setting are tried on a sample of the material until the required result is observed.

6.4.5.1 Analytical Techniques

In order to ablate a single atom from the lattice of a material it must be provided with sufficient energy to exceed the binding energy of that atom. For high laser intensities at short pulse lengths, the material is ionised in the very early stages of the pulse, and hence the remainder of the energy is absorbed by the free electron plasma by 'inverse Bremsstrahlung' and 'resonance absorption' effects. This results in the ablation mechanics for both metals and dielectrics being very similar under these circumstances, Gamaly defines equations 6.7 and 6.8 as first order approximations to the fluence required for ablation.

$$f = \frac{3}{8}(\epsilon_b + \epsilon_{esc}) \frac{\lambda n_e}{2\pi} \quad (6.7)$$

$$f = \frac{3}{8}(\epsilon_b + J_i) \frac{\lambda n_e}{2\pi} \quad (6.8)$$

[13] Where ϵ_b is binding energy, ϵ_{esc} is work function, J_i is ionisation potential and n_e is the number density of free electrons defined by equation 6.9.

$$n_e = \frac{z\rho N_a}{M_r \times 10^{-3}} \quad (6.9)$$

[15] Where z is valency, ρ is density, N_a is Avogadro's constant and M_r is the relative atomic mass (the 10^{-3} is to convert M_r from grams to kg) Aluminium for example has $z = 3$, $\rho = 2712$ and $M_r = 26.98$ hence $n_e = 1.816 \times 10^{29}$. Substituting this along with $\epsilon_b = 6.56 \times 10^{-19}$ $\epsilon_{esc} = 1.16 \times 10^{-17}$ and $\lambda = 800\text{nm}$ into equation 6.7 leaves $f = 1.063 \times 10^5 \text{J/m}^2$ or 10.63J/cm^2 . This value is in the correct region for the ablation of aluminium in ns and longer regimes however much lower thresholds have been observed in shorter pulse regimes to the region of 0.4J/cm^2 due to non-linear effects [16]. Hence these equations could be taken as guides known to overestimate however better equations or experimental methods would be needed to ascertain the exact value required for ablation.

However Schaffer et. al note that experiments so far have only investigated the breakdown threshold at the surface of a sample and not within the bulk, and hence differences between the two have

not been determined let alone equations to forecast such differences. [11] It is hypothesised however that the surface will have more impurities than the bulk which will introduce seed electrons allowing avalanche ionisation to initiate more easily at the surface also resulting in a reduction in the threshold.

6.4.5.2 Experimental Techniques

As noted above, the use of experimental techniques are greatly advantageous to finding the exact ablation threshold for machining due to a wide variety of uncertainties in the response of the material to the laser. However machining is not always performed at or near to the threshold either. Sometimes it is above or below the threshold, with adjusted repetition rates or pulse durations to induce different results from the machining process. An example of this is outlined in the waveguide writing demonstrated by "Chris B. Schaffer" [17].

So for our system experimental techniques will be used to identify the appropriate settings for attaining different effects on various materials while in the factory. These settings will then be stored to a database that the customer can pick the material and process from without needing to know the specific material properties. Furthermore, due to the imaging that our system can also perform, instructions can be given to the customer to do their own experimental tests for more rare materials that may not be already in the database, potentially even with the computer vision processes described later, all without the need for another product.

One of the experimental methods of identifying the ablation threshold is to focus the laser into a sample of the material you wish to find the ablation threshold of. Having performed this for a variety of laser settings that wish to be tested, the focal points can be optically probed. This involves shining a low power laser at where the machining should have occurred, and blocking the directly transmitted light. If the machining laser has caused structural change then this probing laser will be scattered or refracted by the material inconsistency. This scattered light can pass either side of a block to be focussed into a detector. If the light is not scattered then it can be assumed that there is no change to the material for the laser settings being probed. [17]

Another method is to attempt ablation at a range of settings as above then observe the focal points with an imaging technique. Many imaging methods are integrated into this system and are discussed in part 3 of this report. The processes will be able to detect a change in patterning of the material and can also pick up changes in refractive index of the material and hence will allow the detection of the damage threshold as well as the ablation threshold. In addition this will allow the details for waveguide writing to be found experimentally too.

Comparison	Excimer	Ti:Sapphire	Nd:YAG	Fibre
Max Repetition rate	10kHz	GHz	10kHz	GHz
Max Pulse energy	100's of mJ	mJ	100s mJ	nJ
Centre Wavelength	157-351nm	800nm	1064nm	1060nm
Pulse Length	ns	low as 10fs	ps	ps (experimentally fs [18])

Table 6.1: Laser Source Comparison table (Note: the listed capabilities are not always found in the same laser source)

6.5 Set-up

6.5.1 Choice of Laser Crystal and Wavelength

A range of laser sources exist on the market, from pumped fibre through Ytterbium and Neodymium doped crystals, to excimer and Titanium doped Sapphire. Some of the common sources used for Laser Micro-Machining are seen in table 6.1. The main restraint upon the laser is its capacity for generating USPs (Ultra Short Pulses), these are sub 10 pico-second pulses, most often in the femtosecond regime. The reason for needing such a USP laser is outlined in section 6.2.2. This has immediately excluded Excimer and Nd:YAG laser sources from consideration.

For a large variety of materials to be processed, we must be able to reach the ablation threshold of the material, an intensity of around $10^{13} - 10^{14} W/cm^2$ will guarantee ionisation of any target material early in the pulse period [13], hence fibre is also not appropriate (however research is being done on novel methods where fibre can be used). So the remaining option is the Ti:Sapphire option with its high power short pulses, it is ideal for use and widely accepted in the industry as the best for *fs* Laser Micro-Machining.

6.5.2 Repetition rate

Repetition rate is often limited by the need for ablated material to escape the work area before the next pulse arrives. This is in order to avoid absorption by the expelled material which would result in unpredictable outcome of the disturbed pulse. Commonly used repetition rates are in the 1-100kHz range.

There are methods to improve the rate at which the ablated material is expelled. The first is to shorten the pulse length, this results in the material being expelled with more energy and hence greater speed; allowing the next pulse to arrive sooner. The other method is to utilise an 'assist gas'. In this case a gas flow is targeted at the focal point of the laser approaching from the objective lens and spreading out over the surface of the target carrying any ablated material with it. In some cases

it is not possible to have this flow pattern and the gas is just blown across the surface from the side.

Some uses of Laser Micro-Machining however do not require the removal of material, but take advantage of the thermal effects that the laser can induce to change the properties of the material being targeted. The writing of wave-guides in glass for example, requires no removal of material just a controlled refractive index change in the body of the glass. Schaffer et. al showed that this can be achieved by many low power pulses being focussed into the glass at a very high repetition rate of $25MHz$. The f_s pulses provide the high resolution while the high repetition rate results in the energy having insufficient time to leave the focal volume of the material. Hence the focus becomes a point source of heat. The heat build up causes melting of the glass which once re-solidified has a higher refractive index than the rest of the material that has not been affected by the laser. As the only limits to this process are the rate at which the material can be heated with appropriate accuracy, this has a much higher speed than laser micro machining ordinarily does. Wave-guides can be written at speeds of $20mm/s$. [17]

6.5.3 SLM

Having an SLM integrated into the beam path of the Laser Micro-Machining process allows the user control over the shape of the intensity profile of the laser. The method by which this is achieved is covered in chapters 15 and 16 in detail. Here we will look at some of the benefits and uses of this ability.

6.5.3.1 Parallelisation

One of the primary advantages of including an SLM in the beam path is the ability to generate multiple foci. This allows the machining of multiple structures or parts at the same time, hence increasing throughput. The process is covered in more detail in chapter 16.

6.5.3.2 Beam correction

The SLM is also valuable in reducing aberrations and other optical distortions that can affect the focus. This enables smaller focal spot sizes that have improved symmetry while reducing any high intensity areas not at the centre of the focal point. This is discussed in detail in section 19.4.

6.5.3.3 Beam Shaping - Flat Top

While the Laser beam is ordinarily Gaussian in profile, the inclusion of an SLM in our layout enables more than just the creation of multiple foci and correction of aberrations. It enables control over the shape of the beam intensity profile. One of the applications of this is the creation of a 'flat top' beam. A flat top has the benefit of transitioning from zero intensity to high intensity in a very short distance and holds its high intensity level with good uniformity before returning to zero intensity in a short distance.

This has two key benefits, the first being that the 'wings' of the beam (the area receiving laser fluence but not enough to ablate material), are small. This minimises thermal effects and HAZ. The second is that it makes the ablation uniform across the profile resulting in much more easy to work with ablation properties.

It also has its drawbacks however, the beam no longer benefits from the low divergence properties of a Gaussian profile and hence its value for M^2 increases significantly, from near 1 to values exceeding 10 [19] resulting in much larger dimensions for the focal point according to equations 6.14 and 6.17. However it also reduces the Rayleigh range 6.10 resulting in a shallower focus that may also be useful for some applications such as surface and thin film processing, especially when combined with the uniformity that a flat top offers.

6.5.3.4 Beam Shaping - Bessel Beams

Methods for drilling high aspect ratio holes have come some way over the years, with improvements to the processes used such as trepanning and helical drilling seen in figure 6.5.

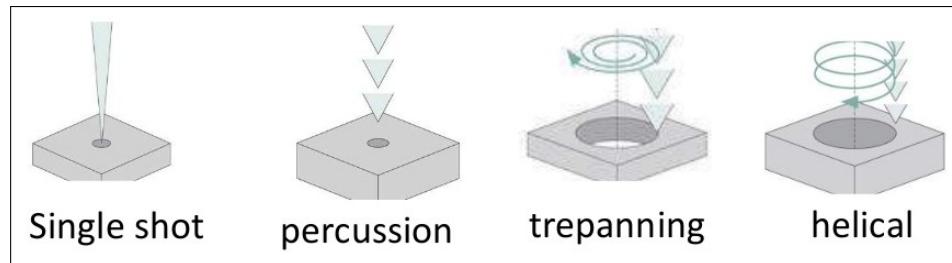


Figure 6.5: Different methods used for drilling holes with Laser Micro Machining [20]

However the point still stands that the depth of the focus is roughly of the same order as the focus size, and can be approximated by the Rayleigh range (equation 6.10) [19] which evaluates as 387nm for our system - an aspect ratio of 0.54.

$$z_g = \frac{\pi}{4\lambda} \frac{d_{dl}^2}{M^2} \quad (6.10)$$

'Bessel Beams' are currently the best way to achieve very high aspect ratio foci. By way of a virtual axicon generated by the SLM a beam profile can be formed for which the electric field can be defined by the solution to the differential equation in 6.11 and is hence a 'Bessel Beam'. specifically for laser micro-machining the zeroth-order solution of the first kind ($\alpha = 0$ and non-singular at the origin) is used (J_0) for which the electric field is profiled by 6.12 [19].

$$x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} + (x^2 - \alpha^2)y = 0 \quad (6.11)$$

$$E(r, \phi, z) = A_0 \exp(ik_z z) J_0(k_r r) \quad (6.12)$$

Intensity is proportional to the square of the electric field, hence yielding a high intensity core with low intensity concentric rings surrounding it. It should be noted that in reality only an approximation to the Bessel function is used, as the intensity continues varying infinitely, asymptotically tending towards zero.

Bessel beams have a focal volume with very high aspect ratio, of the order of 100x (however some specialised systems have exceeded 1000x [21]). The Rayleigh range for the beam can in fact be approximated by 6.13 where D is the diameter of the aperture and d_f the diameter of the central lobe [19]. This is visualised in figure 6.6 where the distinct difference in focal volume can be seen. The drop in intensity on the colour scale should also be noted and comes about as a consequence of energy being transmitted in the concentric rings as well as the main lobe hence reducing the intensity.

$$z_b \approx \frac{\pi D d_f}{4\lambda} \quad (6.13)$$

It is this incredibly high aspect ratio that allows single-shot machining of high aspect ratio holes and even through thickness ablation and cutting while maintaining very narrow cross sections for a variety of potential applications such as wave guide writing without the need for stage translation. The process is now widely used for cutting the toughened glass used in smart-phone screens and our system with its integrated imaging techniques could provide new insights and hence profits into that industry.

6.5.4 Choice of Beam Targeting

6.5.4.1 Galvanometers

A galvanometer mirror system operates by changing the location of the focus with respect to a static sample and objective lens. They are covered in more detail in chapter 23 and can be seen in the optical layout (figure 6.1) on the right side marked x and y. The minimum spot size that a galvanometer system can handle for Laser Micro-Machining is given by equation 6.14. D is the Gaussian beam diameter, taken as FWHM (Full Width Half Maximum), λ is the laser wavelength, f is the focal length of the objective, and M^2 is a measure of the focus-ability of the laser - which is given by the laser manufacturer. The 1.5 value is a coefficient that does in fact vary between about 1.5 and 1.9 depending upon aberrations caused by the more complicated optical structure of using galvos and the

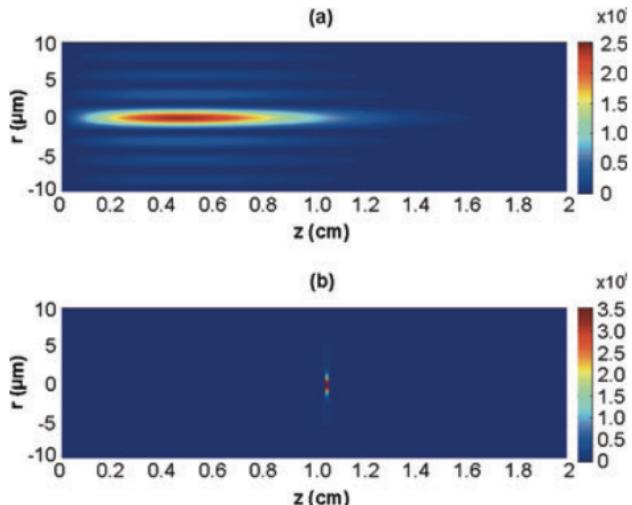


Figure 6.6: Cross section of Bessel (a) and Gaussian (b) intensity profiles [19]

limitations of the beam path, but for the sake of comparison we will take the best case scenario of 1.5.

$$d_{dl} = 1.5 \frac{\lambda f M^2}{D} \quad (6.14)$$

[22] From the data sheet of the laser and amplifier we are using gives $M^2 < 1.3$ and we take $\lambda = 800\text{nm}$. For the purpose of these calculations the same objective as that used in the stage calculations is chosen, due to it being ultimately implemented in the Laser Micro-Machining process furthermore 2 Photon Polymerisation also successfully uses the same objective with galvanometers. These values result in a focal spot of $1.35\mu\text{m}$ larger than the nm scale laid out in the specification of the project.

6.5.4.2 XYZ Stage

Using the stage for targeting (seen in the lower right of figure 6.1 below the objective) operates by relocating the sample with respect to the stationary focus. The focus is no longer limited by aberrations due to clipping of the input aperture or the non-perpendicular angle of incidence. Hence the constant in the governing equation for the focus diameter falls to $4/\pi$ or roughly 1.27 as opposed to the 1.5-1.9 for the Galvanometers. Additionally the value for D increases as the beam now take up the majority of the entrance pupil of the objective.

$$d_{dl} = \frac{4}{\pi} \frac{\lambda f M^2}{D} \quad (6.15)$$

[22] This can be simplified as assuming our laser utilises all of the aperture area, equation 6.16 can be substituted into 6.15 in order to achieve equation 6.17 where n is the refractive index of the material that the lens is operating in and θ is the maximum half angle of the resulting cone of light when focussed through said lens.

$$NA = n \sin \theta = n \sin \left[\arctan \left(\frac{D}{2f} \right) \right] \approx n \frac{D}{2f} \quad (6.16)$$

$$d_{dl} = \frac{2}{\pi} \frac{n \lambda M^2}{NA} \quad (6.17)$$

Substituting in the values in this case with our chosen objective gives $d_{dl} = 716\text{nm}$ a value below the $1\mu\text{m}$ limit that was specified and almost half of the value offered by use of the galvo system.

6.5.4.3 Comparison

Table 6.2 shows a 'Multi Criterion Analysis' of the two methods of beam targeting. It can be seen that stage translation provides us with a number of benefits important to our cause such as small focus size and a large work area, however it comes at the cost of a slower scanning speed. Due to our

-	-	Galvanometer Scanning		Stage Translation	
Comparison	Weight	Points	Score	Points	Score
Scan Speed	3	3	9	1	3
Size of focus	5	1	5	3	15
Max. Work Area	2	1	2	3	6
Accuracy	4	2	8	2	8
Totals	-	-	24	-	32

Table 6.2: Stage VS Galvos Comparison table

priorities of small focal spot size and precise positioning of the focal spot the stage acquires more points than galvanometers and hence is implemented in our system. However due to Two Photon Polymerisation coming to the opposite conclusion, a hybrid system can be implemented between these two processes where the object to be fabricated does not require sub micrometer spot size and would benefit from a high scanning speed, where the stage controls coarse targeting while the galvos deal with a fast scanning process this is used in Two Photon Polymerisation and is outlined later in the report. Due to the optical beam path being identical to that of two photon, this scanning method can also be implemented if speed is of the essence and sub-micrometer resolution is not required.

6.5.5 Choice of Objective Lens

The primary aspect to consider when choosing an objective lens is its Numerical Aperture or 'NA', this is a dimensionless term that characterises the range of angles over which light is incident upon the focal point. However working distance, requirements for immersion and minimisation of aberrations is also important and some of the finer aspects of objective lenses are covered in chapter 17. Following is a summary of what a high NA objective can do for us and likewise for a low NA objective, before they are compared to find the best option for our system.

6.5.5.1 High NA

The relevance of NA upon the focus size has already been noted in part 6.5.4. However there is another benefit that a high value of NA affords it is within bulk machining. Due to the steep angle at which the rays converge on the focal point, the required powers for ablation are only reached very close to the focus. Hence the laser light can penetrate a transparent material without affecting its structure, before non-linear effects take over at the focus to enable machining of material that is otherwise not absorbing of the wavelength. The application of this can be seen in figure 6.7 where the high NA objective enables continuous transverse machining only limited by the the stage travel limits,

and the working distance of the objective. While the low NA objective can only perform longitudinal machining, severely limiting the complexity of structures that can be made and also requiring all of the structures to be exposed to the surface closest to the objective.

At high NA the paraxial approximation breaks down and aberrations become a concern. The paraxial approximation is the same as the small angle approximation with regards to the angle between a light ray and the optical axis. The aberrations come about due to the small angle approximation being exactly that, an approximation. Larger angles involved in high numerical aperture lenses cause deviation from this. However the use of a high quality objective with aberration correction, and the capacity for correction afforded by the pair of SLMs in our chosen assembly results in aberrations being of minimal concern.

The other primary concern of high NA objectives is the reduced working distance that they offer and the requirement for 'immersion'. This limits the versatility and increases the cost of the system. The immersion is more an inconvenience in our design, but the limited working distance means that the penetration of the features that can be machined is limited to the working distance. However as noted above, these features can be more complex and do not have to begin at the surface like the holes and slots created by a low NA objective.

6.5.5.2 Low NA

At low NA self focussing becomes a concern. Self focussing is a non-linear effect that influences the propagation of the LASER through the bulk of a transparent sample. It comes down to the dependence of refractive index upon the intensity, defined as follows [11].

$$n = n_0 + n_2 I \quad (6.18)$$

Where n is the total refractive index of the material, n_0 is the ordinary refractive index and n_2 is the non-linear index. In the event that the pulse is anything other than a perfect flat top then the intensity varies over the cross section and hence so does the refractive index of the bulk material. Typically $n_2 > 0$ so the refractive index is highest for high intensities, this makes the refractive index at the centre of a Gaussian profile higher than in the wings causing a focussing effect.

Despite the relationship between refractive index and intensity (6.18) the self focussing effect is

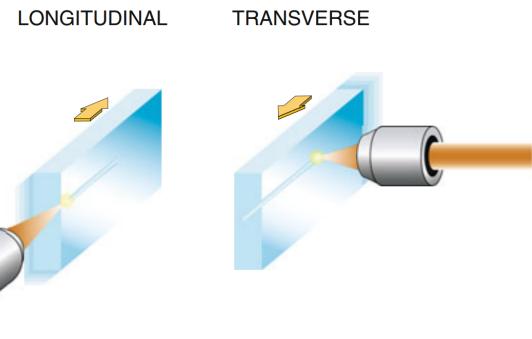


Figure 6.7: Longitudinal vs. Transverse Machining [12]

only dependent upon the peak power of the pulse. Due to the fact that if the same pulse power were to be imparted upon double the area to normal then the intensity would half causing the refractive index to drop by a half as well cancelling out the increase of area and hence the refractive power remains the same. [11]

There is a laser power for any given material which defines the point at which the self focussing effect becomes strong enough to balance diffraction effects, at this point a filament of light is formed.

$$P_{cr} = \frac{3.77\lambda^2}{8\pi n_0 n_2} \quad (6.19)$$

Beyond this point beam collapse is predicted, however this does not occur due to ionised electron gas having a negative refractive index, hence cancelling the self focussing effect. [11] However this electron cloud is caused by ionisation of the material in the same way that ablates the material. Hence staying away from these power levels will avoid unplanned ablation. For laser powers near and above P_{cr} research into the behaviour of the beam is sparse, however below about a quarter of the critical power the effects of self focussing upon laser intensity have been modelled by equation 6.20 allowing the intensity to be forecast. [11]

$$I_{sf} = \frac{I}{1 - P/P_{cr}} \quad (6.20)$$

The reason that self focussing is only a problem at low NA is due to the issue with the size of the focal point. The threshold for bulk optical breakdown is of the order of $10^{13} W/cm^2$ while P_{cr} is of the order of $1MW$ for transparent solids. [11] which assuming a flat top profile means that the focus must be less than $3.57\mu m$ in order to have the intensity to ablate but insufficient power to reach critical self focussing or below $1.78\mu m$ in order to be in the predictable region of self focussing. Part 6.5.4.2 gives us the diameter of the focal point assuming no self focussing and it can be seen that the diameter is inversely proportional to the NA of the objective lens. For an objective with $NA = 1.4$ a focal diameter of $716nm$ can be attained, however at $NA = 0.2$ this diameter increases to $5.012\mu m$, exceeding our specification for focal spot size and also in the region where self focussing can become an issue. Hence the system will not focus on this as an option, however the ability remains to retrofit a low NA objective if the customer desired it.

6.5.5.3 Objectives Comparison

Schaffer et. al performed a comparison between values of Numerical Aperture with respect to the ablation threshold of the target material. Using the experimental technique referenced in part 6.4.4 they varied the NA and tested for change in the material properties. A $100fs$ pulse focussed into 'corning 0211' glass resulted in the results plotted in figure 6.8. [17]

The experiments resulted in equation 6.21 being identified as a good fit to the data in the graph for circumstances with weak self focusing.

$$E_{th} = \frac{I_{th}\tau\lambda^2}{\pi(NA)^2 + I_{th}\lambda^2/P_{cr}} \quad (6.21)$$

[17]

This equation provides us with a simple relationship between the pulse length of choice, the breakdown intensity, and the pulse energy, that also accounts for the effects of weak self focusing and NA. Substituting the fixed values into this equation results in 6.22.

$$E_{th} = \frac{6.4 \times 10^{-13} I_{th}\tau}{6.16 + 6.4 \times 10^{-13} I_{th}/P_{cr}} \quad (6.22)$$

[17]

For the general case of our system with 100 fs pulses into a material with $P_{cr} = 1\text{ MW}$ and a typical intensity of 10^{13} W/cm^2 results in a per pulse energy value of 1 nJ of the order expected for such a material.

Figure 6.3 shows a Multi Criterion Analysis between a high NA objective (approx 1.4) and a low NA objective (approx 0.2). The preference towards the high NA lens for its small focus and ability to machine below the surface of the material have scored it favourably in comparison to the low NA objective resulting in its implementation in our system. It should be noted however that the system also has the capacity for low NA machining due to the Two Photon Polymerisation process putting more emphasis on speed of machining and having no requirement for bulk machining and hence implementing that method of scanning. Due to the identical beam path between the fabrication processes, the methods of beam targeting are interchangeable. This method is discussed in chapters 8 and 22.

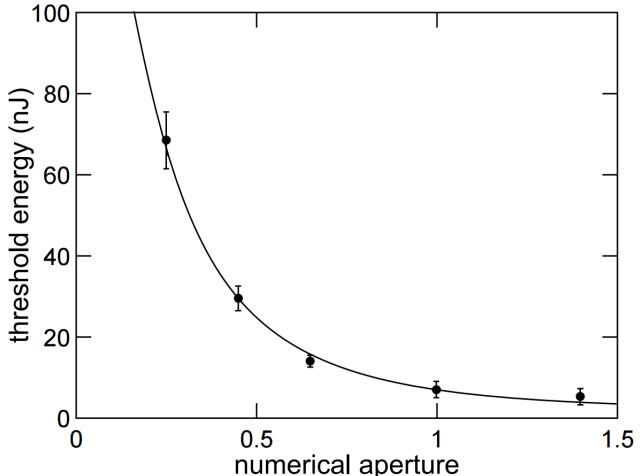


Figure 6.8: Dependence of threshold energy upon NA of objective lens [17]

-	-	Low NA		High NA	
Comparison	Weight	Points	Score	Points	Score
Focus size	5	1	5	3	15
Working distance	2	3	6	1	2
Bulk Machining	4	1	4	3	12
Immersion	3	3	9	1	3
Cost	2	2	4	2	4
Totals	-	-	28	-	36

Table 6.3: High VS Low NA Multi Criterion Analysis table

6.6 Capabilities

6.6.1 Resolution

By using the stage for targeting as opposed to galvanometers and by choosing a high NA objective, part 6.5.4.2 sees the focal spot reduced to 716nm in diameter. This is within the specification that was required and finer manufacturing resolutions can be achieved by careful laser power management in order for only a fraction of the focus to be able to ablate.

6.6.2 Power

As noted in section 20.3.2 from source to objective 53% of beam power is transmitted to the objective, where 70% of that is then passed through to the sample. Hence the maximum pulse energy of $150\mu\text{J}$ generated by the laser falls to approximately $55\mu\text{J}$ at the focus. This results in an intensity of approximately $1.36 \times 10^{17}\text{W/cm}^2$ greatly exceeding the 10^{14}W/cm^2 intensity required to ablate the vast majority of materials. The reason the power exceeds the requirements by such a margin is due to the losses that will be incurred at the SLM when in use. In multiple foci generation the power must be divided by a number of points, each of which need to exceed the ablation threshold. Similarly a Bessel beam has concentric rings around the central useful focus these carry energy and hence the pulse energy must be higher to reach the needed intensity at the focus.

Furthermore, if a different objective were to be retrofitted at a later date then an increased focal spot volume will not reduce the intensity to below threshold levels.

6.6.3 Speed

The speed of fabrication is limited by that of the stage's maximum speed of which is 30mm/s per axis according to chapter 22, however it cannot run at this speed permanently, time for acceleration and retardation is also required. A perhaps more important restriction however is based upon the number of pulses deliverable per second and how they interact / overlap. A single pulse will create a single crater in the sample. If a channel is needed to be machined, then a series of these craters must be overlapped with the correct geometry for a smooth and uniform channel. Figure 6.9 shows the difference between large and small overlap. This overlapping also affects the material removal and the surface roughness of the finished piece. Eleri Williams finds that values of overlap in the $60 - 70\%$ zone offers the best surface finish while 85% provides the greatest material removal [24]. Choosing a pulse overlap between these regimes of 75% sees a separation distance of 175nm which in order to travel at the maximum stage speed would require a repetition rate of 171.4kHz . This exceeds the rep rate for the amplified laser by more than an order of magnitude, but is well within the reach of the seed laser output. The amplified laser can reach 10kHz which corresponds to a maximum travel speed of 1.75mm/s . This is true whether stage or galvos are used (however the greater acceleration of galvos will still save time). As a consequence of this a choice must be made as to whether power or speed is of priority.

6.7 Summary

Given the need for a nanometre accurate reductive manufacturing technique to be included in the fabrication system, femtosecond Laser Micro-Machining was ultimately chosen. The ultra short pulses allow modification to a target material with minimal damage besides that which is intended. An objective with a high numerical aperture allows a focal spot diameter of 700nm , Rayleigh range of 387nm and modification to the bulk of transparent samples without damage to the surface. Metals and dielectrics can be machined, from aluminium and steel, to borosilicate glass, silicon wafers, graphite, and even diamond! Furthermore, by control over the SLM, multiple foci can be generated allowing parallelisation of the machining process as well as creating wave fronts with useful properties such as Bessel beams, further increasing the scope and potential uses that this process can be used for.

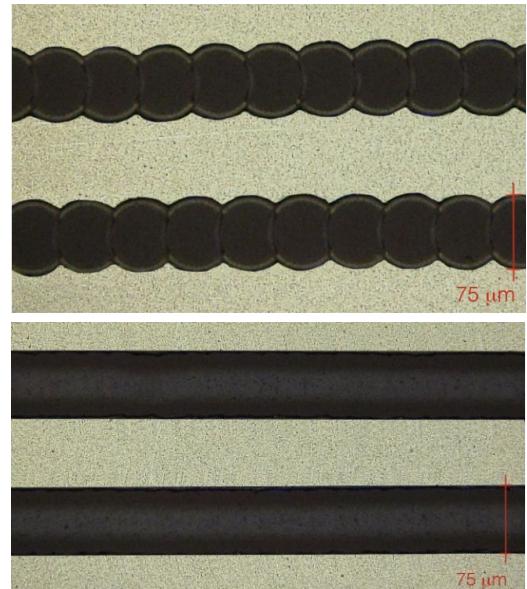


Figure 6.9: Top: 30% Bottom: 97% pulse overlap [23]

7 Stereolithography

Stereolithography (SLA) is the first rapid prototyping technique invented and is still being widely used for its low cost. In our system, it is designed to complete the overall shape of the parts with feature size of 10 micrometers. Fine details will be then completed by Two-Photon Polymerisation (additive manufacturing) or Laser Micromachining (subtractive manufacturing). As the typical overall size of the structure manufactured by TPP(two-photon polymerisation) is below 1mm^3 , including Stereolithography can expand the size range of the machine. Furthermore, the speed of fabrication will be improved by using Stereolithography. For SLA250/50 machine, which is an industrial 3-D printer produced by 3D Systems company, the typical laser scanning speed could reach 750 millimeters per second[25]. The general building time of a model by using Stereolithography is 12 hours[26]. We can therefore achieve large building size and fine details at the same time.

7.1 process overview

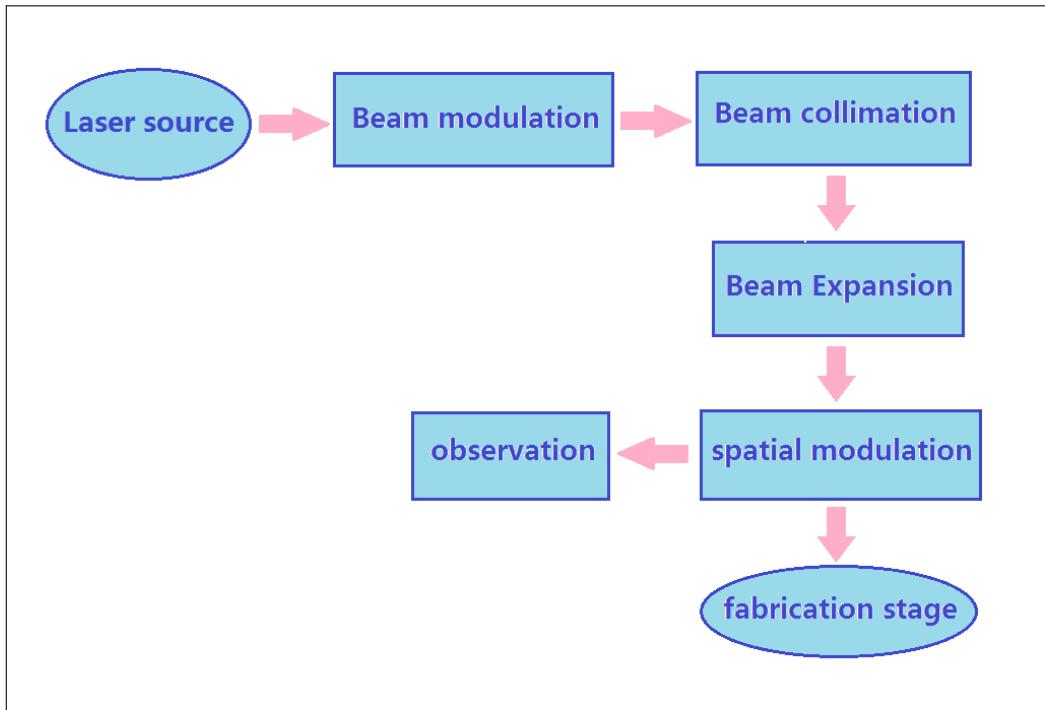


Figure 7.1: Flow Diagram of Stereolithography

The basic steps of SLA can be shown as figure 7.1 above. For beam modulation and collimation, it can be done by an integrated laser generator. An Off-the-shelf product, Genesis CX 355-250STM continuous wave laser from coherent company is chosen to be the laser source.

For beam expansion, there are also plenty of choices of off-the-shelf products. As it can be simply achieved by setting two lenses at a fixed distance, we decided to customize our own beam expanders according to magnification we need. It is explained in Chapter 20.7 (Beam Expander).

For spatial modulation, there are several options available and feasible. The first method is to use

galvanometers and an objective lens to focus the laser beam to the desired place. This is the most widely used method in current commercialized SLA printers. By selecting the components carefully it can be either a cost-effective personal printer or a costly industrial printer with high resolution up to $1\mu\text{m}^2$. The second method is simply to fix the laser source and move the fabrication platform instead. This method is not very common as the working area of SLA can be large. Setting up a moveable stage takes up more than twice of the width and the length of the work space which result in 4 times of the area occupied. Furthermore, the movement of the stage could be really slow as the mass of the stage is large. It greatly reduces the fabrication speed of SLA which makes this set-up undesirable. Another method is to use mask projection. By shining a laser on a computer-controlled digital mask, light does not cure the entire layer but only the desired area of that layer. This digital mask can be realized by Digital Micromirror Device (DMD). DMD consists of arrays of mirrors which can be switched on or off individually. As the size of a single micro-mirror is in the order of micrometers. This fabrication method is also called Projection Micro-Stereolithography.

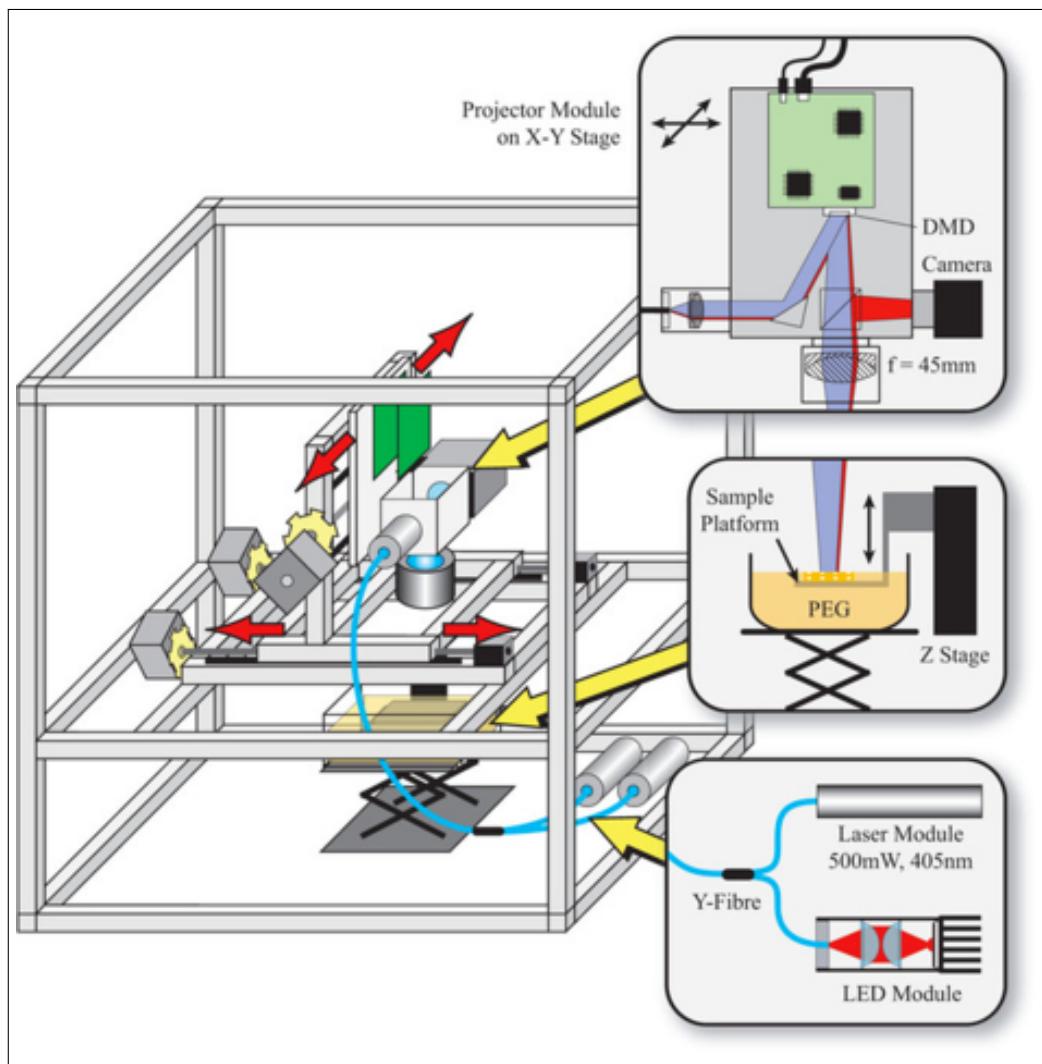


Figure 7.2: Projection Stereolithography with the use of DMD

Figure 7.2 shows one way of masked SLA set-up. As DMD cannot achieve both large working area and small pixel sizes simultaneously. It can be combined with moveable stage set-up: Each is splitted into several subsections. Moveable stage align the laser and DMD to a specific subsection and only polymerise the resin within this section. It repeats until all sections in this layer are polymerised and moves on to the next layer. The finest resolution can be achieved by this set-up depends on the size of each pixel in DMD - which is $10\mu\text{m}^2$ mentioned in the report. The difficulty of this method lies in aligning the DMD and the workplace in moving stages between subsections. It requires relatively high accuracy for stage positioning.

As the use of DMD increase the complexity in configuration and calibration of the stage while it does not have any advantages in resolution. In this system we use the combination of Galvanometers and specific objective lens (explained in the next subsection).

Observation or imaging system is usually excluded from stereolithography as the size of fabrication is relatively large. The vat and protection case is often transparent hence the user can observe the fabrication process with the naked eyes. However, as in our system the desired feature size of SLA is about 10 micrometers, it is better to have live imaging system to monitor the fabrication process therefore the user can actually observe if there are any error occurs. A simple CCD (charge coupled device) camera is used and a simple red LED is added as the light source for illumination.

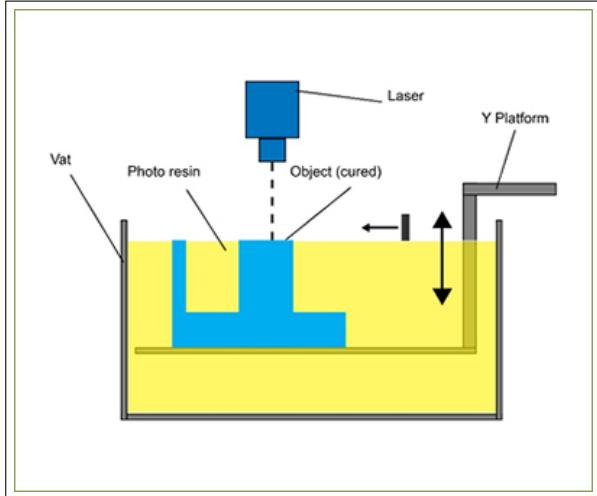


Figure 7.3: The top down set-up of stereolithography[27]

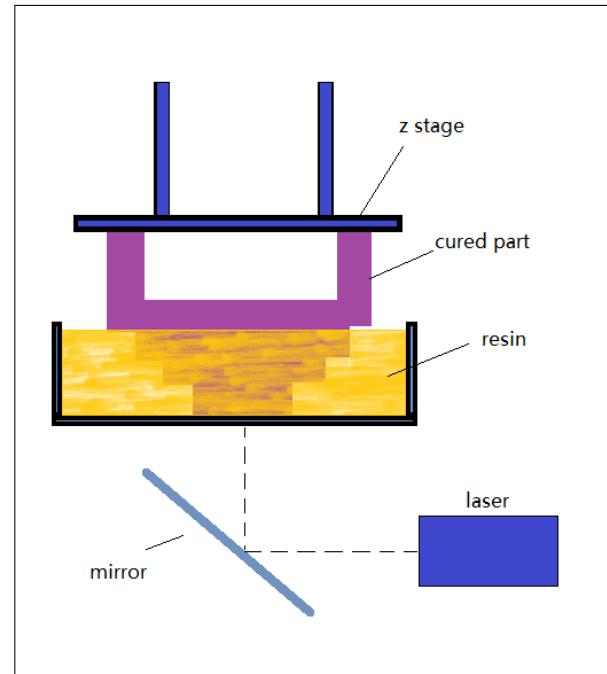


Figure 7.4: The bottom up set-up of stereolithography[28]

For fabrication stage, there are two different types of set-up: top-down (normal) shown in figure 7.3 and bottom-up (inverted) shown in figure 7.4. For normal set-up, laser is shone above the fabrication

stage and polymerisation process occurs at the surface of the resin. Once a layer is finished, the platform in the vat will move down by amount of one layer thickness. The process is repeated until the part is finished. In contrast, inverted set-up has incidence of the laser beam from below. The polymerisation process happened at the bottom of the vat. The platform moves upwards layer by layer in fabrication process. The finished part will be upside down. There is one thing to note: polymerisation happens at the bottom of the vat. To prevent cured polymers from sticking to the bottom, a non-stick film has to be placed at the inner surface of the tank. The most popular material is Fluorinated ethylene propylene (FEP) films which was used in the aerospace industry to keep cured polymers and metal molds separate[29].

To compare these two set-ups, as we can see in figure 7.4 the inverted set-up take less space therefore the resin needed to fill the vat is less than normal set-up. It also avoids the problem of controlling resin level which would be a problem in normal set-up. For these reasons, most of the recent commercialized desktop SLA printers tend to use inverted set-up such as the Form 2 3-D printer from FormLabs or Pegasus Touch 3-D printer from Full Spectrum Laser. However, as the need for transportation between different fabrication techniques, it is easier to be realized using normal set-up. Furthermore, although normal set-up requires more resin to fill the vat, as SLA is an additive manufacturing technique(it means that the component is built by depositing materials rather than cutting materials from a substrate), it does not consume any additional resin in manufacturing process. The size of the machine is also not a huge concern. Therefore normal set-up is chosen.

7.2 Optical Design

Figure 7.5 shows all of the optical components passed by ultraviolet laser beam. The full name of each component can be found in Appendix A. The laser goes through spectrometer **SP** and CCS Cosine Corrector **CC3** for continuous wave correction. It then goes through the half-wave plate **AHWP3**, polarizing beam splitter **PBS3** for power control. Extra laser power is directed to Beam Dump **BD** to avoid hazards from heat accumulation. A shutter **S5** is present as a fail-safe device. It then goes through the beam expander **L19** and **L20**, two 1-dimensional galvanometers **X2,Y2** and an objective lens **F-THL** to reach fabrication stage.

7.2.1 F-theta lens

7.2.1.1 why is F-theta lens needed?

Microscope objective plays significant role in imaging system as it magnifies the sample to create image with high resolution. It is explained in details in Chapter 17. In fabrication system, it is also useful as it focuses the laser beam to achieve fine feature size. It can be classified as three types:

From Figure 7.6 it can be seen that the first type of objective lens produces a curved focal plane.

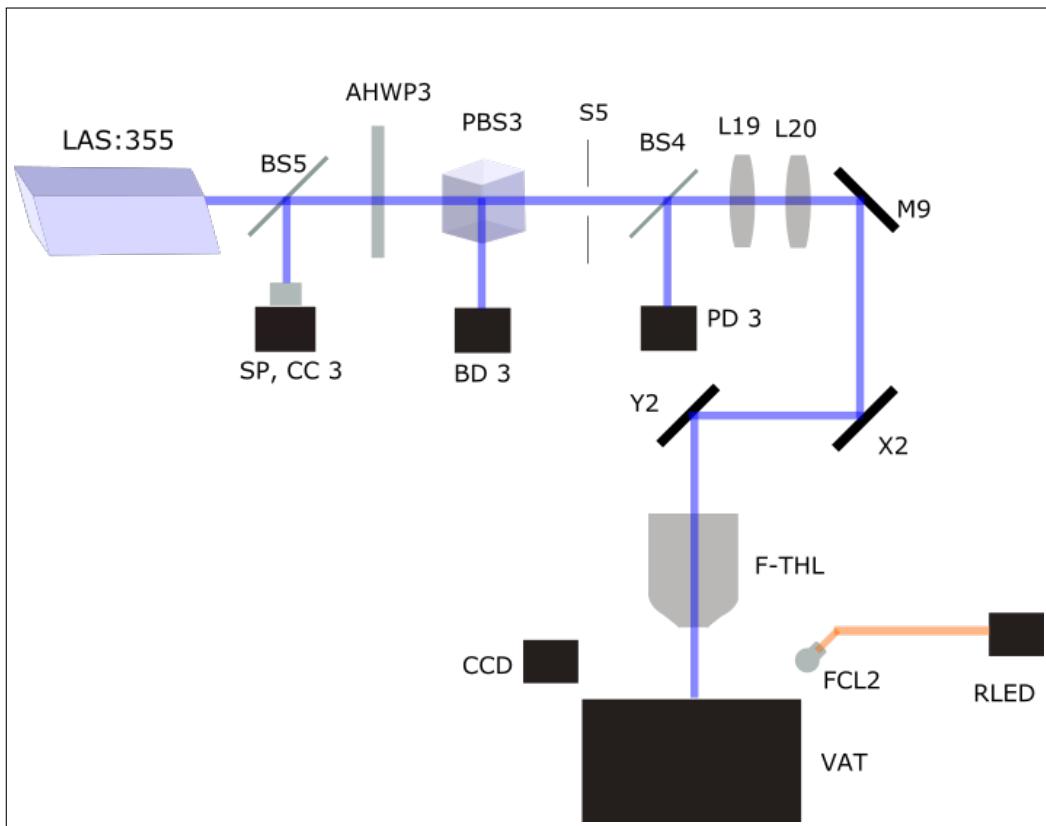


Figure 7.5: Optical Layout for SLA

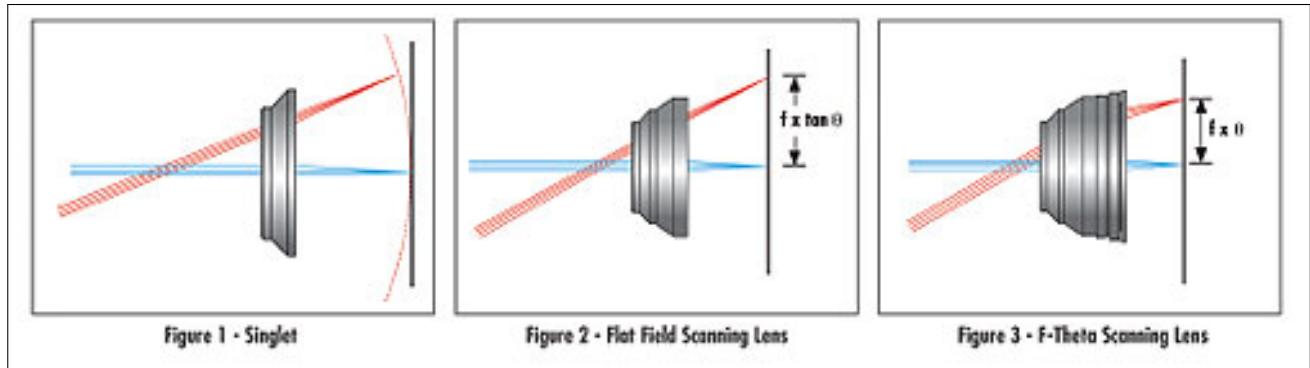


Figure 7.6: Three different types of objective lenses[30]

It is absolutely undesired in SLA as laser beam cannot focus on the platform anywhere except the centre part. It will result in increasing spot size (i.e. feature size) or more severely inability of polymerisation.

The second objectives - flat field lens keep focus on the same working plane. The distance d' from centre to the focus point is given by:

$$d' = f \times \tan \theta \quad (7.1)$$

where f is the focal length of the lens and θ is the incidence angle of the laser beam in radians.

As the angle of incidence is controlled by galvanometers and is directly proportional to electrical

current input (it will be explained further in chapter 7.2.2), the $\tan\theta$ term adds non-linearity to spatial modulation of the beam. In this case, to maintain the same curing time for each spot we want to polymerise, the speed of the galvanometer will be constantly changing which increases the complexity of controlling.

The third lens fixed the problem of non-linearity. For F-theta lens distance from centre to the focus point is directly proportional to the incidence angle. This means that by adjusting the input current of galvanometers, the distance of focal point will vary linearly which makes the control easier.

However, f-theta lens is a type of compound lens. The design and calibration could be very complicated. We decided to select an off-the-shelf product and design the SLA system around it.

7.2.1.2 F-theta lens selection

The first parameter to consider is the feature size that f-theta lens can achieve. As TPP realizes the feature size down to 100nm , it is affordable for SLA to have feature size to around $10\mu\text{m}$. Fortunately there are still few f-theta lenses which can achieve this level of accuracy. One of them is S4LFT4160/075 telecentric f-theta lens from Sill Optics. The table 7.1 shows the specifications of the lens.

design wavelength [nm]	355
effective focal length [mm]	167.2
max. entrance beam-Ø [mm]	15.6
optical scan angle [\pm deg]	15.6
scan length [mm] (1 mirror system)	113.1
working distance [mm]	264.4
aperture stop distance [mm]	37.2
diameter of aperture stop [mm]	10.0
scan area for a 2 mirror scan system	80×80
mirror distance from lens housing for mirror 2 / mirror 1	30.7/43.7
max. telecentricity error [deg]	0.6
total transmission [%]	> 97
lens material	fused silica
damage threshold	$1.0\text{J}/\text{cm}^2$ per 1ns pulse at 50Hz $1.0\text{GW}/\text{cm}^2$ cw
weight [kg]	1.94

Table 7.1: datasheet of S4LFT4160/075 f-theta lens

One more thing to note is the telecentricity of S4LFT4160/075 f-theta lens. The laser beam always strike the image plane perpendicularly no matter what is the beam deflection angle at entrance of the lens. As figure 7.7 shows, when laser beam hit focal plane in some angle, there will be elongation in

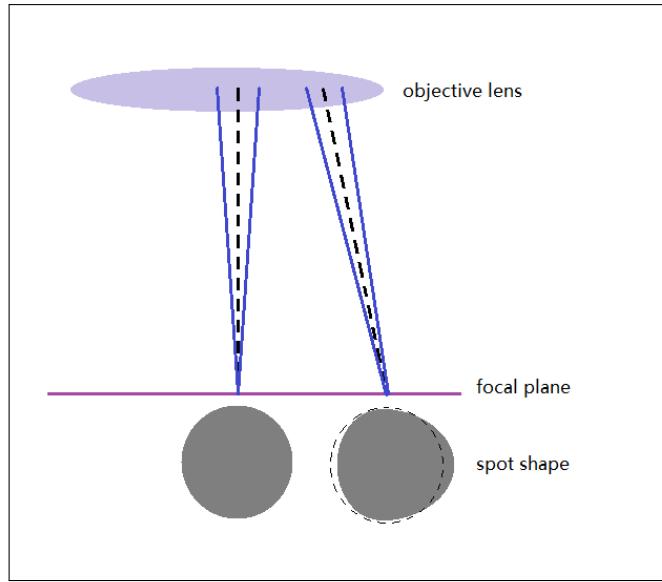


Figure 7.7: Spot shape distortion by angled beam projection

the diameter of the spot in the direction of beam deflection[31]. It not only distorts the shape of the spot but also increase the spot area which increase the feature size that can be achieved. Telecentric lens limits the distortion in spot shape at the edge of the focal plane, which is desirable.

7.2.1.3 calculation of actual spot size

As the numbers appear on the data sheet is always the best case scenario. We need to calculate the actual spot size we can achieve. Assuming that the laser beam has Gaussian profile. The focus spot size can be calculated by following equation[32]

$$\phi_{spot} = \frac{1.83 \times \lambda \times f}{\phi_{beam}} \quad (7.2)$$

where ϕ_{spot} is the spot diameter at focus, λ is the wavelength of laser, f is the focal length of the f-theta lens, ϕ_{beam} is the laser beam diameter at the entrance of the lens. Substituting values from the data sheet, we can calculate the spot diameter to be $10.86 \mu m$, which is an acceptable value. Noting that the spot diameter is inversely proportional to entrance diameter. In order to achieve spot size as small as possible, we need to expand the beam to largest possible value before it enters f-theta lens.

However, the maximum beam diameter can be only achieved with 1 mirror scan system. For 2 mirror scan system, x and y axis scanning mirrors are required to be placed symmetrically on either side of the aperture stop. Doing this will limit the maximum beam diameter running across the aperture. In figure 7.8), purple dotted line shows the chief ray of non-deflected laser beam and

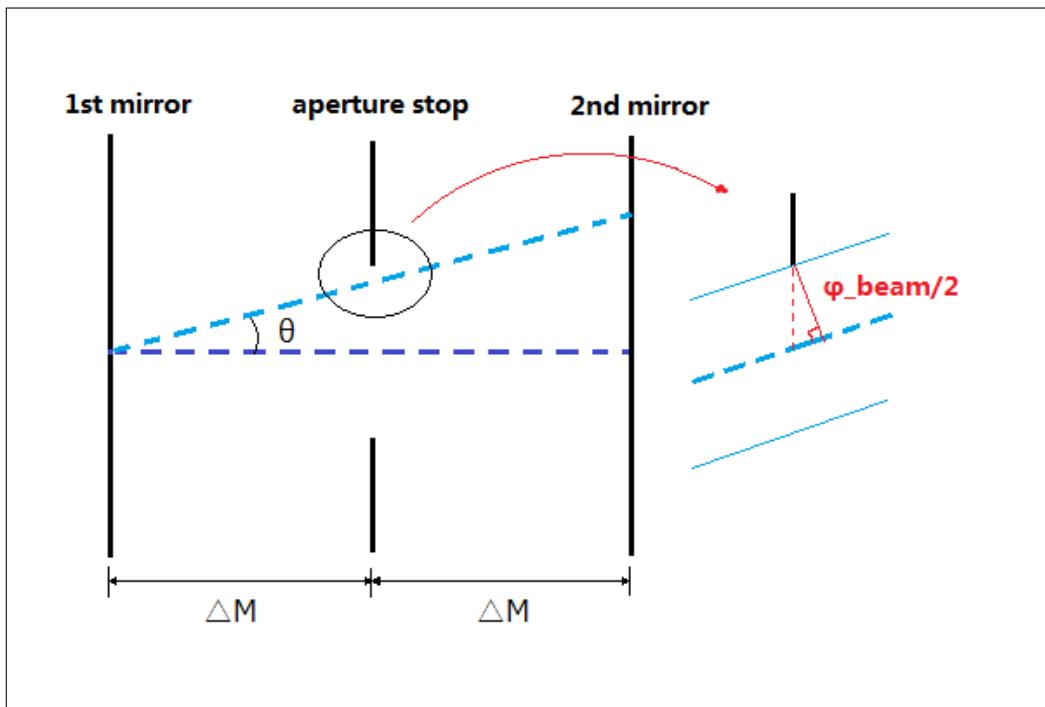


Figure 7.8: Relationship of the maximum beam width and the maximum deflection angle

blue dotted line shows the chief ray of deflected laser beam. In non-deflected beam condition laser goes through the centre of the aperture stop. We can then get the relationship between the aperture diameter, maximum beam deflection angle and the beam width:

$$2 \times (\Delta M \tan \theta + \frac{\phi_{beam}}{2 \cos \theta}) = EP \quad (7.3)$$

where ΔM is the distance between either mirror and the aperture stop, θ is the maximum beam deflection angle (which is optical scan angle on the data sheet), ϕ_{beam} is the maximum beam diameter and EP is the diameter of the aperture stop (also called entrance pupil).

By using equation 7.3 we can calculate the actual maximum beam width to be 9.28 millimeters. Then we substitute this value into equation 7.2. The actual spot size calculated is 11.81 micrometers, it is slightly higher than before but still acceptable.

7.2.1.4 The scan area and distortion

As we can see from the data sheet(table 7.1). The scan length of 1 mirror system is 113.1 mm. For a circular lens, the scanning pattern should also be central symmetrical. This implies that for 1 mirror scanning, the scan area is a circle with diameter equals to 113.1 mm. However, as in 2-D scanning system, each mirror controls deflection of beam in one axis which means that they are independent to each other. It implies that the scan area for a two mirror system is a square. However, the scan area of two mirror cannot exceed the scan area achievable by one mirror system. Therefore the maximum scan area of 2-mirror system is the cyclic square of 1 mirror scan area which is shown in figure 7.9.

We can then easily calculate the scan length for two mirror system:

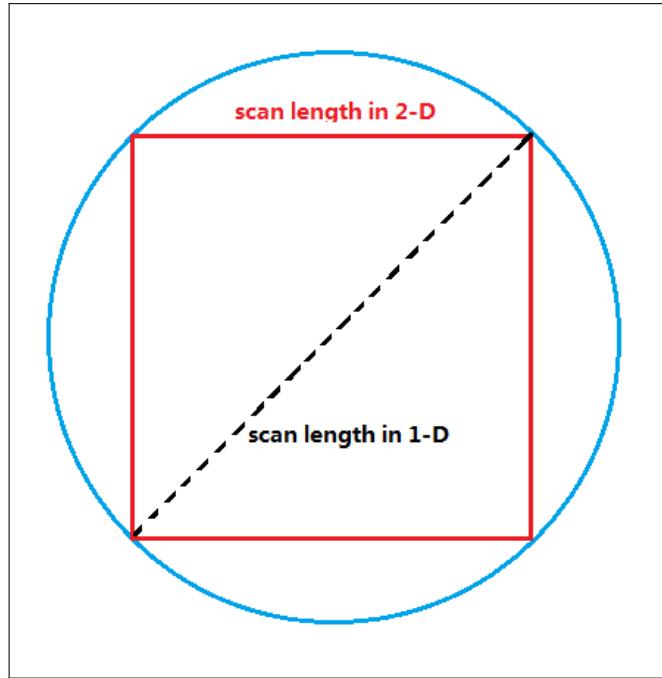


Figure 7.9: Relationship between 1 mirror and 2 mirror scan area

$$\begin{aligned}
 L_{2 \text{ mirror}} &= \frac{L_{1 \text{ mirror}}}{\sqrt{2}} \\
 &= 79.97 \approx 80
 \end{aligned} \tag{7.4}$$

The numbers on data sheet (table 7.1) also proves the result.

In two-mirror system, however, the beam deflected by the first galvo travels farther than the beam deflected by the second galvo. It means that if the angular position of two galvos are the same, the deflection in the first axis will be longer than the second axis. For this reason the scan area is no longer a square. In fact, it looks like the graph on the left of the figure. There is another type of distortion in f-theta lens. The distortion curve of f-theta lens is a quadratics[33] which becomes more and more obvious when the scan angle becomes large. It is shown in the middle of figure 7.10. This is usually called F-theta distortion. Two types of distortion superimposed each other and create a 'barrel-pillow' shaped pattern shown on the right. The distortion is not significant but it has more effect as laser scans to the edge of the working area. Therefore it is better to be processed and filtered by computer.

7.2.2 Galvanometer for Stereolithography

7.2.2.1 Mechanism of galvanometers

Galvanometer is the most common way of steering mirrors in the laser scanning system. It is invented in 1837 by Pouillet and the word is attributed to Italian researcher Luigi Galvany [35]. However, laser

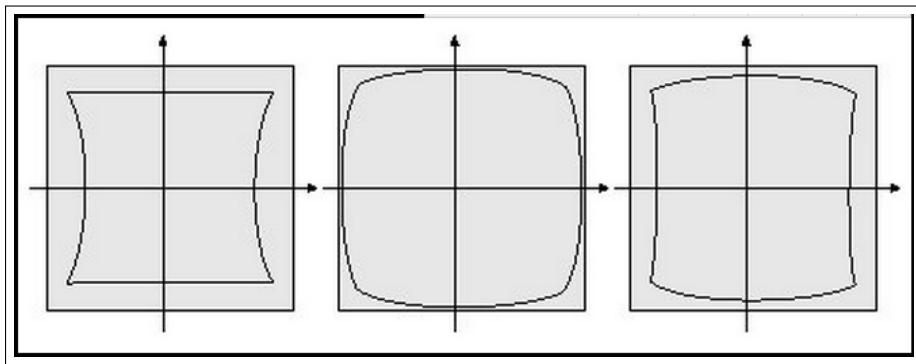


Figure 7.10: The distorted scan area of F-theta lens[34]

scanner is not the original use. As the figure 7.11 shows Galvanometer is first used to detect tiny current by observing the deflection of the pointer installed on the rotational spring. The mechanism is fairly simple: the electrical coil attached to a rotational spring is placed in a magnetic field (normally generated by permanent magnets). When there is electrical current flowing in the coil, a torque will be generated and rotates the coil until it balances with the force generated by torsional spring. As magnetic force is proportional to the input current. The theoretical deflection angle is also directly proportional to the input current.

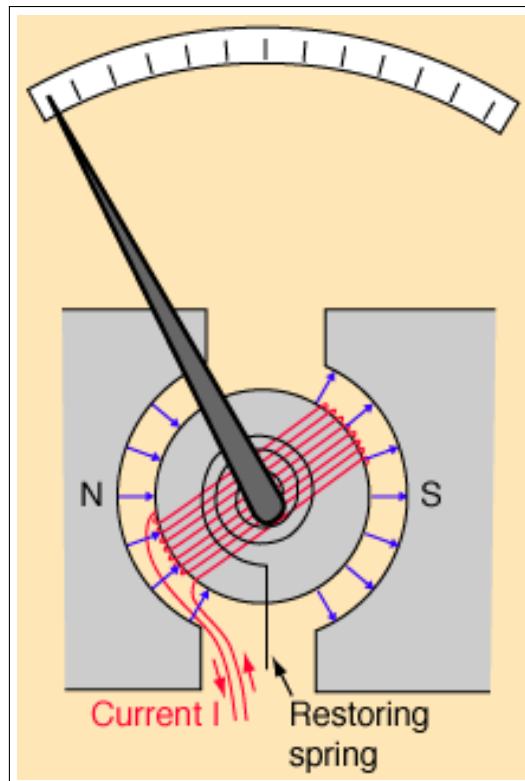


Figure 7.11: The mechanism of galvanometer [36]

Modern galvanometer use exactly the same principal, the pointer is simply replaced by a mirror installed in one end of the shaft so it can direct laser beam to desired angle. Laser scanning galvanometers are often called galvos for distinguishing two different usages. I am going to design

around the moving magnet galvos, which uses magnet as rotary shaft and armature with coils as stator.

7.2.2.2 Control design using PID tuner

In this laser scanning system, the angular position and response time of galvos need to be extremely accurate. Therefore closed-loop control system is necessary. In this section I am going to derive the system transfer function and try to design a controller with accurate response.

We start from the equation describing dynamics of the galvos.

$$T_{magnetic} + T_{elastic} + Torque_{friction} = J\ddot{\theta} \quad (7.5)$$

$$K_T i - k\theta - \mu_f \dot{\theta} = J\ddot{\theta} \quad (7.6)$$

$$\text{where } K_T = n \int B dA$$

Where T is the torque, J is the moment of inertia of the rotor shaft (including the mirror), θ is the angular deflection of the shaft, K_T is the torque constant, k is the rotational stiffness of the spring, μ_f is the coefficient of friction, n is the number of turns of the coil, B is magnetic field strength where coil is placed, i is the electrical current in the coil, A is the area of the coil in the magnetic field.

There are three torques acting on the rotating shaft in the galvos: Magnetic torque>equals to magnetic force multiplied by the lever arm), elastic torque(proportional and opposite to angular displacement) and frictional torque(proportional to angular velocity $\dot{\theta}$ of the shaft). Then we need to find the relationship between input voltage and current generated.

$$V = L \frac{di}{dt} + iR + K_T \dot{\theta} \quad (7.7)$$

Where V represents input voltage, L and R is the inductance and resistance of the coil respectively. On the right hand side of the equation, the first term represents the voltage induced by coil inductance. The second term represents the voltage dissipated in coil resistance. The third term is the back EMF induced by magnetic field flux.

Then we combine these two equations together, take the Laplace transform of the equation and reach the transfer function from input voltage V to angular displacement θ :

$$\frac{\Theta(s)}{V(s)} = \frac{1}{sT + \frac{(sL+R)(s^2J+s\mu+k)}{T}} \quad (7.8)$$

or

$$\frac{\Theta(s)}{V(s)} = \frac{C_1}{C_2s^3 + C_3s^2 + C_4s + C_5} \quad (7.9)$$

$$C_2 = LJ,$$

$$C_3 = JR + \mu L,$$

$$C_4 = kL + \mu L + T^2.$$

$$C_5 = kR,$$

The dynamics of system is input into SIMULINK model as figure 7.13 shows. Then I feed it with closed-loop PID(proportional-integral-derivative) control as Figure 7.12. PID controller is a common feedback control system. A simple feedback loop takes the difference between reference point(desired value) and the output multiplied by a gain as the error and feed it to system input. PID controller however, takes the integral, the first derivative of the output and the output itself, multiply them with different gains respectively and sum them up. Then the sum is subtracted from the reference point to get the error. By adjusting three gains, usually a stable response can be achieved.

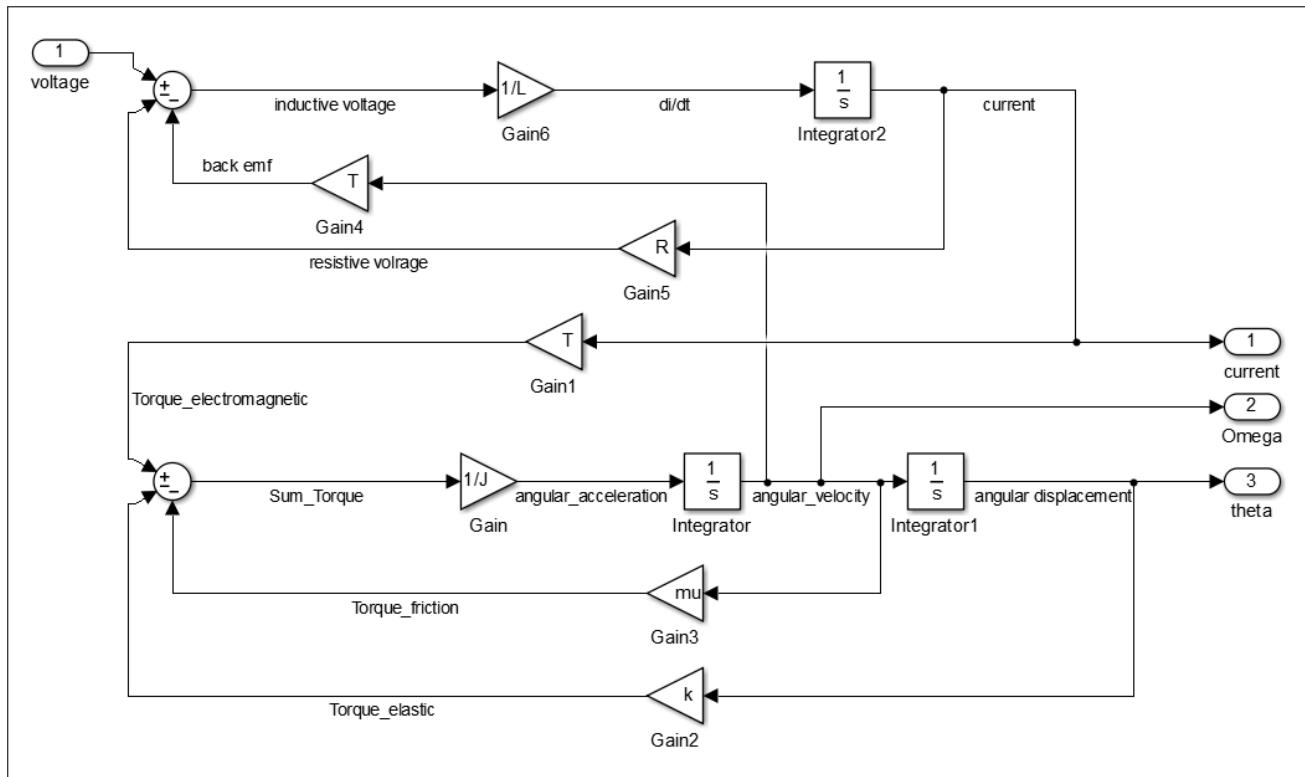


Figure 7.12: The system model of Galvos

By using the tune function in PID controller, I have worked out the parameters of PID controller for the shortest settling time. As we can see in Figure 7.14 , a stable closed loop response is achieved

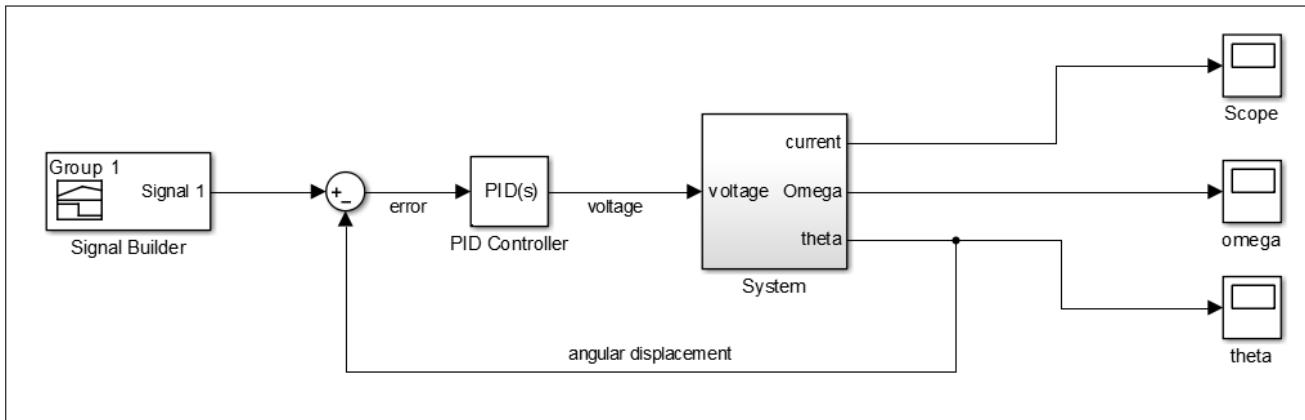


Figure 7.13: Closed loop control simulation of Galvos

and settling time is 0.05s . The overshoot is only 1.69% which is not significant. However, the system is underdamped as there are some oscillations before settling down, and oscillation does remain for long time which makes the control less ideal.

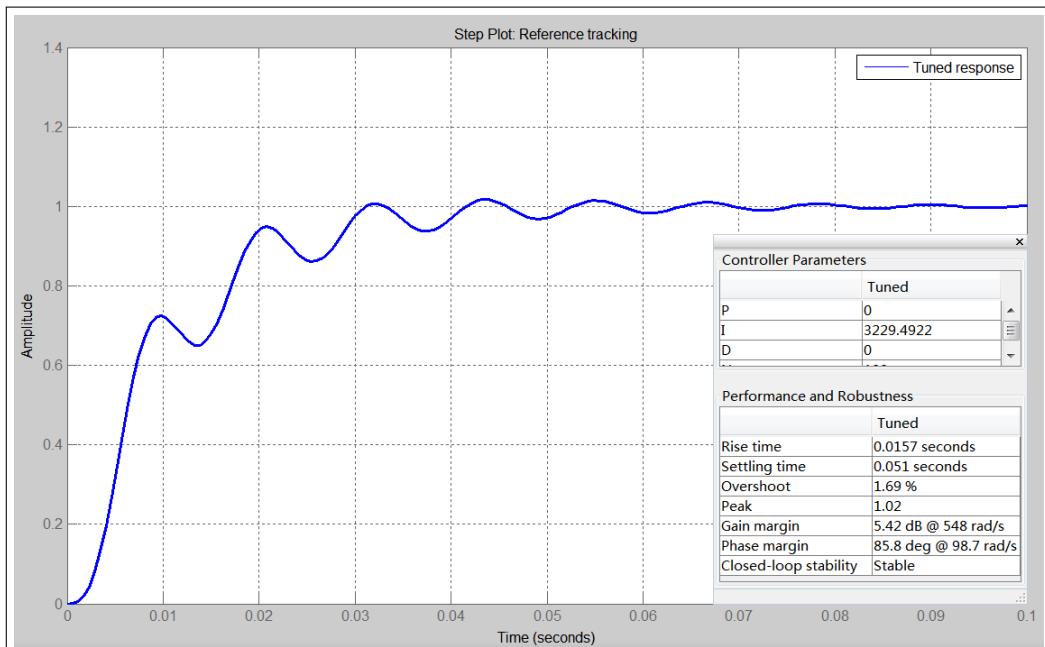


Figure 7.14: Result of tuning PID controller

The next step is to see that if the control is robust enough for noises. For noise simulation I choose the signal-to-noise ratio as 60 dB. This means that the average power of the signal after normalization is 10^6 times of the average power of the noise. This implies that the ratio between root mean square(RMS) value of the amplitude of normalized signal and noise is 1000. This can be seen in the equation 7.10 below.

$$SNR_{indB} = 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right) = 20 \log_{10} \left(\frac{A_{signal}}{A_{noise}} \right) \quad (7.10)$$

Moment of Inertia[J]	kg m ²	5e-7
Torque Constant[K _T]	Nm/A	0.02
Rotational stiffness[k]	Nm/rad	0.15
Coefficient of friction[μ]	Nms/rad	4e-7
Coil Resistance[R]	Ω	4.5
Coil Inductance[L]	H	6e-4

Table 7.2: constants of galvos used in simulation[37]

By adding the noise, the step response becomes figure 7.15. There are some oscillation with maximum value of 2% which is not very much but as we want to achieve accurate control.

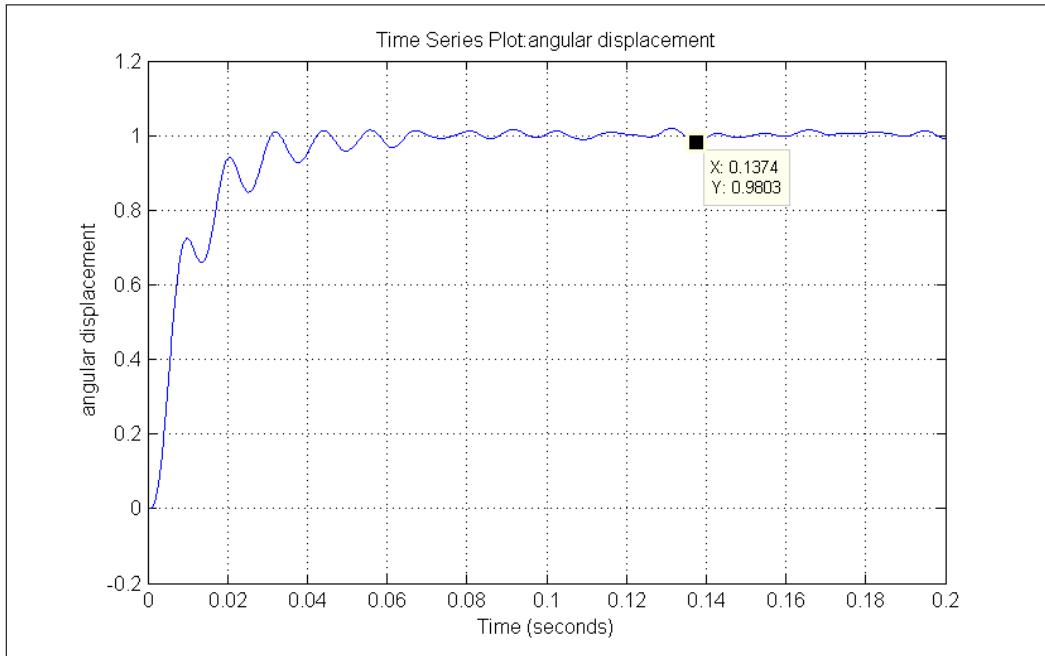


Figure 7.15: Result of tuning PID controller

7.2.2.3 Control design using state-space model

To achieve better control, I decide to use pole-placement method to see if response can be improved. First I am going to create a SS(state-space) model. By observing figure 7.13, we can see that there are three states in this system: angular displacement θ , angular velocity $\dot{\theta}$ and electric current I . From equation 7.6 and 7.7 we can derive equation for each state:

$$\frac{d}{dt}\dot{\theta} = \frac{1}{J}(-k\theta - \mu\dot{\theta} + K_T I) \quad (7.11)$$

$$\frac{d}{dt}I = \frac{1}{L}(-K_T\dot{\theta} - RI + V) \quad (7.12)$$

From two equations above we can write down the state-space equations as follows:

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \quad (7.13)$$

$$\mathbf{y} = \mathbf{Cx} \quad (7.14)$$

$$\text{where } \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{k}{J} & -\frac{\mu}{J} & \frac{K_T}{J} \\ 0 & -\frac{K_T}{L} & -\frac{R}{L} \end{bmatrix},$$

$$\mathbf{B} = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{L} \end{pmatrix},$$

$$\mathbf{C} = (1 \ 0 \ 0).$$

In equation 7.13 and 7.14, the state $\mathbf{x} = (\theta \ \dot{\theta} \ I)$, input $u = V$, output $y = \theta$.

To use pole placement method in a state-space system, it has to be controllable. Controllability can be checked by deciding if the following matrix is full ranked (parameters in table 7.2 is till in use):

$$\begin{aligned} \mathbf{M} &= [\mathbf{B} \ \mathbf{AB} \ \mathbf{A}^2\mathbf{B}] \\ &= \begin{bmatrix} 0 & 0 & 6.667 \times 10^7 \\ 0 & 6.667 \times 10^7 & -5.001 \times 10^{-5} \\ 1.667 \times 10^3 & 1.250 \times 10^7 & 9.153 \times 10^{10} \end{bmatrix} \end{aligned} \quad (7.15)$$

The matrix is full rank(all of the rows in the matrix are linearly independent to each other) therefore this system is controllable.

Now find the open loop pole of current state-space model. It is simply the eigenvalues of matrix \mathbf{A} . It is plotted in figure 7.16. We can see that there are two dominant slow poles near to the imaginary axis. The imaginary part is relatively high therefore implies oscillatory behaviour. To reduce oscillation I decide to put the poles much closer to real axis and shift two dominant poles to the left. We are going to design the controller so that input $u = -\mathbf{Gx}$. Let $\mathbf{G} = f\mathbf{k}^T$, assuming $f = 1$, let desired poles be $p_1 = -5000, p_2 = -200 + j10, p_3 = -200 - j10$. Calculations are shown as below:

$$\mathbf{q}_i = [p_i \mathbf{I} - \mathbf{A}]^{-1} \mathbf{B} f, \quad \text{for } i = 1, 2, 3 \quad (7.16)$$

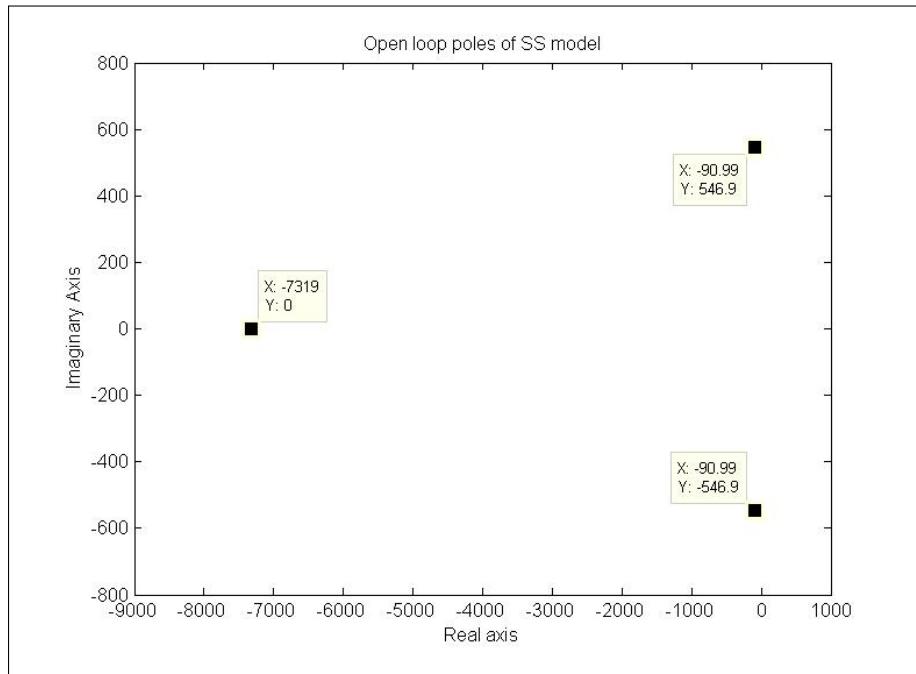


Figure 7.16: Open loop poles of state-space model

$$\mathbf{k} = \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \mathbf{q}_3^T \end{bmatrix}^{-1} \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \quad (7.17)$$

The resulting controller is $\mathbf{G} = [-21.2889 \quad 0.0060 \quad -1.2605]$. In order to verify if it gives a good response, again SIMULINK is used to simulate the step response of the controller. Figure 7.17 shows the simulation plant of the SS model. Step size is set to 0.001 rad (1 mrad), which is typical step size of galvos.

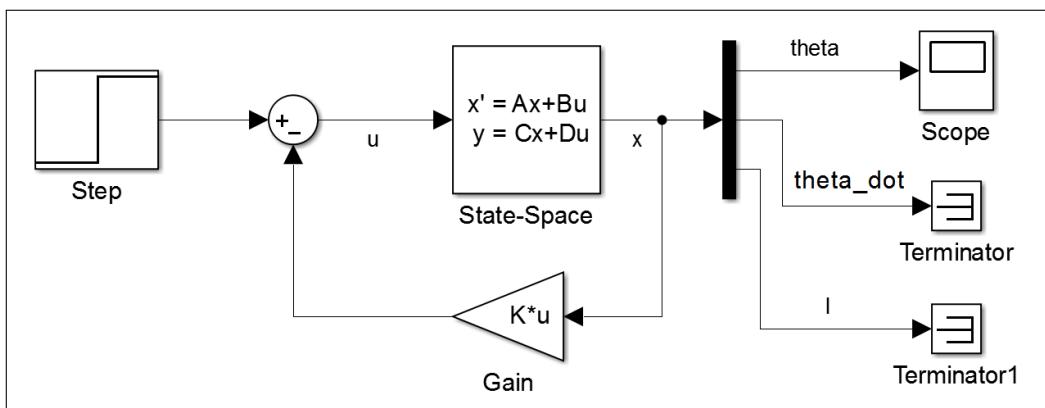


Figure 7.17: Simulation plant for State-space model

As figure 7.18 shows, the response looks great as there is hardly any oscillations and the settle time is approximately 0.04 seconds. However, there is a steady-state error of 66.8%! It means that the galvo could not reach desired angular position. The reason is that by using state controller, the

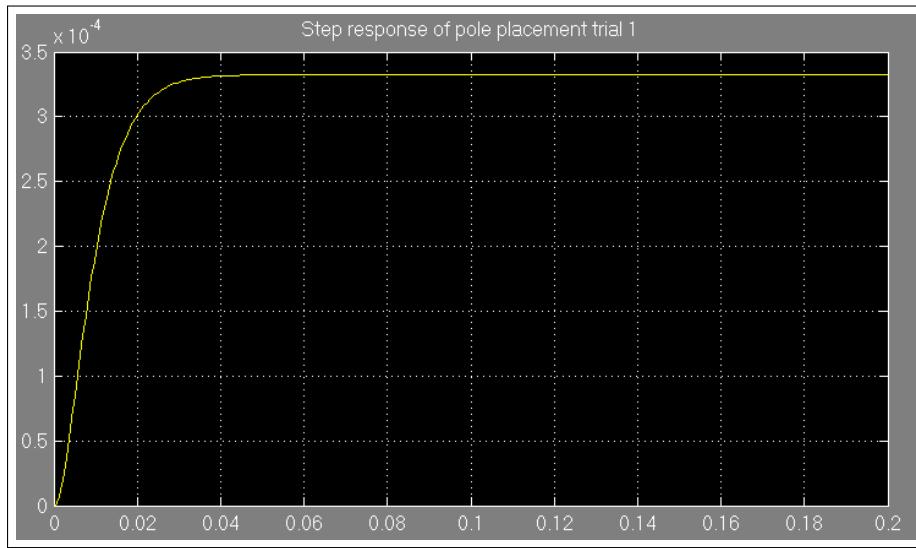


Figure 7.18: step response by pole placement control. X-axis is the time(in unit of seconds) after a step size of 0.001 radians fed to the input while Y-axis shows the angle of deflection of the galvo(in unit of radians).

input is no longer proportional to error in angular position. If there is a state which is not what we want to control is non zero in steady state, the error between reference signal and controller gain will deviate the angular position from desired position. In this case, the state generating steady state error is current I . As there has to be a Torque balancing the torque of torsional spring, the electrical current will be non zero in steady state.

One way to solve this problem is to use pre-compensator in reference signal[38]. It resizes the reference signal to ensure that a state with desired value can be reached. For determining the gain, we are going to find the value of steady state error. Assume that in steady state, there is no change

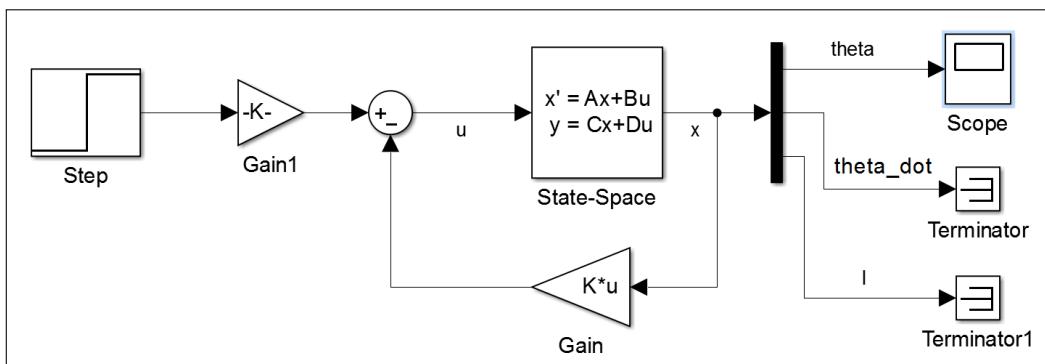


Figure 7.19: Simulation plant for State-space model with pre-compensator

in any state which means $\dot{x} = 0$. We can see from figure 7.19 the relationship between state and the reference signal.

$$u = r - \mathbf{G}\mathbf{x} \quad (7.18)$$

Substitute equation 7.18 into 7.13:

$$\begin{aligned}
 \dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{B}(r - \mathbf{Gx}) \\
 0 &= \mathbf{Ax}_\infty + \mathbf{B}(r_\infty - \mathbf{Gx}_\infty) \\
 \mathbf{x}_\infty &= (\mathbf{BG} - \mathbf{A})^{-1} \mathbf{Br}_\infty
 \end{aligned} \tag{7.19}$$

Then use the definition of steady state error e_∞ :

$$\begin{aligned}
 e_\infty &= r_\infty - y_\infty \\
 &= r_\infty - \mathbf{Cx}_\infty \\
 &= (1 + \mathbf{C}(\mathbf{BG} - \mathbf{A})^{-1} \mathbf{B})r_\infty \\
 &= 0.6675r_\infty
 \end{aligned} \tag{7.20}$$

Therefore the compensator gain should be:

$$\begin{aligned}
 \text{gain} &= \frac{r}{r - e_\infty} \\
 &= 3.0075
 \end{aligned} \tag{7.21}$$

Then we input the result into the simulation, there result is shown in figure 7.20, it settles at 0.001 rad, which is exactly the step size.

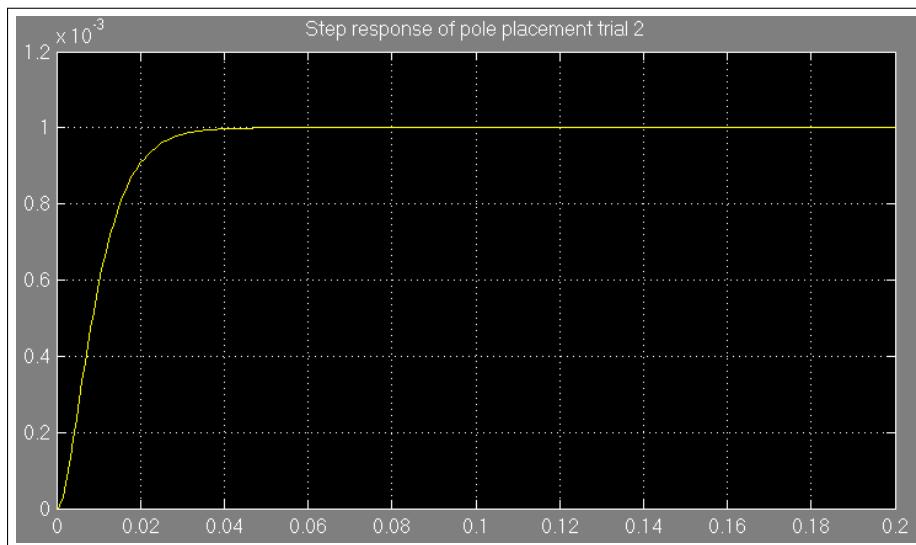


Figure 7.20: step response by pole placement control

However, the weakness of pre-compensator is that it cannot distinguish signal noise from real signal. Any disturbance added into the system will be influencing the steady state accuracy and will

not be cancelled by this pre-compensator. Therefore we want to find alternative method to reduce steady state error. Integrator is always a good controlling method for suppressing steady state error as the error will accumulate along time which forces the error to be zero. Therefore we want to take only the integral part of the PID controller and combine it with full state feedback control as in figure 7.21 [39]. By using this method we can guarantee a smooth response as well as zero steady state error.

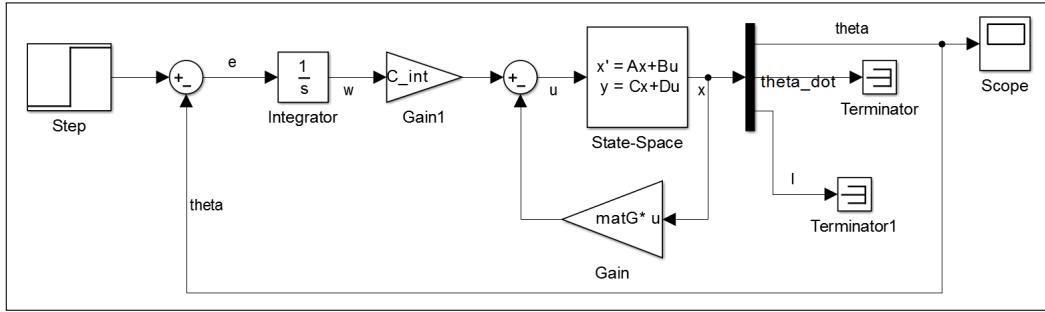


Figure 7.21: Full state feed back including integrator

As an integrator is added in, there is one more state in the system. We can see the whole feedback control loop as a new state space system with 4 states. The input is now reference signal r while the output remains the same. I am going to call the new state w , which is the integration of error. First we need to find relationship between the derivative of the new state and other states.

$$\begin{aligned}
 e &= r - \theta \\
 &= r - y \\
 &= r - \mathbf{C}\mathbf{x}
 \end{aligned} \tag{7.22}$$

For original states \mathbf{x} the original equation still holds(equation 7.13). However the input has now changed to r so we need to express the original input u in other ways.

$$u = C_{int}w - \mathbf{G}\mathbf{x} \tag{7.23}$$

$$\begin{aligned}
 \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}u \\
 &= \mathbf{A}\mathbf{x} + \mathbf{B}(C_{int}w - \mathbf{G}\mathbf{x}) \\
 &= C_{int}\mathbf{B}w + (\mathbf{A} - \mathbf{B}\mathbf{G})\mathbf{x}
 \end{aligned} \tag{7.24}$$

Where C_{int} is the gain of the integration controller. We can then write down the new state-space

equation, let the new state-space variable be \mathbf{z} :

$$\mathbf{z} = \begin{pmatrix} w \\ \mathbf{x} \end{pmatrix} \quad (7.25)$$

$$\begin{aligned} \dot{\mathbf{z}} &= \mathbf{A}^* \mathbf{x} + \mathbf{B}^* u \\ &= \begin{bmatrix} 0 & -\mathbf{C} \\ C_{int} \mathbf{B} & \mathbf{A} - \mathbf{B} \mathbf{G} \end{bmatrix} \begin{bmatrix} w \\ \mathbf{x} \end{bmatrix} + \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} r \\ &= \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -\frac{k}{J} & -\frac{\mu}{J} & \frac{K_T}{J} \\ \frac{C_{int}}{L} & -\frac{g_1}{L} & -\frac{K_T+g_2}{L} & -\frac{R+g_3}{L} \end{bmatrix} \begin{bmatrix} w \\ \mathbf{x} \end{bmatrix} + \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} r \end{aligned} \quad (7.26)$$

Where $\mathbf{G} = [g_1 \ g_2 \ g_3]$. Now the poles for the system can be determined by finding the eigenvalue of matrix \mathbf{A}^* .

$$|s\mathbf{I} - \mathbf{A}^*| = (s - p_1)(s - p_2)(s - p_3)(s - p_4) \quad (7.27)$$

By substituting the poles we want to place and comparing the coefficient of each side of the equation 7.27, we can finally find the coefficient of state feed back control and integration control:

$$\mathbf{G} = [-8.2312 \ 0.0465 \ -0.9605], \quad C_{int} = 1504.5 \quad (7.28)$$

Putting these values into the simulation plant and we get the step response as figure 7.22. As we can see, there is no oscillation which is very close to critical damp. There is no overshoot or steady state error. It is a fairly ideal response.

Again, we use the same technique to check if this response will be affected by the noise. As figure 7.23 shown below, the largest overshoot or undershoot is only 1%. Furthermore, the time taken to settle does not affect by much which proves again that this control system will give an ideal response.

I then tried to put poles in different places to observe the influence of closed loop poles. As we can see in table 7.3, If the real parts of poles are more negative, the settle time will be smaller. However the controller gain will be increased, which increases the control effort. In real life situation the actuator will finally saturate which may not achieve our desired response. On the other hand, imaginary parts of the poles affect the overshoot of the response. By placing poles near to the real axis will obviously reduce or even eliminate the overshoot. Furthermore, all of the controller value tested has no steady state error. It shows that integrator can theoretically eliminate the steady state

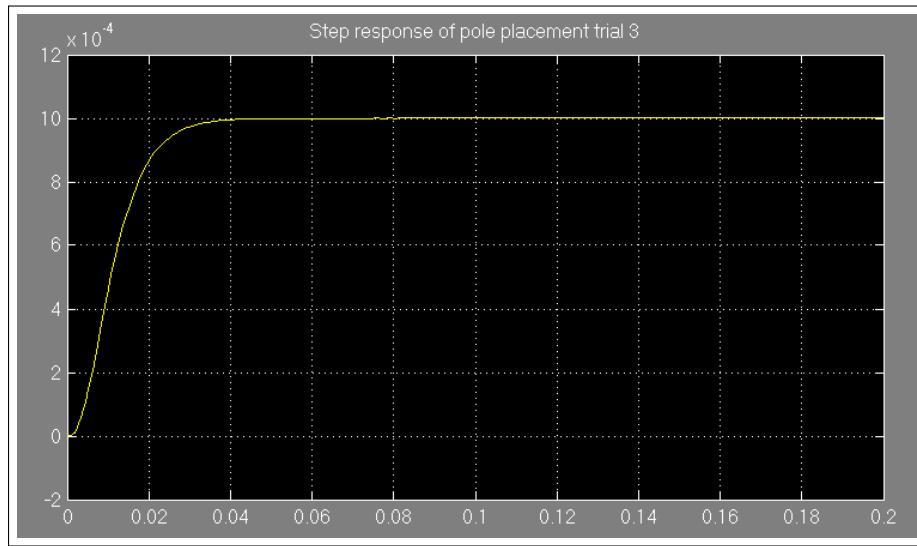


Figure 7.22: step response by pole placement control

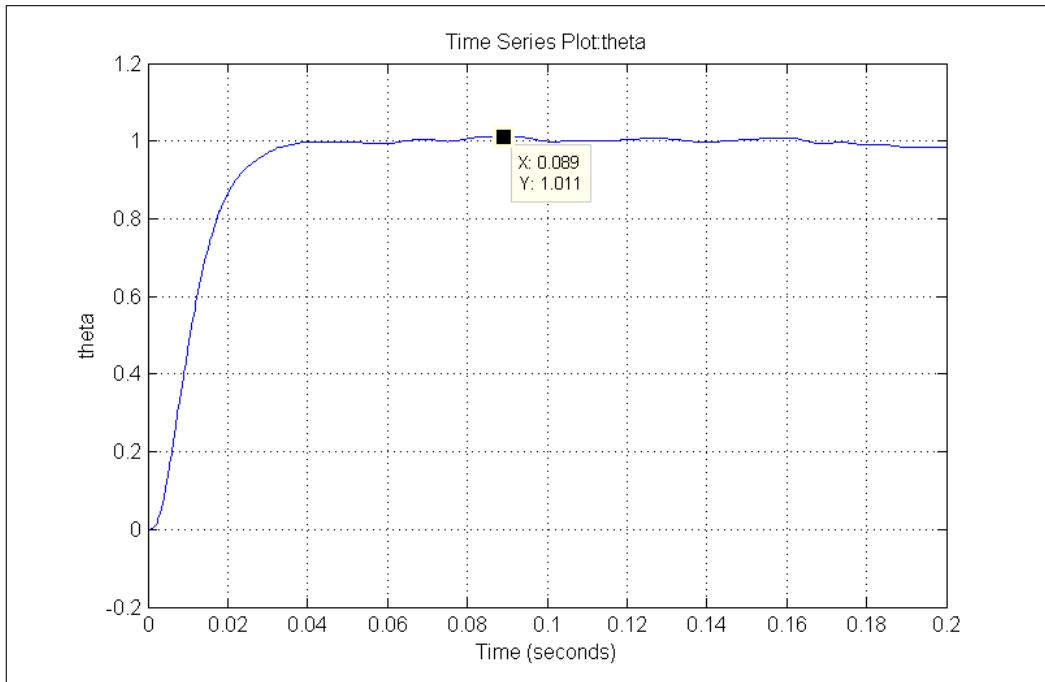


Figure 7.23: step response in presence of noise

error which is desirable in many cases. For our system, the parameters in the second last row of table 7.3 will be used in controlling the galvos as it provides the shortest settling time and no overshoot.

7.2.2.4 realisation of full state control - sensor

From the last section we can see that by using full state feedback decent response will be obtained even when the noise is present. However, to actually achieve full state control, all of the states within the system must be measurable. In this case, we need to find a way to measure the angular displacement, the angular velocity and the current goes through the coil. Measuring the current will be easy as we only need to use an ammeter. For angular displacement, we need a specific sensor called rotary encoder.

Poles	G	C_{int}	settle time(s)	overshoot(%)
$-7000, -100, -10 \pm j10$	[-31.8 -0.012 -0.228]	2.1	0.447	4
$-7000, -100, -150 \pm j200$	[-23.5 0.0188 -0.0605]	656.25	0.0665	0
$-5000, -200, -100 \pm j300$	[-13.5 0.0075 -1.26]	1500	0.0514	9.5
$-5000, -500, -200 \pm j10$	[-8.23 0.0465 -0.961]	1504.5	0.0456	0
$-7000, -1000, -500 \pm j10$	[94.5 0.204 0.900]	2626.5	0.0189	0
$-8000, -1000, -500 \pm j500$	[143 0.238 1.500]	60000	0.0089	3

Table 7.3: The affect of different poles to step response

There are many different rotary encoders in the market. Incremental encoders only measure the change in angular position and the absolute angular position will then be calculated. Absolute encoder, in contrast, directly returns the absolute angular position.

By means of detection rotary encoders can be classified to optical encoder, magnetic encoder and capacitive encoder. Magnetic encoder detect the angular position by measuring the magnetic field strength. It is very reliable and works in demanding environment such as dust and moisture. For optical encoder, the mechanism is shown in figure 7.24, a rotating disk with "windows" is stuck with the object and the light source and the detector is placed at either side of it. When the disk rotates with the object, detector will see a light-dark transition and this means that object rotated for one unit. By using two sets of light source and detectors(A and B) quadrature coding can be done so that both angular speed and direction can be detected. An extra light source(I) is used for detecting multiple revolutions. Optical encoders are mostly incremental encoders. The advantage is that it has almost no extra load for the component as it uses light for observation. However, it is delicate as any dust or moist will affect the performance of the sensor. Capacitive encoder sense the change of the capacitance by generating a high frequency signal. The result can be stored as a 16 bit signal per revolution which means that the resolution can be up to $\frac{2\pi}{2^{16}} = 9.58 \times 10^{-5} \text{ rads}$. Comparing to optical encoder, it can withstand poor environment as well as maintaining the sensitivity.

In conclusion, capacitive absolute encoder is desired in our system as it gives accurate detection of angular position. We can then take the derivative of angular position to calculate the angular speed of galvos. The current can be detected by a simple ammeter.

7.3 Mechanical design

Stereolithography has relatively large working area comparing to two-photon polymerisation and optical tweezing. Furthermore, it does not use the same laser source as other fabrication processes, the

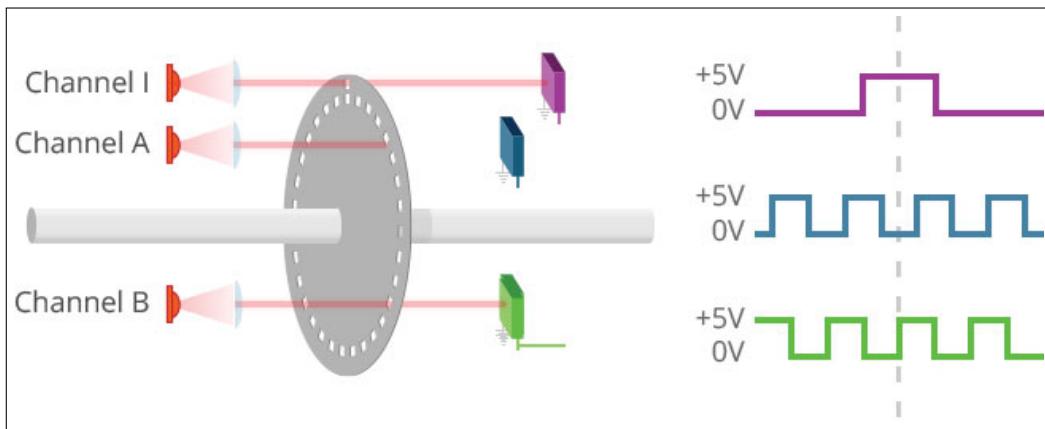


Figure 7.24: Mechanism of optical rotary encoder

workspace is therefore separated from the main stage where other fabrication processes take place.

In designing the resin vat, there are two important points that we need to consider:

1. The post-processes necessary when parts are completed. The post-process of stereolithography includes chemical immersion, baking and possible support removal.
2. The compatibility with other process. It should be sent to the main stage for further process.

Combining the concerns above, my design for the resin vat is shown in figure 7.25.

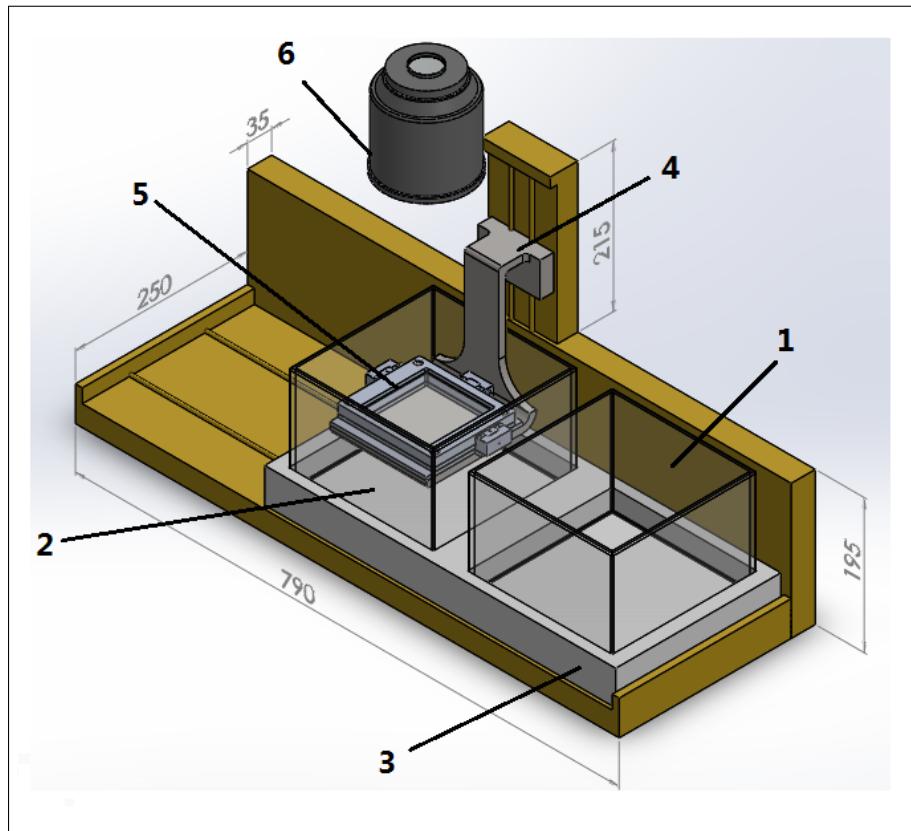


Figure 7.25: The mechanical layout of the resin vat for SLA. 1) post-processing tank for chemical immersion 2) fabrication vat where resin sits in 3) The slider for two tanks 4) Z-level stage 5) The cassette clamp and the cassette 6) the f-theta lens

The SLA process work as following: It starts at the position shown in figure 7.25. As the stage is

lowered, the part is fabricated layer by layer. The height of the tank is 150 millimeters and the height of the stage plus cassette is 32 millimeters which is enough for fabricating parts under 100 millimeters. When the part is finished, the z-levelling stage is raise to maximum(The distance between f-theta lens and top of the tank is 264 mm which allows the entire part to be moved above the tank). The slider underneath will slide to the left so that the part can be immersed into chemicals for further solidifying and for washing liquid polymers remained on the parts. At the meanwhile a sensor is detecting the resin level in the fabrication tank to check if it needs to refill. When chemical immersion is completed, the stage is raised up again. Cassette will be unlocked from the clamp and wait for the robotic arm to transport the part. When completed part is taken away and the new cassette is installed, the system is ready for the next fabrication.

8 Two-Photon Polymerisation

Two-photon polymerization (TPP), is a key fabrication technique in our system design, as it can fabricate three-dimensional free-standing structures. Two Photon Polymerization is the high resolution equivalent of Stereolithography, where polymerization is triggered by a non-linear optical process of two photon absorption, enabling high resolutions of 100nm . It can fabricate structures with a wide variety of polymer composites, using many different custom, or standardized photopolymer resins. This fabrication process was chosen for inclusion in our product to fabricate sub-micron features, utilizing the same materials compatible with stereolithography, to enable users to fabricate large base structures with stereolithography ($10\mu\text{m}$ feature size), and the higher resolution 100nm features with TPP. Incorporating this fabrication process into the system, expands the applications of the system, to be able to fabricate arbitrary micro structures, photonic crystals, microfluidics and bioactive cell scaffolds, of which some can be seen in figure 8.1 below.

To reach an optimal design for the two photon polymerization fabrication system, the literature was reviewed and evaluated to find the design considerations on the apparatus used (section 8.1). Elements of the literature and original designs were combined to create the TPP fabrication design, which was optimally integrated into the entire system (section 8.2). The most significant original design was the live imaging feedback software that used two computer vision algorithm designs to allow accurate live self polymerization calibration, and accurate live spatial closed loop control of the structure being fabricated. These increase the fabrication quality and accuracy, thereby increasing fabrication yields, detailed in section 8.4. The performance of the final TPP fabrication design is reviewed in section 8.5.

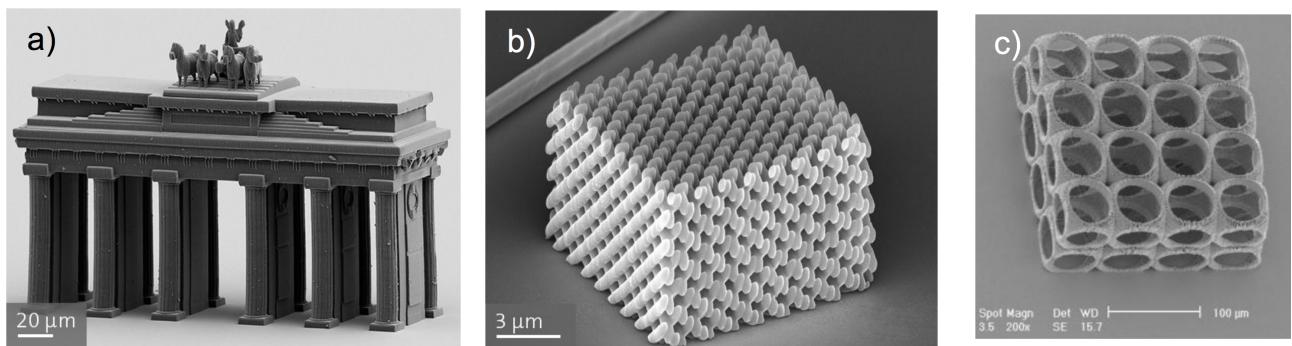


Figure 8.1: TPP applications: **a)** micro structure of Brandenburg Gate [40], **b)** photonic crystal [41] and **c)** bioactive cell scaffold [42]

The TPP process uses the high intensities available from a pulsed femtosecond laser to induce polymerization in a photopolymer liquid resin to create a free standing structure [1]. The structure is created by tracing out the focal point of the pulsed laser so as to polymerize and create a solidified polymer structure [43].

8.1 Literature Review

8.1.1 Theory

Two-photon polymerization fabrication was first demonstrated successfully in 1997, with the work of Maruo et al. [44]. Since then it has become an active research topic, with nearly 5000 papers being published [1]. The TPP process exploits the physical process of two-photon absorption, first theorised in 1931 by Goppert-Mayer [45]. Two-photon absorption (TPA) is a non-linear optical process, which is the absorption of light when the sum of the energy of two photons is equal to the energy gap between electronic states in the molecule, as seen in figure 8.2 [1]. Comparing this to stereolithography, which uses one photon-absorption, TPA uses two photons of a longer wavelength to absorb, and alter the energy state of the molecule [43].

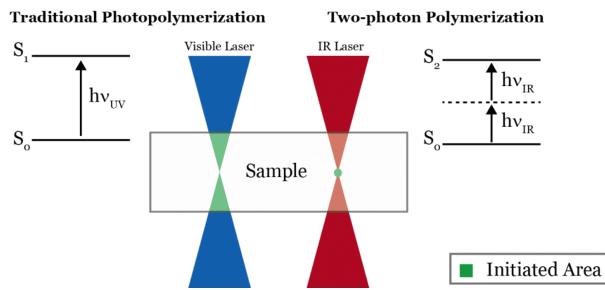


Figure 8.2: Illustrates single photon polymerization, compared to two photon polymerization, which uses two photon absorption, to polymerize a smaller volume inside the photoresist at the focal point [46]

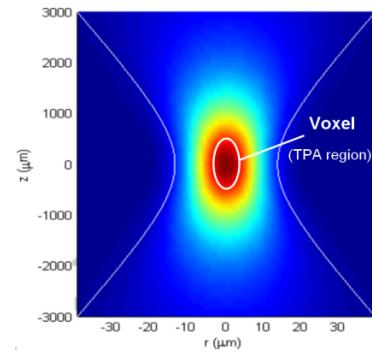


Figure 8.3: TPP Longitudinal beam intensity profile and the polymerized volume [47]

When a photon is absorbed by a molecule, it causes a transition to a virtual state. TPA occurs if a second photon is absorbed whilst the virtual state exists, however the second absorption must occur within $1fs$ of the initial transition [1] [43]. Hence TPA requires high local photon fluxes, generated from a highly intense laser beam at the focal point. These can be produced by femto-second pulsed lasers, which produce pulses shorter than $100fs$, allowing each pulse to have a peak power on the order of a kilowatt [48]. This is combined with a high numerical aperture lens to create an intensely focused beam [43]. The nonlinear intensity dependence ($\propto I^2$) of the TPA is key for fabrication, as it allows absorption to be tightly confined in three dimensions, upon focusing the beam into a small volume, where the voxel (Volumetric element/pixel) has dimensions below the diffraction limit of light. [1]. When TPA occurs in an appropriate photopolymer, with photoinitiators it creates a free radical that can be used for free radical polymerization.

Comparing TPP to one-photon polymerization (Stereolithography), the one-photon polymerization can only polymerize the first layer of the photoresist, since the excitation beam is significantly attenuated by linear absorption before reaching its focal point. Whereas in TPP, the absorption cross

section for TPA is substantially less than the absorption cross-section for single photon absorption, thus polymerization is confined to a small volume (approx 100nm across) where the intensity of the incident laser pulse is sufficiently large, as seen in figure 8.3 [48]. TPP uses the same chemical reaction of polymerization of photoresists as stereolithography. These photoresists, or resins are initiated by an ultra violet light source, hence for TPP the excitation light must be double the wavelength of that of UV, which is near infrared (IR). Therefore in one-photon polymerization it is only possible to create structures using a 2D layer by layer method, alternatively since photoresists are commonly transparent (very low attenuation) to near infrared wavelengths, it is possible to focus inside the photoresist, and only polymerize small targeted volumes at the focal point, where non-linear TPA occurs, illustrated in figure 8.2. This allows fabrication by tracing the focal point in three-dimensions inside the photoresist to create the structure. [49].

TPP systems commonly use two types of photoresists, negative-tone photoresists, where a material exposed to light becomes insoluble, and the opposite being positive-tone photoresists, where exposed to light the material becomes soluble, to the photoresist developer. [1]. The most common of these is negative-tone photoresists, where the material consists of soluble organic monomers, which when exposed with UV light, become cross-linked, and insoluble. TPP cross-linking is induced by generation of free radicals, which are generated from the material's photoinitiators in radical polymerization. A photoinitiator molecule produces a free radical, when stimulated by a UV photon, or when stimulated by TPA of two near infrared photons. The polymerization terminates when all the free radicals produced by the excited photoinitiators, join chains of polymers or when the exciting light is stopped. [44]. TPP systems commonly use commercially available resins such as SCR-500 manufactured by Japan Synthetic Rubber Company, Nopcocure 800 (NOP-800) manufactured by San Nopco and SU-8 manufactured by Micro Chem. These photoresists can be used alone, or mixed with dyes, nanoparticles, special polymers, and other materials, to create a wide variety of polymer composites. This approach leads to structures with specified chemical, biological or optical properties, which are suited to a wide variety of applications, such as photonic crystals, microfluidics and bioactive cell scaffolds. [50]

8.1.2 Resolution

The resolution and feature size of microstructures fabricated by TPP depends upon the average laser power and the irradiation time, which affect the geometry and size of the smallest polymerized volume element, the voxel. Tracing the pulsed laser through the photoresist, forms a series of voxels, which must overlap to form a continuous smooth structure. The smallest structures that can be fabricated, are limited by the voxel size, therefore for a competitive TPP design the voxel size must be reduced

to the smallest practical size. Following the derivation in [51], making the assumptions that the voxel volume element is defined as the region where the density of free radicals ρ exceeds the minimum concentration (threshold value) ρ_{th} , for polymerization to occur, and that a gaussian beam is applied, we can derive the voxel diameter:

$$d(N_0, t) = r_0 \cdot \sqrt{\ln(\sigma_2 N_0^2 n \tau_L / C)} \quad (8.1)$$

$$C = \ln \left(\frac{\rho_0}{\rho_0 - \rho_{th}} \right) \quad (8.2)$$

Where N_0 is the constant photon flux, t the total processing irradiation time, r_0 the beam waist radius at the focal plane ($z = 0$), σ_2 the effective two-photon cross section for the generation of radicals, n the number of pulses (where $n = vt$, where v is the laser pulse repetition rate), τ_L the laser pulse duration and C a constant, in which ρ_0 is the primary initiator particle density. Similarly following [51], the voxel length can be derived:

$$l(N_0, t) = 2z_R \cdot \sqrt{\sqrt{\sigma_2 N_0^2 n \tau_L / C} - 1} \quad (8.3)$$

Where z_R is the Rayleigh length, and N_0 is:

$$N_0 = \frac{2P\Gamma}{\pi r_0^2 \tau_L v \hbar \omega_L} \quad (8.4)$$

Here, P is the average laser power, Γ the fraction of light transmitted through the objective, \hbar the reduced Planck constant or Dirac constant and ω_L the photon frequency ($f = \frac{c}{\lambda}$). The voxel length and diameter are illustrated in figure 8.4, which shows how they define the smallest feature size of a structure fabricated from TPP.

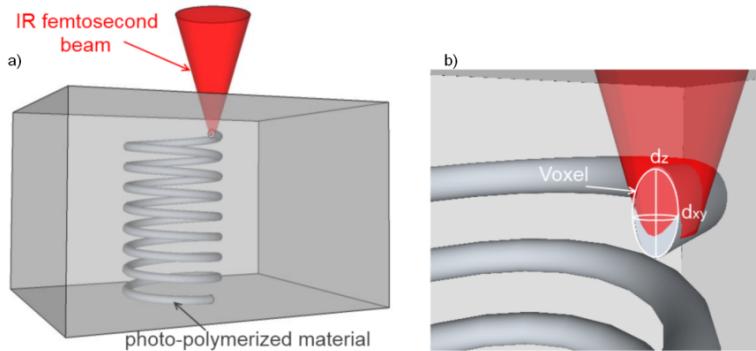


Figure 8.4: **a)** Fabricating a structure, **b)** TPP voxel dimensions, where $d_z = l(N_0, t)$, and $d_{xy} = d(N_0, t)$ [52]

[51] validated the above models, using experimental data recorded, fitting the above voxel equations, to find the unknown parameters of axial resolution and lateral resolution. These values were compared to the resolution limits of the objective [53], where the lateral resolution is $d = \frac{0.61\lambda}{N.A.}$ (where

λ the wavelength of light, and $N.A.$ of the objective), and the axial resolution, $z = \frac{2\lambda n_{oil}}{N.A.^2}$. These resolutions limits were used to validate the voxel model, being accurate to within 6% of the resolution limits of the objective.

Evaluating the model given in [51], it clearly shows that the voxel diameter at the focal point increases as the square root of the laser power. However this model, as indicated in [43], has implicitly assumed that there are enough available photoinitiators to carry out polymerization. [43], suggests that using high laser average powers (P), the resin becomes saturated, since all the available photoinitiators in the focal region are consumed, leaving none to create new polymer chains. This places an upper limit on P , which is proportional to the photoinitiator concentration.

An important feature about TPP resolution, is that it can fabricate structures with sub-diffraction limited resolution, due to the fact that no polymerization occurs below the intensity threshold. This is exploited to achieve the highest resolution, by using an intensity that just exceeds the intensity threshold slightly, allowing the possibility to polymerize regions only 100nm wide [48].

Concluding the resolution of TPP, it is possible to use the given model for designing, and checking the resolution and dimensions of the voxel in the majority of resins in our TPP system design. However in order to achieve the highest resolution, the TPP system needs to have precise control over the laser power, to create a focal point intensity that just exceeds the polymerization threshold.

8.2 TPP Optical System Design

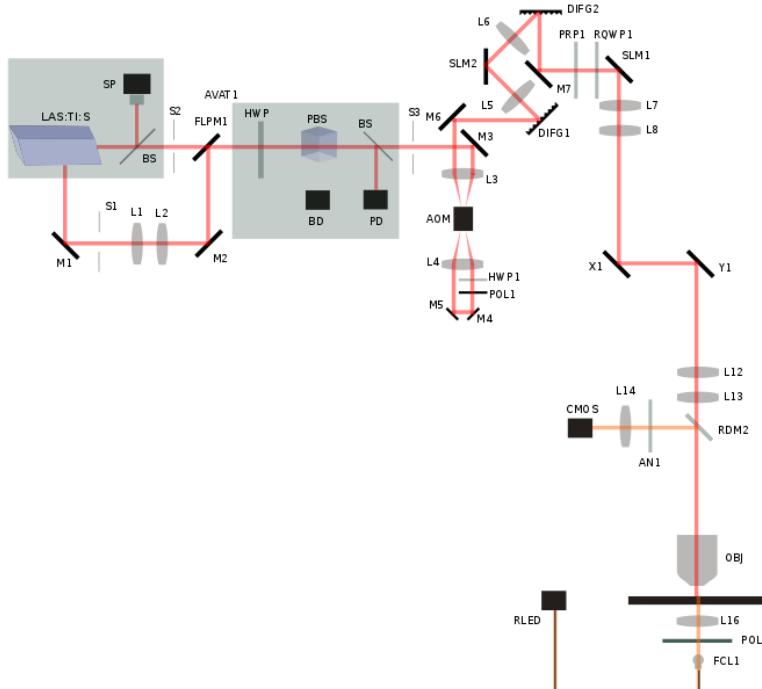


Figure 8.5: TPP schematic design. Red beam path represents the Ti:Sapphire beam path, and the orange beam path represents the red fibre coupled LED beam path, for illuminating the sample.

Figure 8.5, shows the optimal TPP system design. This section will detail the design choices for each component in the system design. Wherever possible, to minimise cost and space requirements, the components are shared with similar fabrication and imaging systems.

The laser source, for best TPP results should be a femtosecond pulsed laser, with a wavelength in the near infrared band. A longer wavelength is superior, since there is more photons for the same laser power ($E = \frac{hc}{\lambda}$), increasing the probability of two photons interacting with a molecule in $1fs$. Therefore **LAS:TI:S** is chosen to be the Ti:Sapphire **Vitara-T-HP** $100fs$ pulsed oscillator laser, which has a $800nm$ wavelength, with a repetition rate of $80MHz$. The laser outputs, either a continuous wave or pulsed wave beam, which are selected between using a motorized Owis flip mirror **FLPM1**, which can turn 45° , with an acceptable repeatability error of $100\mu rad$. For good design, both laser outputs have a ThorLabs electromechanical shutter **S1, S2**, actuated when the laser is in standby or starting up.

The beam power is controlled with a motorized half wave-plate and fixed polarizing beam splitter (PBS); most precise technique for power control for a linearly polarized beam. Rotating the half wave plate, allows the beams polarization direction to be rotated, which when passed through the PBS, splits the beam into two. The output beam of the PBS has an intensity described by Malus' Law:

$$I = I_0 \cos^2(2\phi) \quad (8.5)$$

Where I_0 is the input intensity of the beam into the half wave plate, and ϕ the angle of rotation for the half wave plate. The discarded power is sent to a beam stop. This is implemented as a complete system **AVAT 1**, using a Newport Computer Controlled Variable Attenuator Kit. This uses a PID (Proportional Integral Derivative) closed loop control, which takes 5% of the input beam, using beam splitter **BS1**, measuring the power with a calibrated spectrometer **PD1**. This provides an attenuated output beam power, with an error less than 10%, which is acceptable.

The beam is then expanded through a Galilean telescope, using a negative lens **L1** and a positive lens **L2**. A Galilean beam expander was chosen, instead of a Keplerian beam expander, as the latter focuses the input to a spot between the objective and the image lens, which can easily cause potential breakdown of the air for our high power laser pulses ($E_{AirMax} = 3 \times 10^6 V/m$), and/or heat the surrounding air. This leads to inhomogeneity in the refractive index of the medium, that deflects the light rays from their optical path, causing un-wanted optical aberrations. [54].

The TPP design attempts to eliminate heat accumulation effects, by using a burst excitation mode for the pulsed laser [55]. To implement this a very high speed shutter is required to pick off the $100fs$ pulses from the pulsed laser. This is implemented by an acousto-optic modulator (AOM) design, achieving a switching time of $\approx 25ns$, based from [56] (detailed in section 20.4).

The laser pulses are optimized for TPP, by shaping the $100fs$ pulses in the system. This can shape and shorten the optical pulses, thus increasing the probability of two photon absorption events occurring, by increasing the peak photon intensity. Implemented with a pair of diffraction gratings **DIFG 1,2** and lenses, arranged as a zero dispersion pulse compressor with the spatial light modulator **SLM 2**, used as a pulse shaping mask [57].

The following phase retardation plate **PRP1**, retards the phase in one of two orthogonal components of the polarized input beam. Followed by a **RQW P1**, a rotatable quarter wave plate, to rotate the polarized light, both elements combined help to create the vortex phase modulation, needed for high resolution TPP STED lithography (further explained in section 14.1.7)

To allow TPP to fabricate with either a single focus or using multiple dynamic foci, as demonstrated in [58], the **SLM 1** is used to perform holographic beam shaping, explained further in section 16. To pass the now holographic beam correctly through the system, a 4F correlator comprising of lenses **L7, L8** is used to image the beam onto the galvo scanning mirrors, **X, Y**. This holographic beam is maintained and similarly imaged onto the microscope objective using the same 4F design, with lenses **L12, L13** (explained in section 18). These lenses also re-center the deflected beam from the galvo scanning mirrors, to the center of the objective.

To trace the focus of the beam through the resin, quickly and precisely, a two axis (**X, Y**) galvo scanning system was included. This deflects the beam through two rotatable galvanometer mirrors, that rotate on orthogonal axes with respect to each other. This provides the ability to change the angle of incidence of the incoming beam onto the objective lens, which maps to the spatial *X, Y* position of the focus in the resin.

To maximize the intensity of the beam at the focus, the beam spot area is minimized ($I = \frac{P}{A}$). Investigating the diameter of the beam spot at the focus, $d = \frac{0.61\lambda}{N.A.}$, it is evident that it is possible to either increase *NA* (Numerical Aperture) or decrease λ . λ is already design constrained, therefore the *NA* of the microscope objective **OBJ** must be maximized. The largest *NA* microscope objectives commercially available, have a $NA = 1.4$, which immerse the objective lens and the sample in a liquid with a refractive index ≈ 1.5 , known as oil immersion microscope objectives. The oil immersion objective lens, the Olympus **MPLAPON 100XO** (100 \times magnification, $NA = 1.4$) was chosen for this TPP optical layout. Some TPP designs use oil immersion between the objective lens and a glass cover slip, where the cover slip houses the TPP photoresist resin, whereas other designs immerse the objective directly into the photoresist resin, which produces better results, by reducing optical aberrations that arise from the interfaces of the refractive index discontinuities. Therefore the design immerses the objective directly into the photoresist resin.

Commonly TPP systems, employ real-time monitoring through a transmissive microscope, onto a CMOS camera. This allows manual monitoring of the polymerization reaction and structure fabrication. Implemented in this design by a fiber coupled red LED light source (**RLED**), passing through a polarizing filter and a condenser lens, to illuminate the sample from underneath. This imaging technique is known as bright-field imaging, and is explained further in section 12.1. The sample, consisting of photoresist resin, supported on a glass transparent base. The resins used, and structures fabricated in TPP are transparent to visible light, which allows transmissive illumination. Where the illumination source wavelength is designed to be a visible wavelength λ furthest away from the UV polymerization wavelength, thus a red ($\lambda = 625nm$) light source is used. The design is also constrained to use visible light for illumination, as commercial CMOS cameras are optimized to detect visible light. This transmissive illumination utilizes the same $100\times$ objective, which is redirected to the CMOS camera, by a dichroic lens (**DM2**). This selectively passes the $\lambda = 800nm$ beam, and reflects all other wavelengths including the red $\lambda = 625nm$ illumination light.

8.3 Fabrication Process

The TPP system, is designed to integrate seamlessly with the mechanical sample handling component (cassette). The system can fabricate using TPP, micron structures directly onto the bottom glass surface of the cassette, or it can fabricate micron structures onto existing stereolithography centimeter structures. The system was designed to offer the flexibility of fabricating large ($8cm^3$) structures with nanometre features; the stage design to implement this is explained in section 22.

The designed process, for fabrication is outlined as follows, the operator would import their CAD model using the software suite designed in section 24, which is then converted into machine instructions that control the variable attenuator, shutter, scanning speed and the galvo scanning mirrors, and the X, Y, Z stage. The operator using a clean cassette, would drop a few drops of their chosen photoresist resin, onto the cover glass surface area. This is then placed into the storage rack, where the automated system, would move the cassette to either the stage or the oven then the stage (some resins need pre-baking before polymerization, such as SU8). Once the cassette was clamped securely to the stage (section 23), the stage would move upwards, till the photoresist was in contact with the objective (objective was immersed in photoresist). Finer adjustments would be made so the objective was focused clearly on the solid interface under the resin, either glass or a larger stereolithography polymerized structure. If using a new resin, the system would have to calibrate itself, following open loop calibration procedure (detailed in [59]). The operator could either choose a beam scanning method, or let the system choose the optimal beam scanning method for the given structure, that minimizes the time taken to fabricate the given structure. The polymerization would be-

gin by polymerizing 2D slices of the structure, incrementing the Z axis of the stage down a specified height after each layer was correctly polymerized. The beam focus would be traced through the resin, by the scanning galvo mirrors. During fabrication, closed loop control is implemented to check, and correct that each layer correctly polymerizes both spatially (to the desired dimensions) and the quality of polymerization is on target. Once the single structure was completed, the stage could either move to a new area in the X, Y plane to repeat the processes and fabricate further structures, or it could lower the cassette in the Z direction, where it could be relocated to a storage rack by the automated system. The operator could then access the storage rack, and further process the fabricated sub micron structure to remove the un-polymerized resin with a quick rinse in acetone.

8.4 TPP System Control

This section will explain how existing TPP systems are controlled, analyzing and designing solutions to incorporate closed loop control. Closed loop control systems provide the best solution, as nanometre scale fabrication is significantly sensitive to micro-metre mechanical changes in the system, caused by vibration, thermal expansion and or the precision errors on actuators.

The two key systems to control in TPP, are the spatial fabrication of the 2D slices of the structure, and the resolution of the polymerization voxel, ideally kept at the smallest size possible which can be reliably fabricated.

8.4.1 Spatial Control

Prior literature shows the use of only open loop designs, that use a calibrated galvo scanning model (plant model), outlined in section 22. To calibrate these systems a periodic woodpile structure is fabricated [60], and it is dimensioned using a high resolution (10nm) scanning electron microscope (SEM), to correct the open loop scaling factors.

SEMs are too large, too expensive, and would not integrate well with other systems in the entire product design. A comparable imaging method STED, with a resolution of 55nm was designed in, as it is smaller and re-uses a lot of the expensive optical components found on the optical paths, making an economic entire system product design.

However a closed loop control system is superior, as it allows real-time correction for live disturbances, which include those above, in addition to local temperature changes (which vary the rate of polymerization) and shrinking of the structure shortly after fabrication [1], which is the most significant error. The open loop system model, would be calibrated as described above. However to implement spatial micron closed loop control the actual 2D polymerized slice is accurately extracted from the live transmission microscope images, and then differenced with the computer model 2D slice build plans. This produces a 2D slice that contains areas that were missed during the first slice polymerization

pass. This correction slice could then be further polymerized, and the control loop repeated until the imaged slice spatially matched the build plans (within 2%), where it begins fabrication of the next slice layer.

To achieve this, a Matlab computer vision algorithm was designed that extracts a accurate binary image of the polymerized regions, from the actual 2D polymerized slice image. This algorithm consists of two main sections.

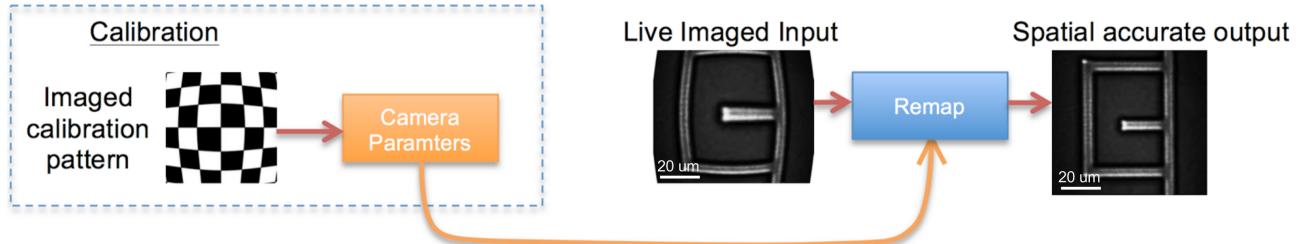


Figure 8.6: First section, spatial accurate mapping

The first section makes the images spatially accurate (no geometric distortion between the image and the actual structure fabricated). Performed with existing geometric camera calibration (camera resectioning) Matlab functions [61], which calculate the camera parameters (intrinsic, extrinsic and distortion coefficients). This calibration only needs to occur once, when the optical setup is changed, such as using a new photoresist resin (with a unique refractive index). The camera parameters are estimated, by imaging a standard known calibration pattern, a chess board pattern on a microscope calibration slide, with square sizes of $10\mu\text{m}$ [62]. Figure 8.6 illustrates, passing the camera parameters and the raw images into the Matlab remap function, outputs undistorted images, that are spatially accurate (rectilinear correction)[63].

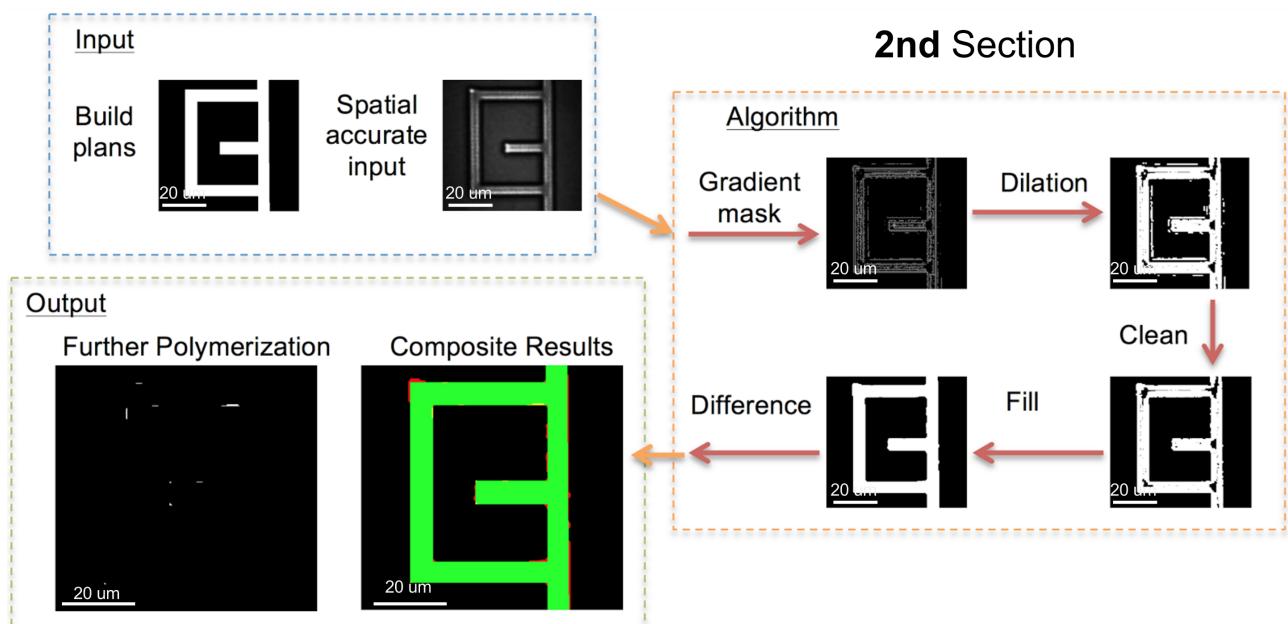


Figure 8.7: Second section, Spatial slice error extraction

The second section extracts the spatial slice error, by extracting a binary image of the polymerized

regions from the accurate, undistorted image, and comparing this with the input build plans. The algorithm was developed, and tested on, in situ live TPP transmissive microscope images, imaged partway through fabrication, and corresponding binary build plans, taken from [64]. The algorithm illustrated in figure 8.7, is designed to convert the colour undistorted image, to gray scale, then pass it through a `wiener` filter to remove image noise, whilst preserving important edge information. Next to detect the polymerized regions, that differ greatly in contrast relative to the uniform background, the gradient of the image is calculated. The gradient image is calculated by the `edge` function, and a '`sobel`' edge threshold applied to return a binary mask containing the segmented structure [65]. The '`sobel`' edge detection method, in the `edge` function, finds edges using the Sobel approximation to the derivative [66].

The next step designed is to dilate the binary gradient mask, to remove the gaps in the image. This is achieved through two passes of dilation, one using a vertical line morphological structuring element ('`strel`'), followed by a horizontal line '`strel`' element. Dilation performed using `imdilate` function, which is a basic operation in mathematical morphology. The following step, is designed to remove any small segmented objects from the image (artefact's from input image noise), to leave the main segmented polymerized structure we are interested in. `bwareaopen`, removes all connected segmented objects that have fewer than 3000 pixels, assuming a default of 8 connectivity in 2D. However the binary image contains holes, within the containing outline, therefore '`imclose`', was used to close the holes using a disk structuring element ('`imclose`' is also a basic morphological function). Finally differencing the output to the build plan binary image, it is possible to calculate a binary image of regions that need further polymerization (undershoot), and erroneous regions that have been polymerized (overshoot).

Here a composite image, showing regions correctly polymerized in green, undershoot regions in yellow, and overshoot regions in red. If the undershoot area is detected to be greater than 2%, a further polymerization mask would be fed into the fabrication software, to polymerize that mask, and repeat the process till the undershoot error was less than 2%. However if the overshoot area was greater than 2%, the next fabricated layer would purposely reduce its line width in the binary build plans. This controller design purposely tries to avoid any overshoot, by operating with an initial undershoot error, and incrementally building up to the specified 2D slice of the structure (when the physical system is disturbed). This technique of control is preferred for the additive TPP process, since undershoot errors can easily be corrected, whereas polymerized overshoot errors cannot.

8.4.2 Polymerization Control

Prior literature for TPP systems, reveals that the polymerization voxel size control is run with open loop control. Typically when a new resin is used for fabrication on a machine, the system will be calibrated for various input power levels, fabricating individual voxels, which are dimensioned using a SEM, to correct for open loop scaling factors [59]. The open loop (plant) system, for this TPP optical system design, is found from equation 8.1, where Γ , the fraction of light transmitted through the objective, is:

$$\Gamma = 10^{-0.1} \cos^2(2\phi) \quad (8.6)$$

The factor of $10^{-0.1}$ was found from figure 19.10, in section 19, which is the fraction of light from the laser source to the objective, without attenuation. Whereas the attenuation factor, given by Malus' Law, from equation 8.5, is $\cos^2(2\phi)$. Therefore to control the output of the voxel diameter/polymerization size we must control the input rotation ϕ . Combining the above equations we get after some algebra:

$$d(\phi, t) = r_0 \cdot \sqrt{\ln(A \cdot t \cdot \cos^4(2\phi))} \quad (8.7)$$

With constant:

$$A = \frac{10^{-0.2} \sigma_2}{v C \tau_L} \left(\frac{2P}{\pi r_0^2 \hbar \omega_L} \right)^2 \quad (8.8)$$

Similarly equations for the voxel length can be found:

$$l(\phi, t) = 2z_R \cdot \sqrt{\sqrt{A \cdot t \cdot \cos^4(2\phi)} - 1} \quad (8.9)$$

Plotting the model for the full range of ϕ ($0 < \phi < 45^\circ$), setting $t = 40ms$ [51], it is evident that the diameter and length of the voxel has a non-linear dependence on ϕ . Optimum TPP designs, aim to minimize the voxel size, however the plant model is extremely sensitive to small changes in ϕ , at smaller voxel sizes. It is possible to see this by calculating the partial derivative of the diameter, and length:

$$\frac{\partial d(\phi, t)}{\partial \phi} = -\frac{4 \tan(2\phi)}{\sqrt{\ln(A \cdot t \cdot \cos^4(2\phi))}} \quad (8.10)$$

$$\frac{\partial l(\phi, t)}{\partial \phi} = -\frac{2 \tan(2\phi)}{(\ln(A \cdot t \cdot \cos^4(2\phi)) - 1)^{3/4}} \quad (8.11)$$

Therefore as $\phi \rightarrow 45^\circ$, $\tan(2\phi) \rightarrow \infty$, $\therefore \frac{\partial d(\phi, t)}{\partial \phi} \rightarrow -\infty$, $\frac{\partial l(\phi, t)}{\partial \phi} \rightarrow -\infty$. Hence it is evident that the gradient of the voxel diameter and length, approaches negative infinity, as the voxel size is reduced. This occurs due to the non-linear optical process utilized in TPP, and means that the polymerized voxel size, is very sensitive to disturbances either on the input ϕ , or in the system model, where the process is operated typically at the smallest reliable polymerized voxel for the system. Due to this sensitivity, and the fact that TPP fabrication typically takes hours to fabricate a single structure,

meaning disturbances are highly likely therefore closed loop control is necessary for robust design.

STED imaging is used to calibrate the open loop model parameters post process, however it is not possible to perform live high resolution STED imaging, therefore the live transmissive microscope is used to image during fabrication. However the polymerized voxel size is set to a constant for fabrication of the entire structure. This leads to disturbances that cause the polymerized voxel to be smaller than requested (undershoot), which leaves un-polymerized gaps in the structure between the polymerized voxels. When viewed on a large scale (few microns), these regions appear transparent compared to the fully polymerized regions (which appear dark), illustrated in figure 8.8. Oppositely for disturbances that cause the polymerized voxel to be larger than requested (overshoot), leave the structure fully polymerized in regions internally, however sacrifice edge resolution. However slight polymerization overshoot error is significantly better than undershoot error, since any undershoot error in the structure, causes it to fall apart when developed, since the structure is not connected.

To implement closed loop control, the transparency of the polymerized regions can be extracted, and compared to known transparencies for complete polymerization. This is the only reliable form of control, using the live imaging system, due to its limited resolution (few microns), thus the voxel size cannot be directly measured, only the result on the micron level, of changing the transparency can be observed and measured. However as described, the structure layers must be fully polymerized, therefore the problem is a observer classification problem to detect if the structure has been correctly polymerized, or not, and if not it can correct itself by increasing the power and repeating the structure layer again. Alternatively on the other scale, when the laser beam power at the focus, is too large ($\geq 2mW$) this results in boiling of the resin, causing damage to the structure. These damaged areas appear bright, due to their difference in refractive index, of bubbles formed relative to the surrounding resin. Therefore there is another classification problem to detect if the structure contains any damaged areas, if found the laser power would be reduced (increasing ϕ), and trigger a high level warning to the operator in the software that the structure has been damaged (As they may wish to abort the fabrication, since some fabrications can take a few hours). The three types of classification of the polymerizes slice can be seen in figure 8.8. Therefore an appropriate controller would operate at the smallest possible power to just polymerize a entire region fully, and if undershoot was detected, it can increase the power (reduce ϕ), and repeat fabrication of that slice again.

To develop this classification algorithm, which takes an input image of a fabricated slice, and outputs if the image is in one of three states, undershoot polymerization, correct polymerization or overshoot (damaged) polymerization. 21 different algorithms designs were created, and evaluated, and the best design chosen. Of these 20 are specific machine learning algorithms adapted for this

problem, and one is a custom designed threshold detector algorithm. The data (images) to calibrate, train and test the algorithms were taken from [67], cropped 16 images, of various live transmissive polymerized slices, where the images show 8 un-polymerized slices, 6 correctly polymerized slices and 3 damaged slices, seen in figure 8.12. The images all undergo simple cleaning filtering, by use of a wiener filter (to remove noise) and conversion to gray scale if they are a color image.

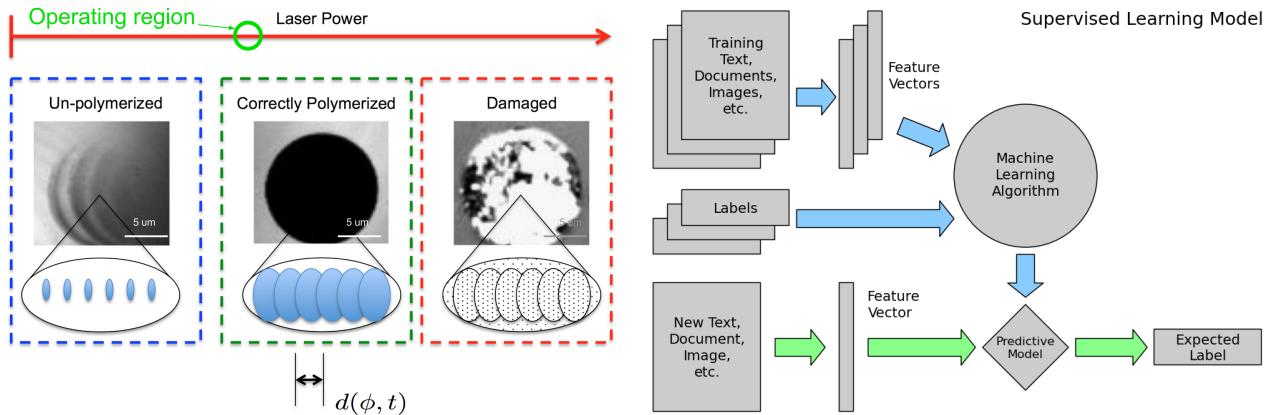


Figure 8.8: Three types of image observed, that must be classified correctly

Figure 8.9: The supervised machine learning algorithm structure, [68]

8.4.2.1 Classification by Machine Learning

Machine learning, algorithms that have the ability to learn complex models without being explicitly programmed [69]. Here 20 different classified supervised learning (learned models from a subset of the data, that were manually labeled, i.e. training data), algorithms were designed. The images themselves contain a lot of redundant information, such as repetitiveness of pixels in the image, therefore each image is reduced to a set of 200 features (features, a measurable property between 0 and 1 of a observable phenomenon), where the pattern recognition problem is easier to solve [70]. The supervised machine learning algorithm structure is shown in figure 8.9. Feature extraction was implemented using ‘bag of visual words’, which is the most robust method for object classification, using the Matlab function `bagOfFeatures`. This outputs a feature vector, which represents a sparse histogram of visual word occurrences in the image. These features were trained, and evaluated using the ‘Cross Validation’ method (Folds the feature space into $n = 5$ folds, where each fold is held out for testing data, using the remaining folds as training data, averaging the result for all folds, to train the model), which is preferred for small data sets, such as in this problem [71]. 20 various models were designed and evaluated using this structure, using the standardized model functions found in the ‘Statistics and Machine Learning Toolbox’ [72]. The table in 8.1 shows the overall accuracy for each model, tested for this problem of classification of structures.

It is clear that some supervised machine learning models are better suited for this problem. Results of comparing all 20 standard models, show that a Quadratic SVM, Cubic SVM and a Fine KNN

Supervised Machine Learning Model	Overall Accuracy	Supervised Machine Learning Model	Overall Accuracy
Complex Tree	81.2%	Medium KNN	50.0%
Medium Tree	81.2%	Coarse KNN	43.8%
Simple Tree	81.2%	Cosine KNN	68.8%
Linear SVM	93.8%	Cubic KNN	50.0%
Quadratic SVM	100%	Weighted KNN	81.2%
Cubic SVM	100%	Boosted Trees	81.2%
Fine Gaussian SVM	43.8%	Bagged Trees	87.5%
Medium Gaussian SVM	50.0%	Subspace Discriminant	81.2%
Coarse Gaussian SVM	68.8%	Subspace KNN	68.8%
Fine KNN	100%	RUSBoosted Trees	56.2%

Table 8.1: Supervised Machine Learning Model Accuracies

have the best overall accuracy of 100% (correctly predicting the true label for the random input test image).

Here the, Support Vector Machine (SVM) models, typically fit a decision hyperplane in the feature space (200 dimensions) which is used to binary classify a support vector, of an image that is represented as a vector with 200 dimensional components. By default SVMs are a linear classifier, however applying a ‘Kernel Trick’, it’s possible to map the input vector to a higher dimensional space (non-linear), where the support vectors are more separable for linear classification. The Quadratic SVM, uses a parameterized the kernel mapping $\vec{K}(x, y) = (x^T y + c)^d$, with $d = 2$. Similarly Cubic SVM’s, use $d = 3$. This has the effect of mapping the linear hyperplane, into a non-linear hyperplane in the input feature space, which may be better suited for dividing the data. Research shows Quadratic SVMs provide the most robust model [73], which was verified in the above results, and is superior since higher degrees (Cubic, Quartic, etc) over fit the model to the training data. Over fitting the model, whilst produces great classification accuracy results on trained input data, performs badly on live, or new data, thus any that significantly overfits is not suitable for this classification problem.

Likewise the k-nearest neighbor model (KNN), showed promising classification accuracy, as it is one of the simplest models, which classifies based on majority votes of the neighbor training feature space points. The results show a Fine KNN worked the best, fine being $k = 1$, which means that classification is based on the closest known feature space point. Although this showed a high accuracy, the model is very sensitive to the local structure of the feature space data, which means that it significantly over fits to the trained structure slice, making it not a robust classification model,

for the problem at hand [70].

Comparing all these designs for this classification problem, the Quadratic SVM model is best suited for classifying the slice correctly, using the feature extraction design of ‘bag of visual words’. This machine learning classification design approach, offers the ability to learn complex classification rules, by optimizing model parameters, which is made possible only by large amounts of prior data (Large sample set), to train a robust model. Commonly with TPP systems, the open loop calibration process would only produce a small sample set. A machine learning approach would only become a reliable indicator when a large sample set had been collected, for reliable training, as any training on a small data set, whilst producing a high accuracy will cause the model to over fit the sample data.

Confusion Matrix for: Support Vector Machine		
		Predicted class
True class	damage	no-poly
	3 100%	0 0.0%
damage	0 0.0%	7 100%
no-poly	0 0.0%	0 0.0%
poly	0 0.0%	0 0.0%
	6 100%	
		TPR / FNR

Figure 8.10: Confusion Matrix for Quadratic SVM

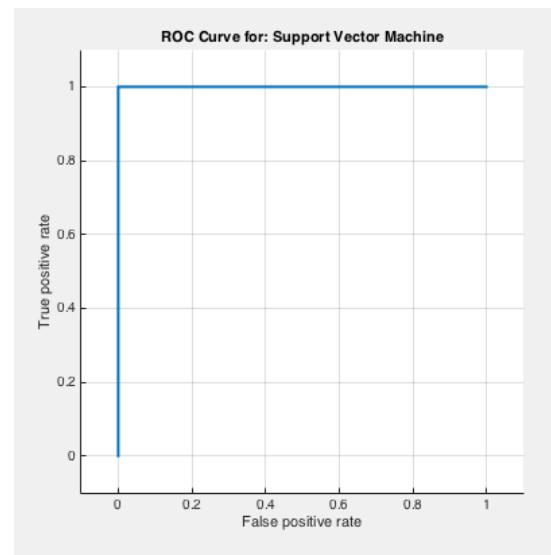


Figure 8.11: ROC curve for Quadratic SVM

Performance of the Quadratic SVM model results, are shown in figure 8.10. The confusion matrix shows 100% correct predictions, as the matrix is only diagonally dominant, meaning that the predicted labels match the true labels. The receiver operating characteristic (ROC) curve in figure 8.11, plots true positive rate, against false positive rate, for various thresholds of the classifier. ROC curve shows a perfect result, indicating no miss classified images.

8.4.2.2 Classification by Threshold

Perhaps a better approach, is to use the prior knowledge of transparency (intensity) information as a single feature for classification. Designing the classification to comprise of a series of binary threshold probability detectors. Where thresholding, is the common computer vision process of converting a single channel (gray scale) image to a binary image, using an input parameter of the intensity level I . The output binary image maps all pixels in the image to white (value 1) if their intensity values is greater than I , and black (value 0) for intensity less than I . The sum is calculated of all the white pixels in the image, divided by the total number of pixels in the image, which outputs a probability of

detection of that classification on the input image.

The images are passed to a binary threshold probability detector, tuned to detect damage with I_{damage} . If the output probability is > 0.5 the image is classified as damaged, otherwise the image is passed to a binary threshold probability detector, tuned to detect correct polymerization with $I_{polymerization}$, which then classifies the image if the output is > 0.5 , as polymerized, otherwise not polymerized.

To find optimal parameter values for I_{damage} , $I_{polymerization}$, prior training images are needed that are correctly labeled. A calibration method designed as follows, taking an input image, that contains majority of the classification, here a correct polymerization image. The histogram of the image is calculated, and the data is smoothed with a moving average filter with 11 samples (`smooth`). The peaks are found in the data, the highest peak represents the most common pixel intensity in the image, which is the uniform transparency intensity of the structure. Therefore the next peak down in the intensity values will optimally separate the background from the structure intensities we want to sample, this intensity value becomes $I_{polymerization}$. A similar calibration process is used to find I_{damage} , using an damaged input image.

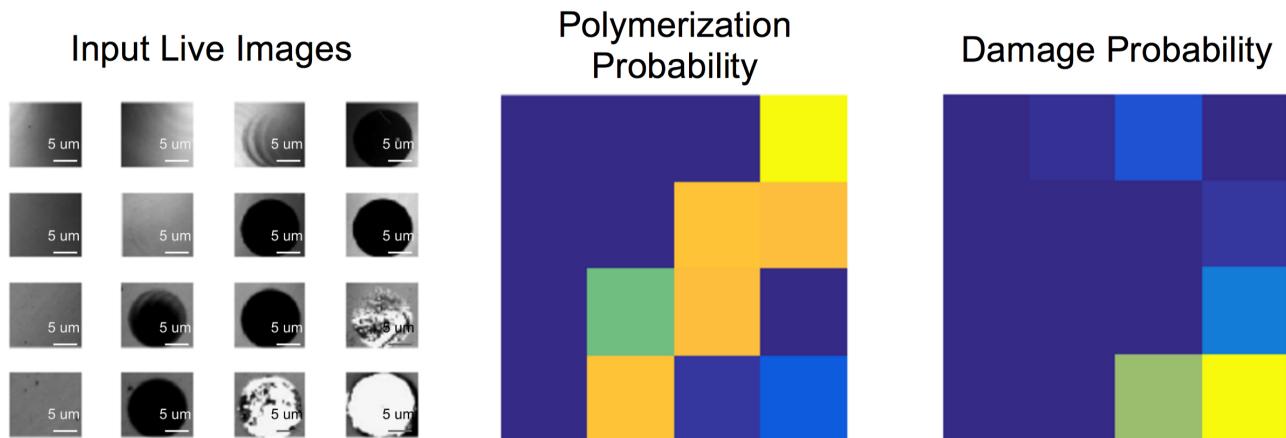


Figure 8.12: Threshold Classification Results

The accuracy of this algorithm is 87.5% using a output probability threshold of 0.5 to classify the images. This is improved to 100% using a probability threshold of 0.1 (i.e. the image has a probability of 10% or more of containing that image classification). The results of the algorithm can be seen in figure 8.12. Here the probabilities from each classification binary stage on the 16 input images are seen. The higher the probability the brighter the color pixel, it is clear that although this method of classification can robustly produce 100% classification accuracy, using a probability threshold of 0.1, it can also be used as a precise transparency arbitrary measurement tool, measuring the quality of the polymerized structure layer (for comparison between layers). This is illustrated clearly where the green pixel in the polymerization probability matrix, shows a structure layer polymerized, with a small undershoot near the edge of the structure.

8.4.2.3 Classification Conclusion

For a robust closed feedback design, the images need a robust classifier of the polymerization quality, ideally polymerizing the structure layers correctly every time. The machine learning approach shows promising image classification, however requires a large training image set. Whereas the custom threshold classification only requires a single image of each transparency type to calibrate itself, initially and then can continue providing robust classification. However it relies on high contrast changes in the transmissive images, to operate optimally.

The best design is to incorporate both design approaches into one algorithm, that is an algorithm that performs classification initially with the threshold design, when starting with the lowest minimal data set of only 3 calibrated images. As the algorithm runs, it stores classified images, building up a sample training set, to train the machine learning Quadratic SVM classifier. When > 20 images have been stored for each classification type, the algorithm would let the machine learning method continue the classification alongside the threshold classifier. The output probabilities for both predicted labeled method, would be averaged, to determine the most likely classified state.

The data set images used to design the algorithm, were of the same structure slice shape, therefore to make the algorithm work on a arbitrary shape, they would operate on a smaller cropped image, taken from inside the structure, which would be determined using the prior accurate spatial mapping design.

8.5 Summary

A optimal TPP optical design was presented, alongside a robust TPP system closed loop control design. Since all existing TPP systems employ open loop control, here a reliable and original robust technique of closed loop control was designed, tested and verified. This improves manufacturing quality for fabricating arbitrary nanometer structures, providing the end user with a reliable nanometer fabrication system, that can quickly adapt to any new photoresist resin.

The TPP optical design, can also be combined with shared optical path systems, to increase the TPP systems specifications. Fabricating resolution can be increased using the stimulated emission depletion (STED) system, to reduce minimum voxel diameter to 55nm [74], discussed further in section 14.1.7. Fabricating time can be reduced by a factor of 5, using holographic beam forming from the spatial light modulator, to create multiple polymerization beams that can be independently controlled [58], discussed further in section 16.1.

9 Optical Tweezing

Highly focused laser beams, like the ones we will be able to create in our system, can be used to exert forces on dielectric objects with mismatching refractive indexes. These **optical tweezers** are commonly used to measure forces in biological samples but will provide extended capabilities to the nano fabrication system and adds a unique selling point.

9.1 Theory

Due to the variation in the shape, size and composition of objects, as well as the complex distribution of the light intensity, the comparative sizes of objects and wavelengths a full theory of optical tweezers has evolved slowly compared to experimental work. [75]

9.1.1 General Principle

The laser beam is focussed by a high NA microscope objective lens (See subsection 17.2.1 on page 186). The beam is brought to a very small focus which contains a strong electric field. This region is referred to as the **beam waist**. Due to the gaussian beam shape (section 3.4), the field is strongest at the centre and the dielectric particle is attracted towards it. This provides the stabilisation of the particle both perpendicular to and in the direction of the beam, due to the shallow depth of field. Photons that are scattered or absorbed by the particle exert a force due to the conservation of momentum which pushes the particle away from the objective, along the direction of the beam. This **scattering force** causes the neutral location of the trap to be ever so slightly below the focus point.

For small displacements, the force can be shown to be linearly dependent on the displacement, according to Hooke's law.[76] More detailed explanation depends on the relative size of the particle to the wavelength of the beam.

9.1.2 Ray Optics Explanation

This explanation is only suitable when the size of the trapped particle is significantly larger than the wavelength of the light used to trap it.

As shown in Figure 9.1 b), the light is refracted by the bead, which has a higher refractive index than the surrounding material. Rays hitting the bead on the right hand side are refracted towards the left and vice versa. This changing of direction of the light requires the particle to exert momentum and an opposite and equal force is exerted on the particle. If the bead is on the left of the centre of the beam, as in the figure, the rays hitting it's right side will be stronger because of the Gaussian intensity profile. This causes a net force towards the right which will push the bead towards the centre of the beam. This provides the transverse trapping force.

Figure 9.1 part c) shows what happens in the axial direction when the bead is at the centre of the beam but below the point of focus. The transverse forces cancel and because of the focussing of the

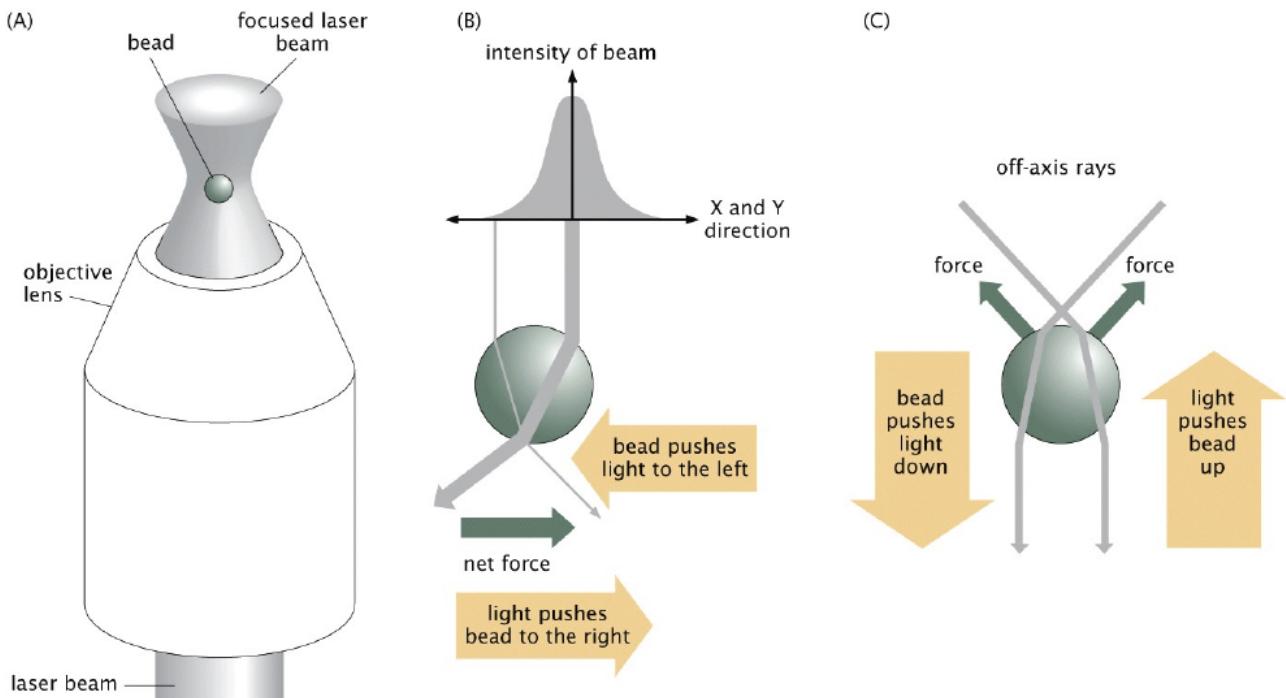


Figure 9.1: Ray Optics explanation of Optical Tweezers. The bead is trapped due to the transfer of momentum from the bead to the photons. Courtesy of [77]

beam, the rays hit the bead at an angle. The rays are refracted downwards which results in a force on the bead against the direction of propagation. If the bead is above the point of focus, the rays will be refracted upwards, leaving the bead at a steeper angle. This results in a force along the direction of propagation. The same happens when the bead is above the point of focus, causing the rays to become more divergent. The net axial force due to these phenomena will cancel the scattering force trying to push the bead away. The stable location is therefore slightly below the beam waist. [76]

9.1.3 Dielectric Explanation

When trapping objects which are smaller than the wavelength of the trapping laser, a different explanation must be used. A polarizable particle in an electric field reduces the net field. The lowest energy state of the system will be when the particle is at the position of greatest field. This field is greatest at the focus which is therefore a potential well which traps the particle. The main assumption is that the particle is small enough that the instantaneous electric field experienced due to the electromagnetic laser is uniform across it.

The **scattering force**, as explained above, is one of the forces acting on the particle. The electric field is oscillating and the dipole follows these oscillations. The oscillating electric dipole then radiates its own electric waves isotropically in all directions. These scattered waves change the overall electromagnetic flux and the net effect is in the direction of the laser beam, the force exerted given by

$$\mathbf{F}_{\text{scat}}(\mathbf{r}) = \left(\frac{n_2}{c}\right) C_{\text{scat}} I(\mathbf{r}) \quad (9.1)$$

where \mathbf{r} is the displacement vector of the particle from the point of focus of the beam, I is the time averaged intensity of the beam and C_{scat} is the scattering cross section given by

$$C_{\text{scat}} = \frac{8}{3}\pi k^4 a^6 \left(\frac{m^2 - 1}{m^2 + 2} \right)^2 \quad (9.2)$$

where $k = 2\pi/\lambda$ is the wave number. Combining these two equations above, the scattering force is proportional to the beam power and the radius of the particle a^6 . [78]

The second part is the **Lorentz force** which describes the electric field acting on the dipole. The **dipole moment** p measures a system's polarity by giving a measure of separation of negative and positive charges. The elementary definition for two point charges, separated by distance d is $p = qd = \alpha \mathbf{E}$ where \mathbf{d} is a vector pointing from the negative charge to the positive charge and α is the polarizability of the media. The polarizability of a round sphere derives from the Rayleigh theory of scattering and is given by a solution to Laplace's equation with surface boundary conditions. The derivation also shows that a spherical cavity in a uniform electric field behaves just like a dipole. [79, p205]

$$\alpha = 4\pi a^3 \epsilon_2 \left(\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + 2\epsilon_2} \right) = 4\pi a^3 n_2^2 \epsilon_0 \left(\frac{m^2 - 1}{m^2 + 2} \right) \quad (9.3)$$

where a is the radius of the sphere, $m = n_1/n_2$ is the relative refractive index of the particle relative to its surroundings. At the frequencies of light, $\mu \approx 1$ so that $\epsilon_r \approx \epsilon_0 n_r^2$. The dipole moment is then given by:

$$\mathbf{p}(\mathbf{r}, t) = \alpha \mathbf{E}(\mathbf{r}, t) = 4\pi n_2^2 \epsilon_0 a^3 \left(\frac{m^2 - 1}{m^2 + 2} \right) \mathbf{E}(\mathbf{r}, t) \quad (9.4)$$

where $\mathbf{E}(\mathbf{r}, t)$ is the instantaneous electric field. [78, p5]

The instantaneous force on the dipole is given by

$$\mathbf{F}(\mathbf{r}, t) = \mathbf{p}(\mathbf{r}, t) \cdot \nabla \mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \alpha \nabla \mathbf{E}(\mathbf{r}, t)^2 \quad (9.5)$$

which needs to be averaged to find the gradient force on the particle.

$$\mathbf{F}_g(\mathbf{r}) = \frac{1}{2} \alpha \nabla \left\langle \mathbf{E}(\mathbf{r}, t) \right\rangle_T = \frac{1}{4} \alpha \nabla |\mathbf{E}(\mathbf{r})|^2 \quad (9.6)$$

The time averaged electric field can be found from the intensity of the beam which is defined as

the average of the Poynting vector. [78]

$$I(\mathbf{r}) = \frac{1}{2} \operatorname{Re}[\mathbf{E}(\mathbf{r}) \times \mathbf{H}^*(\mathbf{r})] = \frac{n_2 \epsilon_0 c}{2} |\mathbf{E}(\mathbf{r})|^2 \quad (9.7)$$

$$\mathbf{F}_g(\mathbf{r}) = \frac{1}{2n_2 \epsilon_0 c} \alpha \nabla I(\mathbf{r}) \quad (9.8)$$

Which shows the Lorentz force is in the direction of highest intensity at the focus. Using Equation 3.28 for the Gaussian beam intensity and Equation 9.3 for the polarizability of a sphere we get

$$\mathbf{F}_g(\mathbf{r}) = \frac{2\pi n_2 a^3}{c} \left(\frac{m^2 - 1}{m^2 + 2} \right) \nabla \left[\frac{P}{\pi w^2(z)} \exp \left(-2 \frac{r^2}{w^2(z)} \right) \right] \quad (9.9)$$

9.2 HOT

Holographic optical tweezing uses a SLM to dynamically change the shape of the beam to create multiple traps or affect the trapped particles in other interesting ways. Our system includes all the components necessary to shape the beam for use with a variety of holographic techniques. This greatly extends the capabilities and applications which will increase the value of the system to researchers. Figure 9.2 Shows 26 silica spheres with a diameter of 1 μm being transformed using

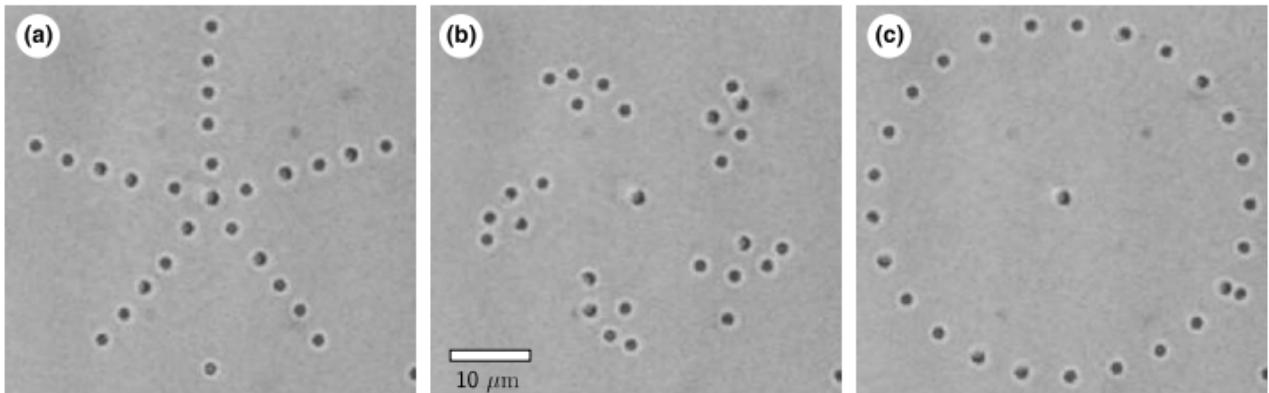


Figure 9.2: Patterns formed with dynamic holographic optical tweezing, from [80]

computer generated holograms. By generating the holograms on the fly or pre-computing them, multiple particles can be simultaneously moved around or larger structures can be trapped in several locations at once. chapter 16 on page 163 describes the parallelisation of fabrication techniques and how the SLM can be used to generate many focal points. In principle, a beam which enters the back aperture straight on will converge at the focal point of the objective. (see section 17.7) If the beam enters the objective at an angle, it will be brought to a focus on the focal plane, to the side of the focus point. A beam that is slightly converging will be focused closer to the objective and a diverging beam further away. This allows multiple traps to be created in three dimensions and by updating the

SLM frame by frame the points of focus can move around, thereby dragging the trapped particles with it.

9.3 What type of objects can be picked up

In extreme cases, objects as small as 30nm can be manipulated by optical traps although in practice the objects are most often comparable in size with the wavelength of the trapping laser. [75, 76] Our live imaging system will be limited to resolutions of 300nm. (section 12.1) While smaller particles can still be trapped, the tracking precision will be limited. For sizes up to the laser wavelength (800nm), the objects can be of arbitrary shape as the aforementioned dipole approximation holds. Larger objects work best if they are round and transparent. As the particle size increases beyond the laser wavelength, the refraction at the surface decreases and more powerful lasers tend to be used. Large beads up to 4.5 μm can be easily trapped with conventional tweezers, with limits of high refractive index beads normally around 10 μm .[81] Using more powerful Laguerre-Gaussian doughnut shaped beams (as explained below), large 12 μm beads can be comfortably trapped.[82]. This is because the scattering force is reduced in the central zero of the beam.

9.4 Simulation

2PP structures have refractive indexes of $n_1 = 1.54$ [83] By suspending them in water with $n_2 = 1.33$ we get a relative refractive index of $m = 1.158$ and using the **LUMPLFLN 40XW** objective lens with a numerical aperture $\text{NA} = 0.80$, chosen in subsection 17.10.5, our system fits well within the most common set-ups for optical tweezing so we can handle the aforementioned range of sizes.

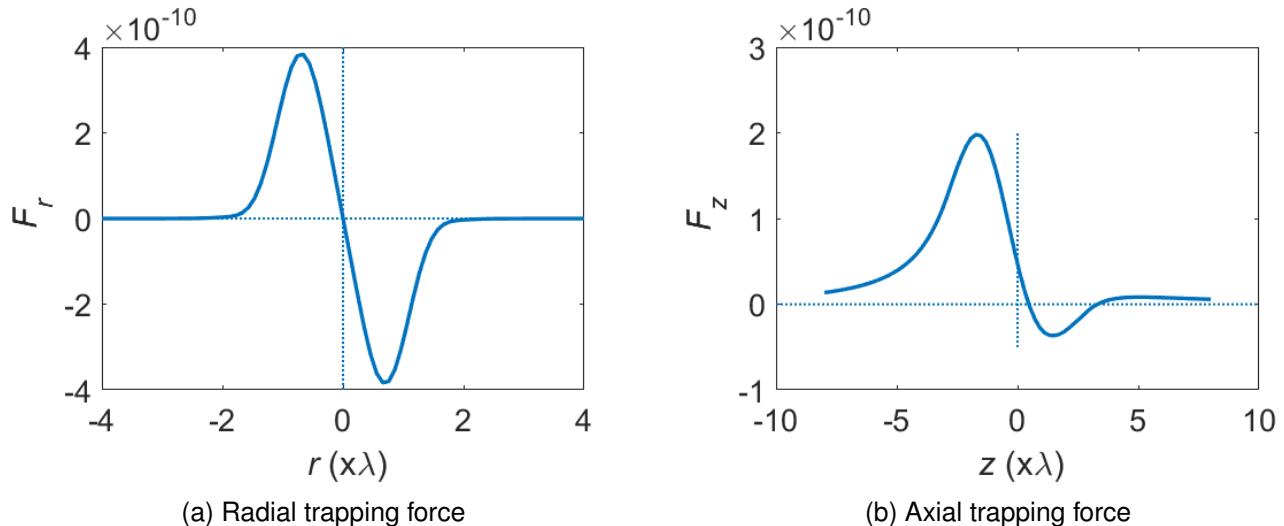


Figure 9.3: Forces on a 1 μm 2PP bead in water using $\text{NA} = 0.8$, $\lambda = 800\text{nm}$, $P = 368\text{mW}$

The theoretical derivation of the dipole approximation gives useful insights into the fundamental principles, but because many particles are of similar dimensions as the wavelength a much more

complicated theory is required. The derivations are too advanced for this report but researchers at the University of Queensland have developed a MATLAB toolbox which implements the latest theoretical and empirical results. [84] Using this we are able to predict the forces for different particle sizes and shapes considering our exact system qualifications like the refractive indexes, NA, and $\lambda = 800\text{nm}$ wavelength. The laser output power is reduced to about 70% throughout the system so $P \approx 525 * 0.7 = 368\text{mW}$. Figure 9.3 are results for a typical sphere with $1\mu\text{m}$ diameter. 9.3a shows a clear linear relationship between force and displacement across the radial section of the beam with a trapping constant of $k = -0.05$. The maximum trapping force is roughly $F_r = 40\text{nN}$ at $r = 800\text{nm}$ displacement which indicates a very stable trap. Along the direction of the beam, the scattering force influences the trap as it can be seen in 9.3b. The neutral trapping position, where the force reaches zero, is slightly below the focal point.

9.5 Alternative Beam shapes

The SLM can not just be used for parallelisation but also to shape higher order beams. Laguerre-Gaussian modes are Laguerre polynomials combined with the conventional Gaussian beam profile. Two such modes can be superimposed to form ring-shaped beams with very interesting properties. By reducing the scattering force in their dark centre, the aptly named "doughnut mode" can be used to trap particles that are absorbing as well as having an increased trapping efficiency meaning they can create more force for the power used. They contain a phase singularity, where the phase changes rapidly by 2π . Each photon in such a beam carries angular momentum associated with the helical structure of the wave around the singularity. This angular momentum can be transferred and set objects into rotation.

The same type of beam is used in STED microscopy in subsection 14.1.2 which includes some diagrams. The ability to precisely rotate a particle is incredibly useful for observation and for powering micromachinery.

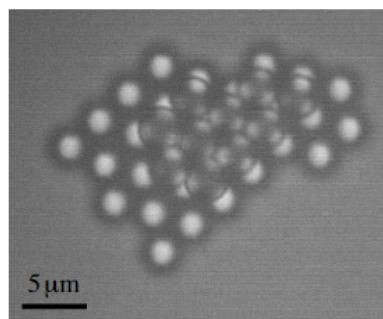


Figure 9.4: Two layers of an FCC crystal assembled with silica beads. From [85]

9.6 Use cases

9.6.1 Fabrication

Holographic optical tweezers can be used to move silica spheres and fix them in place through photo polymerisation of monomers in the sample solution. Figure 9.4 shows $3\mu\text{m}$ diameter spheres who have been assembled into an FCC crystal structure. However, the use cases for fabrication are somewhat limited because we have far superior techniques in our system. More applications are in the assembly of objects which were built using the different fabrication techniques.

9.6.2 Microfluids

With their ability to induce forces on a small scale, optical tweezers are ideal for influencing the flow in microfluidic applications. A large variety of actuators can be driven which redirect flows and particles within. Figure 9.5 shows a variety of such devices. The gate in Figure 9.5a, for example, redirects where particles are flowing, which can be implemented in sorting systems. HOT is used to move the individual beads in 9.5c and drive a flow of liquid.

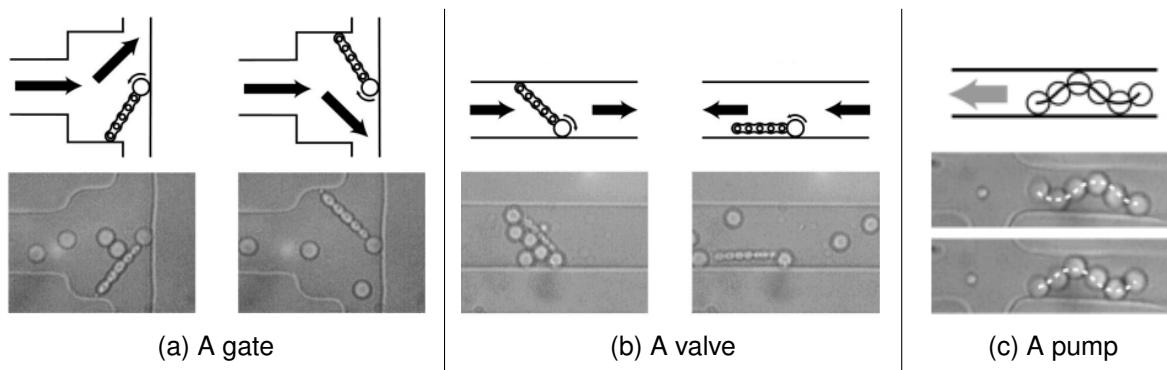


Figure 9.5: Microfluid actuators driven by optical tweezers. Beads are $3\mu\text{m}$ across. Courtesy of [86].

9.6.3 Micromachines

Using the beams discussed before in 9.5, rotational micromachines can be powered easily and effectively. In a landmark paper, Galajda and Ormos built interlinking cogwheels using two-photon polymerization which they powered using a rotor which is held in place and rotated by optical tweezers. [87] The solid arrows are the fixed wheels while the dashed arrow points to the trapped and powered rotor. This system is shown in Figure 9.6a and demonstrates the potential to create complex, light powered mechanical structures.

9.6.4 Lab on a Chip

All the above use cases come together when implemented as a "lab on a chip". As discussed in the business case, chapter 26, this new and exciting area of research is a prime target for this product. Medical, chemical, and biological researchers are hoping for much faster, easier and cheaper analy-

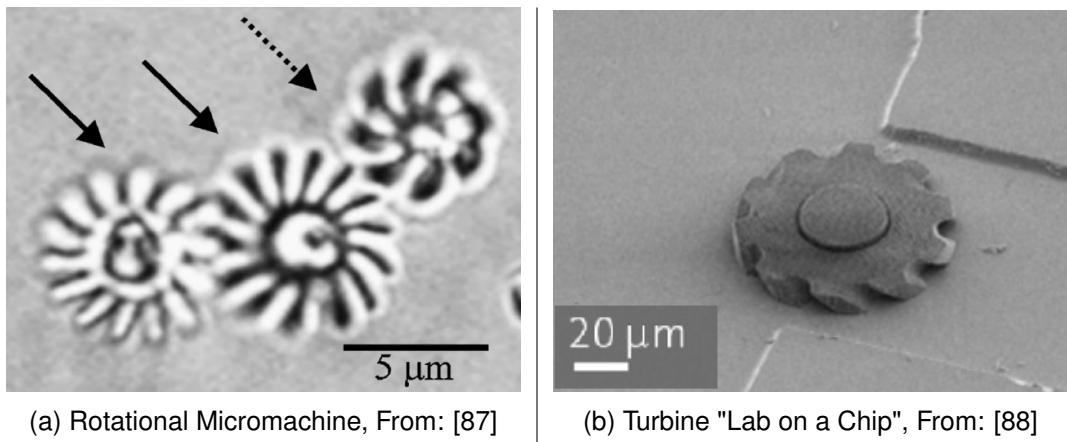


Figure 9.6: Rotational Machines driven by doughnut beams.

sis and diagnosis. Optical tweezers are key, but the whole system really comes together to massively simplify the development. Channels can be cut into a glass substrate with our micromachining capability or 3d printed using stereo lithography. Intricate actuators, turbines, structures, cell enclosures and lattices can be manufactured by two-photon polymerisation and integrated into the channels. Figure 9.6b shows an example use case of a turbine used as a pump. The designs can be tested right on the cassette. Computer automated, high resolution imaging techniques such as confocal, STED, or two-photon microscopy, could be used to analyse specimen and holographic optical tweezers then used to move them between channels depending on certain characteristics. Doughnut beams can drive micromachines and turbines which interact with the samples. The opportunities are endless and all in the hands of our customers. Example articles which use optical tweezing for microfluid research include [89, 90]

9.7 Setup

The high frequency repetition rate (80MHz) Vitara-T seed output of the 800nm Ti-Sapphire laser goes through the Beam Control and Shutter setup as explained in chapter 20. The beam shaper stretches the pulse out so that the pulsed laser becomes quasi-continuous and can be used for optical tweezing - see subsection 20.8.2 . The beam bounces off the SLM (SLM1) which will be used for multiple foci generation and higher order beam modes. As usual, the SLM is imaged through the two 4f systems and the galvanometers onto the back of the objective where it is brought to a tight focus on the stage. If the 532nm laser is used for a single trap, the galvanometers are used to steer the trap location. Imaging is done as explained in section 12.1 with the red LED transmitting through the sample, off the RDM2 diachronic mirror onto a CMOS sensor which is read out by the computer.

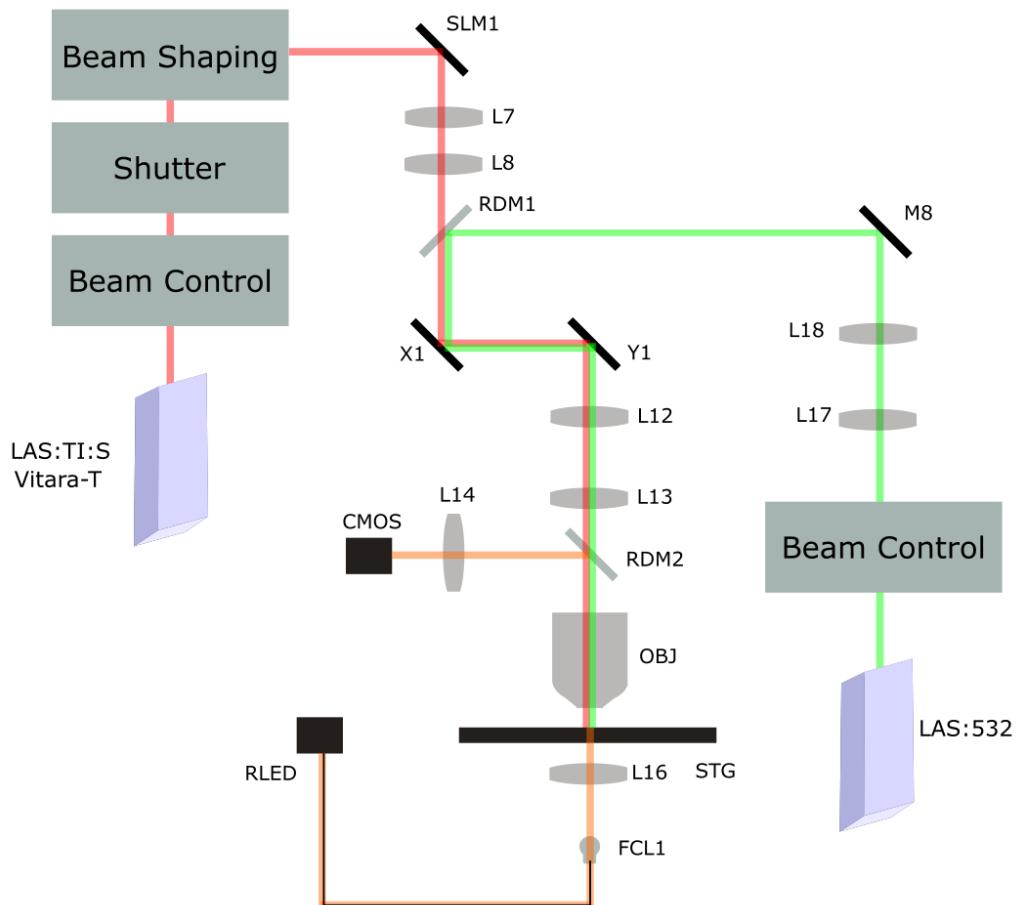


Figure 9.7: Schematic layout for Holographic Optical Tweezing

9.7.1 Conclusion

Adding the ability to move and interact with objects highly increases the utility of our system, especially for post fabrication experiments and assembly. With HOT being able to lift and arrange a large number of objects simultaneously, complex devices can be created and large objects can be manipulated and orientated. These capabilities put our clients to the forefront of multiple areas of research.

10 Excluded Techniques

10.1 Selective Laser Sintering

Selective Laser Sintering (SLS) is one of the additive manufacturing technique invented in 1980s by Dr.Carl Deckard[91]. The process resembles stereolithography as the parts are built layer by layer. The difference arises at raw materials. Sintering turns powdered material into a solid mass by heat or pressure without melting it. Material of fabrication includes metal, plastics, ceramics and even glass. Current application is mainly in aerospace industry and auto racing industry for producing light weight and durable parts. It could be also used to produce model of structures or artworks.

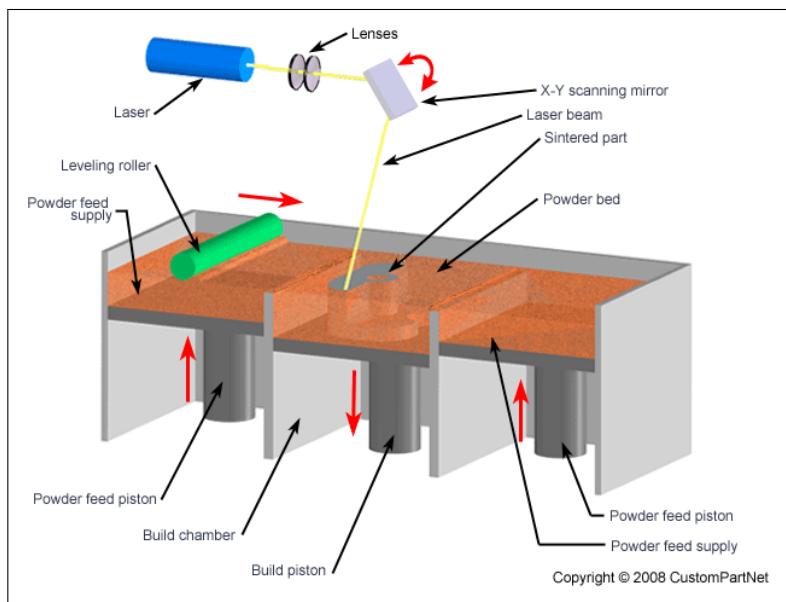


Figure 10.1: Method demonstration for SLS[92]

The reason of excluding this technology are following:

1. Sintering process does not perform well when it comes to nanoscales. In current technology, the finest feature size we can get from laser sintering is approximately 0.005 inch (0.127 millimeters). Furthermore, material in used in SLS is not compatible with two-photon polymerisation. The parts completed by SLS could not be further processed as the parts from SLA hence resolution cannot be refined further.
2. The surface of completed part could be rough and requires surface finish. It might be difficult to realize it in completely automatic design.

Part III

Imaging

11 Imaging Introduction

This section of the report is divided into 3 parts - live imaging modes, post process imaging and super-resolution imaging methods. We first explain in detail the chosen processes and their implementation in the final design before discussing some alternative methods and why they were excluded with brief notes about how our product could be developed in the future.

Users of our system will benefit from an advanced live imaging set-up which makes use of sophisticated polarized light imaging modes in addition to brightfield microscopy in order to provide optimal contrast for different materials. Our live imaging system is required to offer closed-loop feedback in the Two-Photon Polymerisation process as described in 8. Resolution is not optimised for live imaging and the systems have been designed to provide just enough resolution and contrast to let the user know if there are any problems while also being video rate (50Hz)

We also want to provide our users with an advanced set of post-process imaging tools in order to examine samples once they have been fabricated. Using fluorescent confocal microscopy and two-photon microscopy, we obtain resolutions down to 150 nm. Our post-processing imaging tools can also be used as a standalone microscope.

Finally, our unique incorporation of STED into both our two-photon polymerisation and confocal imaging processes allows our users to access lateral resolutions of below 60 nm for certain materials and dyes.

11.1 Requirements and Specifications

For reference throughout this section, we summarise the requirements of our system's fabrication processes in terms of their resolution and whether they require live feedback in figure 11.1.

Process	Resolution	Live feedback required
Stereolithography	12 μ m	yes
Laser Micromachining	700 nm	yes
Two Photon Lithography	100 nm	yes
STED Lithography	600 nm	no
Optical Tweezing	N/A	yes

Table 11.1: Imaging requirements for the different fabrication processes

11.2 Transmitted and Reflected Microscopy

Transmitted light microscopy and reflected light microscopy are two main branches of microscopy modes. Transmitted microscopy is conventionally used for observation in life sciences, while reflected microscopy is more commonly used for material sciences and semiconductor. Our system includes both because of the varying thickness and material of our product. In our system, transmitted light microscopy is used for live imaging and reflected light microscopy for post process and super resolution imaging. In this section, I will first explain the concepts and advantages of two modes, and then moves on to the reasons behind our decision.

As the name suggests, transmitted light microscopy is where the light is transmitted from a light source on the opposite side of the specimen from the objective. Generally, transmitted light microscopy allows more light to pass through and can be used with combination of several techniques. However, the sample must be translucent so that light can pass through the sample. Also, as thick specimen causes more light absorption, thin specimen ($1 - 30\mu m$) specimen is preferred.

On the other hand, reflected microscopy is where the light source lies on the same side as the objective. The objective here acts both as a condenser and an objective. As light is reflected, opaque sample and surface structure of thick sample can be imaged.

Our system aims to produce structure up to $8cm \times 8cm \times 8cm$. The upper limit of the thickness makes it unsuitable for transmitted light microscopy. Thus, the post process and super resolution imaging adopt reflected light microscopy so as to image the product (>30 micro thickness) after fabrication process. Reflected microscopy also enables the examination of surface structure for opaque materials, such as metal and semiconductors.

On the other hand, live imaging is required for two-photon polymerization to provide the closed loop feedback. In two-photon polymerization, high NA lens is used so that spot size is small. On the other hand, live imaging technique needs to capture an area of the sample so as to monitor the process. If live imaging uses the same high NA objective for its condenser, it will also only be able to capture the image of a spot area and thus unable to reveal the progress. Instead, a different condenser should be used to illuminate the whole sample. Thus, transmitted light microscopy setup is adopted, where the condenser and the objective lies on the different sides of the sample. The details of implementation of live imaging will be discussed in the next section.

12 Live imaging modes

The ability to view the sample in real time is highly desirable for an optical fabrication machine. This allows the customers to monitor the fabrication process and input instructions accordingly. Live imaging is also essential in proving the closed-loop feedback for two photon polymerization process and optical tweezing. For our system, we considered different imaging techniques for transmitted light microscopy, including bright field microscopy, polarized light microscopy, dark field microscopy and phase contrast microscopy. Their advantages and disadvantages are discussed respectively and a decision is made based on cost, applicable materials and complexity of set up.

Our system incorporates bright field microscopy and polarized light microscopy. Bright field microscopy is for naturally pigmented or highly contrasted samples. Its advantage includes simple set up with little preparation. If the sample is transparent, bright field microscopy will produce an image of low contrast, thus of little use. To enhance the contrast for transparent samples, polarized light microscopy, dark field microscopy and phase contrast can be used. Polarized light microscopy is suitable in our case because the fabrication product is optically anisotropic [93]. On the other hand, dark field microscopy requires strong illuminating and might cause undesired polymerization. Phase contrast microscopy, produces distorted images for thick specimens. Thus, bright field microscopy and polarized light microscopy are included in our system. Concepts, implementations and considerations of different imaging techniques will be discussed in detail in this section.

Live imaging, designed in our system, aims to provide fair resolution for continuous fabrication process. Illuminating with a red LED (625nm), our setup is able to achieve a resolution of 300nm and scanning speed of 100 frames per second with median readout noise less than 1 electron when a 1.4 NA oil immersion objective is used.

12.1 Technique: Bright-field Microscopy

12.1.1 Basic Principle of Bright-field imaging

Bright-field microscopy is the most commonly used observation technique in optical microscopy. The term "bright-field" refers to the fact that the specimen is dark and the surrounding viewing field stays bright. Most microscopes are capable of bright field imaging. It is widely used for stained specimen or highly contrasted specimens. However, it has disadvantages such as low contrast for transparent specimen and limited resolution.

Bright-field microscopy is adopted in our system for live imaging of two-photon polymerization because it is simple to set up and gives images of fair resolution (300nm). It is a usual practice to use bright field microscopy to monitor TPP process in the real time [94]. The ability to image the sample

in the real time is crucial in determining the position for TPP. It is also essential for spatial feedback control (chapter 8).

12.1.2 Instrumentation of Bright-field imaging

In our system, bright-field transmitted microscopy is used for live imaging of two-photon polymerization (TPP). As shown in figure 12.1, the orange line indicates the optical path for bright field live imaging and the red line indicates the optical path for 2PP. The light emitted from a red LED passes a condenser lens (**L16**), which focuses light onto the sample (**Green block**). Transmitted and scattered light from the sample is then collected by the objective lens (**OBJ**) and the image is magnified. The resultant light reflects off a dichroic mirror (**RDM2**) and is collected by the **CCD**. The polymerization light (red line) passes the dichroic mirror (**DC**) and does not contribute to the image formation.

As our automated stage is not transparent, to achieve transmitted microscopy, an additional cantilever component with a transparent center is designed (**STG**). The cantilever enables support from the side and is empty in the middle (dotted line), which allows light to pass through. The mechanical implementation of transmitted light microscopy will be discussed in detail in Chapter 22.

12.1.2.1 Light source - Fibre coupled red LED

Fibre-coupled red LED (625nm) is selected as the light source in this imaging mode. Compared with conventional LED, fibre directs the light and thus enables additional positioning of the light output. For our integrated system, different fabrication and imaging techniques use the same objective turret. Thus, removal of red LED source and condenser out of the pathway is needed when other techniques are used. Instead of using multiple removable elements (mirrors and condenser) in the system, fibre coupled LED can be compounded with a condenser lens and conveniently moved in or out the optical pathway. This greatly reduces the alignment complexity. Figure 12.2 shows that light source (FCL1) is fixed in positioned in relation to condenser lens (**L16**) and polarizer (**POL1**).

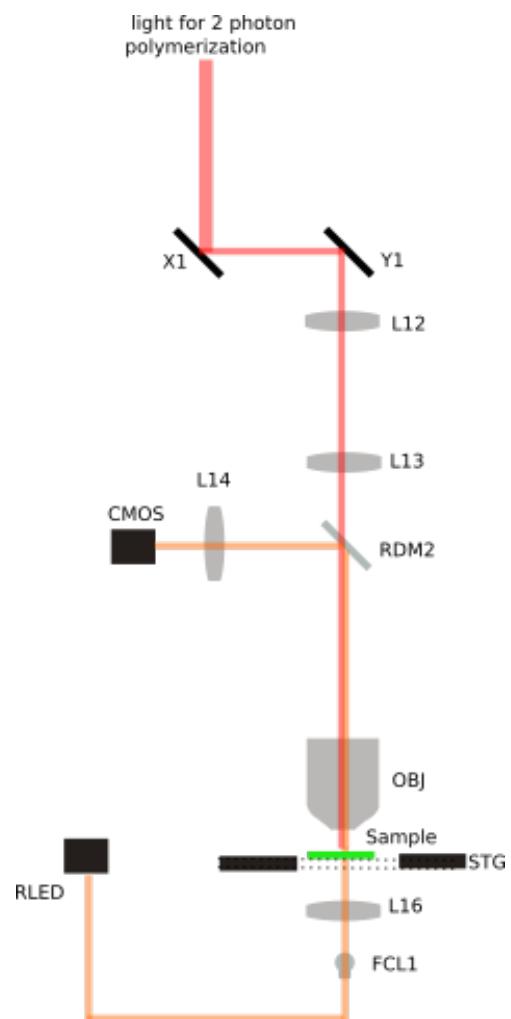


Figure 12.1: Optical Layout for Bright field microscopy

Red LED is chosen as the light source in this imaging mode because of three reasons. Firstly, in 2PP, two photons of wavelength (800nm) are absorbed at the same time to stimulate polymerization. Thus, the excitation wavelength of 2PP process is near the UV region (400nm). To avoid interference of observation light source with the polymerization light, the observation should be far away from UV, which makes red light a suitable option. Secondly, the advantage brought with a shorter wavelength is not that significant, as shown in section 12.1.3 later.

The red LED is also used for the live imaging for stereo-lithography, where its excitation wavelength is 355nm. Using red LED can also ensure that imaging light will not interfere with the fabrication process and a suitable dichroic mirror can be chosen to separate the two rays.

12.1.2.2 Detector - Scientific CMOS

For multiple point detectors, complementary metal-oxide semiconductor (CMOS) and charged coupled device (CCD) are available. Both CMOS and CCD use photodiodes to convert light into electrons. Their difference lies in how they read the value. In a CCD sensor, the charge is transported through a limited number of output nodes to be converted to voltage. While in CMOS, each pixel has its own charge to voltage conversion. The difference in mechanism makes CMOS generally faster but noisier than CCD.

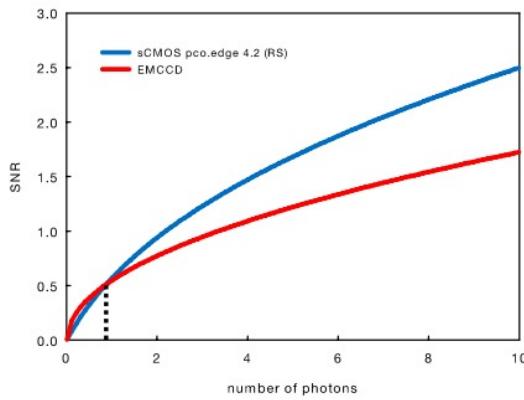


Figure 12.3: Signal to noise ratio for typical CCD camera and selected sCMOS [95]

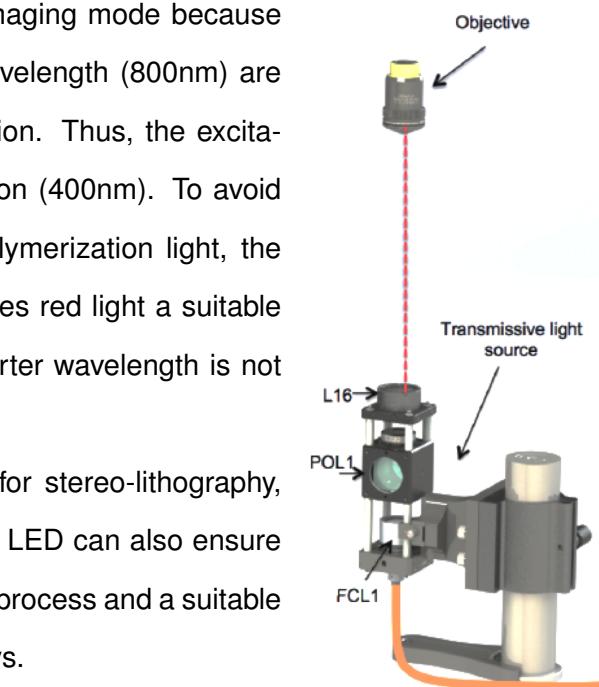


Figure 12.2: mechanical setup of fibre-coupled LED

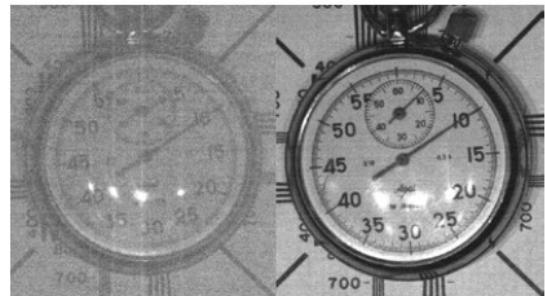


Figure 12.4: image by CCD (left) and image by selected sCMOS(right)[95]

In our system, we opted for a scientific CMOS, which is a new generation for a CMOS camera. Scientific CMOS is able to deliver high sensitivity, fast readout speeds with low noise. Figure 12.3 shows the signal to noise ratio for scientific CMOS is mostly better than that of CCD and this effect is illustrated in figure 12.4.

From the data sheet, our selected scientific CMOS is able to give high quantum efficiency, fast readout speed and high signal to noise ratio. It is able to give a quantum efficiency of 80 % at the wavelength of 625nm. For full frame ($2048\text{pixel} \times 2048\text{pixel}$), the frame rate is 100 frames per second, which means each frame takes 0.01 second. The actual fabrication speed depends on material, but the shortest time to generate a 2D scan with our galvanometer is 7 seconds. Thus, the scientific CMOS can take images much faster than our fabrication process and thus live imaging is achieved. Also, the median readout noise is only 1 electron and the dynamic range is $2^{16} = 65536$, which means our selected sCMOS can give excellent signal to noise ratio. To further reduce the systematic noise, a blank sample should be used before the fabrication process, so that fixed pattern noise can be accounted for.

In addition, we need to check whether the objective chosen (MPLAPON 100XO) is able to give us the optical resolution (300nm). According to Nyquist-Shannon sampling theorem, the sampling rate must be at least twice of the highest frequency component. In this case, two elements can be characterized as distinct if they have at least two pixels per resolvable element. In other words, 300nm must cover at least two pixels. Since each pixel size is $6.5\mu\text{m}$, a total magnification of $6500/150 = 43$ must be achieved at least. Hence, 100X lens will oversample the image and is able to achieve the optical resolution.

12.1.3 Modelling of Bright-field imaging

Even if the microscope is equipped with optical instruments of the highest quality and has the highest numerical aperture, its resolution is approximately half the wavelength of light. This is due to diffraction. For a point source, the spreading of light waves is referred as point spread function, which is discussed in detail in section 3.3.

The resolution of bright field microscopy is given in Rayleigh criterion.

$$\text{Resolution}_{x,y} = 0.61\lambda/NA \quad (12.1)$$

Thus, with NA1.4 oil immersion lens and 625nm red LED, the resolution we can achieve is

$$\text{Resolution}_{x,y} = 0.61 * 625\text{nm}/1.4 = 272\text{nm} \quad (12.2)$$

In the next section, I will introduce the concept of Optical Transfer Function, in order to model the bright-field imaging. This also establishes the background for post process imaging techniques. I will then apply the theoretical model to our bright field microscopy set up and verify the optical resolution.

12.1.3.1 Optical transfer function

Optical transfer function (OTF), is the Fourier transform of the point-spread function (section 3.3). It can be used to describe the lens system.

$$OTF(x) = \int PSF(x)e^{-ivx}dx \quad (12.3)$$

where x is the spatial coordinate in the image plane.

The modulus of OTF is **modulation transfer function(MTF)**, which measures the microscope's ability to transfer contrast from an object to the image plane at a specific resolution. The argument of OTF is **phase transfer function(PTF)**, which measures the phase transfer.

$$OTF = MTF \times e^{iPTF} \quad (12.4)$$

MTP incorporate resolution and contrast data into a single specification. The degree of modulation (contrast) of an output signal is a function of spatial frequency (resolution). **Spatial frequency** describes the number of spacing per unit interval in a specimen. **Modulation** describes the ratio between difference in intensity and sum in intensity.

$$Modulation = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (12.5)$$

where I_{max} and I_{min} refers to the maximal and minimal intensity respectively.

Thus, modulation transfer function at a certain spatial frequency can be calculated as

$$MTF = \frac{Image}{Object} \frac{Modulation}{Modulation} \quad (12.6)$$

12.1.3.2 MTF for bright field microscopy

For incoherent light source (red LED), the point spread function is considered. For a perfect aberration free optical system with a circular aperture and centered optical axis, OTP can be simplified to MTF because the phase is zero for all spatial frequency in all direction. By convention, the MTF is normalized to unity at zero spatial frequency. MTF for an optical system with a uniformly illuminated circular aperture can be expressed as [96]

$$MTF = 2(\varphi - \cos\varphi \sin\varphi)/n \quad (12.7)$$

where

$$\varphi = \cos^{-1} (\lambda v / 2NA) \quad (12.8)$$

In the equation 12.8, v is the frequency in cycles per millimeter, λ is the wavelength of illumination, and NA is the numerical aperture. From the equation, it is obvious that for the same contrast (modulation ratio), spatial frequency is inversely related to the wavelength and linearly related to NA. In other words, to achieve high resolution, short wavelength and large NA is desirable.

We can combine the two equations (12.7,12.8) and plot it with MATLAB for our system. When 625nm red LED and NA 1.4 objective are used, figure 12.5 (blue) shows that modulation ratio decreases monotonously with increasing spatial frequency. The red dot is where the maximal resolution (300nm) is achieved, where modulation ratio is 0.15. When other objectives are used, same modulation ratio corresponds to a lower spatial frequency. Thus, NA 1.4 is able to give the best resolution.

Figure 12.6 shows the modulation transfer function when different wavelength are used. Shorter wavelength will lead to better resolution, although the benefit is not significant. However, due to reason stated in section 12.1.2.1, red LED is chosen.

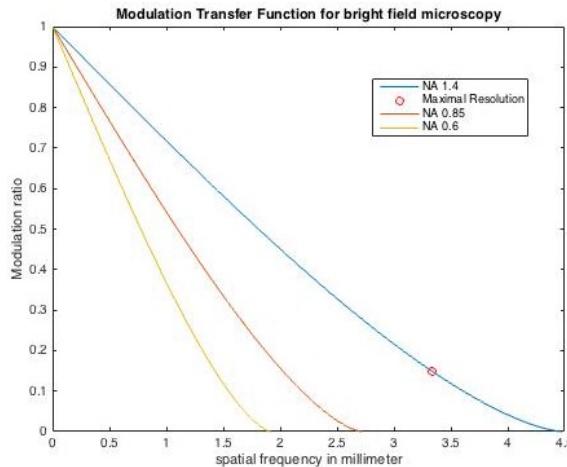


Figure 12.5: MTF for different objectives

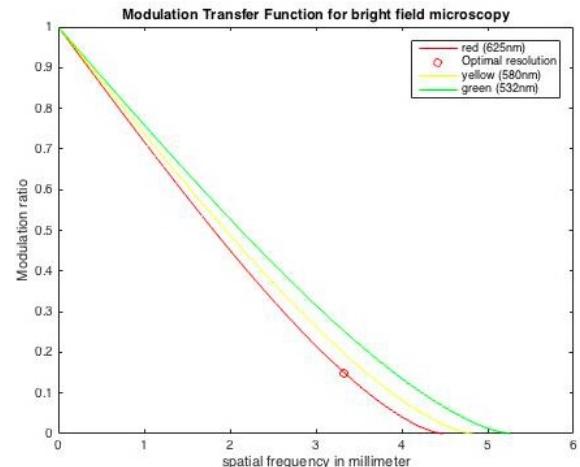


Figure 12.6: MTF for different wavelengths

12.2 Polarized Light

12.2.1 Basic Principle of polarized light imaging

Polarized light microscopy is a widely adopted technique for viewing transparent sample. It takes the advantage of anisotropic optical property of certain specimen to provide contrast-enhanced image. Figure 12.7 is an example of bright field microscopy with a plane polarizer (left) and polarized light microscopy (right). In live imaging, if the product is naturally colorless or transparent, bright-field microscopy will give an image of low contrast. Polarized light microscopy can enhance the contrast and be used to monitor the fabrication progress if the product is anisotropic. [93]

The term **anisotropy** means property is directionally dependent. In this case, the refractive index depends propagation direction of light. The optically anisotropic materials are said to be **birefringent**. As shown in figure 12.8, when light enters the material in a direction that is either parallel (b) or perpendicular (c) to its optical axis, the material behaves like an isotropic crystal. However, when light enters its non-equivalent axis (a), light is refracted into two rays, ordinary ray and extraordinary ray. The two rays are perpendicular to each other and travel at different velocities. **Ordinary ray** is the one whose polarization direction is in the direction of the optical axis. Its velocity is same in every direction within the crystal. Extraordinary ray has a velocity that is depend upon the direction within the crystal.

The difference in refractive index, or **birefringence**, can be expressed as

$$\text{Birefringence} = |n_e - n_o| \quad (12.9)$$

where n_e and n_o are the refractive indices experienced by extraordinary and ordinary rays.

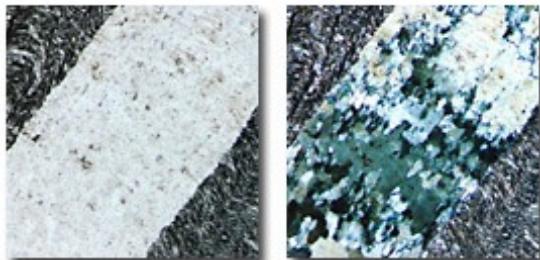


Figure 12.7: phyllite sample. plane polarized(left), cross polarized(right) [97]

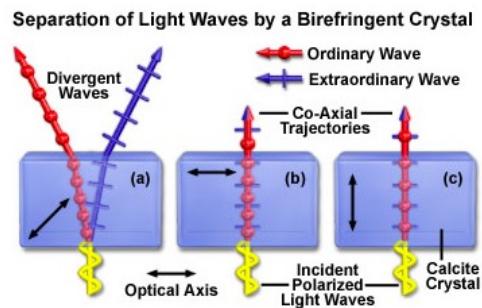


Figure 12.8: Example of birefringence material [98]

The phase relationship and velocity difference between ordinary and extraordinary rays is termed as **retardation**.

$$\text{Retardation}(\Gamma) = \text{thickness}(t) \times \text{birefringence}(B) \quad (12.10)$$

Figure 12.10 shows the concept for polarized light microscopy. Light enters the polarizer (left) and is linearly polarized in the direction indicated by the arrow. The polarized light enters the anisotropic material (sample) where it is refracted into two waves. They pass through the analyzer (right) which only allows the components of light waves that is parallel to the analyzer's polarization direction to pass. As one wave is retarded with respect to the other, constructive or destructive

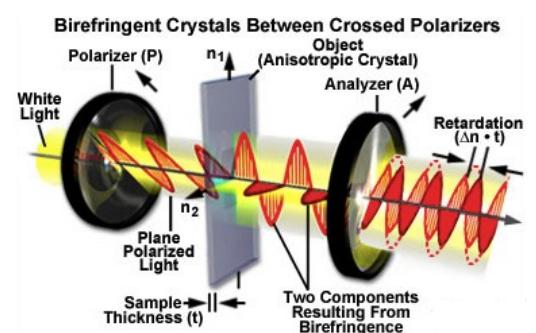


Figure 12.9: Concept for polarized light microscopy[98]

interference occurs and an image pattern is observed.

12.2.2 Instrumentation of polarized light imaging

Polarized light microscopy is a simple modification of bright-field microscopy. From figure 12.10, a polarizer (**POL2**) is positioned before the specimen, and an analyzer (**AN1**) is positioned before the CMOS objective (**L14**). The polarizer is oriented East-west (horizontal) and the analyzer is initially positioned North-South (vertical). Their vibration azimuths are at right angle to each other. Thus, if the sample is isotropic, no direct light passing through the system. For anisotropic material, contrast is related to retardation, from equation 12.10. Thus, with the knowledge of material property (birefringence), thickness can be calculated and the contrast image can be translated to the thickness data. If needed, the analyzer can be rotated to observe the sample in different directions. To increase the brightness of the observation, a retardation plate or a compensator can be included.

The objective most used suitable for polarized light microscopy is UPLFLN 100XOP. It is a strain free objective especially designed for polarized light microscopy. As inherent anisotropic character of glasses leads to natural birefringence and tight fitting frames leads to strain birefringence, manufacturers choose strain-free optical glass or isotropic material to construct the lens which can pass stress test. Other objectives can also be used in polarized light microscopy but might give a compromised image.

12.2.3 Modelling of polarized light imaging

Similar to bright field, Rayleigh criterion is applied to calculate the resolution for polarized light imaging. Instead of NA 1.4 objective (MPLAPON 100XO) used in bright field, a NA 1.3 objective (UPLFLN 100XOP) is used. The lateral resolution is thus

$$Resolution_{x,y} = 0.61 * 625\text{nm}/1.3 = 293\text{nm} \quad (12.11)$$

which is similar to that of bright field (272nm).

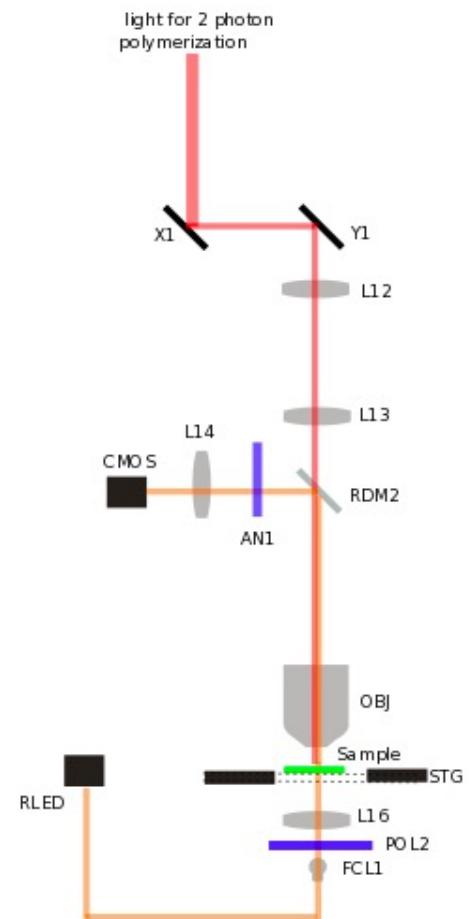


Figure 12.10: Optical Layout for Polarized Light microscopy

12.3 Excluded Methods - James

12.3.1 Dark-field microscopy

12.3.1.1 Theory

The Dark-field imaging mode is generally used to increase contrast and is particularly useful for viewing edges. I modelled the effects of the mode in MATLAB - using an annulus as the Optical Transfer Function to represent the higher spatial frequencies which appear in Dark Field microscopy. This can be useful in highlighting the edges of samples (which manifest themselves as high spatial frequencies). A simulation of the effects of the darkfield mode (with an annulus optical transfer function (OTF) shown in figure 12.11) is shown in an example using a photo of Brooklyn Bridge as the object in figure 12.12 and 12.13. The image is formed by convolving the OTF with the object.

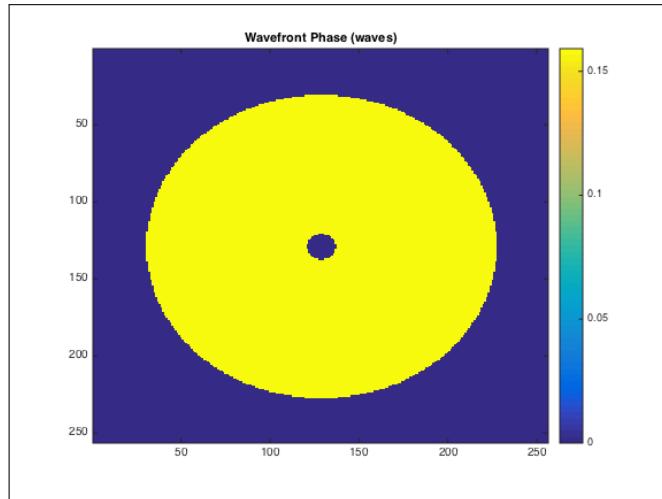


Figure 12.11: The optical transfer function for Darkfield OTF. The patch stop produces the hollow centre.



Figure 12.12: Original object



Figure 12.13: Darkfield Image

12.3.1.2 Implementation

In order to implement dark-field microscopy, we would have needed a patch-stop filter in order to illuminate the sample with a hollow cone of light (as opposed to brightfield which uses a filled cone) as shown in 12.14. Since the objective sits in the corresponding hollow cone section after illumination, if there is no sample present, the field appears completely dark [99]. As a result, when using darkfield microscopy we only see light that has been scattered by the sample. Since there is less illumination to begin with and since we only view the scattered light which falls inside of the cone, it follows that the image formed is significantly less intense for a given intensity of illumination. For this reason when using darkfield microscopy, the best results are obtained by increasing the intensity of illuminating light.

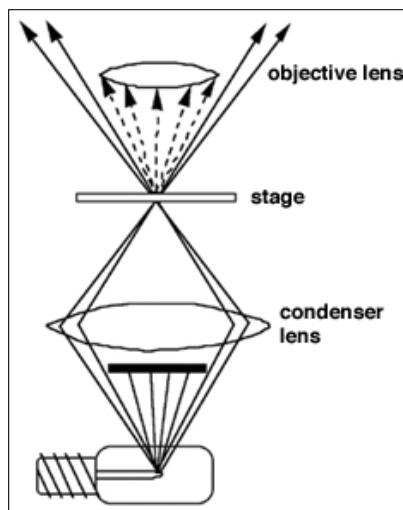


Figure 12.14: [100] Implementation of a Darkfield microscope. A patch stop prevents the centre of the cone of light illuminating the sample while the objective is placed

12.3.1.3 Reason for exclusion

Instead of darkfield we use polarized light microscopy as an increased contrast imaging mode since most of our low-contrast materials are polymers that are naturally birefringent. Polarized light microscopy is more expensive to implement than darkfield due to the additional components required however it is also more efficient in terms of the ratio of intensity of light in to intensity of image out. Also, since our system operates with materials which are highly sensitive to illumination, it is more important to avoid the risk of accidental polymerisation created by using high intensities of illuminating light than it is to save a small amount of money. For this reason we decided that darkfield imaging didn't offer significant benefits over the modes eventually used in our design.

12.3.2 Phase Contrast Microscopy

12.3.2.1 Theory and what it offers

Phase contrast imaging is generally used to view near-transparent objects that would not be easy or possible to view otherwise. Where brightfield microscopy images amplitude objects (objects that affect the amplitude of light passing through or reflecting off them) phase contrast techniques are used to image phase objects (objects that slightly change the phase of light as it passes through them). A useful image must produce a visible change in wavelength or amplitude. By transforming transmitted light, we can make small changes in phase visible and use these to create our image. We illuminate the sample with a ring of light similar to darkfield microscopy. However in phase contrast microscopy we collect both the scattered and unscattered light. Since we know the path taken by the unscattered light, we can use a ringed quarter wave phase plate to manipulate the light so that it destructively or constructively interferes with the diffracted light to form the phase image of the object.

It can be shown that in general and as shown in figure 12.15 the 'D' wave (which represents the difference between the waves passing through the object 'O' and the background unscattered waves 'B') is around one quarter wavelength out of phase with the background. In order to form the phase object image with contrast against the background, we either advance or delay 'B' by a quarter wavelength so that it either constructively or destructively interferes with the diffracted waves. Then the object is viewed as either brighter or darker than the background, with the phase difference converted into a change in amplitude.

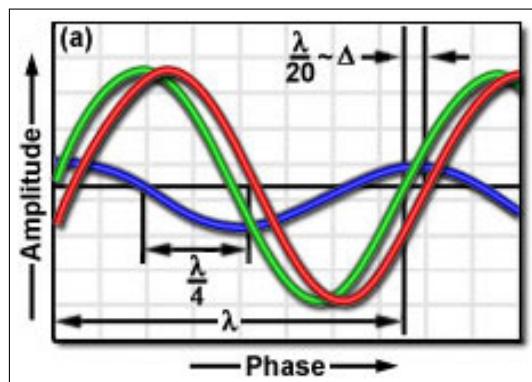


Figure 12.15: [101] In general, B (red) is about one quarter wavelength out of phase with D (blue) as shown. We need to manipulate the Background light B so that the phase difference between O (green) and B is converted to an amplitude difference by advancing B so that it has the same phase (or 180 degrees out of phase) as the diffracted waves D

12.3.3 Implementation

The sample is illuminated by light through an annulus. Unscattered light is imaged at the focal plane of the objective after having passed through a ringed quarter wave phase retarder. Scattered light is then able to interfere constructively or destructively at the image plane allowing the phase object to

be viewed as contrast. In some implementations, the unscattered light is passed through a filter after the sample to reduce the intensity significantly so the proportional change due to interference with the scattered waves is increased.

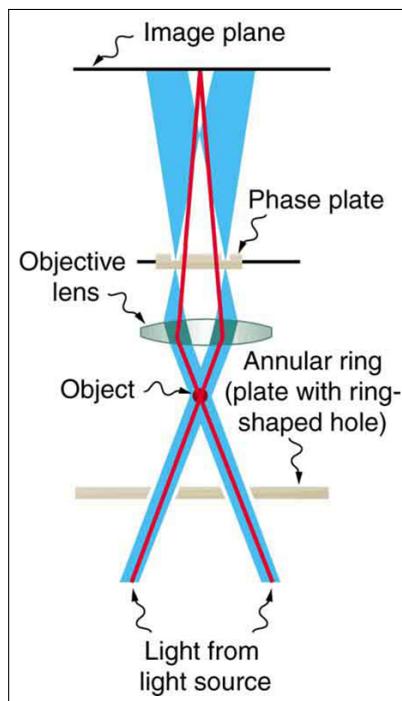


Figure 12.16: Implementation of a phase contrast microscope [102]

12.3.3.1 Reason for exclusion

Phase contrast microscopy is effective when imaging typically thin samples with low absorption which ordinarily require staining in order to view them by brightfield microscopy. The benefits of our fabrication system however are in our ability to build large structures with nanoscale detailing. Phase contrast microscopy is therefore not ideally adapted to the kind of samples which we will likely need to live image. Similarly to the darkfield mode, phase contrast also requires a high intensity of light with the associated risk of accidental polymerisation. Also the polarization imaging modes used in our system are able to image phase objects rendering the phase contrast mode redundant.

13 Post-Process Imaging

In addition to live imaging, we want to provide our users with the ability to examine their product after the fabrication process. Since different materials are used in our various fabrication techniques, we want versatile imaging techniques with high resolution. Also, three-dimensional structures can be produced in our fabrication process. Thus, our post process imaging techniques should be able to obtain three-dimensional volume rendered images.

To meet the above requirements, our system incorporates confocal microscopy and 2-photon microscopy as post process imaging techniques. Our imaging system is able to replace some of the standard imaging systems in the laboratories around the world. Confocal microscopy is able to operate at both reflective mode and fluorescence mode. Reflective mode is compatible with all materials. It is mostly suitable to examine the surface structure of opaque materials, such as the product of laser micro machining. Fluorescence mode uses fluorescence and phosphorescence, which is suitable for transparent product in two-photon polymerization. Compared with confocal microscopy, two-photon microscopy provides improved resolution and unique ability to detect in-depth structure. Both imaging techniques are integrated closely with the fabrication system and many optical instruments are reused. Thus, with little extra costs, customers can take high quality three-dimensional images for different materials.

13.1 Confocal Laser Scanning Microscope

13.1.1 Basic Principle of confocal imaging

Confocal Laser Scanning Microscopy (CLSM) is a fundamental and versatile imaging technique in our system. Compared with bright field and phase contrast imaging, it improves resolution and contrast by placing a spatial pinhole at the focal plane of the lens to eliminate out of focus light. As shown in figure 13.1, confocal microscopy (A) improves both the resolution and contrast, compared with conventional microscopy (B). Also, the detail (white box), which is initially blurred in conventional microscopy, appears with confocal microscopy. In our system, CLSM can provide resolution up to 150nm with 532nm green laser and 1.4 oil immersion objectives.

It is commonly used in biological sciences to image tissue and their movements. In figure 13.2, light from the conjugate plane (red) is imaged by the objective such that it passes the pinhole, whereas light out of focus (orange and yellow) is largely blocked by the pinhole. The position of the pinhole relative to the sample enables serial optical section collected for thick specimens, which is essential for three-dimensional image reconstruction. The presence of the pinhole also reduces the signal to noise ratio and improves image quality. [104]

In principle, spatial pinhole can improve lateral resolution by a factor of 1.4 when the pinhole is

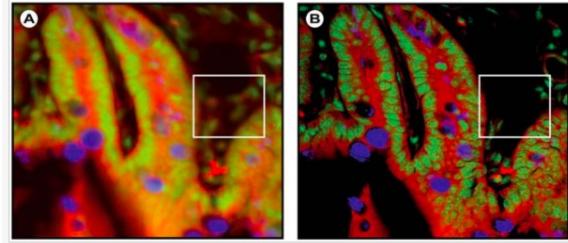


Figure 13.1: Image of triple label cell aggregate. conventional(A), Confocal(B) [103]

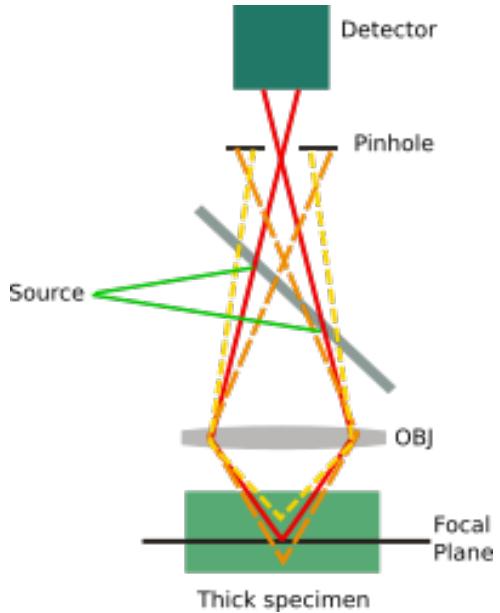


Figure 13.2: Confocal microscopy concept

small ($\ll 1$ Airy disk Unit). Also, in confocal microscopy, the illuminating laser is with shorter wavelength (532nm green laser). Thus, we can derive the lateral resolution ($r_{x,y}$) for confocal reflected microscopy[105]

$$r_{x,y} = \frac{0.4\lambda}{NA} \quad (13.1)$$

$$r_{x,y} = 0.4 \times 532\text{nm} / 1.4 = 152\text{nm} \quad (13.2)$$

Axial resolution in confocal microscopy is worse than axial resolution. A common equation used in calculating the axial resolution(r_z) is given by

$$r_z = \frac{1.4\lambda n}{NA^2} \quad (13.3)$$

$$r_z = 1.4 \times 532\text{nm} \times 1.47 / 1.4^2 = 559\text{nm} \quad (13.4)$$

CLSM can operate in two modes, fluorescence mode and reflective mode. Our system includes both modes for imaging different materials. The former one allows identification of specific molecules/material, which is more suitable for transparent samples, while the latter one allows the detection of surface structure for opaque materials.

13.1.2 Fluorescence microscopy

Fluorescence microscopy is a technique that uses fluorescence to generate an image. **Fluorescence** refers to the absorption and subsequent re-radiation of light by organic or inorganic samples. As figure 13.3 shows, the specimen is illuminated with light of a specific wavelength (red line). Fluorophores absorbs the excitation light and emits light of a longer wavelength (yellow line). The excitation ray

and emission ray is usually separated by a dichroic mirror.

Since illuminating light is filtered out with dichroic mirror, signal to noise ratio is improved and thus image with better contrast can be obtained. Also, fluorescence microscopy adds figure to the sample, by identifying certain structures, bonds or molecules in the sample. In addition, fluorescence microscopy can be used in combination with other techniques, such as confocal microscopy, multi-photon microscopy, simulated emission depletion microscopy and structured illumination microscopy (in section 14.2.2).

However, when sample is illuminated, photobleaching occurs, which means the fluorophores lose their ability to fluoresce. Thus, photobleaching limits the intensity and time for illuminating. In addition, fluorescence microscopy only allows observations of structures that can be labeled. Hence, only specific structures can be viewed or identified. Lastly, dyeing is need for fluorescence microscopy, which complicates the overall viewing process.

13.1.3 Instrumentation of confocal imaging

The optical layout is shown in figure 13.4. The system of confocal microscopy includes a laser source (**LAS 532**), a power control unit (**grey block**), a pair of galvanometers (**X1, Y1**), two 4f lens systems (**L7, L8, L12, L13**), an objective (**OBJ**), a dichroic mirror (**FLPM2**), a pinhole (**PIN**) and a detector (**PMT1**).

The power control unit is needed because imaging needs a much lower power than polymerization process. Beam power is controlled with a half wave-plate (**HWP**) and a fixed polarizer. The detail of power control is explained in chapter 20. The two-axis galvanometer scanning system (**X1** and **Y1**) deflates the collimated laser light (green line) to excite/illuminate the specimen. On the return path, the emitted light (yellow line) is again deflected by the galvo-driven mirrors to the detector through the spatial

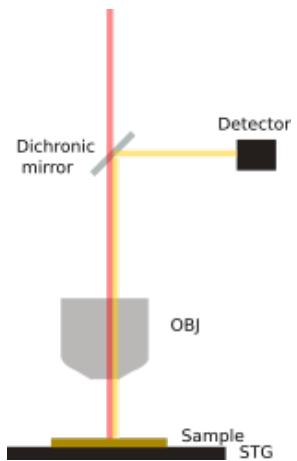


Figure 13.3: Optical layout for Fluorescence microscopy

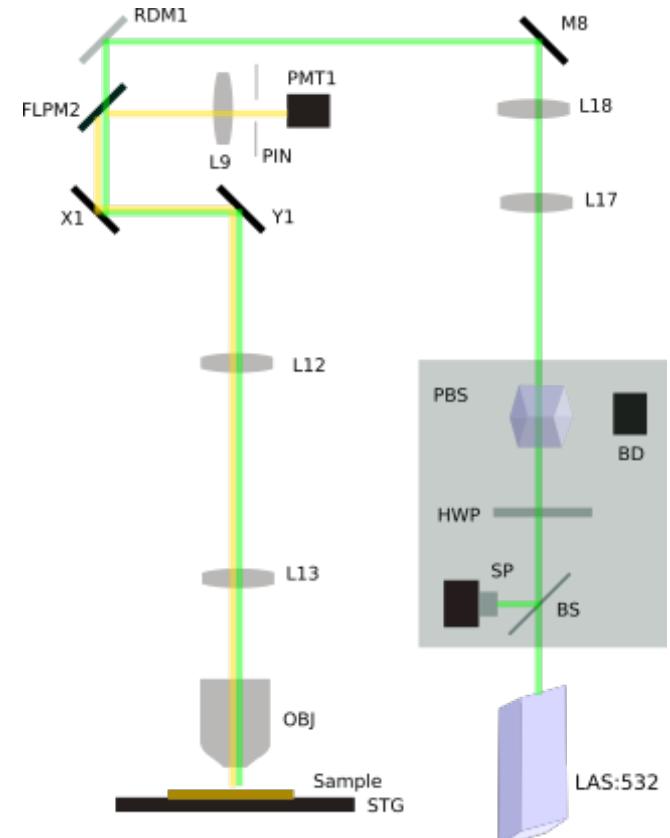


Figure 13.4: Optical layout for Confocal Laser Scanning microscopy

pinhole. The 4f lens system (**L12**, **L13**) is included to re-centre the deflected beam from the galvos scanning mirrors to the centre of the objective. Thus, the sample is illuminated/excited point by point. The reason for choosing the single beam scanning system will be discussed later. An objective turret (MPLAPON 100XO, UPLSAPO 20XO, LUCPLFLN 40X, LCPLN 100XIR) is used in confocal imaging and calculations are based on these four objectives. The optimization of the pinhole size and selection of detector will be discussed in the next section.

13.1.4 laser

Unlike two photon microscopy and STED, Confocal Laser Scanning Microscopy does not have any specific requirement for wavelength. From equation 13.1, a shorter wavelength will be able to give a better lateral resolution. Among the three lasers available in the system, laser 355nm has the shortest wavelength. However, the transmittance of chosen objective at this wavelength is only 40%. Thus, 532nm laser is reused, at which the transmittance is 85%.

With an excitation wavelength of 532nm, a variety of dyes, such as Alexa Fluor 532, Alizarin Red, and Erythrosin can be used for fluorescence mode.[106] Their emission wavelength ranges from 550nm to 580nm. The excitation ray and emission ray are separated by flip mirror (FLPM2).

Reflection mode does not require a dye to be present in the sample. Instead, reflected light interferes and forms the image. The illuminating and reflected rays are also separated by flip mirror (FLPM2).

13.1.4.1 Single beam scanning mechanism

There are various implementation forms of confocal microscopy available, including single beam scanning, line scanning and multiple beam scanning (spinning dish). I will explain their mechanism, summarize their benefit and weakness respectively, and explain our decision.

Single beam scanning is adopted in our system. It is achieved by galvanometer driven mirrors. One of the scanning mirrors directs the light along the fast axis, while the other mirror directs the light along the slow axis. Thus, image is formed point by point, until a two-dimensional image is acquired. When a plane scan is finished, depth (z-axis) can be changed with a stage motor. Multiple optical sections of the specimen can be obtained to construct a three-dimensional image stack. Single beam scanning has the advantage of high resolution and simple set up. However, it takes several minutes to form a 2048×2048 image.

Line scanning confocal microscope uses a line split rather a pinhole. This increases the speed of scanning greatly as a 2D image is formed only with fast scan mirror. However, resolution is compromised because system is bright field in one axis and confocal in the other axis.

Spinning disk confocal microscope (SDCM) uses a disk with multiple of holes rather than a single pinhole. The most advanced spinning disk instruments design is by Yokogawa Electric Corporation of Japan. As seen in figure , the disks (with 20,000 pinholes) are coaxially aligned and a dichromatic mirror is between them. The pinholes are spirally positioned and the image is created in a single sweep. Thus, imaging speed is greatly improved as the disk can operate at a rate of 400 to 1000 scans per second. Also, instead of a single point detector (PMT), a multi-point detector (CCD) is used to record the whole frame at a time.

Spinning disk scanning mechanism has the main advantage of fast scanning rate. However, this is actually limited by camera exposure time, as the emitted light is of low intensity. Even if the most advanced CCD is used, an exposure time of 100 milliseconds is still needed to fluorescence sample. Thus, the actual scanning speed is 10 frames per second. In addition, spinning disk microscope has three main drawbacks. The primary limitation is that scattered light out of focus can still reach the detector by traveling through the adjacent pinholes. This means the signal to noise ratio is hampered, which in turn affects resolution. The second limitation is limited brightness because 44 % to 99% of the illumination light is lost at the disk. Although using a high power laser can overcome the problem, overheating of the disk will occur. Another limitation is the difficulty to adjust various mirrors so that returning pattern of points exactly matches the pattern of point on the detector side.

We select point scanning confocal microscopy because of the following reasons. Firstly, high resolution is desired whenever possible. Since both line scanning microscopy and spinning disk microscopy compromise resolution for speed, point scanning confocal microscopy is preferred. Secondly, as our fabrication processes also adopt single beam scanning mechanism, point scanning confocal microscopy can be integrated into our system easily. On the contrary, spinning disk confocal microscope requires careful calibration of light path and multiple extra instruments, possibly a separate pathway. This increases cost and complexity of our system design. Lastly, different pinholes size are needed when different objective are used. If physically changing the disk is not implemented, spinning disk microscopy can only be optimized for a certain objective. As a result, we think that the drawbacks brought by spinning disk outrun its advantage. Thus, single beam scanning mechanism

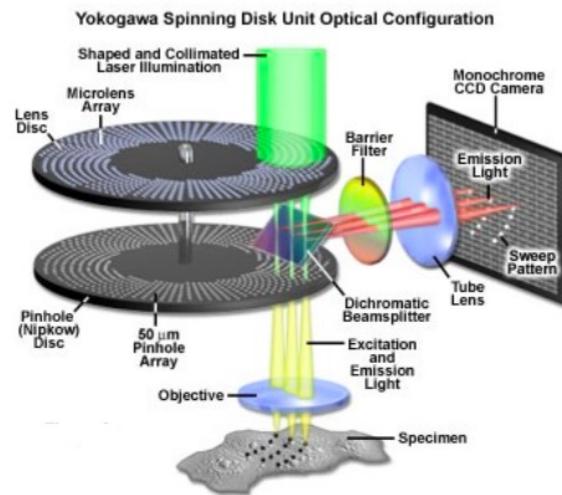


Figure 13.5: Optical layout for Spinning Disk Confocal microscopy [107]

is chosen in our system design.

13.1.4.2 Detector - Photomultiplier tube

For single point detector, photomultiplier tube (PMT) is chosen. Compared with its alternative Avalanche photodiode (APD), PMT has the advantage of fast response, high gain and linear behavior. In our system, single point detector is used in confocal microscopy and two-photon microscopy, where the laser is focused to a small spot and scanned relative to the sample. To obtain an image, a reading is taken for each spot, where the whole image is formed by raster illumination. For an image size of $512\text{pixel} \times 512\text{pixel}$, 262144 readings need to be taken. Thus, fast response is prioritized over high quantum efficiency in order to obtain the image of the whole sample in a reasonable amount of time.

PMT is an extremely sensitive detecting device in the ultra violet, visible and near infrared region. Photomultiplier tube converts the absorption of a photon into emission of an electron by amplifying the electrons generated by a photocathode exposed to a photon flux. As shown in figure 13.6, the incoming light hits the photocathode (top left), which is a photosensitive surface. Electrons are released from the surface as a result of photoelectric effect. Each dynode is at a more positive potential than the previous one. When electrons strike the dynode, their number multiples by the process of secondary emission. The number of electrons collected by anode (top right) can be as much as 100 millions times compared with the incoming photon. The output meter measures the current from anode to ground, which is proportional to the photoelectron flux generated by the photocathode.

For the chosen PMT (PMM01), its maximal gain is 7.1×10^6 and rise/fall time is only $15\mu\text{s}$. If the exposure time for a spot is 1 millisecond, which is over 50 times compared with the rise time, an image size of $512\text{pixel} \times 512\text{pixel}$ will take 4 minutes.

13.1.4.3 Detector - Pinhole size

The resolution of the optical section is determined by the size of pinhole, objective numerical aperture and wavelength of emission/reflected light. From figure 13.7(right), it seems that making the pinhole as small as possible will improve the signal to noise ratio and thus result in better performance. However, as the pinhole get smaller, the number of photons arrives at the detector also decreases, as

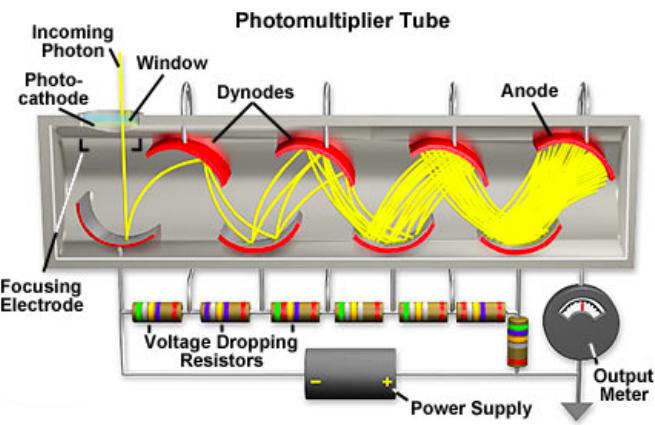


Figure 13.6: Photomultiplier Tube concept [108]

shown in 13.7(left). Experiments and theoretical computations show that if the pinhole size is smaller than that of one Airy unit, resolution is compromised by low contrast. Thus, there is a compromise between signal intensity and signal to noise ratio. Generally, pinhole radius around airy disk first minimum is advised. Also, fluorescent emission is usually at lower intensity compared with reflected light. Thus, a factor term should be included to account for that.

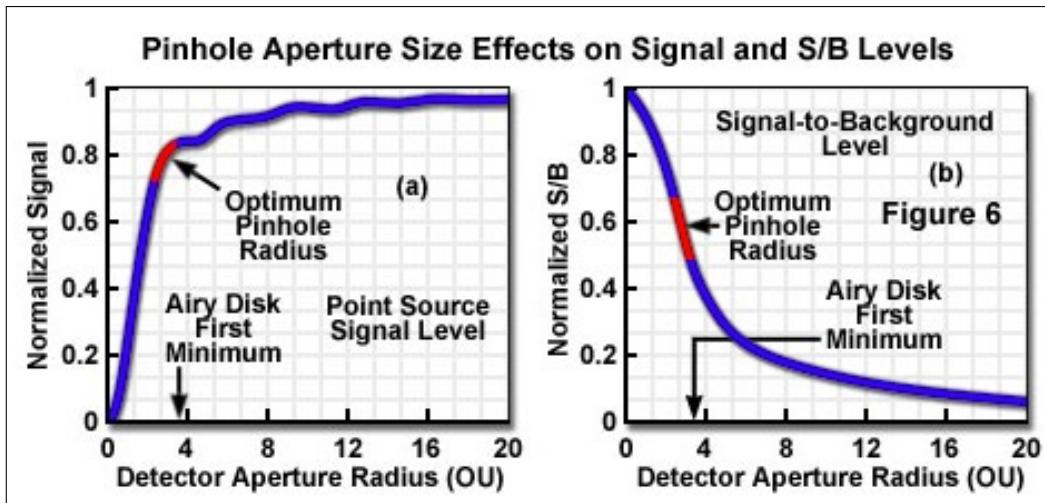


Figure 13.7: Diagram illustrating the effect of pinhole size [109]

In reflected mode, the **optimal pinhole diameter** (D_R) is suggested as [110]

$$D_R = 0.5 \times M_{obj} \times (\lambda_R / NA_{obj}) \quad (13.5)$$

In fluorescence mode, the **optimal pinhole diameter** (D_F) is suggested as

$$D_F = 1.2 \times M_{obj} \times (\lambda_{Em} / NA_{obj}) \quad (13.6)$$

In both equation, M is the objective magnification, NA is the numerical aperture, λ_R is the wavelength for reflected light, and λ_{Em} is the wavelength for fluorescence light. Thus, a turret of pinholes is needed for both modes and different objectives.

Objective	NA	Magnification	Reflection mode	Fluorescence Mode
MPLAPON 100XO	1.4	100X	19	48
UPLSAPO 20XO	0.85	20X	6.3	15.8
LUCPLFLN 40X	0.6	40X	17.7	45
LCPLN 100XIR	0.85	40X	31	79

Table 13.1: Optical pinhole size (in μm for various objective)

Thus, the pinhole turret of $6\mu\text{m}$, $19\mu\text{m}$, $31\mu\text{m}$ and $48\mu\text{m}$ and $79\mu\text{m}$ will be included in the system. This is because maximal resolution is achieved with MPLAPON 100XO (oil immersion) and LCPLN 100XIR (dry condition). Thus, the pinhole size for these two objectives should be optimized. When other objectives are used, compromises can be made.

13.1.5 Modelling of confocal imaging

In the next section, I will derive the optical transfer function for confocal microscopy step by step. I first will make a distinction between coherent and incoherent light source, which leads to different sets of equations. I will start with scattered field theory and apply point spread function (in section 3.3) to both conventional and confocal microscopy. The inadequacy of point spread function in describing the system will be addressed by formulation of Optical Transfer Function (in section 12.1.3.1). I will use the above theory to verify the improvement brought with confocal microscopy.

13.1.5.1 Coherent and incoherent light source

Before any modelling, I want to first distinguish between coherent light source and incoherent light source. **Coherent light** means there is a fixed relationship between the phase of waves in a beam of a single frequency. Two beams of light are coherent if they have a constant phase difference. They are incoherent if their phase relationship is random or changing. An LED light source is an incoherent light source, while laser is usually monochromatic coherent light source. In mathematics term, a coherent system is linear in the complex field (amplitude), while incoherent system is linear in intensity.

13.1.5.2 Coherent system

If the input function $u(\mathbf{r})$ can be represented by a superposition of point-like element, then the output $v(\mathbf{r})$ is given by

$$v(\mathbf{r}) = h(\mathbf{r}) \otimes u(\mathbf{r}) = \int u(\mathbf{S})h(\mathbf{r} - \mathbf{s})d\mathbf{s} \quad (13.7)$$

$h(\mathbf{r}, s)$, is known as the **amplitude point spread function**, or the impulse response of the measurement system.

Apply Fourier transform to equation 13.7.

$$\tilde{v}(\mathbf{m}) = \tilde{h}(\mathbf{m})\tilde{u}(\mathbf{m}) \quad (13.8)$$

where $\mathbf{m} = m, n, q$ is a triplet of spatial frequencies. $\tilde{h}(\mathbf{m})$ is the **Coherent Transfer Function**, given by the Fourier transform of point spread function.

From linear imaging theory, the field E_d at position r_2 , for an object scanned to r_s is given by[111]

$$E_d(r_s, r_2) = \int h_1(r_1)h_2(r_1 + M_2r_2)t(r_s - r_1)dr_1 \quad (13.9)$$

where h_1 is illuminating point spread function, h_2 is detecting point spread function, M_2 a diagonal 3×3 matrix, with its elements representing axial magnification [112]. It describes the linear relationship between the incident field onto the detector and the scattering potential.

The effective PSF can thus be defined as [113]

$$h_{eff}(r_1, r_2) = h_1(r_1)h_2(r_1 + M_2r_2) \quad (13.10)$$

The Fourier transform of $h_{eff}(r_1, r_2)$ is $\tilde{h}_{eff}(m, r_2)$, which is called Coherent Transfer Function (CTF).

For a finite size detector with response function $D(r_2)$, the intensity recorder for each scan position is given by

$$I_d(r_s) = \int (|E_d(r_s, r_2)|)^2 D(r_2) dr_2 \quad (13.11)$$

For a conventional imaging systems with a point detector r_d , $D_{confocal}(r_2) = \delta(r_2 - r_d)$. For the ideal confocal microscope, the delta function gives the detected signal as $(|E_d(r_s, r_2)|)^2$. Thus, $\tilde{h}_{confocal}(m, r_d) = \tilde{h}_{eff}(m, r_d)$. In other words, we can assume the pinhole integrates the amplitude over the back focal plane of the lens. Again we assume that the detector is on the axis (i.e. $r_d = 0$)

$$h_{confocal}(r_1, 0) = h_{eff}(r_1, r_2 = 0) = h_1(r_1)h_2(r_2) \quad (13.12)$$

If we consider the image of a point object, $t(r) = \delta(r - r_0)$, the measured intensity is given by combining equation 13.9 and 13.11

$$I_d(r_s, r_0) = (|h_1(r_s - r_0)|)^2 \int (|h_2(r_s - r_0 + M_2r_2)|)^2 D(r_2) dr_2 \quad (13.13)$$

which is knowns as the intensity PSF.

For conventional microscopy, an infinite detector is assumed, the intensity PSF is given by

$$H_{conventional}(r_s, r_0) = |h_1(r_s - r_0)|^2 \quad (13.14)$$

While for confocal microscopy, point detector is used, and the intensity PSF is given by

$$H_{confocal}(r_s, r_0) = (|h_1(r_s - r_0)|)^2 (|h_2(r_s - r_0 + M_2r_2)|)^2 \quad (13.15)$$

which is just the square of h_{eff} (equation 13.10).

13.1.5.3 Incoherent Confocal Imaging System

Incoherent light is expressed in term of intensity. $I(r) = T(r)I_r(r)$ where $T(r) = (|h_1(r_1)|)^2$ and I_r is illuminating intensity.

The field at position r_2 , assuming the object is scanned to r_s , is given by

$$I_d(r_s, r_2) = \int |h_1(r_1)|^2 |h_2(r_1 + M_2 r_2)|^2 T(r_s - r_1) dr_1 \quad (13.16)$$

which shows the detected intensity is linearly related to the modulus square of the scattering potential.

Similarly, the intensity recorder by a finite size detector is

$$I_d(r_s) = \int H(r_1) T(r_s - r_1) dr_1 \quad (13.17)$$

where $H(r_1) = \int |h_1(r_1)|^2 |h_2(r_1 + M_2 r_2)|^2 D(r_2) dr_2$ is the intensity PSF for a point object at r_1 .

For conventional microscopy (infinite detector),

$$H_{conventional}(r_s, r_0) = |h_1(r_s - r_0)|^2 |h_2(r_s - r_0 + M_2 r_2)|^2 \quad (13.18)$$

For confocal microscopy (point detector)

$$H_{confocal}(r_s, r_0) = |h_1(r_s - r_0)|^2 \quad (13.19)$$

Thus, equations 13.14, 13.15 and 13.18, 13.19 show that the intensity PSF for the coherent and incoherent system are identical. The difference lies in the equation 13.13 and 13.16.

The equivalence of intensity PSF in coherent and incoherent systems makes PSF insufficient description [114]. Thus, Optical Transfer Function description is adopted.

Optical Transfer Function is defined as the Fourier Transform of intensity PSF. Similar to equation 13.8, in the frequency domain,

$$\tilde{I}(\mathbf{m}) = \tilde{H}(\mathbf{m}) \tilde{T}(\mathbf{m}) \quad (13.20)$$

where $\tilde{H}(\mathbf{m})$ is the Optical Transfer function.

Recall that, for coherent microscopy, intensity PSF is given by square of amplitude PSF. However, the incoherent OTF is not given by square of CTF [111]. Rather, it is given by autocorrelation of

coherent transfer function (pupil function).

$$\tilde{H}_{eff}(\mathbf{m}) = \tilde{H}_{eff-1}(\mathbf{m}) \otimes \tilde{H}_{eff-2}(\mathbf{m}) \quad (13.21)$$

13.1.5.4 MATLAB simulation for confocal microscopy

Last section shows that the point spread function for conventional microscopy and confocal microscopy is identical. I will first illustrate the effect of size of pinhole on point image intensity with Matlab. Then, I will show the comparison of PSF and MTF between conventional microscopy and confocal microscopy. To illustrate the difference between fluorescence and reflected mode, two point resolution is plotted. Optical transfer function derived earlier will also be plotted in both cases to support the result. In the end, I will discuss the effect of pinhole on axial resolution.

Spherical aberration was taken into consideration when point spread function is generated. In the calculation, spherical aberration is assumed to have a root-mean-square content of 0.2. The image is 256pixel \times 256pixel. Point spread function is generated from complex amplitude using fast Fourier Transform function. The wavelength is 532nm and NA of objective is 1.4. The effects of wavelength and NA on resolution is discussed in section 12.1.3.1.

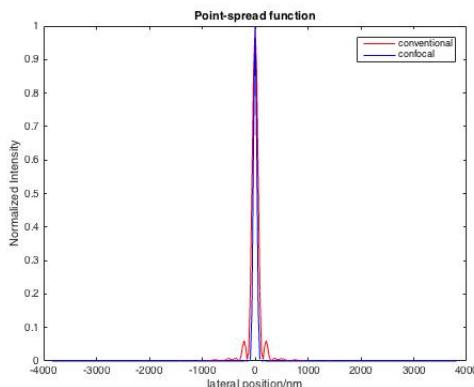


Figure 13.8: PSF comparison (256 pixels)

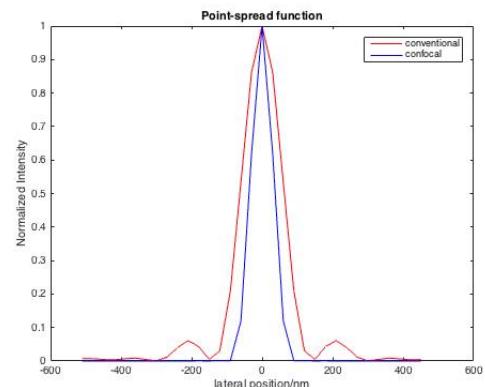


Figure 13.9: PSF comparison (middle section)

The resultant PSF across the image is plotted against its lateral position. Figure 13.8 shows the PSF for conventional and confocal microscopy across the whole sample. Confocal image intensity has a sharper peak, and its outer rings diminish. Figure 13.9 is an enlarged middle section. The first minimum for conventional PSF is about 1.4 times that of confocal PSF. Also, as equation 13.2 suggests, the lateral resolution for confocal microscopy is about 200 nm when wavelength is set to 532nm and pinhole size is set to 19mm. The discrepancy between the theory and image is due to the limited number of pixels and approximation in fast Fourier transform.

If we increase the pinhole size, then resolution gradually decreases, as shown in figure 13.10. The change of pinhole size changes the factor (0.4) in equation $r_{x,y} = \frac{0.4\lambda}{NA}$. Although smaller pinhole

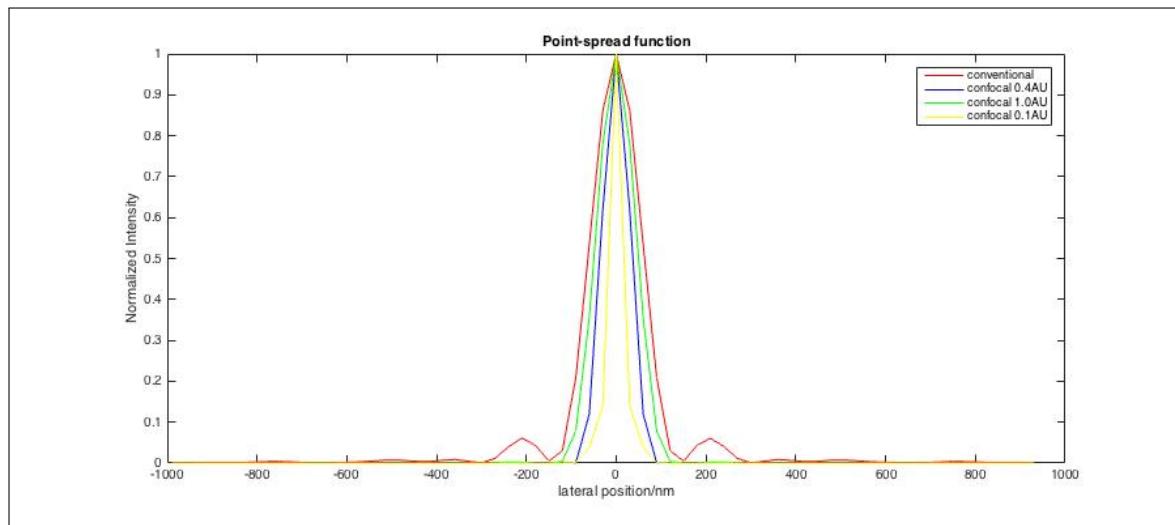


Figure 13.10: Effect of pinhole size on confocal resolution. AU is airy disk unit.

theoretical leads to better resolution, small pinhole size blocks most of the light and signal to noise ratio is affected. With 532nm laser source, NA 1.4 Oil immersion objective, one airy disk diameter is 49 micrometer. In our design, the pinhole size for reflective mode is $19\mu\text{m} \approx 0.4\text{AU}$ and that for fluorescence mode is $48\mu\text{m} \approx 1\text{AU}$.

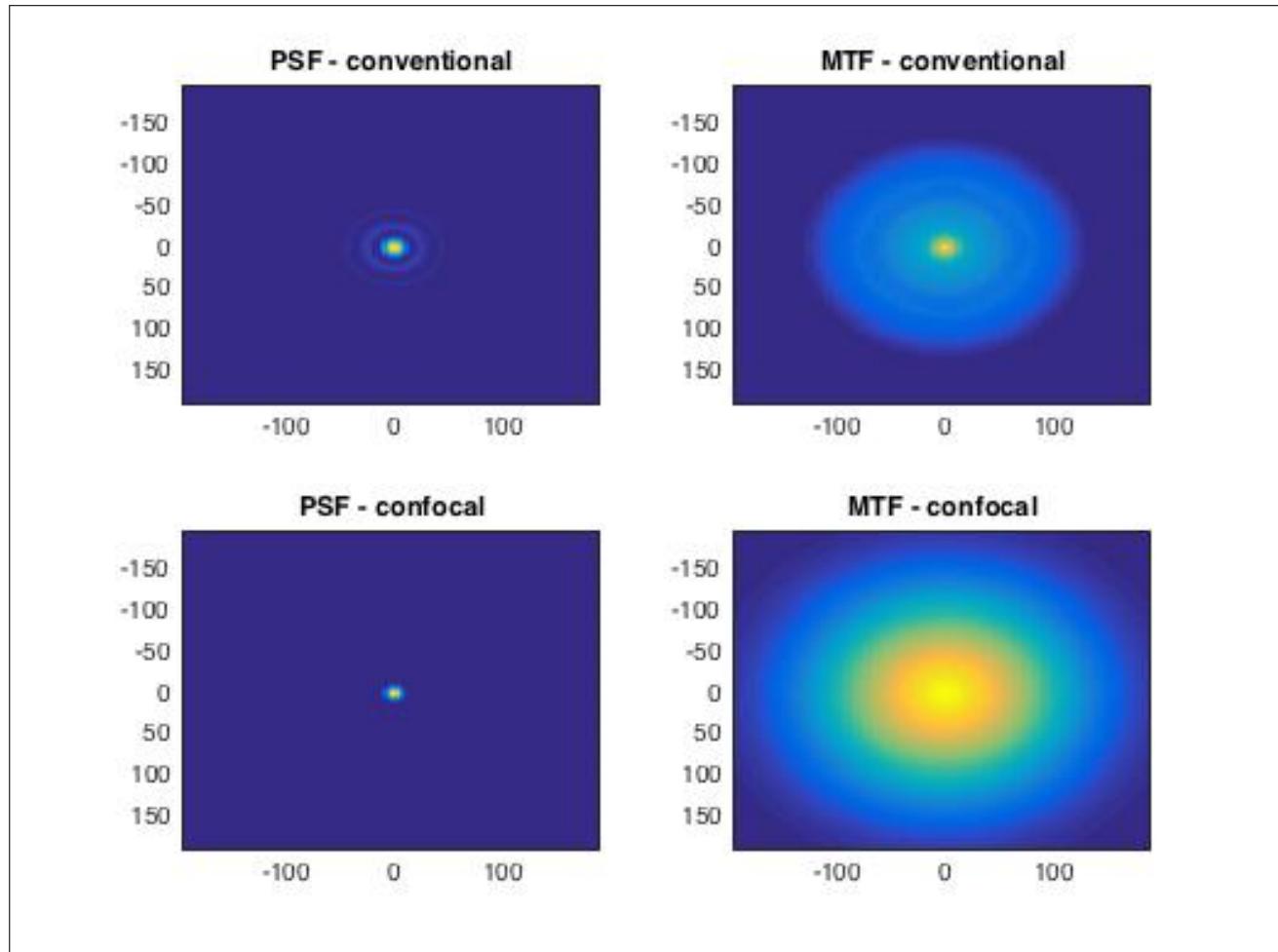


Figure 13.11: PSF and MTF for conventional and confocal microscopy (XY plane)

Optical transfer function is the Fourier transform of point spread function. Figure 13.11 shows the point spread function and modulation transfer function in the xy plane. Comparatively, the central yellow region in confocal PSF is smaller than that in conventional PSF, which means smaller spot size and better resolution. The central yellow region in confocal MTF is much bigger than that in conventional, implying better ability to transfer contrast from object to image.

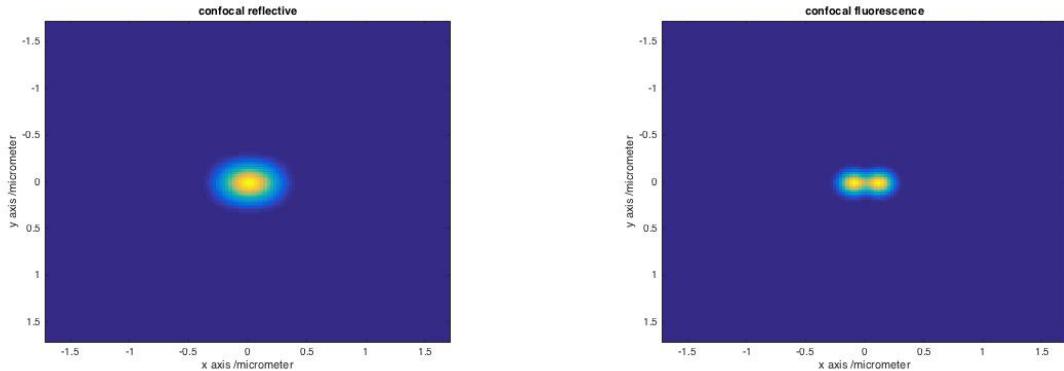


Figure 13.12: Two point resolution in fluorescence mode

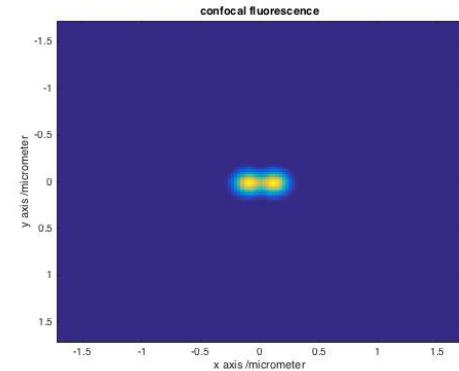


Figure 13.13: Two point resolution in reflected mode

For one point object, the coherence of light source does not affect the image. Instead of one point resolution, we can consider two point resolution.[115] Based on section 13.1.5.2 and 13.1.5.3, equations on two particle resolution can be formulated. For particles 0.75 AU apart in x-direction, simulations is shown in figure 13.12 and 13.13. The pinhole size is set to 1AU, and wavelength is 532nm both cases. Fluorescence confocal microscopy is able to give good resolution as shown by the distinct yellow points. Note that the intensity in the y-axis is almost unaffected. If the distance between two particles is more than 1AU, result for both cases is identical.

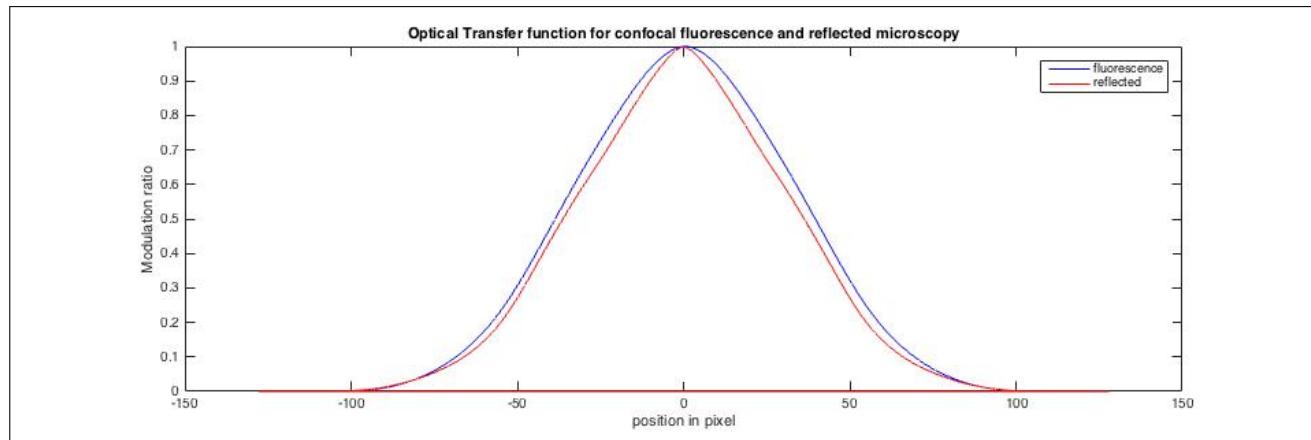


Figure 13.14: Optical Transfer Function for confocal fluorescence mode and reflected mode. Pinhole size is 1AU and wavelength is 532nm in both cases.

In addition, we could compute optical transfer function for both fluorescence and reflected case when the same pinhole is used. Equation 13.15 states that the OTF for coherent source is the

square of Coherent Transfer Function, while equation 13.21 states that OTF for incoherent source is the autocorrelation of Coherent Transfer Function. In figure 13.14, the peak for reflected case (red line) is slightly narrower than that in fluorescence, which shows weaker ability to transfer contrast from object to image. This corresponds to the result in two point resolution discussion.

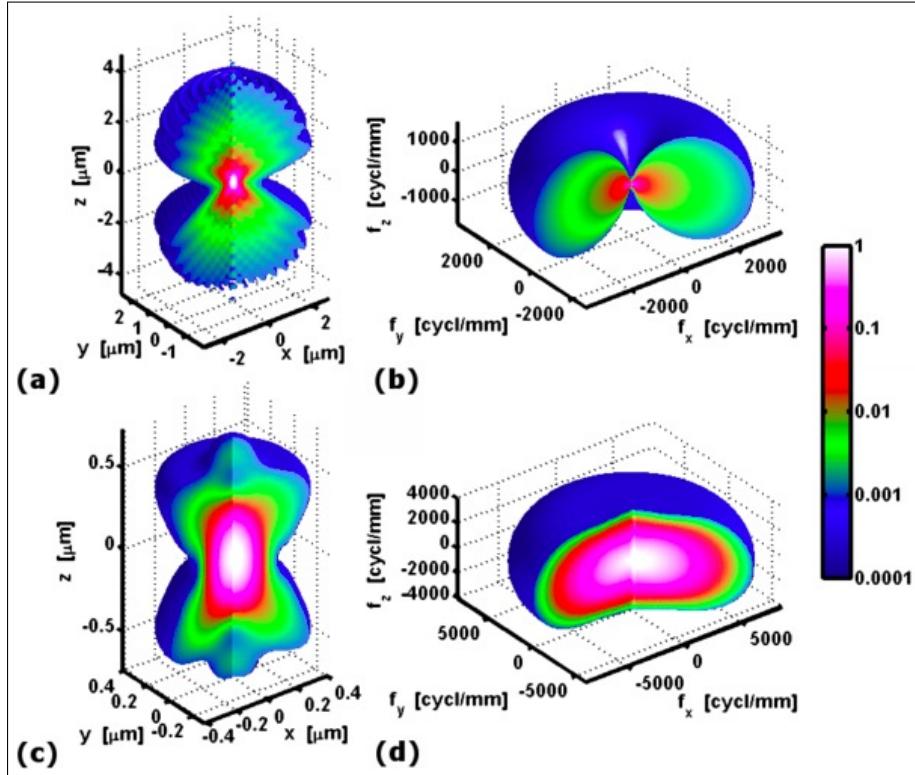


Figure 13.15: Diagram showing the 3D point spread function(left, (a)(c)) and corresponding Modulation Transfer Function(right, (b)(d)). Widespread microscopy is shown at the top (a)(b), and Confocal Fluorescence microscopy is at the bottom (c)(d). [116]

The simulation above is all two-dimensional. Confocal microscopy does not only improve resolution in axial direction, but also in lateral direction. Figure 13.15 is three dimensional simulation of PSF and MTF for conventional microscopy and confocal microscopy. In comparing scale of figure (a) and (b), it is obvious that confocal microscopy has a better resolution in both lateral (x,y) and axial (z) dimensions. The simulation of the modulation transfer function also shows that confocal microscopy is able to maintain a better contrast, as shown by the redder color (higher spectral density) at the same spatial frequency in (d). This also verifies that the resolution for confocal is better than conventional microscopy. In graph (b), along the z-axis, the transfer function is zero except at the origin. The missing cone represents lack of resolution in that direction, which is a problem for wide field microscopy.

In conclusion, confocal microscopy improves resolution in both axial and lateral direction, where smaller pinhole is desirable. At the same time simulations of two point resolution and optical transfer function shows that confocal fluorescence microscopy leads to better resolution than confocal

reflected microscopy, when the pinhole size is identical. However, the emitted light intensity for fluorescence mode is lower than that of reflected mode, which leads larger pinhole is needed. Thus, the advantage of fluorescence microscopy in resolution is compromised. As a result, similar resolution is expected in both modes and the choice of mode should be subject to sample material.

13.2 Two-Photon Microscopy

13.2.1 Basic Principle of 2-Photon Imaging

Fluorescence is the re-emission of light by a substance and offers a variety of advantages to microscopy as explained in subsection 13.1.2. Two photons can be absorbed in a single event as shown in Figure 13.16. A fluorophore that normally absorbs ultraviolet light at 400nm for example can also absorb two infra-red photons at 800nm at the same time—meaning roughly within 10^{-18} seconds. Since it depends on simultaneous absorption, the resulting emission varies with the square of the excitation intensity, rather than linearly. This principle is similar to the excitation in two-photon polymerisation (section 8.1) where two photons need to arrive at the same time to initiate the reaction. To create the same emission as normal fluorescence, you need about one million times as much

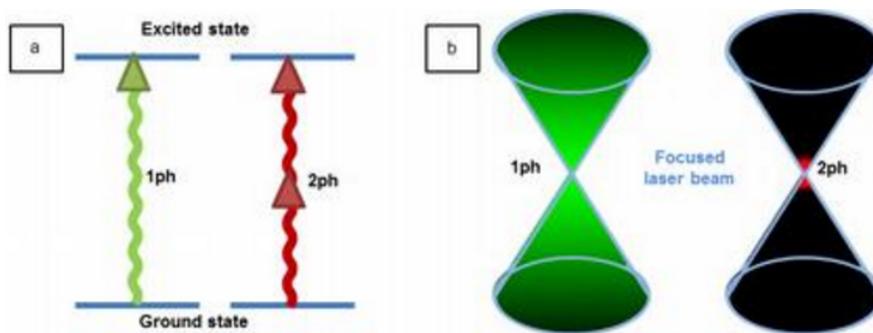


Figure 13.16: Two Photon Excitation, From [117]

illuminating light power. Therefore we use a pulsed laser where the peak of power is high enough but the average power remains low. The laser is brought to a tight focus by the high NA microscope objective.

Confocal microscopy collects the signal only from the focus plane by using a blocking pinhole but the excitation light generates fluorescence throughout the specimen leading to significant photobleaching and phototoxicity—excited fluorophores lose their ability to fluoresce and react with biological matter. Furthermore, scattering and absorption along the beam path significantly limit the penetration depth.

Due to the low probability of excitation and the square law, **two-photon excitation** is only strong enough to generate fluorescence at the focal plane so no pinhole is required. Since there is no absorption in the out of focus plane, it allows for less photobleaching, less phototoxicity, and three times deeper imaging than confocal microscopy. It excels at deep imaging of living cells and is widely used in biology. We will also be able to image inside nano-structures and as such it is a worthwhile technique to have in our system. Another reason for including it is that the added system functionality comes with little extra expense—most of the hardware is shared with two-photon polymerisation.

13.2.2 Instrumentation of 2 photon excitation microscopy

The optical set up for two-photon imaging is mostly similar to that of confocal microscopy. As shown in Fig 13.17, the system of 2 photon microscopy contains a laser source (**LAS:TI:S**), a power control unit (**grey block**), a pair of galvanometer (**X1, Y1**), two 4f lens systems(**L7,L8,L12,L13**), an objective(**OBJ**), a dichroic mirror (**RDM3**), and a detector **PMT2**.

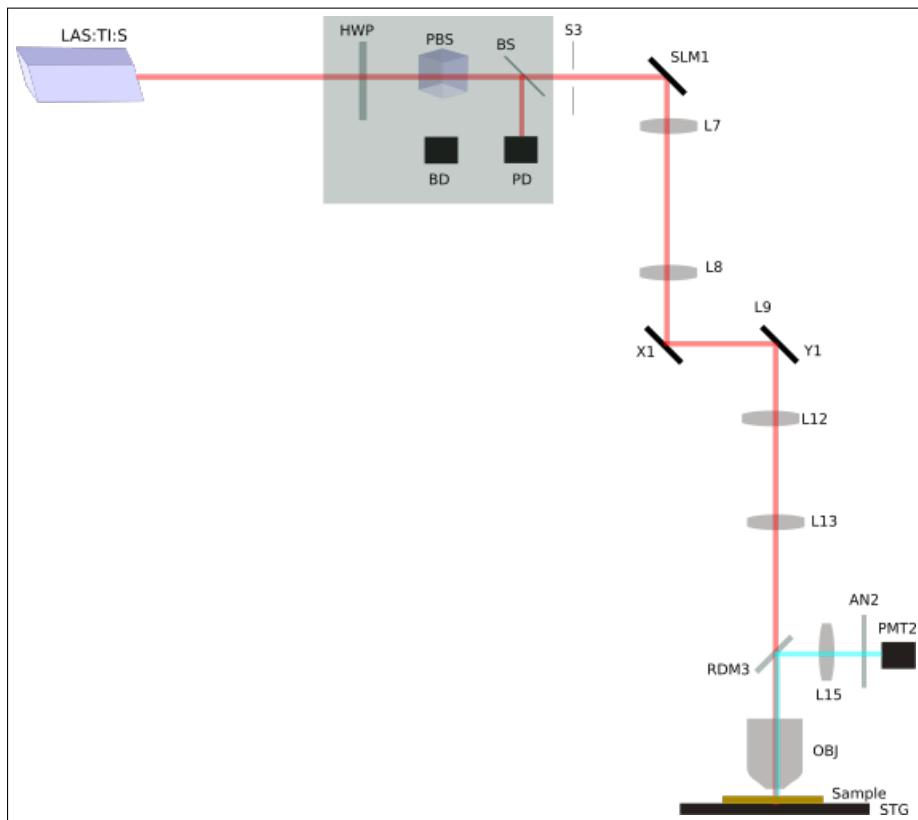


Figure 13.17: Optical Layout for two photon microscopy.
Red line indicates excitation light, blue light indicates emission light.

For two-photon microscopy, high intensity is needed for simultaneous absorption of two infrared (800nm) photons. Thus, we reuse the high intensity 800nm pulsed Ti: Sapphire laser, which gives us high peak power and low average power. For excitation wavelength of 800nm, fluorescent probes such as Amplex Red reagent, Alexa Fluor 594 hydrazide, Laurdan can be used, with an emission wavelength of 520nm to 630nm[118].

The power unit is needed because the intensity needed for two-photon imaging is lower than that of two-photon polymerization. Light (red line) reflects off **SLM1** and pass through the first 4f system (**L7, L8**). The two-axis galvanometer (**X1, Y1**) then directs the collimated laser beam to scan the sample point by point. The second 4f system is to re-center the deflected beam to the center of the objective. Multiple objectives from the objective turret can also be used for two-photon microscopy. On the return path, the emission light (blue line) is separated from excitation light with a dichroic

mirror (**RDM3**).

The difference in optical layout in confocal microscopy and two-photon microscopy is the position of dichroic mirror and PMT. If we consider the return path, for confocal microscopy, the emitted light passes through the 4f system (**L12**, **L13**) before it reflects off the dichroic mirror. While for two photon microscopy, the emitted light does not pass through the 4f systems at all. This is due to the presence of pinhole in confocal microscopy, where the emitted light needs to be re-centered. Comparatively, two photon microscopy does not require a pinhole and thus for better signal to noise ratio, we want as few optical instruments on the return path as possible.

13.2.3 Modelling of 2 photon excitation microscopy

The spatial resolution provided by two photon excitation microscopy is identical to that of a confocal microscopy with a closed pinhole ($\ll 1AU$). The axial resolution ($r_{x,y}$) and lateral resolution (r_z) depends on numerical aperture (NA) and wavelength of light (λ).[119]

$$r_{x,y} = \frac{0.4\lambda}{NA} \quad (13.22)$$

$$r_{x,y} = \frac{0.4 \times 800nm}{1.4} = 230nm \quad (13.23)$$

$$r_z = \frac{1.4\lambda n}{NA^2} \quad (13.24)$$

$$r_z = 1.4 \times 800nm \times 1.47/1.4^2 = 840nm \quad (13.25)$$

Since the wavelength used for two-photon microscopy is longer than that for confocal microscopy, the resolution calculated is worse than that of confocal microscopy. However, the closed pinhole assumption might not be achieved in confocal case, which increases the factor (0.4). While in two photon excitation microscopy, no pinhole is needed and thus theoretical resolution should be achieved.

As mentioned in introduction, two-photon microscopy allows detection in thick specimen. At a great depth, scattering causes excitation photons not to reach the focus and emission photons not to reach the detector. **Depth limit** is when the amount of fluorescence from the outside the focal region that scatters into the detector exceeds the amount of fluorescence from the focal region. In two-photon microscopy, scattering will still cause photons not to reach the focus, but fluorophores outside the focal region will not be excited, as the probability of two scattered photons in the same region simultaneously is negligible. As a result, signal to background ratio is improved. With high frequency pulsed laser, a detection depth up to 1mm can be achieved.

14 Superresolution

Thus far the methods we have used are all limited by Abbe's diffraction resolution limit (3.3.3). However we wanted to give our users the ability to realise feature sizes that went beyond the diffraction limit - increasing the functionality and range of uses of our unit. We also wanted to create a tool which will be able to maintain its value even as feature sizes in nanotechnology become smaller as we anticipate over the next decade. We reasoned that by achieving both of these aims, we would expand the demand for our product considerably and therefore increase its profitability. However, in order to achieve resolutions below the diffraction limit that is possible with available laser wavelength sources, we need to use a different set of methods to conventional microscopy. Our system uses Stimulated Emission Depletion (STED) to obtain superresolution in both fabrication and imaging. With Stefan Hell being awarded a Nobel Prize for the invention of STED in 2014, we believe the process places our product at the cutting edge of optical fabrication technique, giving our customers a product with long-term value.

We see from figure 14.1 the significant improvement that we can make with STED imaging compared to confocal. We aim to achieve resolutions of around 50nm in both imaging and fabrication.

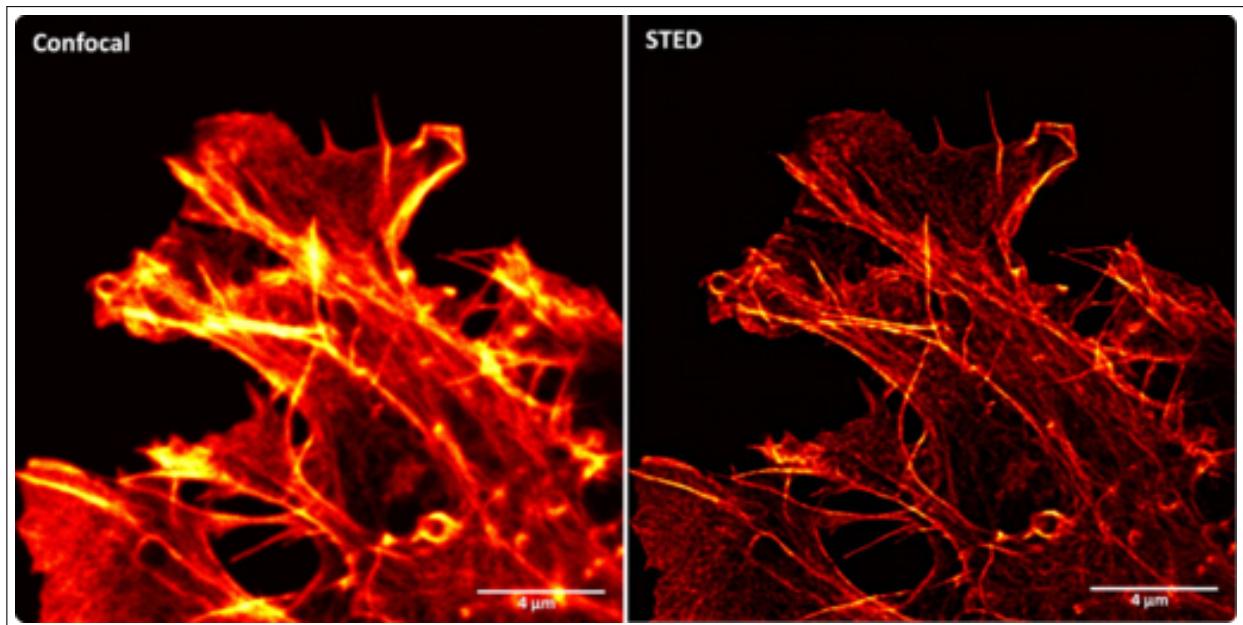


Figure 14.1: [120] The improvement from confocal to STED imaging - scale shown bottom right

14.1 STED

14.1.1 Theory

Where resolution is usually limited by the FWHM (Full Width at Half Maximum) of the PSF - linked to the wavelength of light by Abbe's formula - STED is able to produce a theoretically unlimited resolution by engineering the point spread function of incident light. The idea behind STED is that it is possible

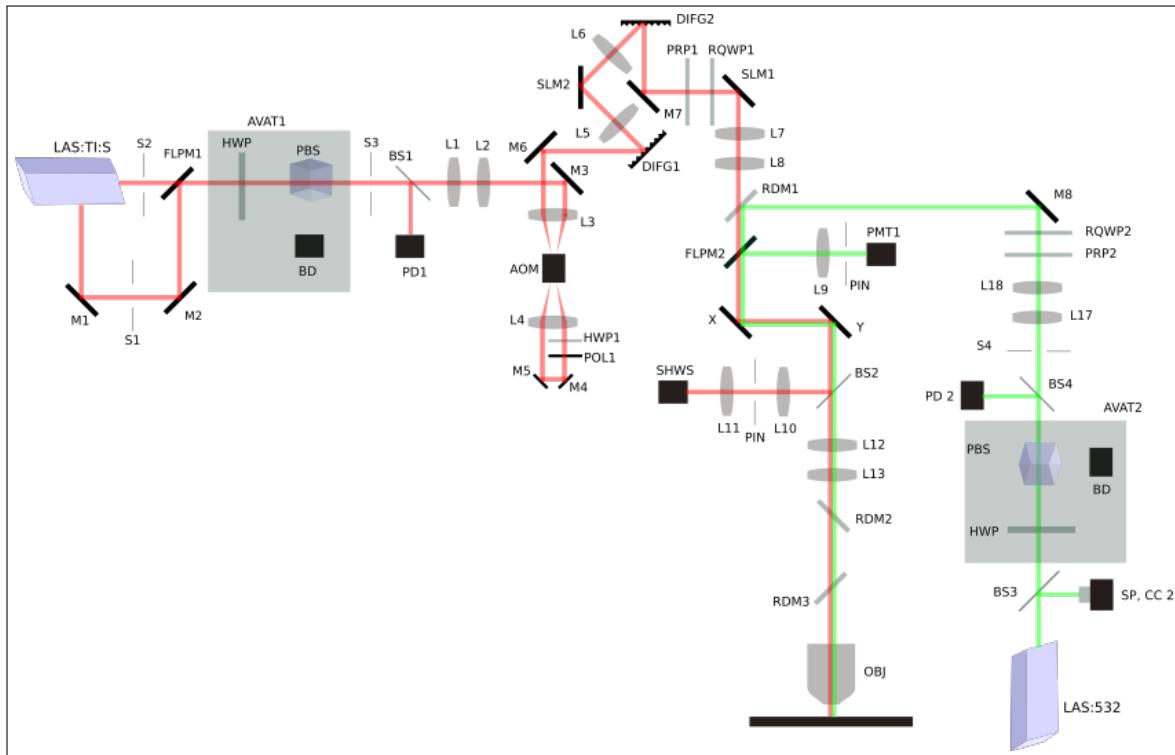


Figure 14.2: The optical layout for the STED system - incorporating both the 532nm and 780nm laser sources. We place PRP1, RQWP1 during STED imaging and PRP2, RQWP2 during STED fabrication

to increase resolution in fluorescent-based microscopy by suppressing emission around the focus and therefore narrowing the point spread function as seen by the detector. A comparison between confocal and STED microscopy is shown in figure 14.1. This is achieved by using 2 different laser beams. The first is the usual excitation beam used in 2 photon microscopy and polymerisation which is used to excite dyes into a state where they will then relax naturally into their ground state emitting photons of a given wavelength. The second beam is shaped to form a torus or 'donut' which, in STED microscopy, is also slightly red-shifted. This beam is used to stimulate the emission of a photon from the excited state which has a different wavelength to the emitted fluorescent photons measured by our detector. Our implementation is shown in figure 14.2 which will be referred to at various points in this section.

By forcing dye molecules that lie in the STED depletion area to relax by a different mechanism to their natural fluorescence we limit fluorescence to a sub-diffraction sized minima at the centre of our torus. By increasing the power and therefore intensity of our STED beam, we reduce the size of the minima and therefore increase our effective resolution. The link between intensity and resolution is discussed further in 14.1.3.

The Jablonski energy diagram in figure 14.4 (showing energy levels for different states of excitation) for a STED-suitable fluorophore generally features a high energy excitation wavelength and

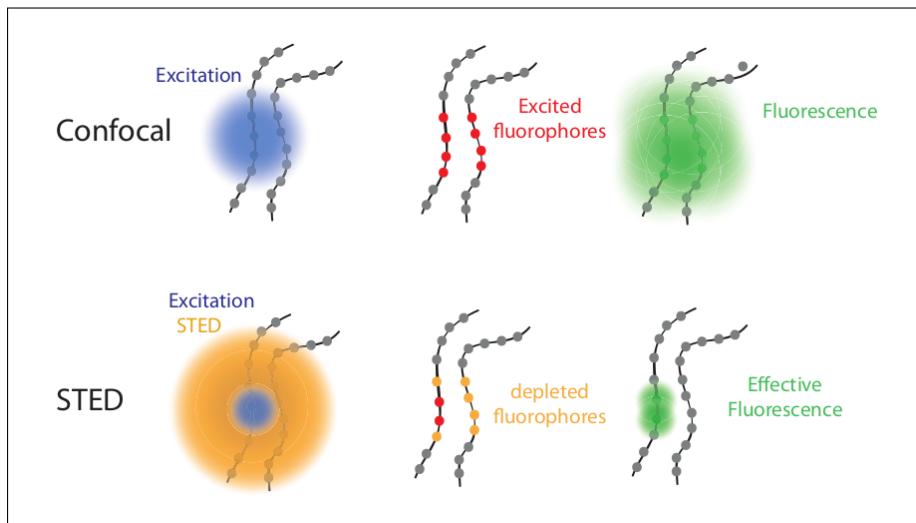


Figure 14.3: Diagram demonstrating how STED improves on the resolution provided by confocal microscopy. Grey dots are unexcited dye fluorophores, red are excited and orange are fluorophores which have been depleted. By overlapping a STED beam with the excitation beam we are able to significantly reduce the number of fluorophores which are able to fluoresce. This results in improved spatial resolution. [121]

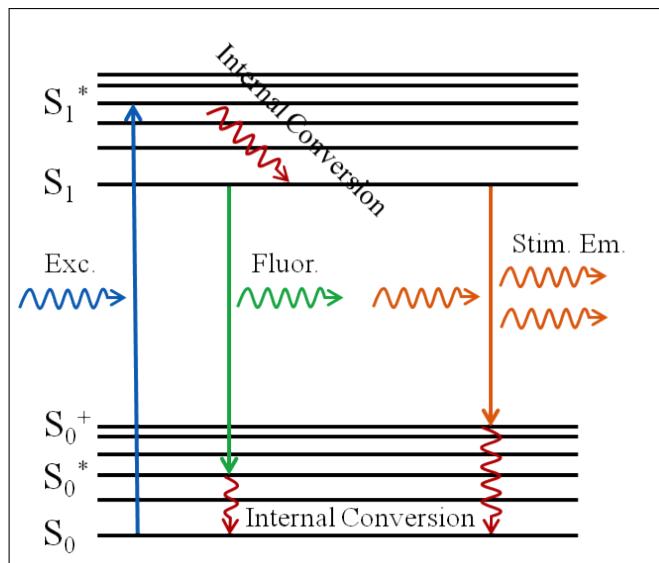


Figure 14.4: Jablonski diagram for STED showing the energy level changes for fluorophore photons. A high energy wavelength is used to excite the fluorophores (blue) and they then return to ground level either by fluorescing (green) or are Stimulated by the STED beam, preventing fluorescence (orange) [122]

at least two different paths for deexcitation: natural 'fluorescent' relaxation and stimulated emission. The STED concept relies on us being able to force the rate of stimulated emission to be greater than the rate of fluorescence. It is clear from the Jablonski diagram that, because we require very specific wavelengths of light, both the excitation and depletion wavelengths of the dyes must match with the available lasers in our optical layout - an important design parameter. This is discussed in more detail in 14.1.6.

The 'donut' beam shape in x-y can be created by using either an SLM, or by using a 2-pi phase

plate to create an optical vortex as in our system. While using the SLM gives us the ability to correct for aberrations as discussed in 19.4, it also results in a lower quality torus when compared with a phase plate. The 2-pi phase plate creates a central zero because the linear (or pseudo-linear depending on the quality of the phase plate used) progression of phase delay from 0 to 2π ensures that light from opposite sides of the entrance pupil interferes destructively at the optical axis in order to create a central zero.

We choose to use physical phase plates to generate the torus in the lateral direction for both fabrication and imaging as not only are phase plates significantly cheaper than using an additional SLM in the 532nm setup (equivalent to SLM1 in figure 14.2), they also produce a very consistent beam shape and we do not employ parallelised STED (for reasons discussed in 14.1.10) so do not require multiple foci generation as in most of the other processes. Another benefit to this setup is that we can use SLM1 to correct for aberrations using adaptive optics (19.4) as well as to implement 3D STED to improve axial resolution where required as discussed in 14.1.9.

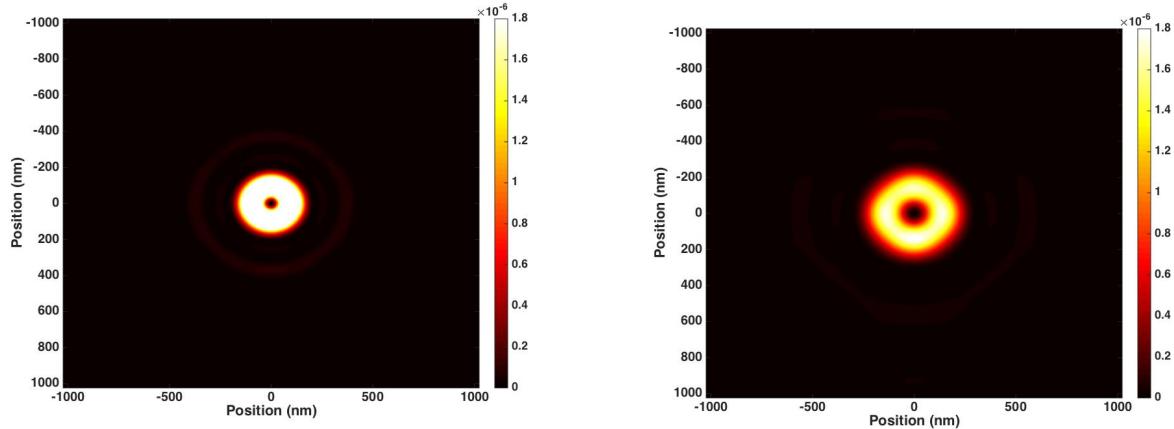
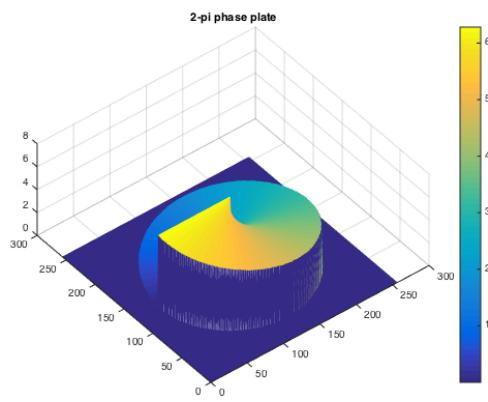
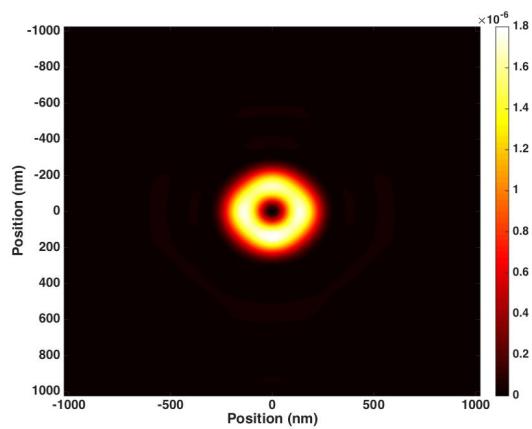
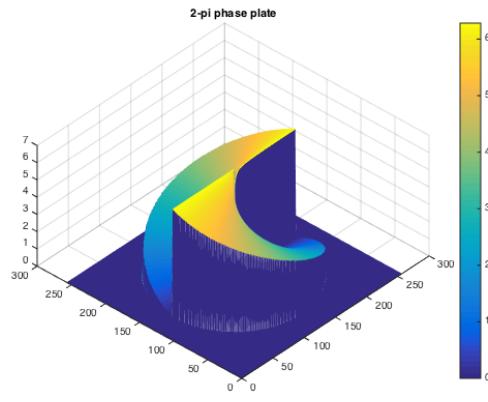
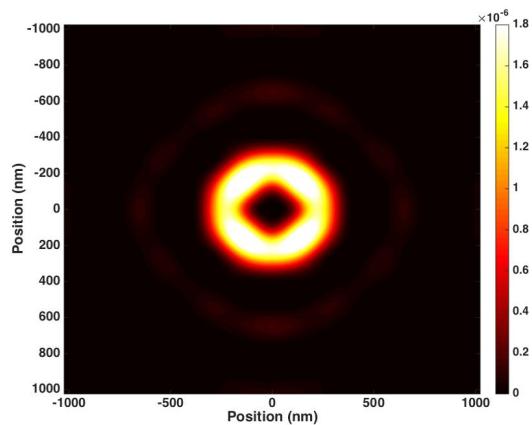
14.1.2 Modelling Vortex Phase Plate in MATLAB

In order to demonstrate how the 'donut' beam shape is formed by using a Vortex Phase Plate, I built a MATLAB simulation. The 2-pi phase plate is modelled as a phase difference in the code. The construction of a 2-pi phase plate in reality therefore requires the optical path difference in a material to vary with the azimuthal angle in order to adjust the phase. Since phase plates alter phase by the optical path difference, we require two different removable helical phase plates, one for each laser wavelength. We are able to purchase helical phase plates from Holoor at a cost of just under £1,000.

When we modelled the 2 wavelengths in MATLAB (with normalized energy) we found that the longer wavelength appears to produce a larger minima (as we would expect) and a larger depletion area. As a result, the intensity (scale on the right) is significantly lower for the 780nm PSF. To produce the same resolution with the 780nm laser in depletion as with the 532nm laser, we must therefore use significantly more power.

Another interesting consideration can be modelled in MATLAB. The topological charge of a vortex phase plate refers to the number of 0 to 2π phase ramps in one 360 degree azimuthal cycle. By increasing the topological charge of our phase plate, we increase the size of the PSF significantly and in particular we increase the size of the central zero (figures 14.6 and 14.7). For our purpose we therefore choose to purchase phase plates with topological charge $m = 1$ to obtain the best possible resolution. From our plot with $R \approx 200\text{nm}$ and $r \approx 25\text{nm}$ we have an approximate STED area of $A = \pi(R^2 - r^2) = 10^{-9}\text{cm}^2$ for $\lambda = 780$ - this is used in 14.1.3.1.

We are also able to use MATLAB to model the effects of increasing intensity. Figure 14.11 shows

Figure 14.5: PSF for depletion at $\lambda = 532\text{nm}$ PSF for depletion at $\lambda = 780\text{nm}$ Figure 14.6: Phase Ramp with $m = 1$ Figure 14.8: PSF with $m = 1$ Figure 14.7: Phase Ramp with $m = 2$ Figure 14.9: PSF with $m = 2$

that doubling the intensity of incident STED illumination has a significant effect on the size of the central zero when compared to figure 14.10. In our design we aim to minimise the area of this central zero but without exceeding intensities that cause permanent photobleaching (destruction of the fluorophores) in the dye.

We also note here a practical consideration in the design: the quarter wave plate. So far we

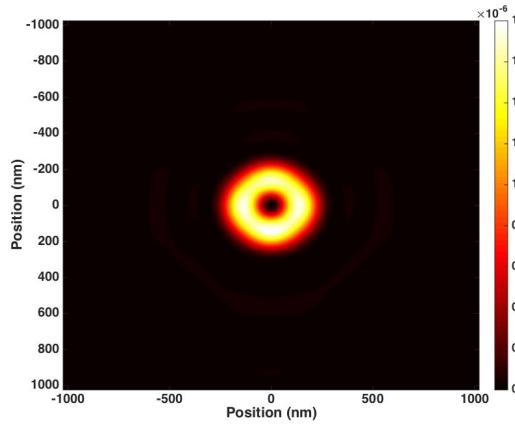


Figure 14.10: Original PSF

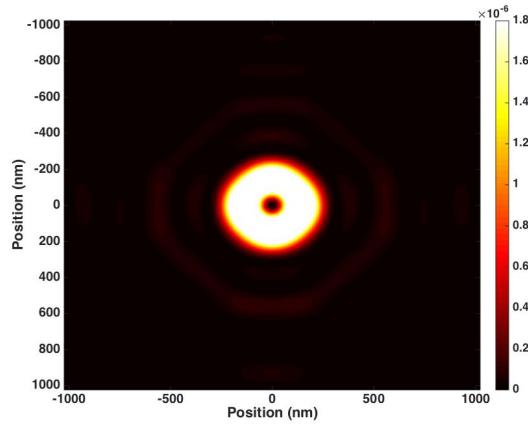


Figure 14.11: PSF with double intensity

have considered linearly polarized light. For low NA this assumption is reasonable because the z component of the electric field is still almost perpendicular to the normal. For the high NA lenses on our objective however, this z-component becomes more pronounced and actually prevents us from generating a clean zero at the focal point. For this reason in addition to the Vortex Phase Plate, we require a quarter wave plate which converts linearly polarized incident light to circularly polarized light.

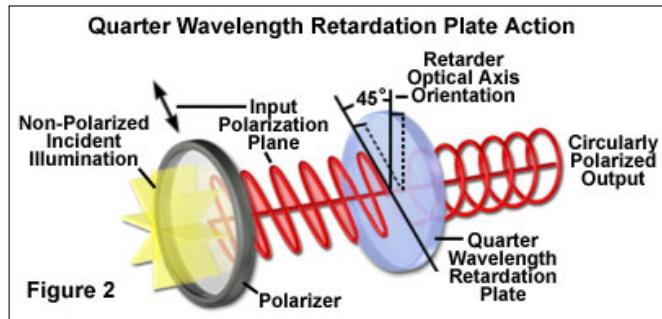


Figure 14.12: [123] Quarter Wave Plate demonstration

A quarter wave plate is made from a birefringent material with a different refractive index on its 'fast' and 'slow' axis. As shown in figure 14.12, when linearly polarized light strikes the quarter wave plate, the electric field is split into two components. The component on the slow axis is then delayed by a quarter wavelength relative to the fast axis leading to a circular pattern on leaving the plate. The handedness of the plate is the direction the electric field appears to travel in (at the 12 o'clock position) as viewed by the sample.

In order to generate the central zero of our donut pattern, we require the handedness of the Vortex Phase Plate to match the handedness of the quarter wave plate. We use fine-tuning rotatable holders available from Thorlabs in order to be able to adjust the Quarter Wave Plates to obtain the best possible STED beam profile. The half wave plate used early in the beam paths for both lasers

described in 20.1 makes it much easier to position the quarter wave plates accurately.

14.1.3 Resolution

When we follow the derivation of the FWHM for a STED microscope given in [124], the link between STED beam intensity and resolution becomes clearer.

We define the saturation factor as:

$$\zeta = \frac{I_{STED}}{I_s} \quad (14.1)$$

where I_{STED} is the maximum of the STED intensity distribution and I_s is the effective saturation intensity (the intensity at which the probability of fluorescence is halved - a function of the dye).

For a STED microscope, the effective PSF is given by:

$$h_{eff}(\bar{x}) = h_c(\bar{x})\eta(\bar{x}) \quad (14.2)$$

This is the confocal PSF h_c multiplied by the suppression factor - the fraction of fluorescence that is still detected with the STED beam applied. This suppression factor is represented by the following equation where σ is the cross-sectional area for STED, t_s is the duration of the STED pulse:

$$\eta(x) = e^{-\sigma t_s I_{STED}(x)} \quad (14.3)$$

$I_{STED}(x)$ could take many forms but requires the property of a central zero to allow the method and is therefore modelled in $\sin^2(x)$ form. The confocal PSF can be similarly modelled in $\cos^2(x)$ form.

Multiplying these PSFs it can be shown that the effective PSF is a narrower Gaussian. We find that for large saturation factors, the optical resolution is inversely proportional to the square root of the saturation factor:

$$\Delta x = \frac{\lambda}{2n\sin\theta\sqrt{\zeta + 1}} \quad (14.4)$$

or in a form that we can compare to confocal microscopy:

$$\Delta x = \frac{d_c}{\sqrt{\zeta + 1}} \quad (14.5)$$

Where d_c is the confocal (diffraction limited) optical resolution.

We find that for large saturation factors, the FWHM is inversely proportional to the square root of the saturation factor. The implication is that, with enough laser power, we would be able to obtain infinitely small resolution. (14.6)

Since $nsin\theta$ is determined by the numerical aperture of the objective ($NA \approx nsin\theta$) we only require further values for the intensity of the STED beam and the saturation intensity of the dye to determine our theoretical maximum resolution using:

$$\Delta x = \frac{\lambda}{2NA\sqrt{\zeta + 1}} \quad (14.6)$$

At this stage a rough calculation using typical values for the saturation factor of 15, and selecting our 1.4 NA objective from the turret gives us resolutions of 45nm for imaging ($\lambda = 532\text{nm}$) and 70nm for fabrication ($\lambda = 780\text{nm}$). Even with a conservative estimate for ζ , the STED set-up produces the highest resolution in both fabrication and imaging in our system. The next few sections largely center around our design decisions in order to maximise the saturation factor, ζ .

14.1.3.1 Saturation Intensity

This section introduces an additional optical concept: the absorption cross section. The absorption cross section is an effective area that is used to measure the probability of absorption under an incident beam. The units of the absorption cross-section (given as σ) are m^2 .

We consider two arrangements discussed in 14.1.4: continuous wave (CW) and pulsed depletion beams.

In a pulsed system, the fluorescent area depletes as $\exp(-\zeta)$ which is equivalent to $\exp(-\sigma\phi)$ [125] where σ is the optical cross-section for STED and ϕ is the number of incident depletion photons in a given area and in a given pulse [126]. For a typical optical cross-section of 10^{-17}cm^2 , we require $\phi > 10^{17}/\text{cm}^2$ in order to obtain $\zeta > 1$ and therefore a subdiffraction sized fluorescent spot. Since the area of the STED donut is approximately 10^{-9}cm^2 we require a STED pulse containing 10^8 photons. Using Planck's constant (h), the number of photons (n) and the STED wavelength we can obtain a pulse energy:

$$E_{pulse} = nh\frac{c}{\lambda} \quad (14.7)$$

$$E_{pulse} = 10^8 \cdot 6.626 \times 10^{-34} \cdot \frac{3 \times 10^8}{780 \times 10^{-9}} = 2.55 \times 10^{-11} \text{J} \quad (14.8)$$

For a pulsed laser at 80MHz, this gives a power of $2.55 \times 10^{-11} \cdot 80 \times 10^6 = 2\text{mW}$.

In a CW STED setup, our calculation is slightly different. In the steady state, the probability of a fluorophore being excited is $k_{exc}/(k_{exc} + k_{fl} + k_{STED})$ [125] assuming that there is negligible excitation caused by the STED beam [127] where k_{exc} is the rate of excitation, k_{fl} is the rate of fluorescence and k_{STED} is the rate of stimulated emission ($k_{STED} = \sigma I$). We require $k_{STED} > k_{exc}$ for fluorophores

to be depleted and $k_{fl} > k_{exc}$ in order to avoid excitation saturation. This allows us to deplete the excited population and obtain subdiffraction sized resolution as before. This requires us to have a STED intensity I greater than I_s . For a continuous laser, this minimum value can be calculated as:

$$P_s = A \times hc \times k_{fl} / (\lambda_{STED} \cdot \sigma) \quad (14.9)$$

Where all symbols have the same meaning as before and A is the cross-sectional area of the STED pulse. If we consider the same STED area and optical cross section as before. Taking a typical fluorescence lifetime for a STED dye of 3ns, we obtain k_{fl} as 3×10^8 and therefore a power of 8.5mW.

14.1.4 Continuous vs. Pulsed

STED systems originally required the use of 2 pulsed lasers which were coupled using triggering. However, recently systems have been demonstrated which use either one continuous mode laser or both laser sources in continuous mode. In this section, we consider the relative merits of these methods and how that relates to our design.

Equation (14.3) showed that the ability of the STED beam to deplete fluorophores is dependent on the duration of the pulse. Our Ti:Saph laser allows us the possibility of either using its quasi-continuous mode or using femtosecond pulses which, even if we used stretching techniques, would have a relatively short pulse length for a pulsed STED system.

STED systems are ideal when the STED laser pulse is significantly shorter than the fluorescent lifetime (around 3ns) but longer than the vibrational lifetime (around 1.5ps) to avoid reexcitation by the STED pulse [121].

There are two main considerations in this design decision. The first is that continuous wave (CW) lasers must have a higher average power than pulsed lasers in STED to achieve the same resolution. This is because the efficiency of a STED pulse is proportional to the flux of depletion photons which is proportional to the intensity of the beam. Since the only effective time of the depletion beam is during the fluorescent lifetime of the fluorophores, a CW laser effectively wastes power between those fluorescent lifetimes.

From figure 14.13 we observe that the average power of a pulsed laser is $P_{pulsed} = \frac{\bar{P}t}{t+T}$, compared to a CW laser with average power \bar{P} where t is the length of the pulse (similar to the fluorescent lifetime of the dye) and T is the time between pulses. $\frac{1}{t+T}$ gives the frequency of the laser. Experiments suggest that for usual values, T is around 3 times longer than t [128] giving a 4-fold increase in average power for CW lasers. This means that CW lasers start to encounter problems with dye photobleaching at much lower effective intensity than pulsed systems. Some of this problem can be

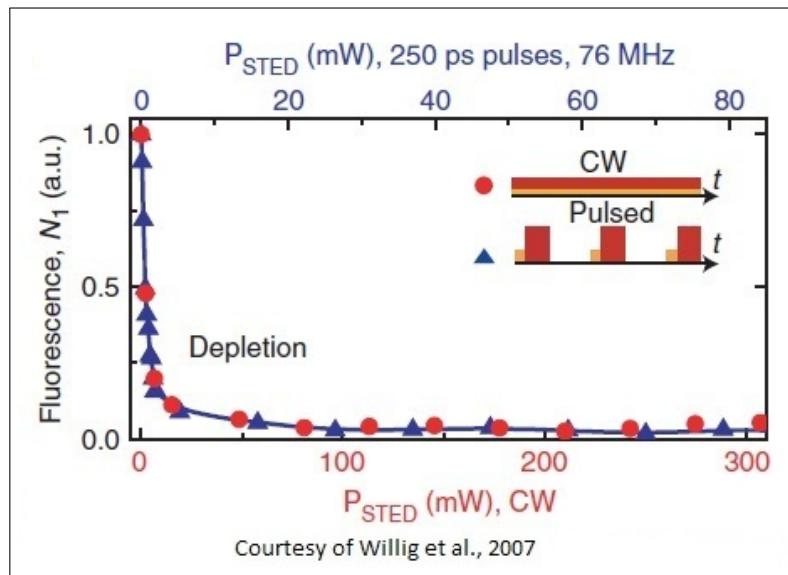


Figure 14.13: Graph from [128] shows the significant power difference (around 4 times) between CW and Pulsed laser power in STED in order to produce the same depletion of fluorescence

avoided by using fast scanning to limit the amount of time the CW-STED beam illuminates the sample as discussed in 14.1.5. Another way of improving continuous systems is to use gated detection [129].

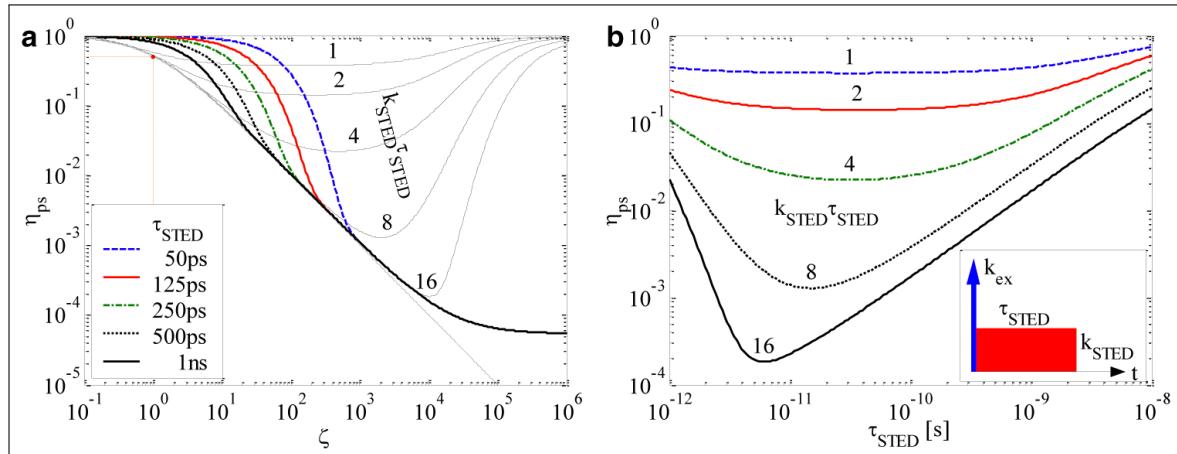


Figure 14.14: From [130] a) the probability of spontaneous emission η_{ps} against saturation factor for an 80MHz source with different pulse lengths b) the same data shown as a function of τ_{STED} . k_{STED} is again the rate of stimulated emission, τ_{STED} is the STED pulse length

Figure 14.14 plots the probability of spontaneous emission (fluorescence) against various STED pulse lengths (where the $k_{STED}\tau_{STED}$ isolines show constant average powers). We notice from figure 14.14 (b) that long pulses (i.e. small ζ) are inefficient for suppressing fluorescence because fluorescent lifetimes (which we are attempting to suppress) are usually relatively short (see 14.1.6). Equally, very short pulses lead to inefficiencies because the rate of depletion is limited by the physics of the particles and we cannot force it any faster. It was shown by [130] that a STED pulse length of $\tau_{STED} \approx 30$ ps provided the ideal conditions for depletion efficiency as intensity is increased.

However, in practice longer STED pulse lengths (of the order of 100 ps) actually yield slightly

better results [130]. There are a number of reasons for this. Firstly, we can minimize polarization effects (because longer pulses give excited molecules time to rotate). Secondly, longer pulses avoid problems of timing jitter [131]. Where CW lasers suffer from needing to use lower power to avoid photobleaching of the fluorophores, the benefits of reducing STED pulse length below the fluorescent lifetime of the dye are limited by the maximum power being applied. Since photobleaching probability scales non-linearly with higher orders of intensity, we look to avoid using very high powers.

In our design, using the 80MHz repetition rate of the Vitara T with a 100 fs pulse width at output would require us to use our pulse shaping (20.8.1) to stretch these pulses to obtain pulses in the picoseconds in order for us to be able to consider the source as pseudo-continuous with respect to the dyes.

Ultimately, the design decision between CW-STED and pulsed STED for imaging came down to factors elsewhere in our product. In general, lasers which had the ability to switch between pulsed and continuous wave modes produced lower energy-per-pulse. Since the quality of micromachining in particular is limited by energy-per-pulse, we made the decision to use the Coherent Libra with its Vitara seed source output. This adds a small degree of complication with regard to pulse shaping when imaging as mentioned.

For STED fabrication, we use the Ti:Saphire without significant pulse shaping and the Continuous mode 532nm laser is used to deplete the photoinitiators.

From experiment, there is little difference in fluorescence inhibition if the excitation pulse is pulsed or continuous [130]. We use continuous excitation in imaging due to the single mode of our 532nm laser however we use a pulsed excitation laser in STED fabrication for the thermal efficiency reasons discussed in 6.2.2.

14.1.5 Power Considerations and Photobleaching

When illumination is relatively weak as in one photon excitation, photobleaching is linearly dependent on intensity. However, for two photon excitation the dependency can increase to 3rd order or more [132]. For high levels of illumination, even single photon processes have a non-linear dependence on intensity. Since signal-to-noise ratio is one of the main limiting factors for STED resolution and photobleaching significantly reduces our signal strength it is of great importance that we minimise the effects.

It was shown by [132] that the primary cause of photobleaching in STED microscopy was the action of the STED beam on excited fluorophores. One of the easiest ways of reducing photobleaching then is to increase the scan speed. By increasing the total number of scans, we can actually maintain the dye's exposure time and the image acquisition time but with a much higher fluorescent yield. As

shown in 14.15 the fluorescent yield is improved by increasing the dark time between scans.

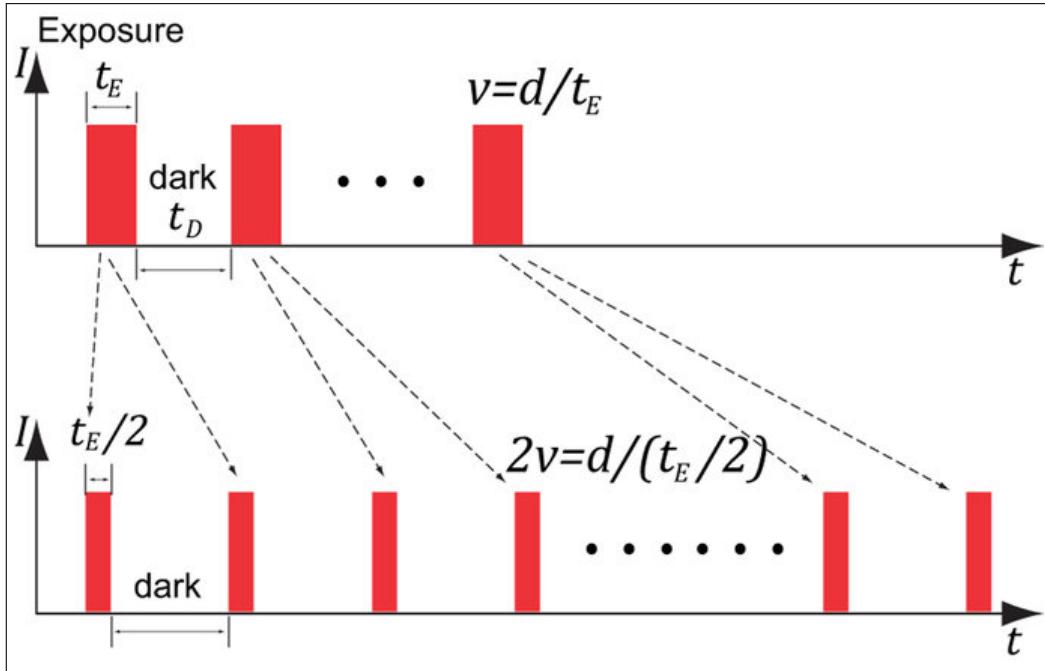


Figure 14.15: [132] a graphical demonstration of the effect of faster scan speeds and scanning twice

In order to obtain an image, we illuminate each area of the sample with an equal dose so that [132]:

$$nFL \frac{I}{v} = \text{constant} \quad (14.10)$$

Where F is the number of frames, L is the number of lines in the y direction (with a given separation), I is intensity of illuminating light and v is the scan speed and n is the number of scans. From equation 14.10 we see that in order to maintain the dose of illumination for a given area as we double scan speed, we must either double intensity or double the number of scans. As already discussed, we wish to limit using very high intensities due to non-linearities in the likelihood of photobleaching. We therefore must aim to double the amount of scans by going over the scan area twice. Crucially however, by employing a system with a faster scan rate, we increase the time between successive excitations. This gives the fluorophores in a given area more time to relax from their excited triplet state (when they are most vulnerable to photobleaching). This means we can all but remove the advantages of pulsed laser implementation of STED microscopy in avoiding triplet state build up by using a very fast scan speed. In fact, fast-scanning allows us to continuously excite fluorophores which theoretically should lead to a higher fluorescence count rate and therefore easier and better detection at the PMT [133].

All of these reasons mean we choose to combine the excitation and depletion beams early on (at

RDM1) so that we can use our system's galvanometers to change the position of the beam rather than the piezo stage and obtain the scan speeds necessary to improve our fluorescence efficiency.

14.1.6 Dyes

The dyes that are used in STED microscopy need to have some important properties. In order to reduce the amount of lasers in our system and therefore increase cost-efficiency we reuse the pulsed Ti:Saph laser for excitation in STED fabrication and depletion in imaging while using a 532nm laser for depletion in fabrication and excitation in imaging. In order to be able to use the same pair of lasers in both imaging and fabrication, we require a compatible polymer and dye. We conducted a multicriteria analysis in order to decide which dyes to use in our system. The following factors were taken into account:

- Excitation Wavelength (nm). Earlier in the design process we considered using the same laser for STED excitation as for confocal illumination and our live imaging modes. This required us to aim for as short a wavelength as possible in order to improve the resolution from those methods. Since we altered the design to give us a versatile LED light which can be used in all of our live imaging methods, this was no longer a consideration. 532 nm lasers are used in the early STED fabrication systems [134] as the depletion beam so we noticed that if possible we should aim to match the required laser wavelengths for STED imaging and fabrication. In order to give us this versatility in our system, we aimed for dyes with an excitation wavelength of 532 nm. All of the candidates considered met this criteria so it is not weighted in the multi-criteria analysis.
- STED wavelength (nm). In order to be able to reuse the same Ti:Saph laser in Two Photon Polymerisation and STED, we require a STED wavelength that falls inside the 790-820nm FWHM of our laser power (at full bandwidth). Dyes are therefore penalised for deviation from this range. This is weighted significantly because we aim to minimise I_s in order to obtain maximum resolution for a given intensity and dyes become significantly less efficient when excited away from their peak absorption.
- Saturation Intensity. The effective saturation intensity I_s gives us a way of estimating the optical resolution of the process as it is required to find the saturation factor (14.1). The lower the Saturation Intensity, the greater the score (note: this is not completely reliable as a measure because we may have a different upper bound for power if one dye is more susceptible to photobleaching).
- fluorescent Wavelength (nm). In order to be able to detect the fluorescence, we need to be able to separate fluorescent light from the reflected light from the sample. Since the intensity of

light is necessarily very low - the better resolution we obtain the fewer fluorophores we must be detecting - we need to be able to use a dichromatic mirror that filters fluorescent light from reflected incident light. We therefore score positively dyes which have a larger difference between incident and fluorescent light

- fluorescent lifetime. We are scanning the sample very quickly so the shorter the lifetime of fluorescence, the better the quality of our imaging.
- Price (P). Price is weighted relatively little in our multicriteria analysis because we are aiming to make a profit margin on selling dyes. A higher priced dye with more ideal properties could be sold onto the customer for perhaps an even greater margin than a cheaper dye.

With all of these factors in mind, we shortlisted the candidates down to 3 options: MR121 - a commonly used dye in STED systems, Nitrogen vacancies (implemented by doping with nanodiamonds) - which have recently been used to obtain some of the highest available resolutions (down to 10nm [135]) because of their very high photostability, and Abberior STAR. Their properties are summarised in table 14.1 and our multi-criteria analysis including weightings is completed in table 14.2.

Criteria	MR121	Nitrogen Vacancies	Abberior STAR
STED Wavelength (nm)	775	793	750
Fluorescence Wavelength (nm)	676	630-700	627
Fluorescent Lifetime (ns)	1.85	12	3.9
Saturation Intensity (MW/cm^2)	40	6	80
Cost/mg	280 [136]	2.75 [137]	486 [138]

Table 14.1: STED Dye Properties

From our MCA we decide that our default dye will be the MR121. However as part of our business plan, we may offer users the opportunity to use other dyes. In what is a rapidly developing field, we also expect the creation of alternative dyes over the coming years.

With a saturation intensity of $40MW/cm^2$, we can now calculate the possible saturation factor and therefore our resolution. We can obtain a STED area of $10^{-9}cm^2$ meaning if we use all of our available laser power we can obtain an intensity of $525MW/cm^2$ through the Ti:Saph and a saturation factor of 13 - around the value calculated earlier and giving an optical resolution for imaging of 49nm.

14.1.7 STED Fabrication

The system was originally designed to provide high resolution imaging. However, during the research and design process, it became clear that with relatively little additional components and a carefully

Criteria	Weight	MR121		Nitrogen Vacancies		Aberrior STAR	
		Rating (5)	Score	Rating (5)	Score	Rating (5)	Score
STED Wavelength	3	2	6	3	9	1	3
Fluorescence Wavelength	1	3	3	3	3	3	3
Fluorescent Lifetime	5	3	15	1	5	2	10
Saturation Intensity	3	4	16	4	16	3	12
Cost	1	2	2	3	3	2	2
Total		32		29		21	

Table 14.2: Multi-Criteria Analysis (MCA) for STED Dyes

chosen alternative laser source (the 532 nm laser) to the Ti:Saph laser we could implement STED fabrication into our system and improve both fabrication and imaging lateral resolution to 10s of nanometres. Figure 14.16 shows the similarity in the set-up. The main consideration is that we need drastically different pulse lengths for imaging (picoseconds) and fabrication (femtoseconds) from our Ti:Saph laser. We employ pulse shaping to achieve this requirement 20.8.1. We switch easily from STED microscopy and fabrication in our system by removing RQWP1 and PRP1 and placing RQWP2 and PRP2 in our optical layout (see figure 14.2).

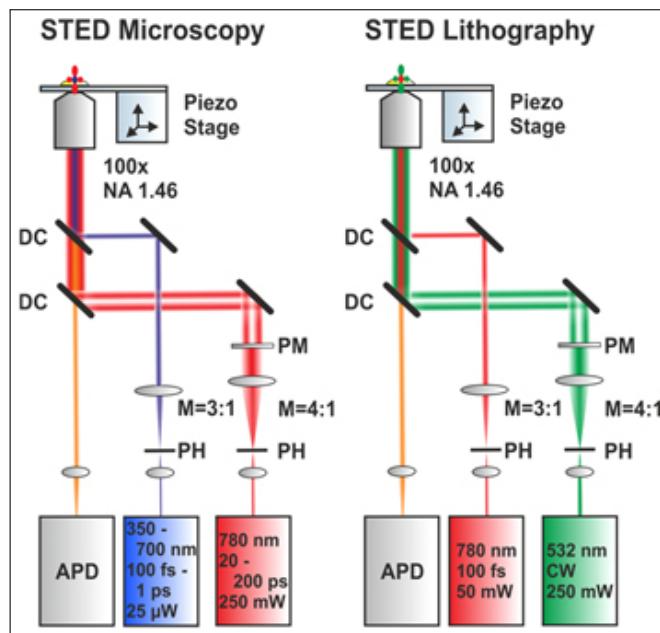


Figure 14.16: [139] A comparison of the set-ups for STED imaging and fabrication

Just as we needed dyes with particular properties in order to image using STED, in lithography we require a photoinitiator with a number of important properties (from [140]).

- The photoinitiator must not absorb the STED wavelength - a large difference between the ab-

sorption and emission spectra (this is known as the Stokes shift)

- The photoinitiator must remain stable when illuminated with the depletion beam and therefore resistant to photobleaching even at high intensities. The more resistant to photobleaching, the higher the intensities of depletion that may be used and therefore the more reliable the depletion is - even of very short singlet states.
- The lifetime of the excited singlet state of the photoinitiator should be long enough to allow the photoinitiator to be depleted and therefore avoid causing polymerisation.

With the technology being relatively primitive at this stage, there are only 2 photoinitiators that fulfil the criteria, isopropyl thioxanthone (ITX) and 7-diethylamino-3-thenoylcoumarin (DETC). We choose to use DETC as its excitation wavelength (780nm) most closely matched our Ti:Saph laser source. The 532nm laser source was actually chosen in order to match the depletion wavelength of DETC and has been one of the main constraints in the design process. Future iterations of our project are likely to have more options for photoinitiators as the chemistry develops and therefore more options. It should be stressed that our laser sources are popular design wavelengths and therefore we expect our system to be compatible even as the technology develops and other photoinitiators become available.

The resolution of STED fabrication is unlimited in the same sense as STED microscopy - we can increase the depletion intensity to reduce the size of the central zero infinitely - however this also requires a photoinitiator with a threshold of polymerization which is similar to the peak of the incident PSF [134].

Users will be able to achieve resolutions of just over 50 nm giving feature sizes that are at the cutting edge of nanoscale fabrication as shown theoretically in 14.1.3 and experimentally [134]. An example of what could be achieved is shown in figure 14.17

14.1.8 Beam Alignment and Drift

Beam alignment in the STED process is perhaps more important than anywhere else in our product. Not least because we are aiming for resolutions close to 50nm so drifts of 5-10nm now represent a 10-20% error whereas in the confocal system they might have represented 2-4% and been acceptable. This is exacerbated because misaligning one or both of the beams used in STED doesn't just lead to a shifted pattern, it removes all of the resolution benefits that the process provides. Figure 14.18 shows the impact a misaligned STED beam could have on the effective fluorescence area. Note that if the beams are completely misaligned with no overlap, we will simply obtain the confocal image. The better aligned our beams are, the better the efficiency and therefore the quality of the image. Since

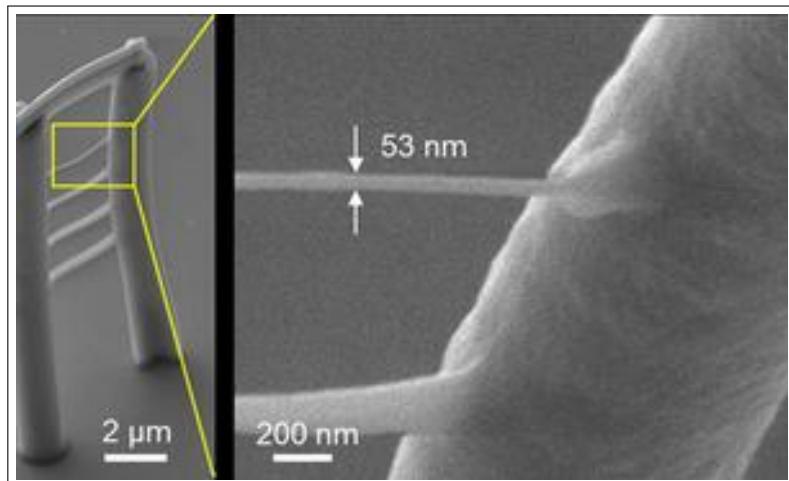


Figure 14.17: An example of the level of detail in fabrication that our system can achieve [141] (scale bottom left)

we do not use adaptive optics in our STED processes, we cannot use the SLM to correct for drift [142]. This means we need to develop a method for aligning the beams whenever the STED system is to be used.

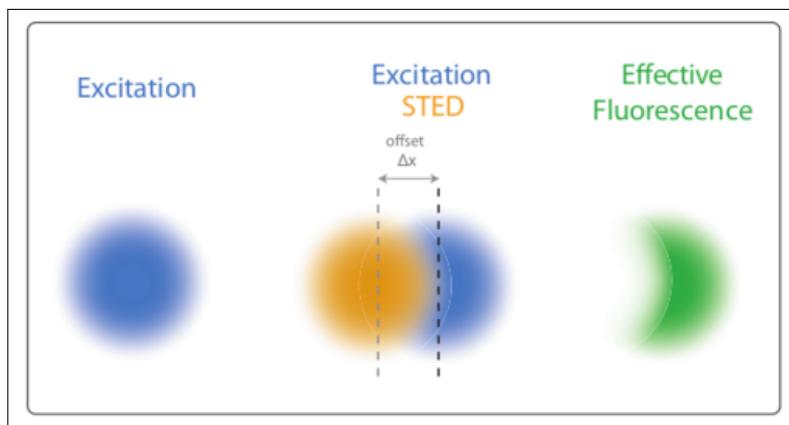


Figure 14.18: [121] The affect on the effective fluorescent spot caused by a misaligned depletion beam

The problem when aligning a STED system is that we are aiming to achieve the maximum STED brightness - which occurs when the central zero of the depletion beam matches the excitation peak - but a very misaligned beam produces effectively the confocal brightness which is significantly brighter than the STED case. We must be more careful than simply searching for the alignment which causes the maximum brightness

Due to the very high sensitivity to poor alignment, our system will automatically check its alignment whenever the STED processes are used and every 2 hours if there is prolonged use of the STED mode.

As in the system described in [143], we use a pseudo-sample of gold nanoparticles spincoated onto a sheet of glass to image the PSFs of both the excitation and depletion beams by viewing the

scattered light. Gold nanoparticles are highly reflective so we use low powers in both lasers. We use a dichroic mirror in place of DM3 and our CMOS camera in order to view the scattered reflection of the nanoparticles. A single gold nanoparticle with good reflectivity is chosen to be used for alignment by moving the galvanometers in order to place different particles at the focus. We obtain the PSFs by alternating scans with the 2 beams and make small adjustments to align their centres to within 5nm (10% of our desired resolution) which we consider to be an acceptable error.

14.1.9 3D STED

So far, the theory has revolved around lateral resolution because VPPs give us an easy way to implement the required torus in the lateral direction. The process can also be implemented in the axial direction but requires a slightly different implementation. We either use a 4-pi set up (discussed in section 14.2.1) or we must implement a torus in the axial direction. The ideal set-up uses a variable beamsplitter to implement a 3D torus at the same time as a lateral torus. This allows us to send some of the beam through the usual quarter wave plate while the rest of the beam instead passes through a phase ring (as shown in figure 14.19)

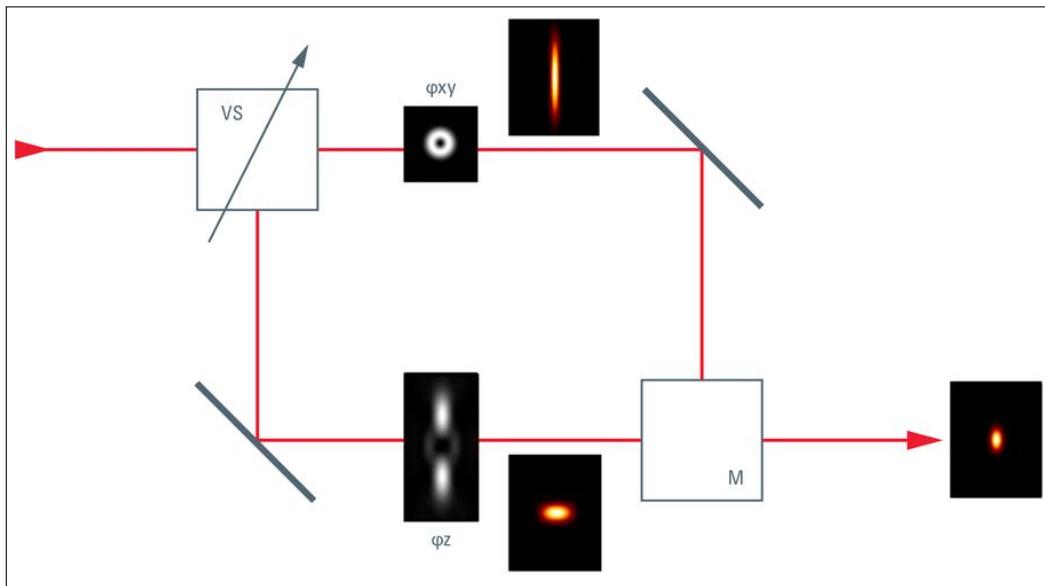


Figure 14.19: [139] A schematic diagram of a 3D STED set-up which allows tuning of the STED donut using a variable beam-splitter

The phase ring actually generates a torus in both lateral and axial directions as shown in figure 14.20. We use the 2-pi phase plate in our STED system because the quality of the torus when improving lateral resolution is better. The reason for using the set-up shown in figure 14.19 is that we can obtain the benefits of both the 2-pi phase plate and the phase ring and are able to tune our beam (with better axial resolution coming at the cost of worse lateral resolution). The x-z plot of the PSF produced by a pi phase plate is shown in figure 14.20 with the familiar 'donut' shape now visible in the axial direction.

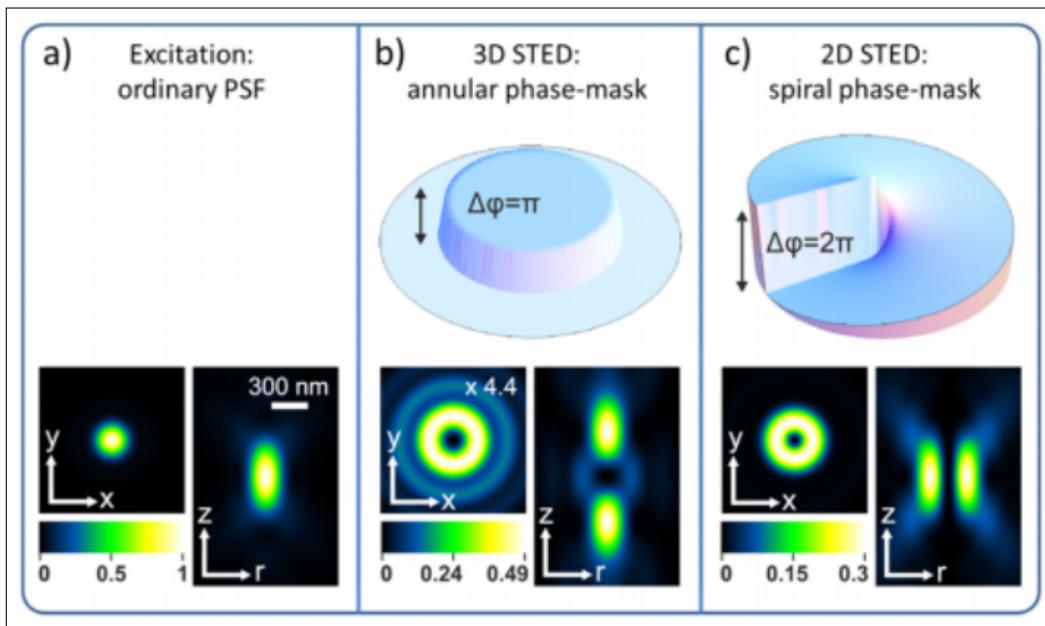


Figure 14.20: The PSFs in axial and lateral directions for the ordinary (no phase plate), 3D (Phase Ring) and 2D (Vortex Phase Plate) cases [144]

The problem with this set-up is that the focal points of the two path lengths must be the same (to within 25 nm, half the resolution) or the separate zeroes will be misaligned and the two paths will interfere to destroy both zeroes [145]. We therefore choose to just implement a standard axial STED system with best possible axial resolution but without the ability to tune. To achieve this, the system removes the helical phase plate PRP1 and introduces an axial 'donut' by using SLM1 to generate the annular phase-mask shown in 14.20 (centre). From experiment this will allow us to obtain imaging resolutions of 100 nm [146].

STED is also compatible with a 4-pi set up (discussed in 14.2.1) to improve axial resolution but we decided against this configuration for the reasons given later.

14.1.10 Future Developments - Parallelisation

A possible future development for our system would be to enable the parallelisation of the STED processes in the manner described in 16. In the current design, the speed is still limited by the fluorophore lifetimes of the fluorophores and the the kinematics of the galvanometers. It has recently been shown that the STED process can be performed with multiple foci [147].

There would be a few challenges in implementing a parallelized STED design in our system. We would need to be able to employ either removable diffraction gratings or another SLM in order to generate the multiple foci required. We would also need to reconsider the limitations on power. Since more of the sample is instantaneously illuminated in a parallelized scheme, the incident power is also greater. Since many of the problems with early parallelized systems have been with high levels of photobleaching we would need to carefully design the scan rate to allow fluorophores and

photoinitiators to relax out of vulnerable excited states.

14.1.11 Summary

Our incorporated STED fabrication and microscopy system provides our users with the unique ability to fabricate and image features at below 60nm scale. We believe that as commercial uses are developed for nanotechnology in the 'tic toc' fashion seen in electronic hardware and software, there will be increased demand for smaller feature sizes. Our product keeps us at the cutting edge of optical fabrication and we do not expect feature sizes in the market to decrease further without significant changes in the available technology.

14.2 Excluded Methods

14.2.1 4-pi Illumination - James

14.2.1.1 Theory and what it offers

4-pi illumination could have given our system improved resolution in the z-axis. In its simplest form 4-pi illumination uses 2 objectives to illuminate the sample from both directions. Providing there is no phase difference between the 2 sides at the focal point, the light will constructively interfere and generate a sharper PSF and also help to suppress sidelobes. Effectively, 4-pi illumination increases the aperture of the objective. The limit is of course a full sphere around the sample but this isn't possible in practice - this is demonstrated in figure 14.21. Even high NA objectives only transmit a segment of the perfect spherical wavefront - usually a maximum of around 140 degrees giving a total 2.6π wavefront in 4-pi set up [148]. The term '4-pi' is used to refer to principle for simplicity.

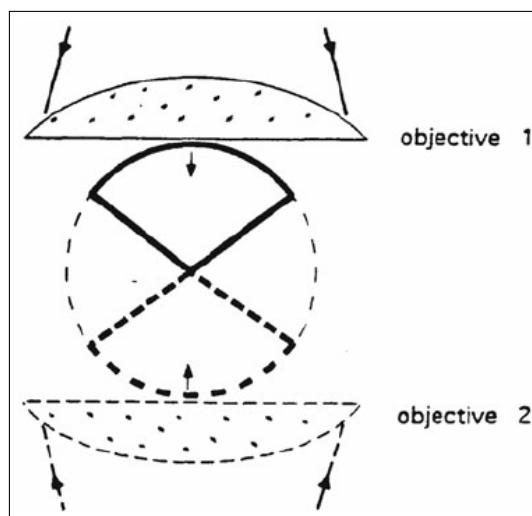


Figure 14.21: [148] Demonstration of the near-spherical wavefront produced by 4-pi illumination

In order to achieve constructive interference we need to be able to manufacture two optical path-lengths which are within a few microns of each other. In order to work out our margin for error we consider the position of one pulse given by:

$$\Delta l = \text{pulse length} \times \text{speed of pulse} \quad (14.11)$$

For a 50 femtosecond pulsed laser this is:

$$\Delta l = 50 \times 10^{-15} \text{ s} \times 3 \times 10^8 \text{ m s}^{-1} = 15 \times 10^{-6} \text{ m} \quad (14.12)$$

This of course represents the limit at which there will be any constructive interference at all. We would in practice need to aim for path lengths with an even smaller difference.

14.2.1.2 Resolution

The improvement on confocal microscopy resolution is shown by the Optical Transfer Function. A brightfield confocal microscope has transfer function for a single object given by:

$$I = |h_1(r, z)h_2(r, z)|^2 \quad (14.13)$$

Where h_1 is the PSF of the illumination and h_2 is the PSF of the detected light. The three different configurations for confocal microscopes have different optical transfer functions.

If we illuminate from both sides but only detect from one side, we have the type A set-up. We consider $h_1(r, z) = h(r, -z) + h(r, z)$ while we take $h_2(r, z) = h(r, \pm z)$ - we are able to use the same function $h(r, z)$ because illuminating and detecting light passes through the same objective. This gives a transfer function:

$$I = |h(r, -z) + h(r, z)|^2 |h(r, \pm z)|^2 \quad (14.14)$$

The type B set-up, with illumination from one side but detection of the constructive interference between scattered light in both directions gives the same overall transfer function as the type A except this time $h_2(r, z) = h(r, -z) + h(r, z)$.

In the type C set-up (14.23) we both illuminate and detect from both sides of the sample, giving the transfer function:

$$I = |h(r, -z) + h(r, +z)|^4 \quad (14.15)$$

Type A microscopes effectively enlarge the illumination aperture, Type B microscopes increase the effective detection aperture while Type C microscopes enlarge both apertures.

In order to obtain the response in the z direction, we must integrate the intensity across r . The theoretical plots are given by [149] in figure 14.22.

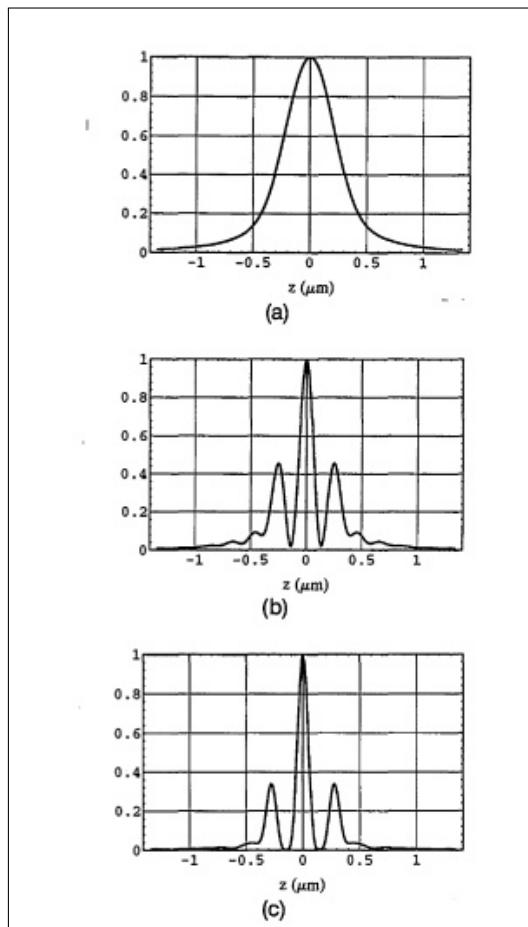


Figure 14.22: [149] a) the z response for a confocal microscope b) the response for a type A 4pi microscope c) the response for a type C 4pi microscope. b) notably sits inside the envelope of a)

Since in reality 4π illumination is not possible we cannot obtain a spherical illumination and detection PSF. As a result we have small sidelobes which must be suppressed in order to access the full benefit of 4-pi microscopy. Typically, we use two photon emission (so that intensity depends on the square of illumination) and a pinhole as in confocal microscopy in order to reduce the sidelobes to an acceptable level. We note from these plots that the FWHM is reduced to around a quarter of the confocal case for type A 4pi and a fifth for the type C microscope. Two photon emission can be used to improve the resolution up to sevenfold on the confocal case (the commonly cited number in literature [150]).

14.2.1.3 Reason for exclusion

There were a few reasons for excluding 4-pi illumination from our system. Firstly, since 4-pi illumination requires us to manufacture 2 optical paths with a difference of less than 15 microns it adds significant complication to the optical and mechanical design. The second and perhaps more important reason is that 4-pi illumination would limit the depth of objects that we would be able to work with. Since the benefit of our system is in being able to take larger objects with microscopic detail manufactured quickly by stereolithography and apply nanoscopic detail with the main system, we

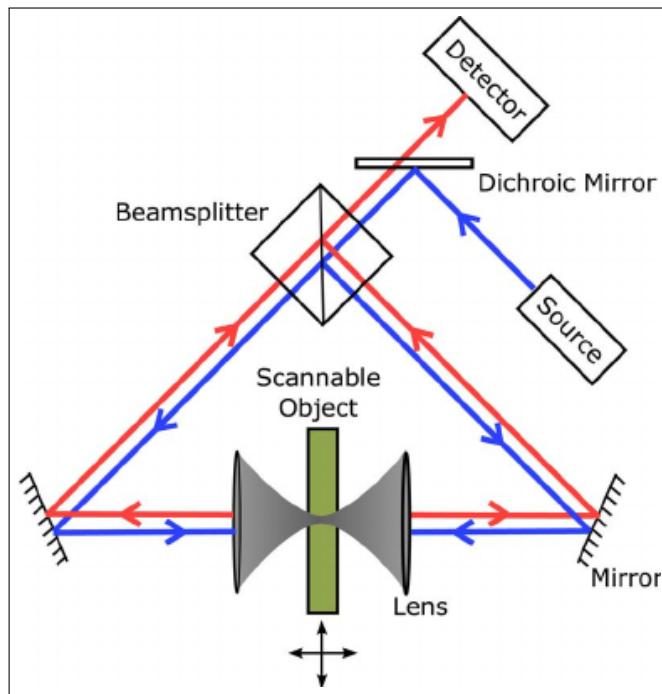


Figure 14.23: [151] a simplified optical layout for a 4pi microscope shown in type C configuration (with both coherent illumination and detection from both sides of the sample)

expect users to manufacture objects with a depth larger than double the working distances of our objectives. A removable 2nd objective for low depth objects was considered but moving parts make it even more difficult to build with 15 micron accuracy. The compatibility with our existing STED and 2PP systems and the significant improvements the setup offers make 4-pi an area to consider for future development of our product.

14.2.2 Structured illumination Microscopy - Xiaonan

Structured illumination Microscopy (SIM) is a super resolution 3D illumination technique. As shown in Figure 14.24, strip pattern is projected onto the sample to produce high frequency variation of sample fluorescence. A lower frequency Moire interference pattern can be captured, which contains information that is otherwise non-resolvable. By taking a series of images at different orientations, the sub-diffraction information could be computed and reconstructed with computational processing. Thus, a super resolution image can be obtained. It is commonly used in imaging cell activities.

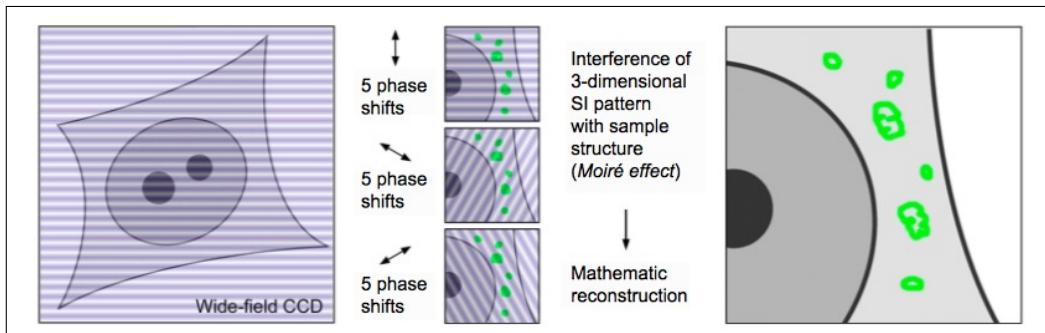


Figure 14.24: Illustration of structured illumination [152]

From literature [153], SIM demonstrates an improved resolution in both lateral and axial direction. With a 1.4 NA oil immersion objective, it can achieve axial resolution up to 100nm and lateral resolution up to 300nm. Also, it allows rapid image collection.

14.2.2.1 Reason for Exclusion

The main benefit of SIM is live imaging ability with improved resolution. If we want to implement SIM as a method of live imaging, the objective for SIM should have a lower NA than that for fabrication process, so SIM process can capture an area of image. At the same time, SIM is implemented in the reflected microscopy set up. Thus, two objectives with different NA need to be on the same side of the sample and focus onto the same area of sample. This is extremely difficult to realize and thus two processes cannot be carried on at the same time.

SIM would improve live imaging resolution, but as a post-processing imaging technique it is not as good as two photon microscopy and STED. Also, the system must be calibrated accurately in order to obtain faithful reconstruction of the image. Hence, the benefits brought can be achieved with our techniques in the system. As a result, we exclude the idea of Structured Illumination Microscopy.

Part IV

Optics

15 Spatial Light Modulators

15.1 Motivation

A Spatial light modulator (SLM) is a device that modulates light according to a fixed spatial (pixel) pattern.[154] A central feature of our facility is the ability to create multiple foci simultaneously during fabrication and hence decrease the time taken by the technique to fabricate a unit. Aberrations developed during beam propagation over time are corrected using adaptive optics, as accuracy is key in the design. Both of these key applications need an element to modulate the light and hence we need a dynamic modulator which was chosen as an SLM. Spatial light modulators allow programmable control of the wavefront of an incident light beam.[155] They can be used to modify the amplitude and phase of a light wave. The requirement in our application is for phase-only modulation. There are 2 SLMs used in our system, one for the implementation of adaptive optics and multiple foci generation and one for pulse shaping discussed in section 20.9 with the rest of the pulse shaping system. They are positioned in the optical layout as shown in Figure 15.1.

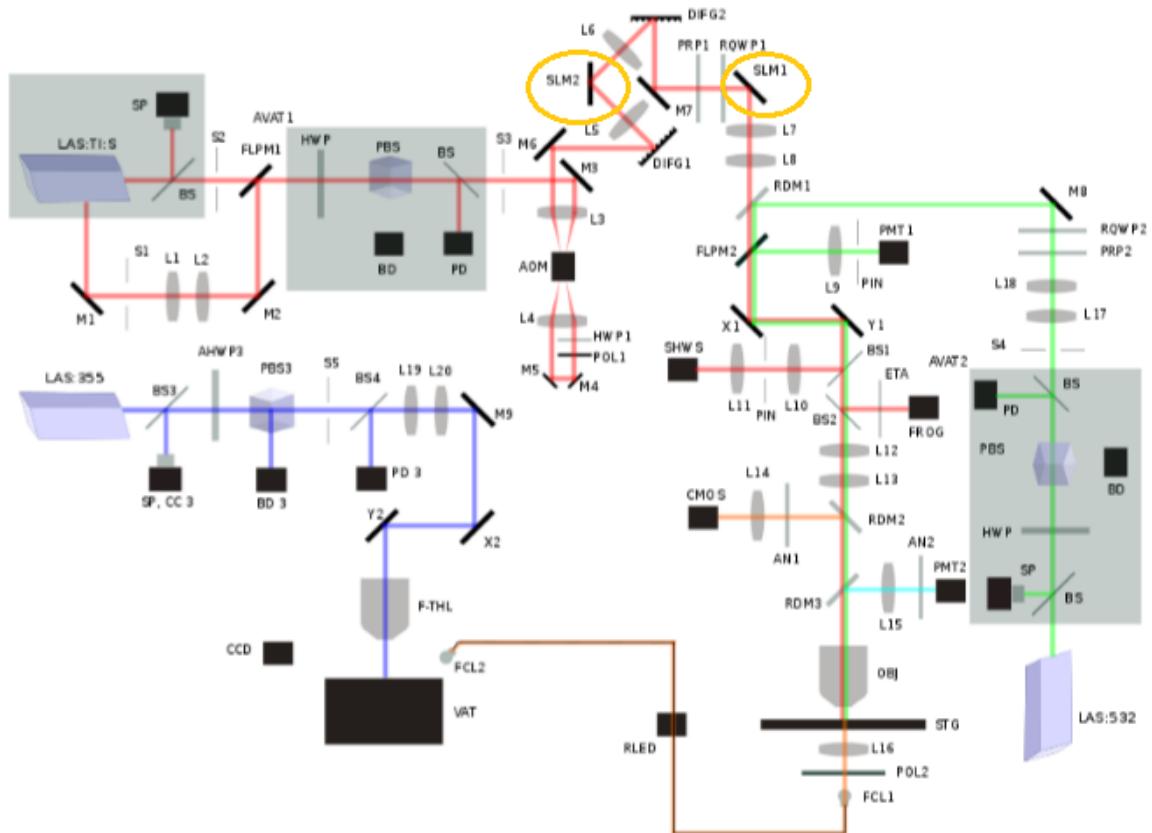


Figure 15.1: Position of SLM in optical layout – SLM1 is used for adaptive optics and parallelisation and SLM2 is used for pulse shaping discussed in section 20.9

Phase-only modulation is achieved by changing the optical path of the beam which can be done either by changing the refractive index or the thickness of the medium for a given wavelength. This would also better accuracy. The implementation of multiple foci generation and adaptive optics re-

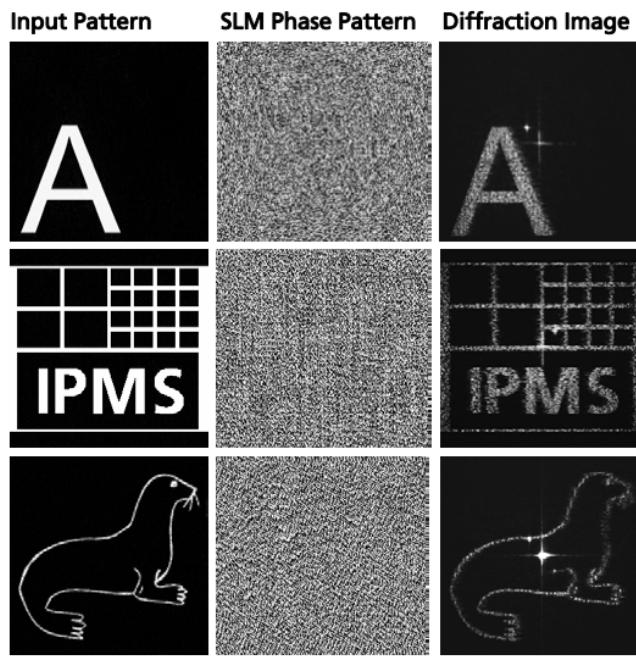


Figure 15.2: SLM phase pattern retrieval by iterative fourier transform algorithm[156]

quire the range of phase modulation to be atleast 2π , however, the higher the better. Similarly, the aim is to reach high refresh rates and response time as well, so as to then increase the overall speed of the process. This should however be balanced with the efficiency of the system and the level of control available.

For parallelisation, defined intensity patterns need to be created at the working area which is done by phase alterations pixel by pixel to create the desired pattern. Spatial light modulators are used in adaptive optics essentially as a phase- conjugate generator. The light containing the information about the image is passed through the SLM. The optical field of the source containing the image information gets distorted due to the various aberrations that have developed on the path. The phase conjugate of the distortions created are first algorithmically calculated and then mapped onto the light beam through the SLM. The use of the SLM can also be extended to holographic optical tweezing to trap objects at various foci simultaneously. This is done by focussing a beam of light and the SLM can then be used to create multiple foci, which allows simultaneous control through electronics and software of multiple points in 3D space. SLM serves the function of a device capable of controlling the phase and amplitude of light passing through the system at a pixel by pixel level without the need to move anything around, being completely electronics based. [157] This makes them extremely easy to integrate, and in the design the laser TiS beam always passes through the SLM.

15.2 Types of SLM

15.2.1 Optically and Electrically addressed SLM

To prevent interference with the already existing laser beams, an electrically addressed modulator would be used instead of an optically addressed one.

15.2.2 MEMs

MEMS(microelectromechanical systems) are miniaturized devices that can control optics and are fabricated on microscopic scale, like on silicon wafer.[158] For the fabrication of MEMs, thin layers are deposited on a base and then selectively etched away on a microscopic level.[159] It consists of both electrical and mechanical parts, where the electrical part processes the data and the mechanical part acts in response to it. They have revolutionized the field by replacing large electrical or mechanical systems with just a small chip. A typical example can be seen in Figure 15.3 DMD are deformable mirror device is a dynamic modulator based on MEMs technology. It consists of microscopic hinged moving mirrors, which corresponds to one pixel in a projected image. They either tilt towards the light source to reflect or away to block, changing the physical path of light and hence the phase. The length of time the mirror faces the light determines the brightness of the dot. [159] MEMS are manufactured by building up on a substrate sequentially and there can be residual stress from the manufacturing process which may change the curvature of the surface.[160]

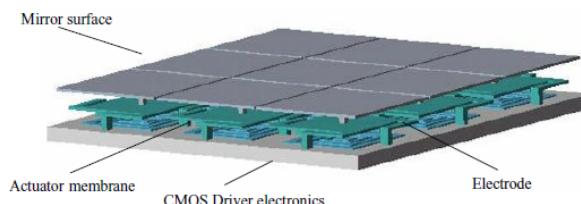


Figure 15.3: MEMs structure [161]

The basic functioning of a MEMs device based on the alternative arrangement of the thin strips of mirror. Half of these are fixed and the other half movable. The movable part can be anywhere between completely out of phase and in phase. The otherwise reflected light, on application of voltage the actuated strips are pulled down by certain length x . This creates a path difference of $2x$ and hence a change in the phase as there is a delay in physical path. The modulator response H can then be given by, [162]

$$H[\phi_1, \phi_2] = \frac{\exp(j\phi_1) + \exp(j\phi_2)}{2} \quad (15.1)$$

15.2.3 Multiple Quantum Well SLM

Multiple Quantum Well (MQW) SLM are a rather niche kind of SLM. Using a solid state based SLM usually poses the problem of having low diffraction efficiency while offering fast response. Liquid crystal based SLM have much higher optical efficiency but at the cost of response time. MQW achieves both high diffraction efficiency and speed although the spatial band width is now limited. MQW structures are made up of thin layers of narrow band gap semiconductor (well) sandwiched between layers of large gap materials. [163] As the name suggests, there is quantization in the movement of electrons and holes in the semiconductors. On application of electric field, the electrons and hole are pulled apart. However, the barrier layer restricts this effect and prevents ionization. This causes change in the energy levels of the electron hole pairs and hence allows the high absorption edge to shift into the spectral region of the sample. [163] They do however pose many problems such as high power dissipation which may lead to temperature changes, high crosstalk and poor contrast ratio. Further, they are not available easily and are very expensive.

15.2.4 Liquid crystal on silicon SLM

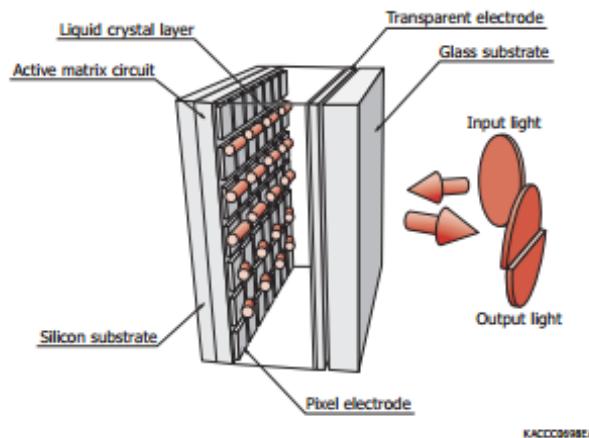


Figure 15.4: LCoS SLM structure[164]

Liquid crystal on silicon (LCoS) based SLM is a phase modulator in reflective mode. It is controlled electronically by varying the voltage passing through the device at pixel by pixel basis. The nematic liquid crystal based SLM, is made of a silicon backplane and aluminium electrodes that make up pixels. The voltage is controlled electronically and independently by the pixels. The liquid crystal is filled in the space between the silicon and glass substrate layer. This can be seen in Figure 15.4. The liquid crystal molecules are originally in parallel to the substrate and on application of electric field there is a tilt in the molecules which leads to a change in the optical path. The polarization and direction of propagation of light hence change the refractive index of the material, a property known as birefringence. This leads to phase delays enabling phase modulation. High light utilization, better

reflectivity and aperture ratio and lower diffraction noise are the key advantages of using an LCoS SLM.[164] Further properties of liquid crystals are discussed in section 15.4.

15.3 Comparison of various SLM types

The required application of phase modulation can be achieved by either changing the distance of physical path travelled in comparable values to the wavelength of light, or by changing the optical path. [157] A change in the physical distance would require movement of the optical elements and even in a completely electronically controlled system, it is very difficult to achieve accuracy. The device would be very sensitive to any small perturbations or errors of calibration. Hence, systems are developed to change the optical properties and hence the optical path which in turn would change the phase of light, rather than physical paths as done by MEMs. This is done by changing the refractive index of the medium the light is travelling through and liquid crystals through their property of voltage controlled birefringence offer a good option.

While MEMs based SLM have an advantage of faster frame rate, bigger spectral range and efficient use of unpolarised light[165], they are quite niche for phase modulation which makes it rather challenging and expensive. The range of phase modulation offered is between 0 and 2π . Further, the resolution is quite low and the pixel size achievable is not as small as in an LC based SLM. The ability to control polarization rotation electrically through birefringence, LC based SLM can modulate phase and intensity both.

Liquid crystals are the preferred choice for SLMs for more reasons. They offer the ability to put a large number of pixels in a small panel keeping the fill factor ratio high due to the reflective arrangement of the silicon backplane. [166] Further, in the past decade there has been considerable research on LCoS SLM leading to many technological advancements. These include but are not limited to, faster response time, better resolution and larger fill factor which increases frequency. It can produce several phase levels to generate high quality phase holograms with low signal to noise ration.[155]

Nematic liquid crystals allow grayscale modulation which means 8 bits would give 256 addressable levels and analogous control, whereas ferroelectric liquid crystals allow binary modulation. Nematic crystals have a lower response time around 20ms, whereas ferroelectric liquid crystals have much higher response times close to 50μ s. Further, nematics have a switchable retarder meaning that one polarisation may see no change while the other would experience a phase shift. Ferroelectric liquid crystals have each pixel behave like a waveplate on which two different orientations can be performed. This leads to the need to spend equal time in each state, failing which over a period of time would lead to the liquid crystal decomposing and hence making the refreshing more com-

plicated. LCOS nematic SLM are better, giving almost analogous phase modulation, as they have higher diffraction efficiency of multilevel hologram structures, FLC (ferroelectric crystal) on the other hand can only achieve binary holograms leading to basically a loss of 50% light in system in addition to 50% duty cycle in FLC. [165] The reason for this would be the appearance of a conjugate image in ferroelectric liquid crystals along with the intended. This happens due the fact that on reconstruction the complex conjugate of the hologram is similar to itself. Phase, amplitude or modulation in both phase and amplitude together is achieved by different configurations dependent on the alignment of molecules and the polarization of light in a particular direction. Using parallel arrangement as shown in Figure 15.5 is the most common and can be thought of as a ‘voltage controlled waveplate’[166]. Further, it allows a faster response and greater range of phase that can be modulated. It is also gives a higher intensity transfer, a better efficiency.

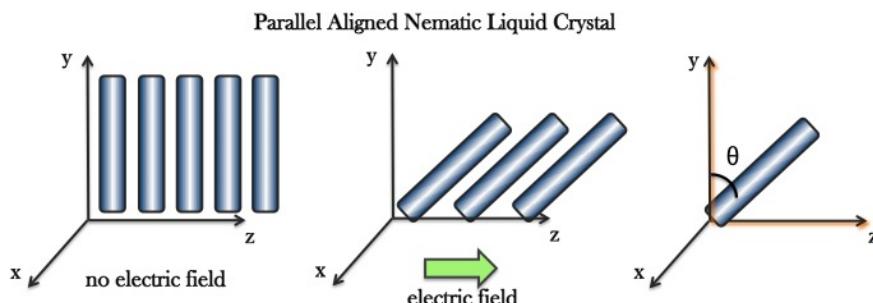


Figure 15.5: Parallel aligned nematic liquid crystal [167]

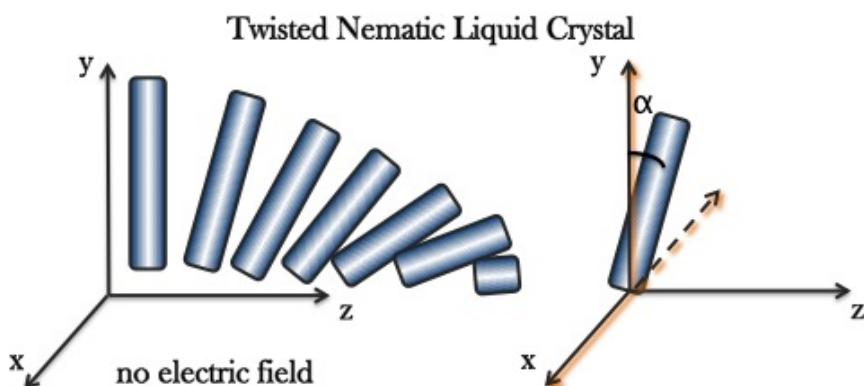


Figure 15.6: Twisted nematic liquid crystal [167]

Reflective SLM is better than a transmissive one for the simple reason that light will pass twice through the LC and this would mean better modulation range for the beam. Further, the fill factor is much higher, which increases the diffraction efficiency. For example, HOLOEYE has a transmissive SLM with 63% fill factor while one with reflective has fill factor at 93%. [168] Further, using a reflective LCoS SLM allows higher number of pixels in a small sized panel, increasing resolution and keeping fill factor high. Hence, an electrically addressed, reflective nematic LCoS SLM would best serve our

purpose.

An improvement can be that the back surface of the SLM can be covered with dielectric mirrors so that the pixelated structure is not visible and reflectivity of the SLM and light utilization is increased. However, this decreases the spectral range and needs higher voltage application due to increased crosstalk in adjacent pixels. [165]

15.4 Liquid crystals

Liquid crystals are in a state between liquid and solid state and hence they possess properties that are intermediate between both. They are fluid-like but still possess structural order in arrangement of molecules. Depending on how the phase is induced, there are two main types of liquid crystals: thermotropic and lyotropic. As the name suggests, thermotropic liquid crystals have temperature-dependent properties and can be further divided into nematic, cholesteric and smectic phase.[157] The nematic phase of the liquid crystal possesses long range orientational order, meaning that the long (extraordinary) axis of the molecules generally align in the same direction. However, nematics do not have positional order. The axis to which the molecules align is called the director which is represented as a unit vector. The molecules can rotate about the long axes and the orientation of the molecules are the same in either directions. A nematic optically behaves like a uniaxial material with symmetry.[169] Cholesterics are a phase of liquid crystals, similar to nematics in that they have long range orientational order and no long range positional order, with the difference that the direction of the director varies along the material. Twisting a nematic phase liquid crystal along the x axis in a medium where the director is aligned to the y axis would result in a cholesteric phase. There is periodicity in the way the alignment of the axis varies along the medium. The distance traversed by the twist axis over which the director rotates through a full circle is called pitch, hence nematics can be thought of as cholesterics with infinite pitch. [169] Smectics have further even positional order in addition to orientational order and are characterised by their stratification. Ferroelectrics are another type of liquid crystal, they are electrically polarized even in the absence of an external field and hence show instant switching on application of an electric field.[170] All these states are depicted in Figure 15.7.

Key properties of liquid crystals are[169]: 1. Anisotropy, which means to have a physical property that has a different value when measured in different directions [oxford dictionary] This becomes key by enabling birefringence discussed later

2. Long axis of the molecule has a rigid backbone with double bonds

3. They are easily polarizable due to the presence of strong dipoles

The direction of the molecular axis determines the orientation of the LC. This in turn is determined

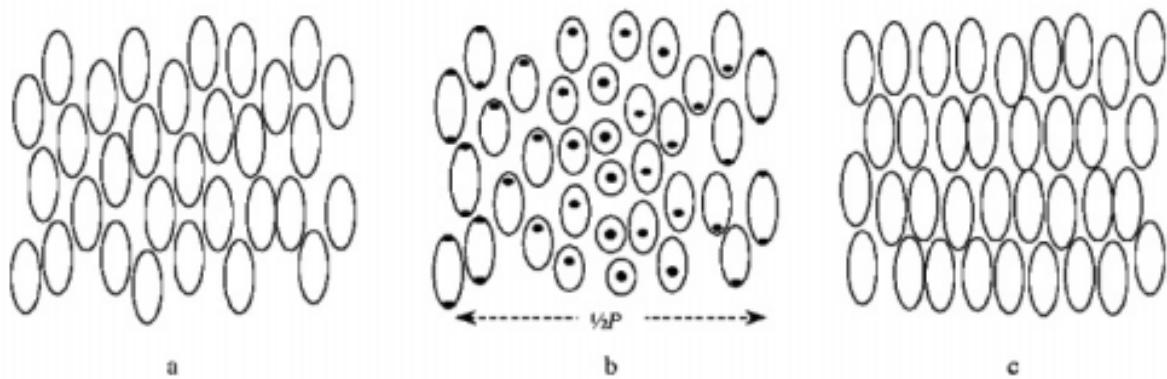


Figure 15.7: Liquid crystal arrangement of molecules a. Nematic b. Cholesteric c. Smectic [169]

by the alignment of the LC layers. There are three main alignments: parallel, vertical and twisted. In twisted cells the orientation between the top and bottom layers varies by 90 degrees, creating a helix-like structure. Parallel aligned and vertical aligned have the LC layers aligned in parallel, meaning that the molecules have the same orientation[168] Twisted alignment is used during amplitude modulation as it creates the required polarization effect while phase modulation is usually done in parallel. To modulate the amplitude of an incoming beam (which is polarized) it is passed through the LC and immediately after by a polarizer at 90 deg angle to the incident polarization.[165] A vertical or parallel aligned LC could also be used at angle of 45 degree. The attenuation provided by the polarizer is controlled by changing the birefringence which is controlled by the voltage. Phase modulation is achieved through a parallel aligned cell by linearly polarizing the parallel to the extraordinary axis of the liquid crystal. The birefringence (property that the refractive index can be changed by polarization) of the liquid crystal creates retardance and hence a phase change takes place. Twisted arrangement could have been used too but a polarisation effect on the light beam causes changes in intensity which is not required in pure phase modulation and may even interfere by causing discrepancies in expected and actual result of modulation. Hence parallel aligned liquid crystals are used for phase only modulation. Further, the Jones matrix of a twist alignment is quite complex to solve and requires acquiring optical measurements which are error prone and further add more variables to the system. Jones calculus is used to understand polarised light. Quite simply, taking the product of the Jones matrix of the optical element (which is the SLM) and Jones vector which denotes the polarized light, the emerging light's polarization can be calculated. The Jones matrix is used to measure the effect of an LC cell on a polarized light and for twist alignment is given by, [15]

$$W_{TN-LC}(f, g, h, j) = \exp(-i\phi) \begin{pmatrix} f - i.g & h - i.j \\ -h - i.j & f + i.g \end{pmatrix} \quad (15.2)$$

$$f^2 + g^2 + h^2 + j^2 = 1 \quad (15.3)$$

$$f = \cos\gamma \cos\alpha + \frac{\alpha}{\gamma} \sin\gamma \sin\alpha \quad (15.4)$$

$$h = \cos\gamma \sin\alpha - \frac{\alpha}{\gamma} \sin\gamma \cos\alpha \quad (15.5)$$

$$g = \frac{\beta}{\gamma} \sin\gamma \cos(2\psi - \alpha) \quad (15.6)$$

$$j = \frac{\beta}{\gamma} \sin\gamma \sin(2\psi - \alpha) \quad (15.7)$$

$$\gamma = \sqrt[2]{\alpha^2 + \beta^2} \quad (15.8)$$

Here, α is the twist angle, ψ is the front director orientation, the parameters f, g, h, j are determined from the above relations. For a PA or VA cell the Jones matrix is given by, [171]

$$W_{PN-LC}(f, g, h, j) = \exp(-i\phi) \begin{pmatrix} \exp(-i\beta) & 0 \\ 0 & \exp(i\beta) \end{pmatrix} \quad (15.9)$$

$$\beta = (n_{eo} - n_o) \frac{\pi d}{\lambda} \quad (15.10)$$

$$\phi = (n_{eo} + n_o) \frac{\pi d}{\lambda} \quad (15.11)$$

Here, β is the birefringence, ϕ is the phase offset, λ is the wavelength of the field, d is the thickness of the cell, n_e and n_o are the refractive index in the extraordinary and the ordinary direction respectively.

15.5 How it works

LCoS SLM consists of an array of pixels and a liquid crystal layer that is placed between two electrodes that control the alignment of the molecules of the liquid crystal. One of the electrodes is transparent while the other is silicon in reflective arrangement. This backplane is used for optical interaction with the aluminium pixels individually, where the aluminium acts as a mirror and an electrode.[155] The voltage applied changes the refractive index by reorienting the molecules resulting in variable phase modulation. The initial alignment of the nematic molecules are set due to a thin rubbed polymer layer. [154] The complex field of the beam propagating through the medium can be given by, [155]

$$U = a \exp(i\phi) \quad (15.12)$$

Application of the phase change by the hologram turns the field into

$$U = a \exp(i(\phi + \beta)) \quad (15.13)$$

One of the key properties of Liquid crystals was seen to be the ability to change the refractive index experienced by the light electronically. The refractive index depends on the angle between the molecular orientation of the liquid crystal and the direction of propagation of the electric field of the light passing through. The optic axis defines the two axes of the molecular arrangement, the ordinary and extraordinary. For a planar aligned nematic liquid crystal though, the optic axis and the director are co-aligned. The extraordinary axis is the one which has polarization in direction of optical axis, similarly, the ordinary axis is in perpendicular. As the ray propagates through the SLM it experiences a phase delay due to changes in optical path caused by the changes in refractive index, which is electronically controlled by the voltage applied at the backplane. The group velocity v_p and refractive index η are related by,

$$\eta = \frac{c}{v_p} \quad (15.14)$$

As a result, light travels slower in the higher refractive index regions of the liquid crystal inducing a phase delay. This phase delay is given by,

$$\phi_e = 2\pi \frac{d_m(n_e - n_o)}{\lambda} \quad (15.15)$$

where ϕ is the delay, d is the thickness of the LC layer of the SLM lying usually within a few microns, λ is the wavelength of light passing through and n_e and n_o are the refractive index in the extraordinary and the ordinary directions respectively. The complex field of the beam propagating through the

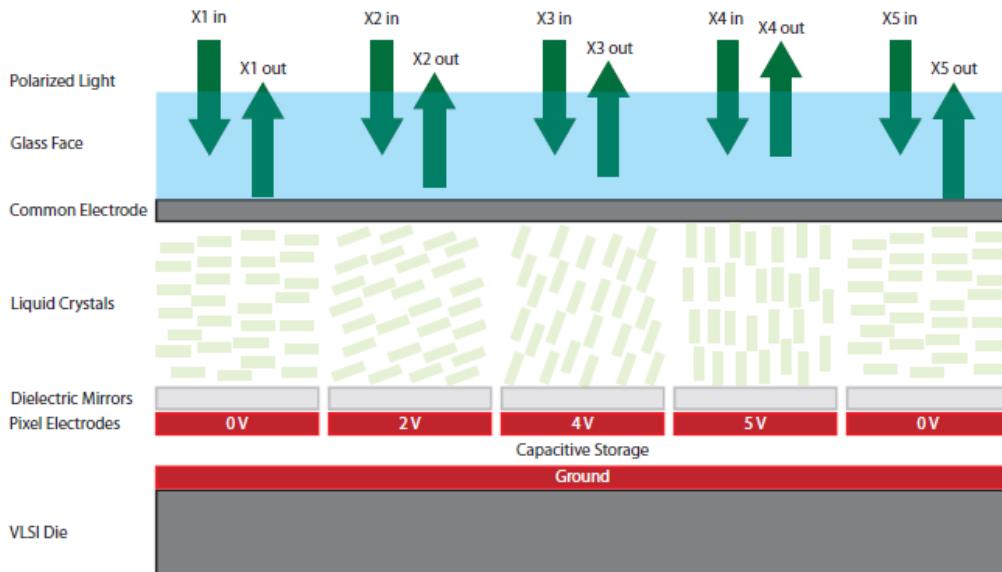


Figure 15.8: Effect of applying external field [155]

medium can be given by, [155]

$$U = a \exp(i\phi) \quad (15.16)$$

Application of the phase change by the hologram turns the field into

$$U = a \exp(i(\phi + \beta)) \quad (15.17)$$

There are many methods to determine the value of β according to the requirement of the beam after passing through the SLM. One could be using the direct binary search algorithm, which involves minimising the error function over iterations at every pixel. The error function being the difference between the recorded value of the field and the reference value. This method, however, is understandably time consuming and iterating over each pixel would make it computationally inefficient. The commonly used method the Gerchberg-Saxton algorithm which derives the phase of the hologram according to the intensity needed at the focal plane, based again on principles of fourier transforms. Iterations of measured values and aimed values of intensities is done in the fourier domain, with constraints both in the fourier and object domain which leads to convergence in solution. [172] This has been explained in detail in section 19.4.

15.6 Key parameters that influence SLM choice

There were certain requirements from the SLM and making a choice needed several design decisions based on constraints. Firstly, the resolution required was to be at least in a few hundred nanometres so as to be able to produce complex beam profiling and be able to correct aberrations at all levels. Secondly, accuracy achieved by the SLM is key as it is used in both Laser micromachining and 2 photon polymerization. The scale of manufacturing goes as low as the nanoscale and hence at that stage the accuracy of control is crucial. Thirdly, the only laser beam passing through is an 800nm TiS and hence the SLM should be able to function at that wavelength. Finally, using a pulsed laser gives the added possibility of much phase flicker so the SLM should be capable of minimising the effects of that. Having discussed the advantages of using a liquid crystal on silicon (nematic), the key properties that affect the performance of the SLM on various parameters were looked at. Firstly, the number of pixels determines the resolution of the SLM and in our application this becomes key for accuracy as the smallest size of production goes down to a few hundreds nanometres. Pixel comes from ‘picture element’ and refers to the smallest basic unit that is programmable on a display while resolution is the number of pixels contained in the monitor[166], with the number of pixels on horizontal axis and vertical axes as height and width of display monitor. Hence the screen is like a matrix of pixels each of which control the voltage and hence the phase difference. Having a higher number of pixels would hence lead to a higher resolution and is desirable for the system. Diffraction efficiency is higher in phase-only modulation rather than amplitude modulation and for bigger values

of fill factor.[173] Diffraction efficiency can be defined as the ratio of power sent to designated direction over the power that was actually incident,

$$\eta = \frac{P}{P_o} \quad (15.18)$$

Better efficiency is achieved by an analogous phase modulator than a binary phase modulator. This is because of the 50% duty cycle, as the drive requirements split the energy in two, the true image and the inverse image of the half cycle.[7] These two different images are created on reconstruction of a binary hologram, because the representation of the original and its complex conjugate is similar.

Having a higher fill factor would prevent high order diffraction. Slight overfill ensures the beam is incident upon the entire SLM display.[155] Phase noise and flicker can be a rather dominating problem as we are working in fourier plane and in short time scales. This happens because fluctuations during the refresh of the system make fourier transforms of higher orders as well. A reflective SLM would be preferred over a transmissive one as the beam would pass through the device twice, leading to higher efficiency and better isolation of the input and output beam. However, using a reflective SLM, there is a chance that the beam gets reflected straight back to its original path so it should be placed at a slight angle to the laser path. [155]

15.7 Choice of SLM

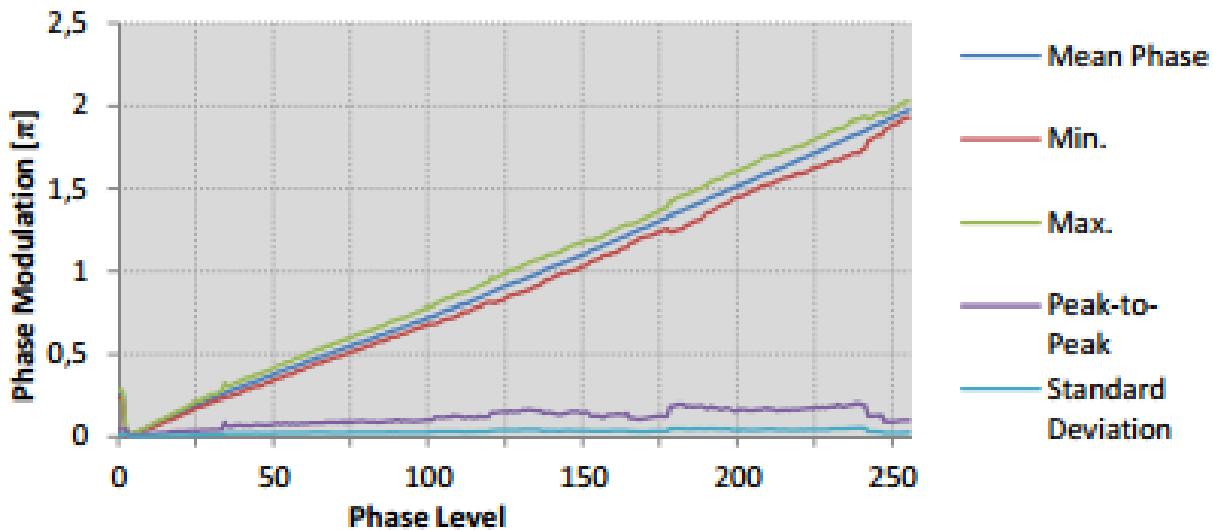


Figure 15.9: HOLOEYE PLUTO NIR-011 phase flicker [168]

Constraints to the choice of the SLM would be that it is operated at 800nm by a pulsed TiS laser and is done in phase-only mode, as required for adaptive optics and multiple foci generation (explained in section 19.4 and chapter 16). As established before it would be best to use a nematic liquid crystal based SLM. Further, an SLM operating in reflective mode is better than one in transmissive mode as the beam passes through the liquid crystal twice and results in more accurate control

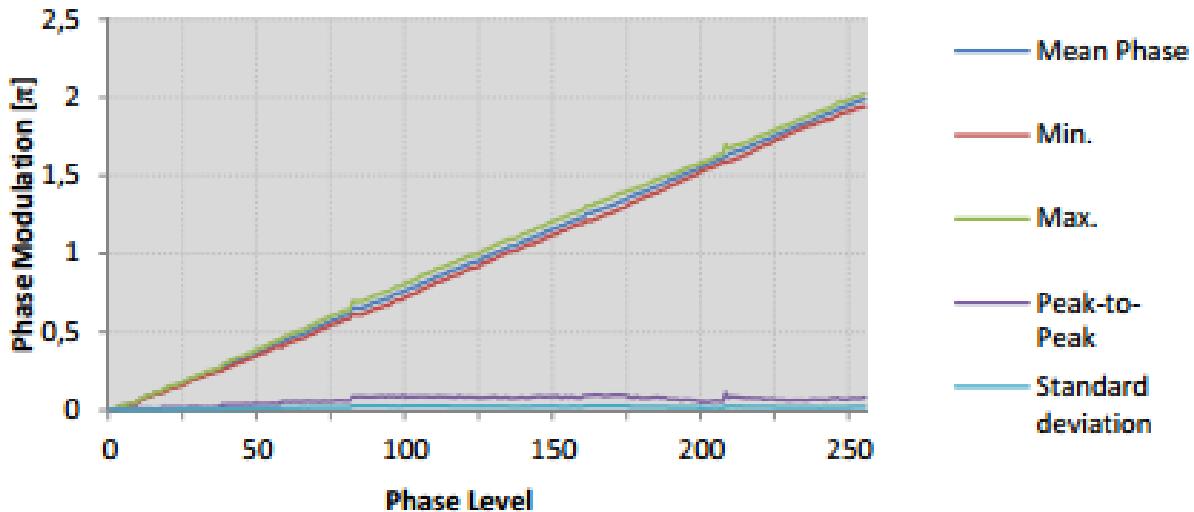


Figure 15.10: HOLOEYE PLUTO NIR-015 phase flicker [168]

of the beam. The SLM would be addressed electrically. HOLOEYE has 2 options for reflective SLM operating in phase-only at 800nm: PLUTO-NIR-11 and PLUTO NIR-015.

Pixel pitch is the centre to centre spacing between adjacent pixels. Inter-pixel gap describes the edge-to-edge spacing between adjacent pixels. [174] To keep adjacent cross talk low it is better to have low pixel pitch. Both the SLM have the same quite low pitch of $8\mu\text{m}$. They both have a fill factor of 93% which is quite high and maintains efficiency. The NIR-015 has slightly better reflectivity as compared to NIR-011. However, the key difference as also visible in the graphs in Figure 15.9 and Figure 15.10, is in the phase flicker which is also crucial for accuracy. Having a pulsed laser in the system, the phase flicker becomes an issue and PLUTO NIR-015 resolves this by providing the highest flexibility in terms of phase stability and maximum phase shift. Phase flicker is a temporal phase fluctuation. In PLUTO NIR-015 the error in the phase level would definitely be within 5% and is hence much better than PLUTO NIR-011. A comparison of experimental measurement for phase flicker done by HOLOEYE at 633nm can be seen in Figure 15.11.

Panel version	Max phase @ 633nm	Flicker Std. Deviation 2π @ 633nm	Flicker Peak-to-Peak 2π @ 633nm
PLUTO-NIR-011	2π	0.04π	11 %
PLUTO-NIR-015	6.2π	0.01π	3.5 %

Figure 15.11: Phase flicker measured at 633 nm for PLUTO NIR-011 and NIR-015 [168]

A further multi criterion analysis with HAMAMATSU reveals HOLOEYE to be a better choice. As seen in Figure 15.13 the phase flicker results of HAMAMATSU x10468-02 are also slightly worse off than that of HOLOEYE NIR-015 in Figure 15.10

	-	HOLOEYE PLUTO NIR-015			HAMAMATSU x10468-02		
Factor	Weight	Value	Points	Score	Value	Points	Score
Number of pixels	4	1920x1080	4	16	792x600	2	8
Pixel pitch	3	8 μm	4	12	20 μm	2	6
Fill factor	3	93%	4	12	98%	5	15
Area	2	15.36x 8.64 mm	4	8	15.8 x 12 mm	3	6
Frame rate	4	60Hz	3	12	60Hz	3	12
Maximum phase shift	3	5 π	4	12	>2 π	3	9
Cost	2	£10,646	4	8	£14,000	3	6
Total	-	-	-	80	-	-	62

Figure 15.12: SLM Multi criterion analysis

The chosen HOLOEYE PLUTO NIR-015 has 1920x1080 pixels and an active area of 15.36x8.64 mm. The SLM is controlled by an 8 bit addressing system meaning it has 256 addressable gray levels (2^n , where n is number of bits). The maximum retardation is 4.4π rad and hence the maximum phase precision is $\frac{4.4\pi}{256} \approx 0.053996$ rad. HOLOEYE provides a software for the control of the SLM and it will be connected to a PC for control. Further in section 19.4 and chapter 16 much development has been done to implement adaptive optics and parallelisation. The SLM as visualised in Figure 15.14 would be placed in system through the mechanical design as in Figure 15.15 which is explained in chapter 21 and is customised.

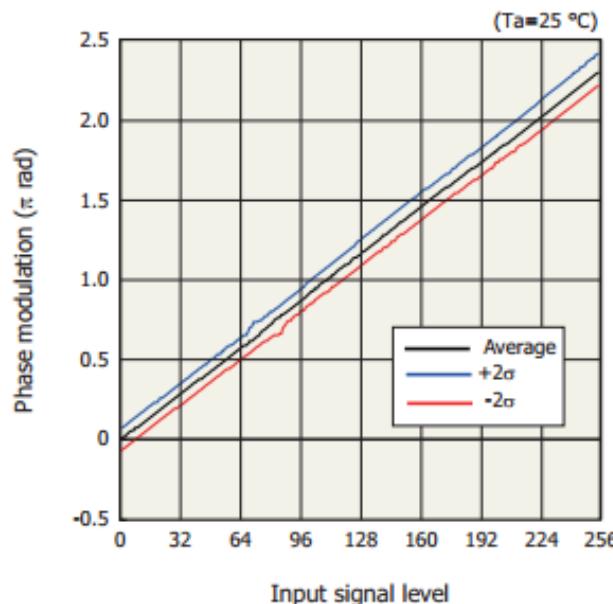


Figure 15.13: Hamamatsu phase flicker [164]

Email correspondence and their webpage reveal that HOLOEYE is planning to release another



Figure 15.14: HOLOEYE PLUTO SLM [168]

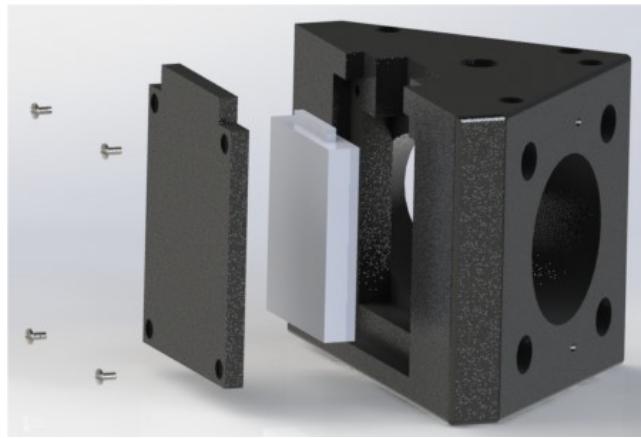


Figure 15.15: Corner mount for SLM [156]

SLM fulfilling all our criteria in the second half of 2016. It would be even better than our current chosen SLM as they plan to use a dielectric coating which will minimize absorption and maintain organic alignment layers. It is deemed to be sold under the product category of 'GAEA' by HOLOEYE.

15.8 Power Considerations

For high laser power applications, the behaviour of SLM with respect to temperature is important and is defined by the viscosity of the liquid crystal. Viscosity is a measure of the friction offered by fluid. Higher temperatures usually decrease the response time and increase phase modulation at the cost of increasing temporal noise. Hence, a safe distance from the damage threshold should be maintained, or else the organic alignment of the liquid crystal is harmed which in turn affects the properties of the liquid crystal.

Section 20 of the report notes that the maximum power leaving the laser would be 0.15mJ. Although there will be power losses along the system and the power reaching the SLM would be even lower, assuming a safety factor we can take it to be 0.15mJ. The width of the SLM is 15.36 mm and the height is 8.64mm. Given that the average repetition rate is 10kHz, the fluence should be

11.3nJ/cmsqr which is much less than the damage threshold of 50nJ/cmsqr. Hence the SLM is safe from any damage due to temperature changes or high peak powers of the laser.

15.9 Conclusion

The spatial light modulator is a key component of the system which enables 3 key techniques applied in the design. These are pulse shaping, parallelization and adaptive optics. They really set our system apart, increasing both accuracy and speed of fabrication and imaging. Our system uses a nematic liquid crystal on silicon SLM for its high optical efficiency, ease of integration and control. This decision was made after a comparison with other technologies. Further, the system needs are for phase-only modulation that is achieved by causing a change in the optical path of the light beam. This happens due to the change in the refractive index, caused by the change in voltage which is controlled in pixel by pixel basis. This provides the ability to create much more complex beams and have more robust control.

16 Parallelisation

The parallelisation capacity of the optical facility refers to its ability to perform a single function at multiple locations in the working area simultaneously. This acts to optimise the time taken for a number of fabrication processes and can also provide completely new capabilities to other processes within the facility. Some light control techniques are necessary to achieve parallelisation, but using Fourier Optics theory (introduced in Section 3.2) and holography, it has been demonstrated that it is possible to fabricate at multiple locations in [175], [176] and [177].

The theory and use of holography is explained in more detail in 16.2, but using Diffractive Optical Elements (DOEs) to create holograms provides the ability to define light intensity distributions in the working area. The most important profile for our implementation of parallelisation is an array of points, which models a laser focus at each individual point. By carefully managing laser power it can be ensured that necessary intensities are reached in each of these focus points, enabling work to be done at all foci concurrently rather than at a single focus. Other intensity distributions could be applied to the working area for various applications. For example, it could be interesting to see the effect of 'changing the drill-bit' and applying different shapes of foci to the fabrication processes to observe their effectiveness. Other applications like this aren't investigated in this report, however.

16.1 Applications

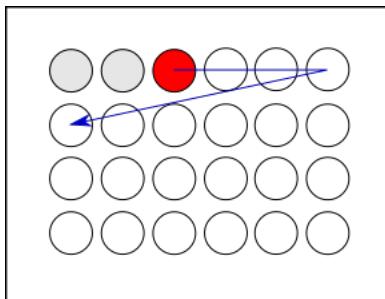
Parallelisation using holography simply involves distributing power across the working area in defined distributions, so the implementation would be compatible with the majority of the optical system's fabrication processes; Two-Photon Polymerisation (TPP), Laser Micro-machining (LMM), Stereolithography (SLA) and Optical Tweezing (OT). TPP and LMM fabrication speeds are slow in comparison to SLA due to relative scanning speeds in addition to the difference in resolution (TPP and LMM finest resolutions are two orders of magnitude smaller than that of SLA). Because of this, it was decided that a subsystem should be designed for the parallelisation of the TPP and LMM fabrication processes since they contribute most significantly to fabrication time. In addition, OT increases in usefulness with parallelisation, discussed in Section 16.1.2.

16.1.1 Two-Photon Polymerisation and Laser Micro-machining

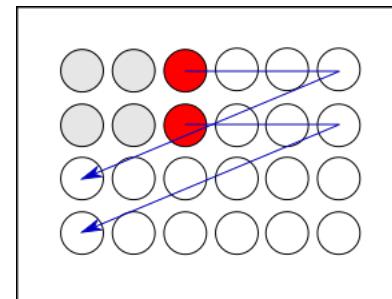
Both Two-Photon Polymerisation and Laser Micro-machining can be observed and treated as similar processes from the perspective of parallelisation, because they are successful in fabrication when intensity is within a threshold defined by the working material.

As described above, the main application of parallelisation is to be able to improve the speed of fabrication processes. This can be achieved in a number of ways depending on the requirements. With careful design of holograms, holographic foci can be used statically to create lattice structures

[177], updated periodically to fabricate defined 3D shapes, raster-scanned using galvanometers to increase the volume of fabrication at each pass, or parallel processed in the classical sense; scanned as in the case of a single focus, creating multiple copies of a single fabricated structure. Each



(a) Raster scan with a single focus



(b) Raster scan with two foci

Figure 16.1: Comparison of single- and multiple-focus raster scanning (grey points are already scanned, red are currently scanning and white are yet to be scanned)

of these modes present interesting control problems, and individual control schemes would have to be designed for each mode. Demonstration of the generation of a tool path and control of a single scanned focus is found in Section 24.6. This method is compatible with the classical parallel processing mode if care is taken to ensure that foci are sufficiently spaced with reference to the size of the fabricated features to avoid overlap. From Figure 16.1 it is clear to see that with each horizontal pass more area is scanned in the parallel case. By managing power to ensure fabrication is possible at each scanning point, this increases the speed of fabrication by a factor of n , the number of foci. Of course, there are shapes that cannot be fabricated at the same scale when using the raster scanning method with multiple foci (e.g. curves and non-right-angled vertices), but for rapid fabrication of simple structures like cuboids this method could be useful. Figure 16.2 shows the mechanism of the classical parallel fabrication technique, where a number of foci trace out the same shape in different locations within the working area. By controlling one of the focus points, the others are also controlled.

16.1.2 Holographic Optical Tweezing

Optical Tweezing can also benefit from parallelisation. The theory and physics of Optical Tweezing are explained in Section 9, but simply, the term refers to the mechanism of holding and manipulating the position of small particles with the use of lasers. As discussed in Section 9.3, this optical facility is designed to have the ability to optically tweeze objects between 30 nm and 12 μm , giving the possibility to hold, move or rotate individual features of fabricated parts and/or full fabrications.

Holographic Optical Tweezing (HOT) is a well-established technique to create multiple coexisting optical traps in three dimensions to hold a number of nano and micro-scale objects at the same time [80]. This greatly increases the capabilities of the Optical Tweezing since it allows controlled objects

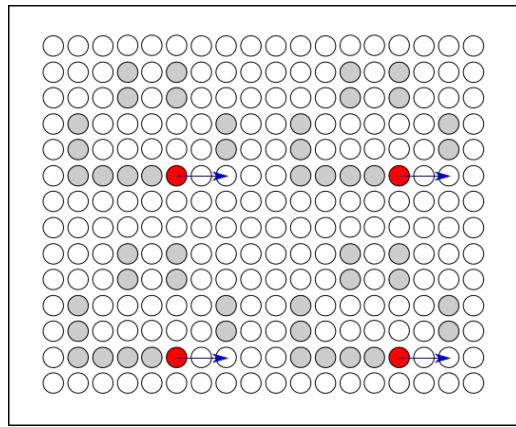


Figure 16.2: Classical parallel fabrication with four focus points (grey points are already scanned, red are currently scanning and white are yet to be scanned)

to interact with one another, adding to the possibilities of manipulating the fabricated products or aiding fabrication.

The intensity profiles (and thus holograms) necessary for HOT are similar to those required in the parallel fabrication techniques; arrays of focus points form the optical traps. However, with HOT it is unlikely that we would raster scan the foci, instead the usefulness of the technique is realised when the locations of the optical traps are changed relative to each other for interaction. To enable this, holograms must be updated in frames to define the movement of each optical trap. This requires fast updating of holograms, which is achieved in this system by Computer Generated Holograms (CGHs). Some important theory of Computer Generated Holography is introduced in Section 16.2.2 and generation of CGHs is explored and demonstrated in Section 16.4.

16.2 Holography theory

A hologram is a total recording and reconstruction of a defined field of light at a target location. A *total* recording refers to all information being known about the light field; its amplitude and phase at all points. We can generate holograms that are useful for parallelisation as described above.

Easily available and widely used recording media for light reacts to intensity only, so information about amplitude (since intensity is proportional to the square of amplitude) can be gained but not phase. By summing a mutually coherent reference wavefront of designed amplitude and phase to the incoming wavefront at the recording plane of the hologram, we can infer information about the phase difference of the wavefronts and therefore gain information about the phase of the incoming wave. A more thorough mathematical explanation of this can be found in [178], but is not particularly useful here. Reconstruction of the hologram can then be achieved by shining light through (or reflecting from) the exposed medium.

There are many suitable photographic media for recording holograms, but this traditional method

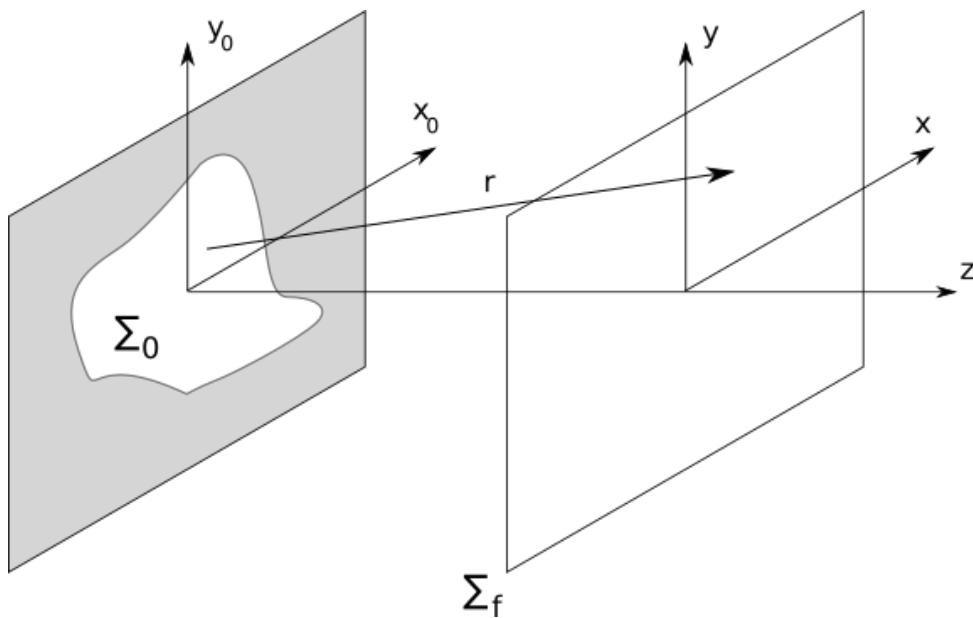


Figure 16.3: Geometry for diffraction expressions, Σ_0 is the aperture surface in the DOE image plane and Σ_f is the hologram reconstruction (HR) plane.

of recording by interference is clearly limited. Firstly, an individual recording is necessary for each hologram, which limits the speed that holograms can be changed by necessitating some physical mechanism for switching between holograms. This also requires numerous exposures of photographic media, increasing the time and resources used. Secondly, the light fields we can record are limited to light from objects that are physically available to be recorded (i.e. they exist and are accessible). These limitations can be overcome if we record and reconstruct holograms electronically, but this presents other challenges in both mathematical definition of holograms, computation and display of holograms.

16.2.1 Fourier Optics Revisited

By defining a hologram mathematically we can create a DOE to act as the recorded hologram, without needing to actually record a light field. Fourier Optics enables an understanding of how light will be diffracted through different apertures and how the light field will appear at varying distances. Considering the geometry in Figure 16.3, where DOEs are modelled as an aperture (Σ_0) in the image plane, $z = 0$, and the hologram reconstruction (HR) plane is at some distance z , the complex fields at the two planes are related by the Fresnel-Kirchoff formula (from [178]), Equation 3.5. This is valid for $r \gg \lambda$, but further approximations can be made to simplify the expression into a more usable form. In doing so, the two most common hologram types; Fresnel and Fraunhofer, are formed

These approximations are described in Section 3.2.2. The first approximation is the Fresnel approximation, leading to the Fresnel diffraction formula in Equation 3.6, which is valid in the near-field to the aperture. Fresnel Holograms are then simply the reconstruction of the light field due to

diffraction through a DOE at a plane where the Fresnel approximation is valid. The complex amplitude of a field within this region is said to be the Fresnel Transform of the complex amplitude immediately after the aperture.

The alternative approximation, the Fraunhofer approximation, leads to the Fraunhofer diffraction formula in Equation 3.9 which is valid in the far-field from the aperture. Fraunhofer Holograms, analogous to Fresnel Holograms, are light field reconstructions where the Fraunhofer approximation is valid.

In both hologram types it is clear that by changing the complex amplitude, $U(x_0, y_0)$, at Σ_0 , the complex amplitude, $U(x, y)$, at the HR plane Σ_f can be changed.

16.2.2 Computer Generated Holography

Computer Generated Holography creates the opportunity to digitally define and generate complex fields at the working area, without the need for photographic media or physical objects for recording. Instead, a desired complex amplitude $U_d(x, y)$ can be defined in the focal plane Σ_f , then the Fresnel or Fraunhofer diffraction formulae can be used to relate this to a complex amplitude $U_0(x_0, y_0)$ in the DOE image plane Σ_0 ; then it simply needs to be physically ensured that the laser has this complex amplitude at Σ_0 by modulating its amplitude and phase.

However, although the two planes are related by two-dimensional Fourier transforms, the calculation of U_0 is not as straightforward as an inverse Fourier transform. This is because there are no physical recording media known that can modulate both amplitude and phase of light independently; so even if the ideal U_0 could be calculated, we would not necessarily be able physically modulate this on to the laser. Phase-only modulation is preferable to amplitude-only modulation since modulation of the amplitude necessarily reduces the energy of the laser pulse reaching the sample. Pixelated phase modulators (discussed in Section 15) can perform the modulation of a laser beam to a calculated phase distribution which, under a Discrete Fourier Transform (due to pixellated modulator), represents the complex amplitude U_d with minimal error. The calculation of the phase distribution is discussed in Section 16.4.

16.3 Optical Layout of a holographic setup

Figure 16.4 shows the beam path for the Ti:Sapphire laser fabrication processes; Laser Micro-Machining (LMM), Two-Photon Polymerisation (TPP) and Optical Tweezing. These techniques share a beam path since they require the same components in order to parallelise.

Fresnel Holograms are not realistically applicable in our optical facility since we still require some beam processing between our DOE (denoted **SLM1** in Figure 16.4) and our working area; most significantly the microscopic objective, which makes the near-field approximation invalid.

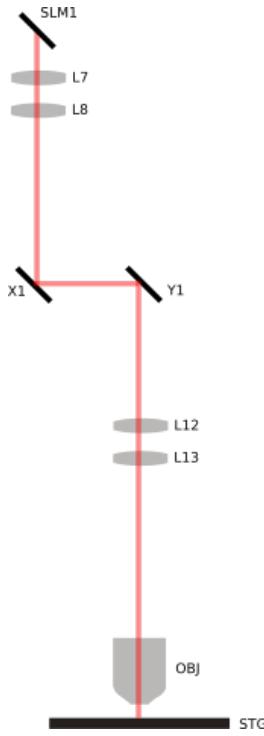


Figure 16.4: Optical path for the parallel techniques

Fraunhofer holograms can be utilised in our optical facility by bringing the far-field to a workable location; the far-field is brought forward to the focal plane when light travels through a converging lens. Analysis in [178] shows how the field at the focal plane of the converging lens can be described simply by Equation 3.9, with $z = f$ where f is the focal length of the lens. Therefore, if the DOE is imaged onto the microscopic objective, the Fraunhofer diffraction pattern forms the Fraunhofer Hologram in the working area. A simple layout of this is shown in Figure 16.5. Since the intensity

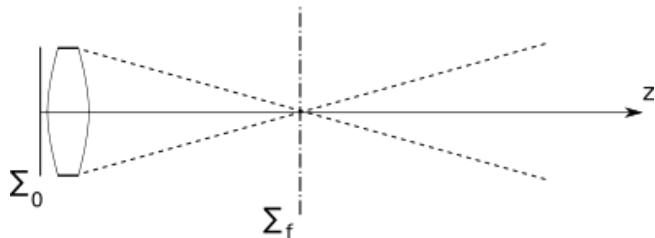


Figure 16.5: Side-view of simple setup for reconstructing Fraunhofer holograms with a converging lens, Σ_0 is the imaged DOE plane immediately before the converging lens and Σ_f is the focal plane of the lens.

distribution of our hologram field is the only important property for all of our parallelisation methods, we can disregard the phase factors and gain an expression for intensity at the focal plane of the lens:

$$I_f(x, y) = \frac{1}{\lambda^2 f^2} \left| \int \int_{\Sigma_0} U_0(x_0, y_0) e^{-j \frac{2\pi}{\lambda z} (xx_0 + yy_0)} dx_0 dy_0 \right|^2 \quad (16.1)$$

so that if we have some material defined threshold intensity for success in fabrication, I_{th} , then we

can be confident of success at a point $(x = i, y = j)$ at the focal plane of the modelled hologram that satisfies the success inequality,

$$I_f(i, j) \geq I_{th}. \quad (16.2)$$

The parallelisation infrastructure begins with Spatial Light Modulator 1 (**SLM1**), which is explained in detail in Section 15 and its specific usage within holography is briefly recounted in Section 16.3.1. The beam then propagates through two 4F systems (discussed in Section 18) and a pair of scanning galvanometers in order to image the display of **SLM1** onto the back of the Objective Lens, **OBJ**, which focuses the light onto the sample on the stage, **STG**. The light field at the input to **OBJ** can then be considered to be modulated according to **SLM1** and the situation is simplified to match the geometry showed in Figure 16.5.

16.3.1 Use of a Spatial Light Modulator (SLM)

A modulation device is a vital component of a holographically capable optical facility. As described in Section 15.2, there are numerous components that can modulate properties of light fields, including phase, amplitude and polarisation (not necessarily independently). As discussed in Section 16.2.2, to create defined intensity patterns at our working area we modulate the phase of the complex field. Phase-only SLMs have been demonstrated to accurately modulate the phase of a light field and are available to purchase. Our facility will include a reflective HOLOEYE PLUTO NIR-015 phase-only SLM [179], with 1920×1080 , $8 \mu\text{m}$ pixels. Each pixel is 8-bit, so has $2^8 = 256$ levels of phase modulation up to a maximum retardation of 4.4π rad, or a maximum phase precision of $\frac{4.4\pi}{256} \approx 0.053996$ rad.

When connected to a PC, the SLM software allows it to act as an external monitor with its native resolution, converting 8-bit pixel values into discrete phase alterations. The image displayed on the SLM forms the DOE that modulates the phase of the complex amplitude and therefore all that is necessary to display a desired phase pattern is a bitmap grayscale image with 8-bit encoding corresponding to variation in phase. These DOEs are the CGHs introduced in 16.2.2 and modelled in 16.4.2.

16.3.2 Imaging system

The imaging system is the infrastructure allowing the phase changes displayed at the SLM to be applied at the back of the objective lens, in the Σ_0 plane in Figure 16.5. The 4f imaging system (discussed in Section 18) is a lens configuration useful in this application due to its input and output plane relation,

$$U_{out} = FT[FT[U_{in}(x, y)]] = U_{in}(-x, -y), \quad (16.3)$$

which shows that the output field is simply the input field reflected about spatial coordinate axes. This reflection can be ignored since it doesn't affect the hologram reconstruction. The only difference is then the size of the image at the output plane, which is determined by the focal lengths, f_1 and f_2 of the lenses in the configuration. So by determining the necessary scaling factor from SLM to galvanometers and again from galvanometers to the objective lens, the SLM display can be imaged at the back of the objective lens.

16.4 Computer Generated Holograms

As described in Section 16.2.2, generating the DOE to be displayed on **SLM1** to generate a desired hologram is not trivial. We can calculate an intensity distribution at our DOE image plane (immediately before the objective lens), $I_0(x_0, y_0)$, and define a desired intensity profile at the focal plane, $I_d(x, y)$, and we know that the two can be related through Equation 16.1, the intensity expression derived from the Fraunhofer Diffraction expression, Equation 3.9. However, it is clear that $I_f(x, y)$ depends on the complex field of the DOE image plane, $U_0(x_0, y_0)$, defined by the laser's intensity and the phase change applied by **SLM1**. So the problem reduces to phase retrieval; using I_0 and I_d to calculate the unknown phase of U_0 to be modulated on to the laser by **SLM1**, so that $I_f(x, y) = I_d(x, y)$. There are a number of algorithms that have demonstrated this phase retrieval, which are described and evaluated below.

16.4.1 Phase Retrieval Algorithms

A large number of Phase Retrieval algorithms need to utilise the Fourier Transform in order to propagate between different optical planes. The propagation function in each algorithm is the model for light propagating from the DOE image plane to the HR plane. The Fourier Transform is applicable in this case since $U(x, y)$ is proportional to the Fourier Transform of $U_0(x_0, y_0)$ for both Fresnel and Fraunhofer holograms, which is seen clearly in Equations 3.6 and 3.9 respectively. However, the algorithms are clearly designed for computation and display on the pixellated SLM and therefore it is necessary to use the Discrete Fourier Transform (DFT), defined as

$$F[n] = \frac{1}{N} \sum_{k=0}^{N-1} f[k] e^{-jk\frac{2\pi n}{N}}, \quad (16.4)$$

resulting in an N point spectrum, where $f[k]$ is the original discrete signal. This is easily extended to two dimensions;

$$F[n, m] = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f[k, l] e^{-j2\pi \frac{kn}{N} \frac{lm}{M}}, \quad (16.5)$$

where $f[k, l]$ is the original 2D discrete signal. There are large number of DFT algorithms, but the most commonly used is the Fast Fourier Transform (FFT).

16.4.1.1 Gerchberg-Saxton

The Gerchberg-Saxton Algorithm, proposed by Gerchberg and Saxton in [180], is an iterative phase retrieval algorithm which acts to reduce the square error between I_f and I_d . The algorithm requires inputs $I_0(x_0, y_0)$, $I_d(x, y)$ and a first approximation to the phase mask, $\phi_0(x_0, y_0)$ in order to calculate better approximations to $\phi_k(x_0, y_0)$. Figure 16.6 shows a basic flowchart for the algorithm and the process for a single iteration is detailed in Equations 16.6 to 16.10.

- Create approximation to complex amplitude at the DOE image plane:

$$u_k(x_0, y_0) = \sqrt{I_0(x_0, y_0)} e^{j\phi_k(x_0, y_0)} \quad (16.6)$$

- Propagate the approximated complex amplitude to the focal plane (from (x_0, y_0) to (x, y) plane):

$$U_k(x, y) = \text{FFT}[u_k(x_0, y_0)] \quad (16.7)$$

where FFT refers to the Fast Fourier Transform.

- Combine the phase of the propagated field with the desired intensity profile:

$$U'_k(x, y) = \sqrt{I_d(x, y)} e^{j\theta_k}, \quad \theta_k = \arg(U_k(x, y)) \quad (16.8)$$

- Propagate the new approximation for the focal plane complex amplitude back to the DOE image plane (from (x, y) to (x_0, y_0) plane):

$$u'_k(x_0, y_0) = \text{FFT}^{-1}[U'_k(x, y)] \quad (16.9)$$

- Calculate the phase to form the improved approximation to the phase mask:

$$\phi_{k+1}(x_0, y_0) = \arg(u'_k(x_0, y_0)) \quad (16.10)$$

To determine whether an approximation ϕ_k is usable as a phase mask, the intensity related to complex amplitude U_k is compared to I_d according to some defined error metric(s). If the error is sufficiently small with respect to some error criterion then we can stop iterating and use the k^{th} approximation, otherwise the algorithm continues until the the criterion is satisfied.

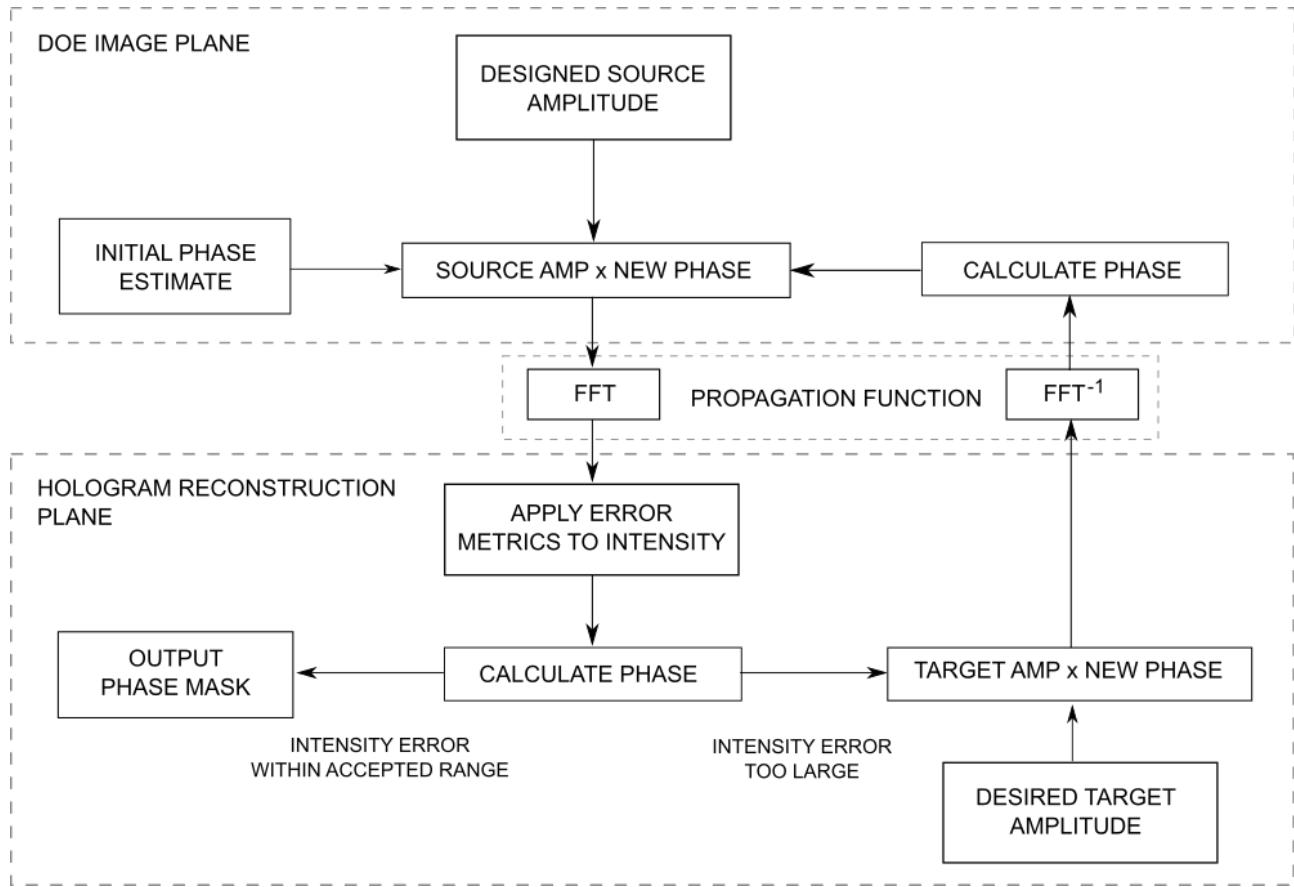


Figure 16.6: Flow chart for the Gerchberg-Saxton Phase Retrieval algorithm [180]

A proof that the mean squared error (MSE), defined as

$$e = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \left| I_d^{m,n} - |U_k^{m,n}|^2 \right|^2 \quad (16.11)$$

(where m and n are pixel indices) must decrease (or stay constant) with iterations is provided in [180]. This is important since it ensures that the result of the algorithm will not diverge, suggesting that an acceptable approximation to the desired phase mask will be obtained with enough iterations, or more accurately; theoretically, the approximation to the phase mask cannot be made worse with more iterations. If there are no limits on computation or time, we can expect the algorithm to retrieve the phase mask ϕ_k that corresponds to $I_k \approx I_d$ with reasonably small error (demonstrated in Section 16.4.2), but cannot guarantee that the result will converge to ϕ_d .

The output, $\phi_k(x_0, y_0)$, is a discrete phase distribution, but each individual pixel has a continuous possible range $-\pi < \phi_k(i, j) \leq \pi$ radians. Therefore it is necessary to quantise these pixels to be displayed on the SLM because it is a digital output device. The SLM has 8-bit addressing, which is equivalent to $2^8 = 256$ discrete phase output levels or a smallest phase interval of $\frac{2\pi}{256}$ radians. The quantisation of each pixel clearly leads to errors in the system that are not predicted when the holograms are computed, but it is possible to account for this by including the quantisation within

each iteration of the algorithm. Any change such as this could cause the non-divergence proof of the algorithm to be rendered invalid, but this method of quantisation was demonstrated in [181] with success.

16.4.1.2 Weighted Gerchberg-Saxton

The Weighted Gerchberg-Saxton (WGS) algorithm [182] is a modification on the traditional Gerchberg-Saxton algorithm. The WGS algorithm measures the intensity at the focal points in $U_k(x, y)$ during each iteration and then weights the focal points of $I_d(x, y)$ in an attempt to make the intensities uniform (i.e. attenuate the intensity of a point in the target if its corresponding approximation is larger than the other focal points and vice versa). Formally, for each iteration k , we give each of n foci in I_d an amplitude weighting factor, $(w_k^n)^2$, based on the ratio of its amplitude compared to the average of all the foci amplitudes. For the k^{th} iteration of the algorithm and n^{th} focus point;

$$w_k^n = \frac{\langle |U_{k-1}^n| \rangle}{|U_{k-1}^n|} \quad (16.12)$$

and

$$\sqrt{I_{d(k)}^n} = w_k^n \sqrt{I_{d(k-1)}^n} \quad (16.13)$$

Again, changing the iterative routine of the tradition GS algorithm potentially invalidates the non-divergence proof. However, it has been shown in [183] that using the WGS algorithm produces more efficient phase masks that generate more uniform intensity peaks in the HR plane than the regular GS method, with only slightly larger computational complexity due to the calculation of the weightings (which is negligible compared to the FFT).

16.4.1.3 Direct Binary Search

A binary hologram differs from the phase masks created in other algorithms (such as Gerchberg-Saxton) by only having two states per pixel; 0 or 1. Each pixel that takes the value 0 can be said to be transparent in transmittance and each pixel with the value 1 is opaque. This creates a diffraction grating with a number of rectangular apertures. The light field at the HR can be modelled using the same diffraction equations as non-binary phase masks, Equations 3.6 and 3.9.

The Direct Binary Search (DBS) algorithm has been shown to successfully generate binary holograms that result in approximations to $I_d(x, y)$ at the focal plane [184]. The DBS algorithm is iterative and, similar to other algorithms, requires inputs of I_0 , I_d and ϕ_0 . A flow chart for the algorithm is shown in 16.7 and an iteration follows this process:

- Propagate the input complex amplitude to the focal plane (the same as the first two steps in GS; Equations 16.6 and 16.7)

- Calculate the mean square error (MSE) between $\sqrt{I_d}$ and $|U_k|$ using Equation 16.11
- Invert the pixel value indexed by (m, n)
- Propagate the new approximation of complex amplitude to the focal plane
- Calculate the MSE, e' , of the new approximation at the focal plane
- If $e' < e$, retain the inverted pixel's new value, otherwise revert to its original value
- Increment the column index, m , if possible (due to the dimensions of the image). If m cannot be increased, set $m = 0$ and if possible, increment the row index n . If m and/or n were changed then return to the pixel inversion step. This process raster scans the image. If m and n cannot be incremented (no pixels left to check within this iteration), then the iteration is complete and the $(k + 1)^{\text{th}}$ approximation to the phase mask is the output.

The DBS repeats this process until an iteration is performed without a single pixel value inversion, corresponding to a minimum error configuration.

The original DBS raster-scanning scheme is as described above, but other schemes could be used, for example, scanning from the centre outwards or selecting pixels probabilistically (pseudo-random or weighted in certain areas). It is likely that scanning scheme time-effectiveness depends on the desired intensity profile, so it would make sense to determine advantageous schemes empirically if we were to implement the DBS algorithm.

It should be noted that this minimum error configuration is not necessarily the global minimum, because not all permutations of binary phase masks were generated and checked (for a 1024×1024 display, this would require $2^{(1024^2)} \approx 6.7 \times 10^{315652}$ different phase masks to be generated and errors calculated, which is clearly an unreasonable amount of computation). A modification of this algorithm, Simulated Annealing (described in Section 16.4.1.4), modifies the algorithm so that it becomes convergent to the global minimum error in probability.

Time and number of iterations to arrive at a minimum varies depending on the original phase estimate and the scanning scheme. In theory, an iteration of the DBS algorithm requires a two dimensional FFT over the entire image and one calculation of MSE for each pixel in the image, or N^2 FFTs and N^2 MSE calculations per iteration for an $N \times N$ pixel image. For large N , this is more computationally intensive than the GS variant algorithms, whose number of FFT and MSE calculations do not scale with image size.

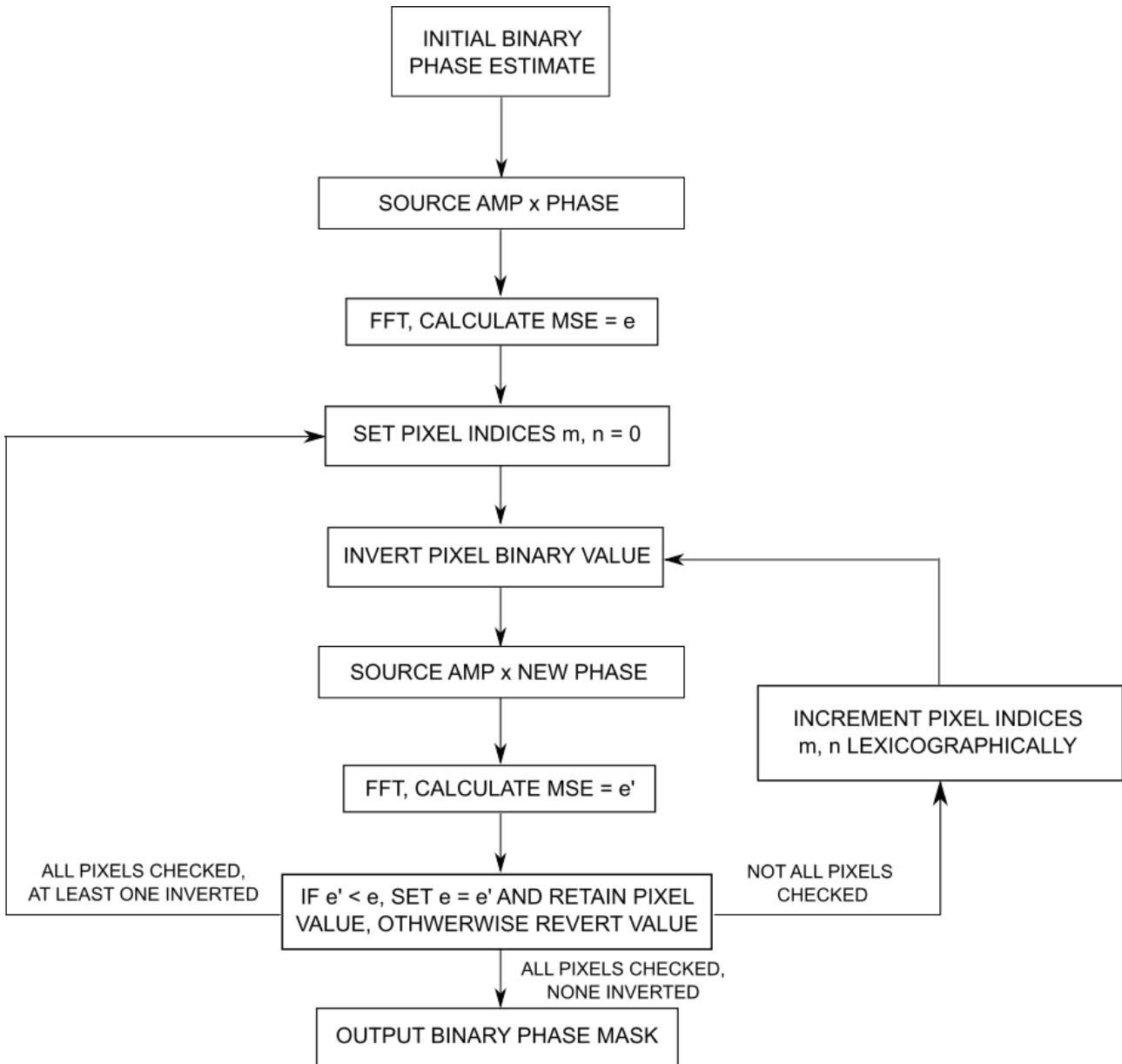


Figure 16.7: Flow chart for the Direct Binary Search algorithm

16.4.1.4 Simulated Annealing

The Simulated Annealing (SA) algorithm is another example of a binary phase retrieval algorithm. It builds upon the DBS algorithm so that it converges probabilistically to the global optimum binary phase mask to produce the lowest MSE reconstruction of an approximation $|U_k|^2$ to I_d . The difference between the DBS and SA algorithms is the SA allows the possibility of retaining a changed pixel even if it increases the MSE, with some probability. The SA algorithm includes a new parameter referred to as the temperature, T , which is varied to control the probability that a MSE increasing pixel change is accepted. Generally, as the algorithm progresses and the error reduces, T decreases (or 'cools') and this probability decreases, guiding the binary phase mask to the global optimum.

In [185] the probability p of accepting a switched pixel value is defined as

$$p = \begin{cases} 1, & \text{for } e' \leq e \\ e^{-\frac{e'-e}{kT}}, & \text{for } e' > e \end{cases} \quad (16.14)$$

where k is the Boltzmann constant. It is clear that if the error increase due to a pixel value change is larger, the probability of accepting the change is smaller and that if T is smaller then the probability of accepting the change is smaller. Therefore it is more likely to accept an error-increasing change if error increase is small and temperature is high. By design, T is made to decrease monotonically with number of iterations so that we reach a point such that SA resembles DBS within the vicinity of the global optimum minimum error, with very little probability of accepting a pixel value change that increases the MSE.

Although probabilistically convergent to the global optimum binary phase mask; our facility is not dependent on this optimum if the local optima are also sufficient for fabrication. The SA algorithm is also more computationally intensive than DBS due to the calculation of the control probability at each pixel check. As a result, SA is also much slower in computation than GS variants.

16.4.1.5 Extension to Three Dimensions

In the above algorithms we have only described their applicability to 2D hologram reconstructions using the Fraunhofer approximation to create an intensity profile at the focal plane of a converging lens (Equation 16.1). We can extend all of these methods into three dimensions by modifying the propagation function to approximate the inclusion of the propagation from the focal plane, Σ_f , to some planes spaced by distance dz around Σ_f . This can be achieved by modifying the z distance in the phase factor of the Fraunhofer diffraction formula (Equation 3.9), our propagation function). To propagate to the focal plane we use $z = f$ where f is the focal length of the lens, but to propagate to other close by planes we use $z = f \pm dz$. A geometry for the case with 3 planes of interest is shown in 16.8.

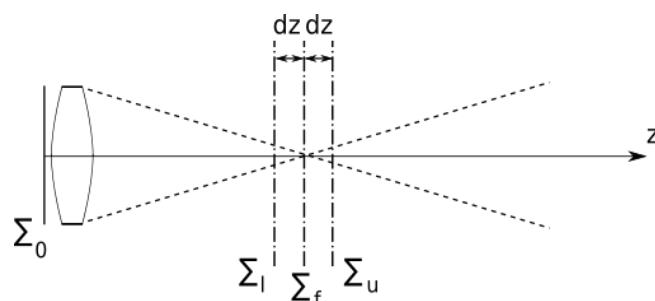


Figure 16.8: Geometry for modelling propagation of light fields to multiple planes

If we follow the same algorithm procedures for each plane then we generate a DOE phase mask for each plane, which can be summed to create a single DOE phase mask to be modulated by **SLM1**.

This approach will at least multiply computation time for a single hologram by N_p number of planes. However, other approaches may reduce this computation by combining each plane's DOE within iterations, which is explored in [186].

16.4.2 Modelling

In this section we demonstrate our facility's ability to generate CGHs to apply to the SLM for parallelisation purposes. All following modelling and computation was scripted and performed in MATLAB. The modelling aims to show that we can produce two dimensional arrays of foci capable of reaching intensities necessary to fabricate in common materials, using 1024×1024 pixels of the SLM's 1920×1080 (since the FFT is more efficient when performed on power-of-two size images) to display up to 8-bit holograms.

16.4.2.1 Error metrics

In order to determine how successful our CGH generation is, we define some error metrics that can be used to evaluate the CGHs and act as conditionals within the generation algorithm.

Mean Square Error (MSE) provides a quantitative measurement of how well matched the entire intensity approximation $|U_k|^2$ is to the desired profile, I_d . In this implementation, where $M = N = 1024$, the MSE is defined as

$$e_{MSE} = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left| I_d^{m,n} - |U_k^{m,n}|^2 \right|^2. \quad (16.15)$$

MSE is used as an overall error metric to assess how impactful further iterations of algorithms are.

We also use a measure of uniformity for the peaks of the approximated intensity distribution. Uniformity is defined in [183] as:

$$e_{UNI} = 1 - \frac{\max(I_j) - \min(I_j)}{\max(I_j) + \min(I_j)} \quad (16.16)$$

where I_j are the approximated intensity peaks. By normalising the difference in peak height using the sum of the peak heights, the uniformity metric will be 1 for a perfectly uniform peak distribution ($\max(I_j) = \min(I_j)$) and will be 0 for a distribution with a missing peak ($\min(I_j) = 0$). Uniformity is important particularly in TPP, where the fabrication process is non-linear. Therefore it is necessary to keep the focal intensities as close to designed as possible to avoid large variations in fabrication effects at different focus points.

Another metric we use is to determine the effectiveness of the CGH in machining at our focal points. If we define P as the number of intensity peaks I_j within $|U_k|^2$ that satisfy a discrete version

of the fabrication success inequality (continuous version is Equation 16.2),

$$I_j \geq I_{th} \quad (16.17)$$

and Q as the number of peaks that don't (where $P+Q$ must equal the number of designed peaks), then we can calculate the Fabrication Success Ratio as

$$e_{FSR} = \frac{P}{P+Q}. \quad (16.18)$$

This metric gives us a direct percentage of the number of designed focal spots that will fabricate using the tested CGH and for a successful CGH, $e_{FSR} = 1$. However, this is a necessary but not exhaustive success criterion, since a CGH with a greater number of focal spots than designed would also maintain $e_{FSR} = 1$. As a result, using this metric in conjunction with e_{MSE} and a visual representation of the intensity approximation is important.

A final metric we can use to assess the efficiency of our use of light in hologram generation is the proportion of power e_{PRP} contained within the foci in the HR plane. Since $P = \int_S I \cdot dA$, the power within a pixel is $I \Delta x \Delta y$ where Δx and Δy are pixel dimensions. However, pixel dimensions cancel in the ratio, leaving:

$$e_{PRP} = \frac{\sum_j I_j}{\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} |U_k^{m,n}|^2}. \quad (16.19)$$

Our desired profile has $e_{PRP} = 1$, but this is unlikely to be attainable due to limited number of pixels and computation time.

16.4.2.2 Choice of algorithm

The important criteria for the algorithm provided in the software suite are defined by the error metrics and amount of time taken to generate. Desirable results for e_{MSE} and e_{UNI} have been demonstrated for both GS and WGS by Di Leonardo et al [183], who found that WGS algorithm clearly improves the CGH's uniformity. It is also clear from discussion in Section 16.4.1 that DBS is improved by the probabilistic approach of SA. SA by definition would eventually provide a satisfactory CGH by the e_{MSE} metric, but e_{UNI} is not directly optimised as in WGS.

The the number of FFTs required for two-dimensional WGS and SA (using the built-in MATLAB `FFT2` function), n_{WGS} and n_{SA} , are

$$n_{WGS} = 2 \times i_{WGS}, \quad n_{SA} = N^2 \times i_{SA}, \quad (16.20)$$

where i_{ALG} is the number of iterations of algorithm ALG and N is the number of pixels of the square image in one dimension. If we assume that $i_{WGS} = 50$ and $i_{SA} = 4$ for usable CGHs, then for the designed 1024×1024 pixel image $n_{SA} \gg n_{WGS}$ and the time elapsed for SA will be much larger than WGS, as shown in Table 16.1. It has been suggested in [184] that a FFT is not necessary for each pixel in the image for DBS and this could be extended to SA; but this would require a large amount of increased complexity in software development that WGS does not.

Algorithm	No. of MATLAB FFT2s required, n_{ALG}	Predicted Time Elapsed (s)
WGS	100	5.5
SA	5242880	286601

Table 16.1: Table of predicted time to generate a usable CGH with different algorithms, assuming all the computation time is taken by the FFT2 function. The time for a single FFT2 was found by timing the FFT2 function on a comparable complex field model image of 1024×1024 pixels.

For its reliability in generating high uniformity, low error CGHs in a short time period compared with other algorithms, our optical system is designed to use the Weighted Gerchberg-Saxton algorithm.

16.4.2.3 Results

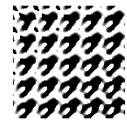
Using the 0.5 mJ pulse energy and 100 fs pulse length, the maximum calculated source intensity before the objective lens is a gaussian of peak intensity $I_0 \approx 8.3 \times 10^9 \text{ W/m}^2$. In the following modelling, the laser power is reduced to produce optimal intensity distributions in the focal plane. Throughout this modelling, we will use the fabrication operating point threshold $I_{th} = 4 \times 10^{17} \text{ W/m}^2$ (the typical LMM ablation intensity from Section 6.5 increased by a factor of 4 to compensate for power dissipation not modelled) since this is the largest intensity required from any fabrication technique.

Table 16.2 shows the error metrics at selected interesting iteration numbers of the WGS MATLAB model for a 5×5 foci array of foci modelled as 200 nm squares (our maximum resolution with TPP). Iterations 1 and 2 are not usable approximations due to the low uniformity. Reduction in e_{MSE} is much quicker in earlier iterations and iteration 12 generates an approximation that could be used. It can be decided by the user how many iterations they would like to perform before using the generated phase mask (an example section of which is shown in Figure 16.9), depending on time constraints (discussed in 16.4.3) and fabrication performance required; calibration would be certainly required.

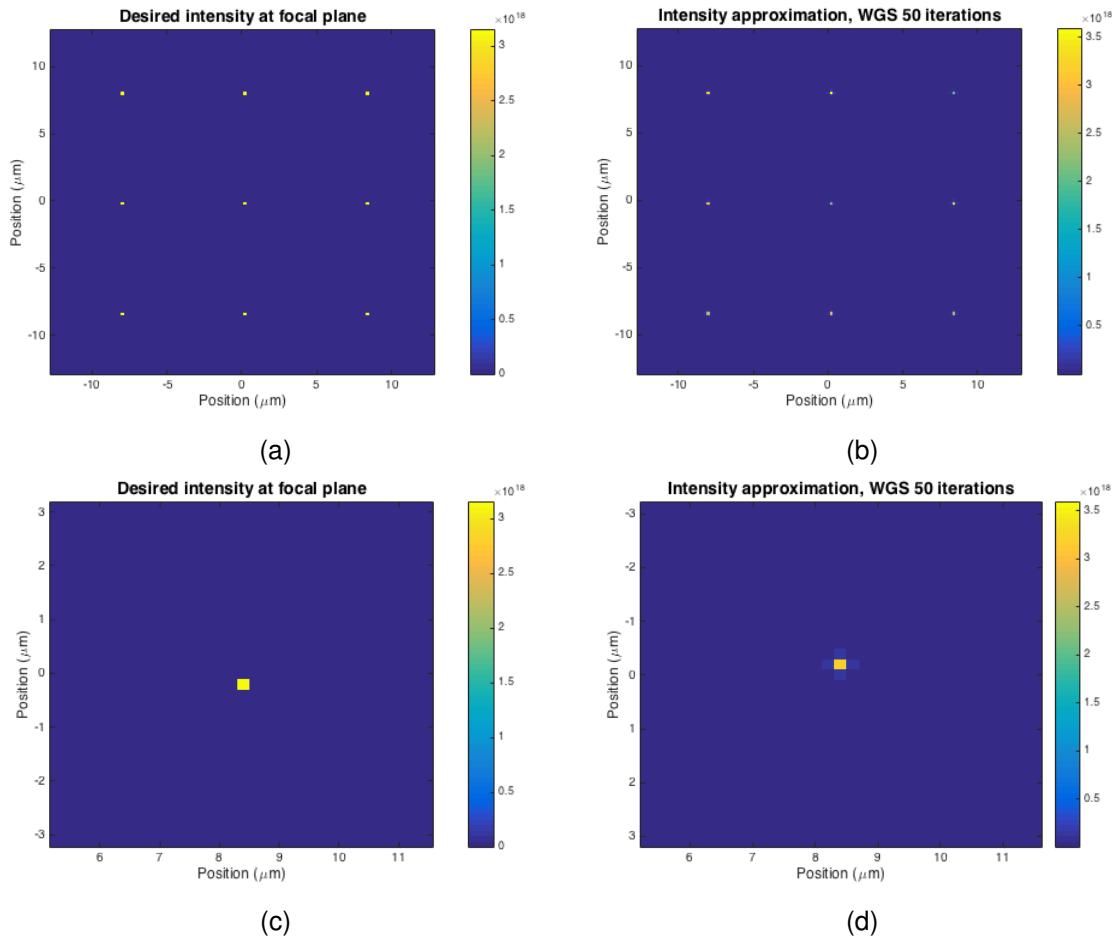
Figure 16.10 shows the intensity profile results of the model after 50 iterations, zoomed into the central 3×3 foci for clarity. The approximation definitely resembles the desired profile and the intensities of the focal points often exceed the desired intensities. The surrounding areas of the foci in the approximation have larger intensities than desired, but not large enough to fabricate, so the

Number of iterations	e_{MSE} (W/m^2) ²	e_{UNI}	e_{FSR}	e_{PRP}
1	4.03×10^{33}	1	0	0
2	3.87×10^{31}	0.3	1	0.76
12	1.75×10^{31}	0.61	1	0.78
50	1.44×10^{31}	0.65	1	0.81

Table 16.2: Table of error metrics at different iterations of WGS

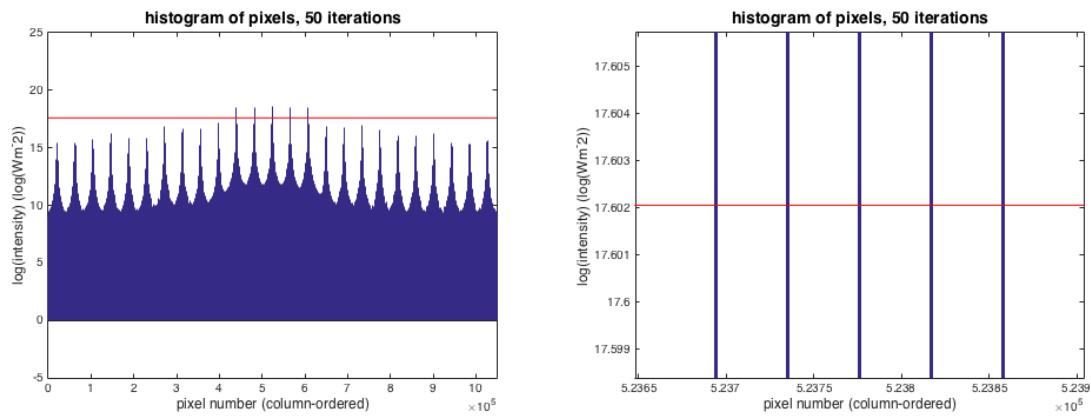
Figure 16.9: 128×128 pixel section of WGS generated phase mask

resolution in this case is theoretically the same as when only using one focus. This is confirmed by plotting the pixel histogram and I_{th} , seen in Figure 16.11.

Figure 16.10: Comparison of the target focal plane intensity distribution and the approximated reconstruction after 50 iterations of the WGS algorithm in MATLAB, for 5×5 , 200 nm square foci. The colour bar has units of W/m^2 . (a) and (b) show the 3×3 foci middle section of the larger array, desired and approximated, respectively. (c) and (d) show views of a single focus point, again for desired and approximated, respectively.

16.4.2.4 Number of foci

Changing the number of foci greatly changes the outcome of the algorithm due to the sharing of energy between focus points. By managing laser power carefully, 100 foci can be achieved without reduction in resolution. However, low uniformity is not easily corrected by changing the laser output power. Generally, approximations with less foci have significantly better uniformity, as seen in Table



(a) Histogram of all pixel intensities of the 50 iteration WGS approximation

(b) Zoomed view of one of (a)'s visible peaks

Figure 16.11: Histogram of the 50 iteration WGS approximation to a 25 foci array. The red horizontal line is I_{th} . 25 pixels exactly cross I_{th} , showing that we have no reduction in resolution. The difficulty in seeing the 5 pixels in visible peaks in (a) is caused by the small amount of pixels between each focus, and it can be seen in (b) that only 5 pixels cross I_{th} per visible peak.

16.3.

Number of foci	e_{UNI} after 50 iterations
4	1.00
9	0.89
16	0.98
25	0.65
49	0.66
100	0.73
144	0.59

Table 16.3: Table showing how uniformity changes with number of foci generated

16.4.2.5 Off-axis intensity

Parallel fabrication performance in the physical system may deviate from the modelled results above due to the effects of the continuous spatial Fourier Transform on the discretised phase signal that **SLM1** pixels define. In one dimension, the form of this phase signal can be seen in Figure 16.12, generated by convolving the sampled phase distribution with a pixel model function.

Since the model output is a convolution in the spatial domain, we can calculate the effect in the frequency domain (focal plane) by applying the *Spatial Convolution-Frequency Multiplication* property of the Fourier transform:

$$\mathcal{FT}[f(x) * g(x)] = \mathcal{FT}[f(x)] \mathcal{FT}[g(x)] \quad (16.21)$$

where $f(x)$, $g(x)$ are the spatial signals. From [187], the Fourier transform of a delta pulse train is

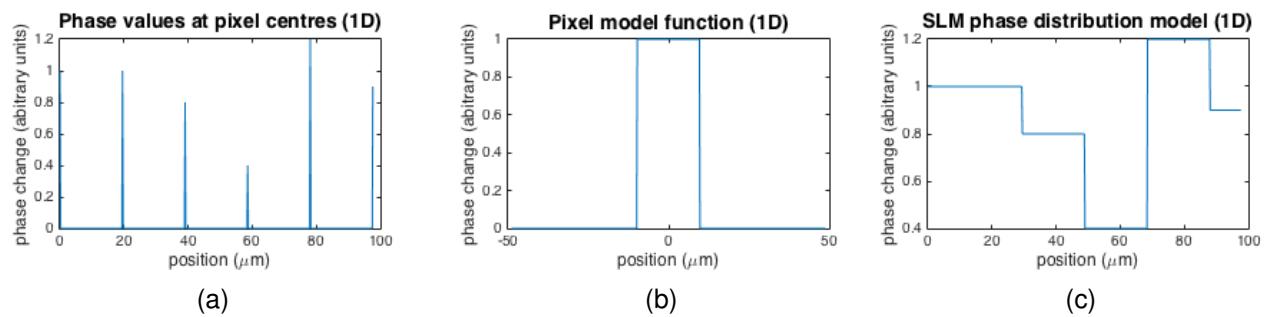


Figure 16.12: MATLAB model of the physical **SLM1** display in one dimension. (a) shows the continuous phase distribution sampled at the pixel centres, (b) shows the pixel model function used and (c) is a convolution of (a) and (b) to create the model output of **SLM1**

simply another pulse train and the Fourier transform of the model pixel function of width X is a sinc function:

$$\mathcal{FT} \left[\text{Pxl} \left(\frac{x}{X} \right) \right] = X \text{sinc} (fx), \quad (16.22)$$

where f is the spatial frequency. The Fourier transform of the convolution is the multiplication of the delta pulse train with the sinc function, and this causes the amplitude of off-axis points in the focal plane to be reduced according to the sinc function (and therefore intensity reduction by the square of this). The two-dimensional case simply extends from the one-dimensional case.

When X is the pixel length at the objective lens input (approximately 19.5 μm with square pixels) and using theory in Section 3.2.3 we find that the first zeros of the sinc function occur at a radius $r = \frac{\lambda f}{X} = 41 \mu\text{m}$ in the focal plane. Figure 16.13 shows the intensity reducing function in one dimension.

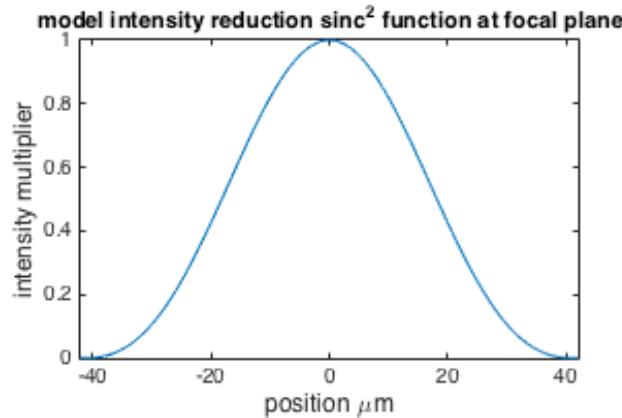


Figure 16.13: MATLAB model of the one-dimensional intensity reduction in the focal plane due to the discrete phase signal of **SLM1**.

As a result of this reduction in intensity, it is clear that the foci should be located as close as possible to the optical axis. Modelling of the 200 nm square foci showed the beginning of intensity interference between neighbouring foci when they were spaced by 600 nm centre-to-centre. The WGS algorithm also allows outer foci on the target intensity to be set higher, generating higher intensity

approximations in these foci.

16.4.3 Speed

The time taken to generate a phase mask in the MATLAB model of the WGS algorithm is linearly related to the number of iterations. For 100 foci, approximately 20 iterations is needed for a usable phase mask. In the model this takes 1.75 s, which would limit the hologram generation frequency to 0.57 Hz. The physical limiting factor of frequency is the refresh rate of **SLM1**, 60 Hz [179]. Therefore, if multiple holograms are needed, the phase masks are calculated before fabrication and stored in memory so that **SLM1** updates its output with a new phase mask on each refresh.

16.4.3.1 Considerations of GPUs, FPGAs

Computation speed could be improved by a large degree by changing the hardware used. All the modelling above was performed on MATLAB using a 2.2 GHz Intel Core i7 processor, but using a more powerful processor would obviously reduce the time needed for similar computations. Cullinan et al [188] performed benchmarking tests on various CPUs, GPUs and FPGAs for the FFT function and found that the GPU execution time was four orders of magnitude smaller than the fastest CPU. Our usable holograms can be generated in approximately 1.75 s, so with a GPU this could theoretically decrease to 0.175 ms, where hologram generation frequency becomes 5714 Hz; limited by **SLM1** to 60 Hz. This opens the possibility to update holograms during fabrication in real-time.

16.5 Parallelisation Conclusion

Parallelisation enables the optical facility to utilise HOT so that objects can be brought together to interact. By computation using GPUs it could be possible for our system to update holograms in real-time and use feedback from the live imaging techniques in order to optimise the locations or intensities of foci. The parallelisation system modelling demonstrated that the facility is capable of generating arrays of up to 100 foci in desired locations with satisfactory uniformity. This allows the optical facility to significantly increase speed, as quantified in 16.4. This reduction in time is clearly significant and would be very desirable for companies or researchers under time-pressure.

One focus TPP and SLA	100 foci TPP and one focus SLA
7 days	14.4 hours

Table 16.4: Table comparing time taken to fabricate a model R2-D2 (as seen in Figure 24.18) using a raster-scanning scheme with 1000 layers of high-resolution TPP detail with and without parallelisation.

17 Microscope Objective Lenses

As explained in section 3.1, objectives enable the optical system to generate a real and magnified image which can be viewed through an eyepiece or in our case an image sensor. As such, they critically influence many system performance characteristics including resolution and the quality of images. They can achieve angular magnifications of up to 200x allowing even small details to be sufficiently magnified. In order to achieve this, a complex, multi-element design is required. This makes the objective the most difficult component to design and manufacture in optical systems. Several manufacturers offer a wide range of designs, targeted for specific applications. We will rely on the decades of research which has enabled them to produce highly corrected and high performance objectives, rather than designing our own. There are many types of objective lenses and in order to choose the correct ones for our requirements, it is important to understand how they are classified and what differentiates the types.

17.1 Classifications

17.1.1 Refractive Objective Lenses

The majority of objectives are refractive, meaning that the light passes through and is refracted by the glass lenses inside the objective. Achromatic doublets and anti-reflective coatings are used to reduce back reflection and aberrations and improve light throughput. There are a range of refractive objective types with increasing amounts of aberration correction and subsequently increasing number of elements, complexity and cost as shown in Figure 17.1.

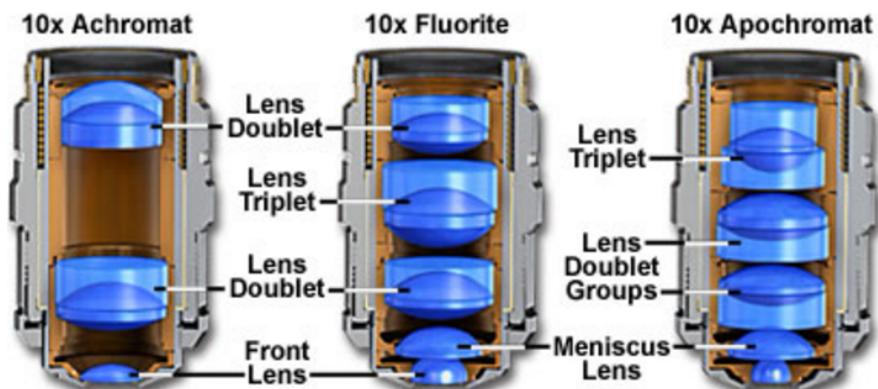


Figure 17.1: Types of Microscope Objectives. Courtesy of [189].

17.1.1.1 Achromats

Achromats are the cheapest and most common type of objectives. They are chromatically corrected for two colours, meaning that two wavelengths will focus in the same point - often red and blue wavelengths. They are also spherically corrected for a third colour - typically green. The limited chromatic correction makes these objectives unsuitable for colour micrography.[190] Since we will

use a wide range of wavelengths, this type of objective is not suitable in this system.

17.1.1.2 Fluorites

These objectives are more complex, also chromatically corrected for two but spherically corrected for up to three wavelengths. They also have better light-gathering abilities than the achromats. This type is useful in low strain applications which are explained later.

17.1.1.3 Apochromats

Apochromats are the best type of refractive objectives in terms of correction. They almost eliminate chromatic aberrations by spherically and chromatically correcting for up to four colours even at large magnifications which fits our requirements well.

17.1.1.4 Flat Field - "Plan"

All types described above experience a significant curvature of field, especially at high magnifications. This means a planar object results in a non-planar image. In effect, when captured on a flat surface like an image sensor, only the centre will appear sharp. Modern objectives are often flat-field corrected. This is indicated with the prescript "Plan". A Plan apochromat, therefore, is the most expensive but also most aberration free refractive objective. In scanning applications, like two-photon polymerisation or confocal microscopy (chapters 8 and 13.1), a flat field means that the point of focus is the same distance away from the objective across the field of view. This is important to reduce the complexity in control.

17.1.2 Reflective Objective Lenses

In contrast to the transparent lenses used in refractive objectives, reflective ones use mirrors. Similar to a reflecting telescope, a parabolic secondary mirror with a central aperture reflects light onto a centrally mounted, smaller primary mirror which reflects the light through the aperture in the secondary mirror, as shown in Figure 17.2. Mirror based designs don't suffer from chromatic aberrations and can be used over a much broader spectrum of wavelengths, often from ultraviolet all the way to infrared. However, the backside of the primary mirror and its mount cause obscurations - some light is reflected back rather than transmitted through the objective. Due to the mirrors, a wavefront error is introduced which we would be able to overcome with adaptive optics (section 19.4). We won't require reflective objectives because the modern refractive ones have broad enough transmission ranges for our requirements and the resolution and light gathering power of reflective objectives is limited because the highest achievable numerical apertures are around 0.5.

17.1.3 Finite Conjugate Objectives

This is the old style of objective which is used with fixed tube lengths in combination with an eye piece or image sensor. As described in section 3.1 on the basic principles of microscopes and shown

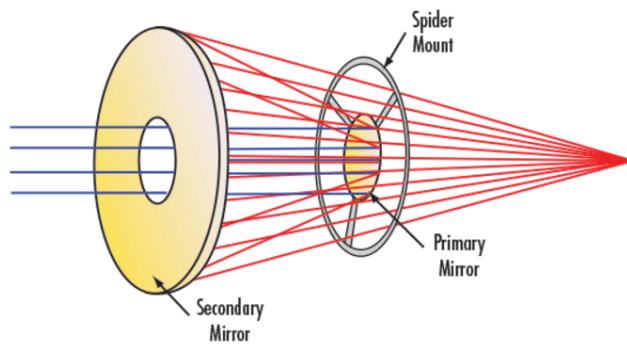


Figure 17.2: Reflective Objective Principle: Light from the focus point on the right is reflected by the secondary mirror onto the mounted primary mirror. The light leaves the objective through an aperture in the secondary mirror. Courtesy of [191].

in Figure 3.2 on page 5, the object is placed just outside the front focal plane and a real image is projected at a fixed length inside the tube. The rays leaving the objective are converging in the intermediate image plane. This setup is simple but it does not allow for intermediate optical elements to be inserted into the beam path which is necessary in our system.[192]

17.1.4 Infinity Corrected Objectives

These objectives project the image to infinity, rather than a fixed point. The rays coming out of the objective are travelling in parallel through what is termed the **infinity space**. Optical elements such as polarisers and dichroic mirrors can be inserted with minimal effect. A **tube lens** is then used to focus the parallel rays and form the intermediate image, as shown in Figure 17.3.

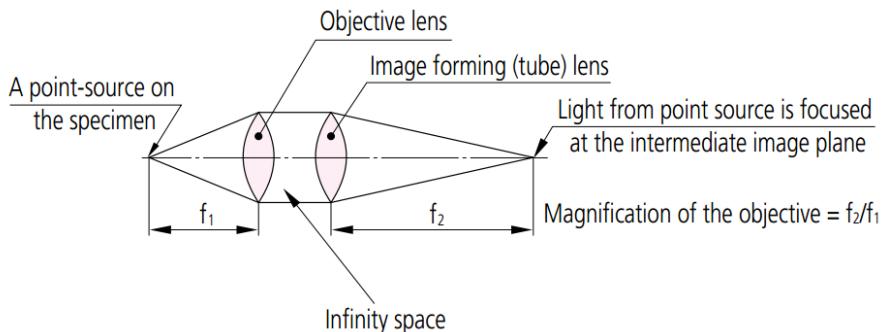


Figure 17.3: Infinity Optical System, the point source is at the focal plane so the image is at infinity, the emerging rays are parallel before being focussed by the tube lens. Courtesy of [192].

17.2 Specifications

17.2.1 Numerical Aperture

One of the most important attributes of an objective lens is its angular aperture - the maximum angle over which it will accept and transmit light rays. The numerical aperture is a dimensionless number which characterises this angle and is defined as the sine of the angle θ , the half-angle of the maximum cone of light which can exit or enter the objective, multiplied by n the refractive index of the medium

between the objective and the object.

$$NA = n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \quad (17.1)$$

Incorporating the refractive index of the medium allows the numerical aperture to remain constant as the light passes through different mediums as shown in Equation 17.1 and Figure 17.4. The numerical

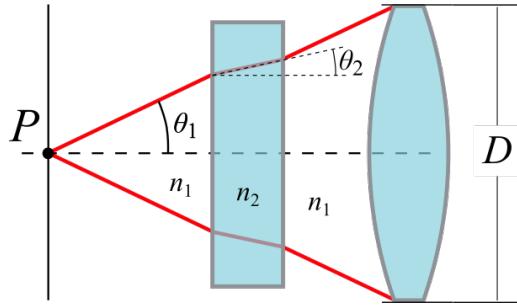


Figure 17.4: Numerical Aperture is unchanged, Adapted from [193]

aperture sets an upper limit on the **resolving power** of the optical system. It also determines the **focal depth**, and the **brightness** of the obtained image.

There is an obvious limit to the angular aperture at 180 degrees. In practice, high performance objectives reach a half-angle of up to about 72 degrees with a sine of 0.95. To further increase the numerical aperture and improve the objective's capabilities, a fluid with a higher reflective index than air may be used in between the objective lens and the object. Objectives have been designed for a variety of immersion media, most commonly water, glycerine and oil with refractive indexes of 1.33, 1.47 and 1.51 respectively. [194]

An objective's numerical aperture is related to its **entrance pupil** diameter, D in Figure 17.4 and its **focal length**, as in Equation 17.2. Larger pupils and shorter focal lengths increase the numerical aperture. The focal length in turn determines the magnification of the objective, as explained in subsection 17.2.2. Generally, objectives with higher magnifications have larger numerical apertures.

$$NA = n \sin(\theta) = n \sin \left[\arctan \left(\frac{D}{2f} \right) \right] \approx \frac{nD}{2f} \quad (17.2)$$

17.2.2 Magnification

17.2.2.1 Finite Conjugate Objectives

The magnification given by an objective lens is dependent on the placement of the object, according to equation (17.7). International standards ensure that the magnifying power engraved on the objective has a meaning. As explained in section 3.1, in a finite compound microscope, the object is placed just outside the objective's focal plane. This results in a real, magnified image, inside the

tube of the microscope which connects the objective and the eyepiece. The distance between them is referred to as the **mechanical tube length**, L in Figure 3.2 on page 5, and is standardised by manufacturers. In reality, the tube length is much larger than the focal lengths of either the objective or the eye piece. The location of the object is only a few microns from the focal point of the lens so $f_o \approx d_o$. The magnification of the objective can be calculated from the tube length and the objective focal length.

$$M_{\text{Objective}} = \frac{d_i}{d_o} \approx \frac{L}{f_o} \quad (17.3)$$

17.2.2.2 Infinite Conjugate Objectives

In an infinity corrected system, the objective projects the image to infinity and the tube lens creates the intermediate real image. The magnification is given by the ratio of the two focal lengths. [195]

$$M = \frac{f_{\text{tube lens}}}{f_{\text{objective}}} \quad (17.4)$$

The tube lens focal length is also often given as the **reference focal length** which varies depending on the manufacturer between 160-200mm. Typical lengths are found in Table 17.1.

17.2.3 Parfocal Distance

The distance between the flange, the end of the objective furthest away from the object, and the object under examination, is known as the parfocal distance. Having equal parfocal distances of objectives mounted on the turret allows objectives to be swapped without significant refocusing. The turret design is discussed in the mechanical section, chapter 21. The parfocal distance has therefore been standardized although different manufacturers adhere to different standards, as seen in Table 17.1.

17.2.4 Working Distance

The working distance is the spacing between the front of the objective and the object or the cover slip, if the objective is designed for such. The parfocal length is the sum of the working distance and the length of the objective. For a given parfocal length, longer objectives have shorter working distances. This means that more complex arrangements, especially plan-apochromats with high NA and magnification, often have working distances less than 0.2mm. This limits the depth at which we can image and polymerise. Building objects layer by layer avoids this problem. It does mean that detailed imaging needs to be performed after a few layers of polymerisation and the resulting images stored to create a final 3d image of a large object.

17.2.5 Cover Slip Thickness

Many specimen are covered by a thin glass for protection and avoiding contamination. These cover slips may introduce aberrations which the objective has to compensate for. While the refractive index

of cover slips is relatively constant, the thickness may vary even within a common batch. The effect is most pronounced with steep angles and large differences in the refractive indexes between the immersion medium and the glass. Some dry, high NA objectives have an a **correction collar** which can be adjusted for the spherical aberrations caused by varying thicknesses. The accepted standard is a size 1.5 cover slip which is usually 0.17mm thick although we will avoid cover slips where possible.

17.2.6 Design Wavelengths

An uncoated air glass interface typically reflects four to five percent of the incident light. Modern objectives with many elements and interfaces reduce back reflections by using multilayer coatings with proprietary formulations. By controlling the thickness of several layers, reflected light interferes destructively, leading to extremely high transmission values over a broad range of wavelengths. Detailed transmission data can be obtained from manufacturers to check the suitability of objectives.

Table 17.1: Manufacturer Standards

Manufacturer	Thread	Tube lens focal length	Parfocal distance
Mitutoyo	26mmx0.706mm 36TPI	200mm	95mm
Olympus	RMS	180mm	45mm
Nikon	M25	200mm	60mm
Zeiss	RMS	165mm	45mm
Leica	M25	200mm	45mm

17.3 Resolution

Resolution refers to the minimum distance in the object plane where two points are still observed as being separate. Any sample can be viewed as a collection of complex diffraction gratings. The light hitting the specimen interferes destructively or constructively to form the final image. According to Abbe's theory, at least two orders of the refractive pattern need to be captured to reconstruct the original shape that caused it, although the more orders pass through the objective the more accurate the image will be. Immersion oil with high refractive index, which increases the numerical aperture, compresses the diffracted rays and thereby increases resolution. Similarly, if a shorter wavelength light is used, it will refract by lesser amounts, also improving the resolution.

A small detail in the object will be projected as an **airy disk** on the image plane, the final image being made up of a selection of those airy disks. Higher NAs allow more diffraction orders to be captured, thereby reducing the size of the disks. More detailed work is provided in section 19.3.

A frequently used definition is that given by Rayleigh which assumes the limit to be where the peak of one airy disk overlaps with the first zero of the next one. This gives the often quoted Rayleigh equation for resolution:

$$d = 0.61 \frac{\lambda}{NA} \quad (17.5)$$

17.4 Magnification

17.4.1 Thin Lens Magnification

An object outside the focal point of a lens creates a real image on the other side. Figure 17.5 shows

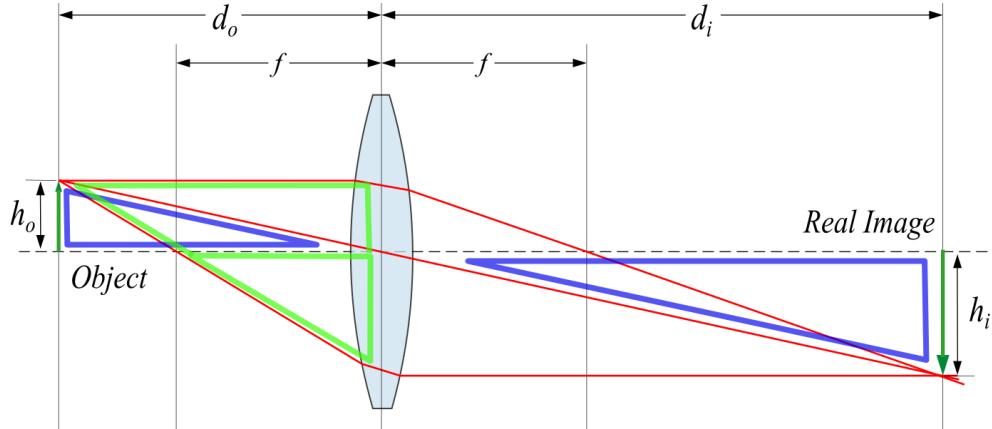


Figure 17.5: A real image formed by a lens

the three principal rays leaving the object at a distance d_o away from the lens. As usual, rays crossing the focal points on either side of the lens are parallel to the central axis on the other side and the ray passing through the midpoint of the lens is not refracted. Looking at the blue right angled triangles, the magnitude of the ratios will be equal.

$$\frac{h_o}{d_o} = \frac{h_i}{d_i} \quad (17.6)$$

$$m = \frac{h_i}{h_o} = -\frac{d_i}{d_o} \quad (17.7)$$

From the two green triangles, a similar argument shows:

$$\frac{h_i}{f} = \frac{h_i + h_o}{d_o} \quad (17.8)$$

$$\frac{1}{f} = \frac{h_i + h_o}{h_i d_o} = \frac{1}{d_o} + \frac{h_o}{h_i d_o} \quad (17.9)$$

$$\text{Using Equation (17.7): } \frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \quad (17.10)$$

17.4.2 Empty Magnification

Increasing the magnification without improving the resolving power of the system may not uncover any more details. Similar to digital zoom - using only a part of the camera sensor - high magnifying powers with limited numerical aperture results in empty magnification and should be avoided.

If the tube lens images directly onto an image sensor, the sensor needs to be large enough to cover the full image. Otherwise empty magnification will result in a darker image. A reduction lens

may be used to decrease the image size. These considerations are important for the live imaging system in section 12.1.

17.5 Rays, Pupils, Apertures

While these concepts apply to all optical systems, they are important for the following discussion on objective lenses. **Stops** are physical barriers that stop rays which are far away from the optical axis from passing through the system. They can be lenses of a particular size or apertures in an opaque material.

17.5.1 Aperture Stop

The aperture stop limits the angle of acceptance of the marginal rays which can just about pass through the system. This angle determines the angular and therefore numerical aperture. It can be located at any point within the optical system. If there are several stops, whichever limits the angle of marginal rays the most is the aperture stop. Reducing the size of the aperture stop reduces the numerical aperture and the amount of light that passes through — the image gets darker.

17.5.2 Marginal and Chief Rays

When analysing an optical system we normally restrict ourselves to rays that are on the same plane as the **optical axis** which goes through the centre of the system. These rays are called **meridional rays** and the following two rays fall under this category.

The **chief ray** or **principal ray** originates at an object point off the optical axis and passes through the centre of the **aperture stop** [196]. There can be an infinite number of chief rays but we normally use the one leaving from the outer edge of the object.

The **Marginal Ray** originates at an object point on the optical axis and passes through the edge of the aperture stop. This means that any rays which enter the system at a steeper angle than the marginal ray will hit the aperture stop and not pass through. Both rays are shown in Figure 17.6

17.5.3 Pupils

As shown in Figure 17.7, the aperture stop can be at any point within the system as long as it limits the marginal rays. When looking into the system from the front, one can see an image of the aperture stop. This is the **entrance pupil**. The image of the aperture stop on the back side is the **exit pupil**. Like any image, pupils can be real or virtual. As you would expect, the pupils are located where the

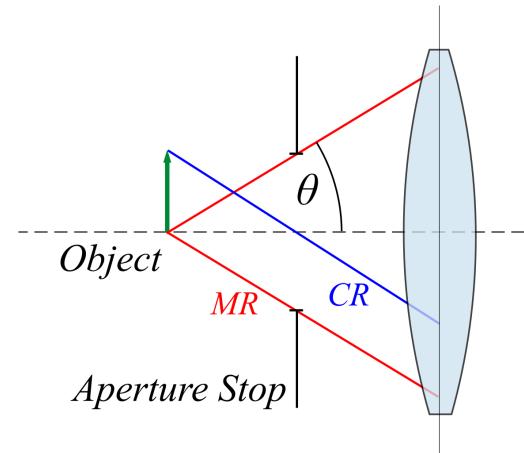


Figure 17.6: Aperture Stop, Marginal Ray and Chief Ray

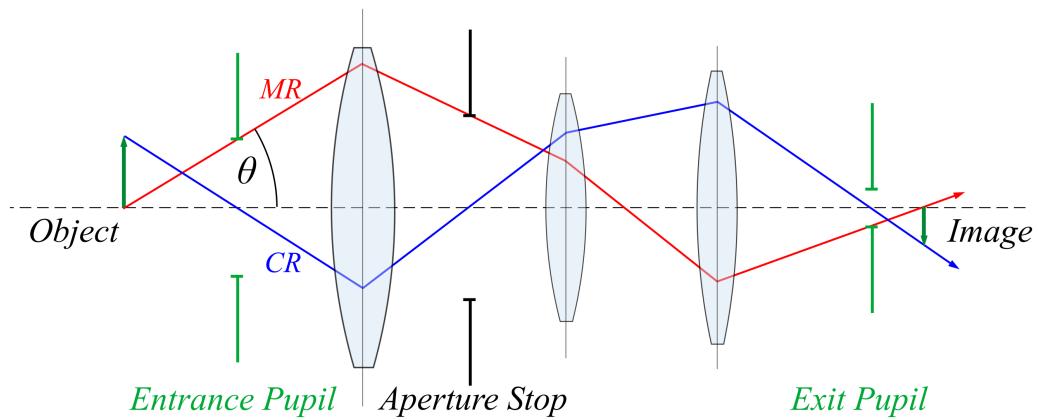


Figure 17.7: Aperture Stop and Pupils

chief ray crosses the optical axis, just like at the aperture stop.

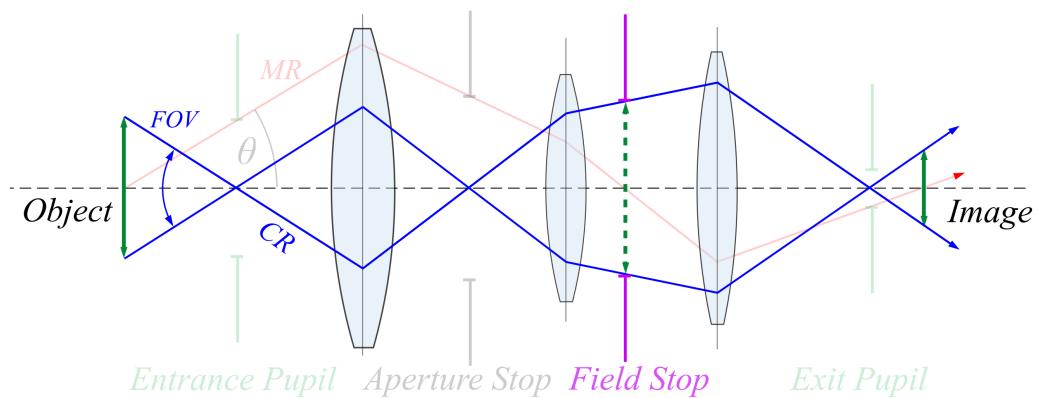


Figure 17.8: Field stop and angular field of view

17.5.4 Field Stop

Figure 17.8 has the previously discussed marginal ray, aperture stop and pupils greyed out. The **field stop** is the element in the optical system which limits the angle of incidence of the chief rays. It has to be placed at the location of an image, in the figure above it is at the intermediate image. In many systems, the sensor at the final image acts as the field stop. The marginal ray crosses the optical axis and the chief rays just touch the field stop, as one would expect at an image location. The two chief rays go through the centre of the aperture stop and the pupils. The angle subtended at the object gives the angular **field of view**. The larger the field stop the larger the field of view and more of the object can be imaged at once.

17.6 Field of view

The field of view is the size of the area imaged by the system, sometimes also given as the diameter of the circular area. It is determined by the magnification of the objective and the size of the field stop which is normally given by the **field number** of the eyepiece or size of the sensor used. A rectangular

camera sensor often images only a proportion of the full, round field of view.

$$\text{FOV}_{\text{Camera}} = \frac{\text{Sensor Size}}{M_o} \quad \text{FOV}_{\text{Eyepiece}} = \frac{\text{Field Stop}}{M_o} \quad (17.11)$$

The obtainable field of view is also limited by the objective but the maximum usable field number is often not easily obtained from manufacturers. Modern flat field objectives may reach field sizes of 28mm or more. [196]

17.7 Rear Aperture and Back Focal Plane

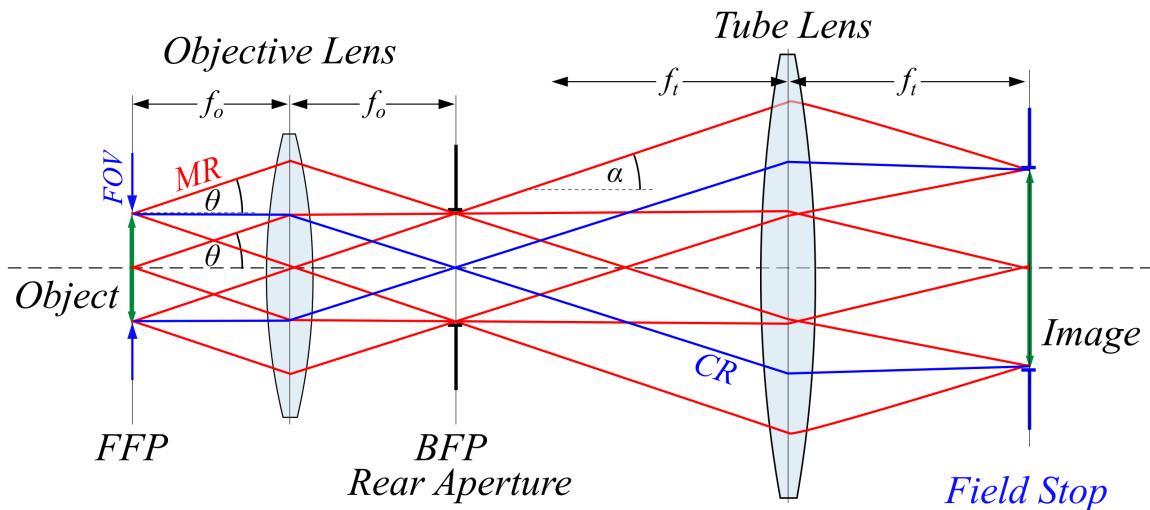


Figure 17.9: Rays in an infinite conjugate objective

Figure 17.9 shows a simplified ray diagram of an infinite microscopic system. All rays leaving the object which is situated at the **Front Focal Plane (FFP)** in parallel, are brought to a focus in the **Back Focal Plane (BFP)**. The **Rear Aperture** acts as the aperture stop of the objective and is placed at the BFP. Therefore, it also acts as the objective's exit pupil. As before, the angle θ of the red marginal rays determines the numerical aperture. In this ideal system, the NA is the same across the whole field of view which means the light gathering power is constant. This is because all rays that make it through the aperture stop are used to form the Image. If this was not the case, because the tube Lens is too small for example, the edges would appear darker and have less resolution, an effect called **vignetting**.

Having the aperture stop at the back focal plane causes the entrance pupil to be at infinity. This means the objective is **telecentric** in the object space. The chief rays are running parallel to the optical axis and intersect in the centre of the aperture, as expected. The **Field of View** is therefore of a fixed size, rather than an angle and Equations 17.11 become obvious. The field stop in this case is the size of the screen or sensor in the image plane.

The position of the BFP is often within the objective barrel and the diameter of the aperture d_{BA}

is not straightforward to calculate. From the diagram, one might expect that

$$\tan(\theta) = \frac{d_{BA}}{2f_o} \quad d_{BA} = 2f_o \tan(\theta) \quad (17.12)$$

However, the thin lens approximation breaks down in this case and it turns out that it is actually

$$d_{BA} = 2f_o NA = 2f_o n \sin(\theta) \quad (17.13)$$

Basically, this is because objectives which are designed to reduce coma, spherical and other aberrations fulfil small angle approximations even for large angles so that $\sin(\theta) \approx \tan(\theta)$. [197] Also note that contrary to what the diagram might suggest the maximum angle of divergence of a parallel ray bundle $\alpha \neq \theta$, the maximum angle of acceptance. What is true however, is that that an angular movement at the BFP of the objective translates to a lateral movement along the FFP according to $h = \tan(\alpha)f_o$ so that

$$\tan(\alpha) = \frac{\text{FOV}}{2f_o} = \frac{N}{2f_o M} = \frac{N}{2f_t} = \frac{d_{TL} - d_{BA}}{2L} \quad (17.14)$$

where N is the maximum field number as defined in section 17.6, M is the objective magnification from Equation 17.4, d_{TL} is the diameter of the tube lens entrance pupil and L is the distance between the BFP and the tube lens as in Figure 17.9.

This can be rearranged to give the diameter of the bundle of rays emerging from the objective.

$$d_{TL} = \frac{N}{f_t} L + d_{BA} \quad (17.15)$$

Equation 17.15 is useful in designing our optical system because it helps to avoid cutting off rays at the edge of the field which would lead to vignetting.

17.7.1 Scanning and Epi-Illumination

For scanning applications like in Two-Photon or Confocal Microscopy as well as Two-Photon Polymerisation, see chapter 13 and chapter 8 respectively, a laser beam hits the back focal plane and is brought to a focus in the front focal plane, like in Figure 17.9 but going the other way. It is important that the aperture is completely filled with light, otherwise the effective numerical aperture is decreased. [198, Chapter 9] Equation 17.13 should be used to calculate the minimum beam diameter and Equation 17.14 the maximum scan angle to reach all of the objective's field of view.

17.8 Depth of Field

The **depth of field** is defined as the distance between the closest and furthest point of the object in the z direction that are in focus at the same time. This axial resolution, like the horizontal resolution, is determined by the numerical aperture of the objective. Since the point spread function is three-dimensional, the location of the focus is not precisely defined. A higher numerical aperture will reduce the depth of focus. This can be useful in 3d imaging application or two-photon polymerisation.

Depth of focus refers to the location where the image is formed. It depends on the numerical aperture as well as the magnification of the objective. If it is small, the effects of the image sensor being slightly out of focus will be more severe than when it is larger. [195]

The real depth of field depends on a range of parameters but there are two main contributions. One is distance between the first minima of the three dimensional airy disk, the physical depth of field. The second is the geometric depth of field which depends on the resolution of the detector. The total is the sum of both terms as given:

$$d_{\text{tot}} = \frac{\lambda n}{\text{NA}^2} + \frac{n}{M \text{NA}} e \quad (17.16)$$

Where e is the smallest distance resolved by the detector, λ is the imaging wavelength, M is magnification, and n the refractive index of the imaging medium. The depth of field depends on the method of illumination and may be considerably smaller than Equation 17.16 suggests. [196]

17.9 Light Gathering Ability

Again, the light gathering ability of a microscope objective is primarily determined by its numerical aperture. This is intuitive, a higher NA means more oblique angles can be captured and therefore more light is used to create the image. A higher magnification objective, with the same NA, will capture a smaller FOV and spread it over the same distance at the detector leading to a reduction in brightness. These two effects are captured in equations 17.17

$$\text{Brightness}_{\text{Transmitted}} \propto \left(\frac{\text{NA}}{M_O} \right)^2 \quad \text{Brightness}_{\text{Epi}} \propto \left(\frac{\text{NA}^2}{M_O} \right)^2 \quad (17.17)$$

When the objective is used to illuminate the sample as well as imaging it (Epi-Illumination) the Numerical Aperture is even more important. Other factors including the absorption in lenses and transmissivity of reflective coatings also influence the objective's light gathering ability.

17.10 Choices and options

For the many different use cases of our system we will need to offer a range of objectives that the surrounding optics are designed to work with. Specialized objectives are available and required for some aspects.

17.10.1 Confocal, STED, Brightfield

For normal bright field observations and STED, we want high NA objectives which are directly immersed into the oil or the polymer fluid which has a similar refractive index to the oil. There will be no cover glass and the objective choice reflects this. For effective digital image analysis, it should be flat-field and well corrected which makes plan apochromat objectives the ideal candidates. A high magnification one for detailed work and a low magnification one for a large field of view will be useful for fast-nano. These type of objectives will be referred to as "Universal" and "Low Mag" respectively.

17.10.2 Two Photon Microscopy, Two Photon Polymerisation, Micro Machining

Ideal objectives for multiphoton applications are made of special glass with high ultra violet to low infra-red transmission characteristics as well as low self fluorescence. They are often fluorite objectives without planar correction which makes them non-ideal for brightfield observation. Since we will not use low infra-red, just up to 800nm, we do not need special multi photon objectives which often transmit up to 1600nm. The same universal objectives as for brightfield above can be used, as long as the transmission up to 800nm is adequate.

17.10.3 Polarized Light

For polarized light observations special objectives which minimize the strains on the elements are required. They are normally fluorite objectives with as few elements as possible and therefore not well corrected. One high magnification polarized light should suffice for our requirements. They are normally designed to be used with cover glasses but since the refractive indexes of the oil and polymer fluids are matched to the glass they will work without introducing aberrations.

17.10.4 Non Oil Immersion

Being able to work and image with different immersion media will extend the versatility of the system. A water immersion lens will be used for optical tweezing applications and micro fluids experiments because the refractive index is sufficiently different from the polarized particles as explained in section 9.3. Water is also useful as a solution to study living cells in. A dry objective will be required for use without fluids. Initial brightfield and confocal imaging with less resolution can be done quickly this way.

17.10.5 Manufacturer selection

While it is possible to use objectives from different manufacturers, doing so would come with a few drawbacks. Firstly, the non parfocality means that either we need to reset the stage to different heights with every objective change which would increase the fabrication time or we use parfocal extenders—adapters which are put between the turret and the objective. Secondly, different screw thread sizes although they can also be changed with simple adapters. In both cases, adapters change the location of the back focal plane considerably which has a negative influence on the surrounding optical systems such as the scanning and the 4f distances. Thirdly, in many cases tube lenses are designed to work with the objectives of a particular manufacturer.

I will be focussing on **Olympus** and **Nikon** because they do all the correction within the objectives themselves, rather than using tube lenses for corrections like Leica and Zeiss do.

17.10.6 Objective selection

Table 17.2 has a comparison of relevant objectives for the two possible suppliers. The second row contains important specifications include the numerical aperture (NA), the working distance (WD), the maximum usable field number (FN), and the specified cover glass thickness (CG), all in millilitres. The third row contains calculated values for the maximum field of view (FOV/mm) and resolution (R/nm), using Equation 17.11 and Equation 17.5 respectively. The last value is a figure of merit (M) which is calculated from the FOV, the resolution and the working distance, and important trade off. FOV and R are weighted roughly equally and WD half as much by using $M = 100\text{FOV} + 0.1R + 50 * WD$. The larger values are highlighted in green, the smaller ones in red. In many cases, the resulting figures are similar with Olympus having a slight advantage due to a marginally larger field number.

Looking at the "Low Magnification Oil Immersion" category, there is no Nikon equivalent to the **UPLSAPO 20XO**. The smallest oil immersion is 60X with an obviously much reduced FOV.

I was only able to find transmission statistics for the **MPLAPON 50X** up to 700nm and we can therefore not guarantee it working with our IR laser.

The **MPLFLN 40X** is only loosing to the **CFI S Plan Fluor ELWD 40X** because of the reduced working distance which may not be critical for dry imaging. The **LUCPLFLN 40X** has a longer working distance with slightly less correction both these objectives will work well.

The **XSLPLN 25XGMP** is an interesting dedicated multi-photon objective with a huge working distance and excellent transmission. However, it has a longer parfocal distance (75 rather than 45) and a different thread than the other Olympus Objectives. Further, the relatively small field number 18 limits the obtainable FOV.

In the water immersion sections, Nikon's objectives have a slight advantage with longer working

Table 17.2: Objective comparison. The recommended choices are highlighted in green.

Use	Olympus	Nikon
Universal Oil	UPLSAPO 100XO NA1.4, WD0.13, FN26.5, CG0.17 FOV0.3 R349 M68	CFI Plan Apo Lambda 100X Oil NA1.45, WD0.13, FN25, CG0.17 FOV0.3 R337 M65
Universal Oil	MPLAPON 100XO NA1.4, WD0.1, FN26.5, CG0 FOV0.3 R349 M66	CFI Plan Apo NCG 100X Oil NA1.4, WD0.1, FN25, CG0 FOV0.3 R349 M65
Low Mag Oil	UPLSAPO 20XO NA0.85, WD0.17, FN26.5, CG0 FOV1.3 R574 M198	CFI Plan Apo Lambda 60X Oil NA1.4, WD0.13, FN25, CG0.17 FOV0.4 R349 M83
Polarized Oil	UPLFLN 100XOP NA1.3, WD0.2, FN26.5, CG0.17 FOV0.3 R375 M74	CFI Plan Fluor 100X Oil NA1.3, WD0.16, FN25, CG0.17 FOV0.3 R375 M71
High NA Oil	UAPON 60XO TIRF NA1.49, WD0.1, FN22, CG0.13-0.19 FOV0.4 R328 M74	CFI Apo SR TIRF 100X Oil NA1.49, WD0.12, FN25, CG0.17 FOV0.3 R328 M64
2PM Silicone-Oil	XSLPLN 25XGMP NA1, WD8, FN18, CG0 - 0.23 FOV0.7 R488 M521	-
No CG Dry	MPLFLN 40x NA0.75, WD0.63, FN26.5, CG0 FOV0.7 R651 M163	CFI S Plan Fluor ELWD 40X NA0.6, WD3.6, FN25, CG0 FOV0.6 R813 M324
LWD Dry	LUCPLFLN 40X NA0.6, WD4, FN22, CG0-2 FOV0.6 R813 M336	CFI S Plan Fluor ELWD 40X NA0.6, WD3.6, FN25, CG0 FOV0.6 R813 M324
CG Dry	UPLSAPO 40X2 NA0.95, WD0.18, FN26.5, CG0.11-0.23 FOV0.7 R514 M127	CFI Plan Apo Lambda 40X NA0.95, WD0.25, FN25, CG0.11-0.23 FOV0.6 R514 M126
High NA Dry	LCPLN 100XIR NA0.85, WD1.2, FN22, CG0-0.7 FOV0.2 R574 M139	CFI LU Plan Apo 100X NA0.9, WD0.51, FN25, CG0 FOV0.3 R542 M105
Med Mag Water	LUMPLFLN 40XW NA0.8, WD3.3, FN26.5, CG0 FOV0.7 R610 M292	CFI Apo 40XW NIR NA0.8, WD3.5, FN25, CG0 FOV0.6 R610 M299
High Mag Water	LUMPLFLN 60XW NA1, WD2, FN26.5, CG0 FOV0.4 R488 M193	CFI APO 60XW NIR NA1, WD2.8, FN25, CG0 FOV0.4 R488 M230

distances and better corrections. Both the 40x and 60x will work well however and the **LUMPLFLN 40XW** was chosen because of a larger field of view in HOT applications. (See section 9.2)

Overall, Olympus offers the better selection but it is the advantage of the 20X Plan Apochromat Oil immersion that makes the main difference. The recommended objectives have been highlighted in green although customers may choose to employ a different set from this list, depending on their

requirements.

Table 17.3: Selected Objectives use cases and characteristics.

	Universal Oil MPLAPON 100XO	Low Mag Oil UPLSAPO 20XO	Polarized Oil UPLFLN 100XOP	Water LUMPLFLN 40XW	Dry LUCPLFLN 40X
Fab	Stereo Lithography				Seperate F-Theta Lens
Micromachining	✓	✓			✓
Two-Photon Poly	✓	✓			
HOT				✓	
Image	Brightfield	✓	✓	✓	✓
Polarized			✓		
Confocal	✓	✓		✓	✓
Two-Photon Micro	✓	✓			
STED	✓				
Characteristics	NA	1.40	0.85	1.30	0.80
WD (mm)	0.10	0.17	0.20	3.30	4.00
FN (mm)	26.5	26.5	26.5	26.5	22.0
f_O (mm)	1.80	9.00	1.80	4.50	4.50
BFP diameter (mm)	5.04	15.30	4.68	7.20	5.40
FOV diameter (mm)	0.27	1.33	0.27	0.66	0.55
α (degrees)	4.21	4.21	4.21	4.21	3.50
Resolution (nm)	349	574	375	610	813
Make: Olympus, Parfocal d: 45mm, Thread: RMS, 400-800nm Transmission > 80%					

Table 17.3 lists the fabrication and imaging techniques in which the selected objective lenses will be used as well as their calculated characteristics. The headings are defined in subsection 17.10.6. The objective focal length (f_O) is calculated using Equation 17.4, the back focal plane / exit pupil diameter (BFP) with Equation 17.13, the maximum field of view diameter using Equation 17.11 and the resolution (RES) using Equation 17.5. The maximum angle of divergence (α in degrees) is important for the requirements of the scanning system and comes from Equation 17.14.

17.11 Conclusion

The microscope objective lens as one of the most complex and important parts of every optical system are our "eyes into the nano world". By understanding the classifications and specifications of objectives as well as analysing how they effect key system characteristics we can select a versatile set which satisfy the multitude of requirements. Designing the processes around these characteristics means we provide a highly optimised solution to our clients.

18 4f

The 4f imaging system is a key optical component in our system. It provides a method for producing an image in another plane, as well as magnifying or demagnifying the image. The configuration can also be used to remove spatial frequencies but, in the case of our system, this is not used. The 4f system consists of finite conjugates — one focal distance to the left of the objective lens and one focal distance to the right of the collector lens respectively, see Figure 18.1.

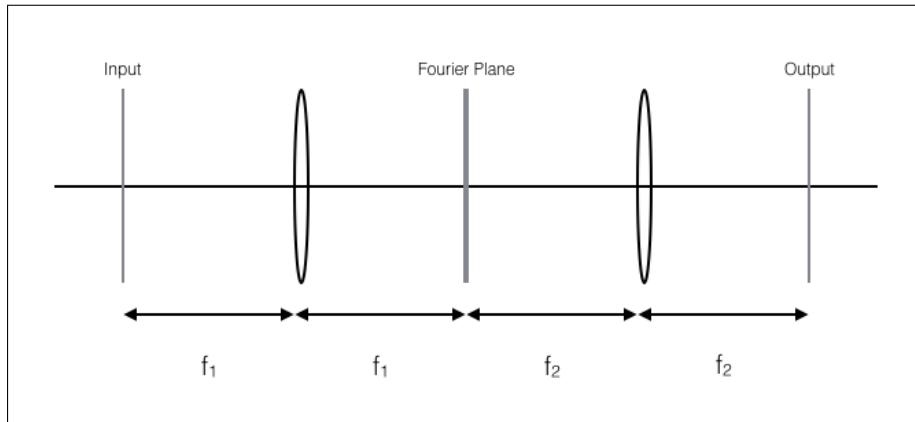


Figure 18.1: The 4f imaging system's spatial arrangement

From the Fourier transforming property of lenses derived in subsection 3.2.3, it can be seen that the 4f system produces a cascade of two Fourier transforms, thus producing the input image in another plane. It must be noted from Fourier transform theory that taking the Fourier transform twice produces an inverted image Equation 18.1.

$$\mathcal{F}[\mathcal{F}[g(x, y)]] = g(-x, -y) \quad (18.1)$$

18.1 Magnification and Demagnification

As previously described, the 4f system provides a method for magnifying and demagnifying the input. Using Figure 18.2 and a geometrical optics view, the magnification capabilities of the 4f can be shown. Light rays arrive from a distant point in parallel. It can be deduced that at a distance f_o from the first lens the focal point is a distance d from the centreline. Using simple geometry,

$$\theta_o \approx \frac{d}{f_o} \text{ and } \theta_e \approx \frac{d}{f_e} \quad (18.2)$$

The transverse magnification is then

$$M = \frac{\theta_o}{\theta_e} = \frac{d/f_o}{d/f_e} = \frac{f_e}{f_o} \quad (18.3)$$

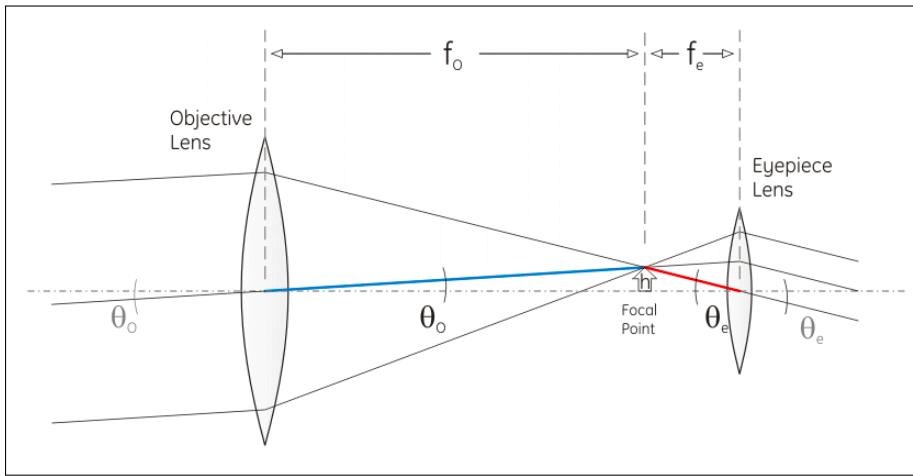


Figure 18.2: Geometrical layout of the magnification of the 4f imaging system.[199]

The ability to magnify or demagnify the input is extremely useful and section 18.4 further describes how it is utilised in our system.

18.2 4f as a Spatial Filter

The 4f system can also be used to implement spatial filtering. This is executed by introducing a transparency at the Fourier plane. Spatial frequencies which hit opaque portions of this transparency are removed from the output. The transparency may be grayscale (partial block) or a phase mask (introducing a phase delay). The system is able to produce high quality filters as the masks can be sharp edged thus producing sharp roll offs. Despite the remarkable spatial filtering abilities of the 4f, it was decided to not include spatial filtering within the 4f system. The reasons for this is that our project brief is to design a versatile system. Despite it producing effective results, the number of masks and mechanical setup (in order for it to be versatile) would require a complicated and bulky layout. In addition, the filtering can actually be produced with the SLM and using post imaging processing both of which have greater adaptability.

18.3 4f vs a Single Lens

This section provides an analysis as to why using the 4f system is of greater advantage than using a single lens imaging system. Starting with Figure 18.3 which shows the 4f system and single lens system with a clear aperture. Using geometrical optics, the numerical aperture is defined as the angle that is subtended towards the optical system if a point source is placed on axis. The numerical aperture, therefore, is limited by the physical apertures of the system. In the case of the 4f system in Figure 18.3, the pupil mask is the limiting factor. The numerical aperture can be computed as the radius R of the mask over the focal length f_1 . This is only an estimation due to the paraxial approximation defined in Fourier optics, subsection 3.2.3. In the case of the single lens system the numerical aperture is R over z_1 . If the single lens system is to produce a real image (as apposed to

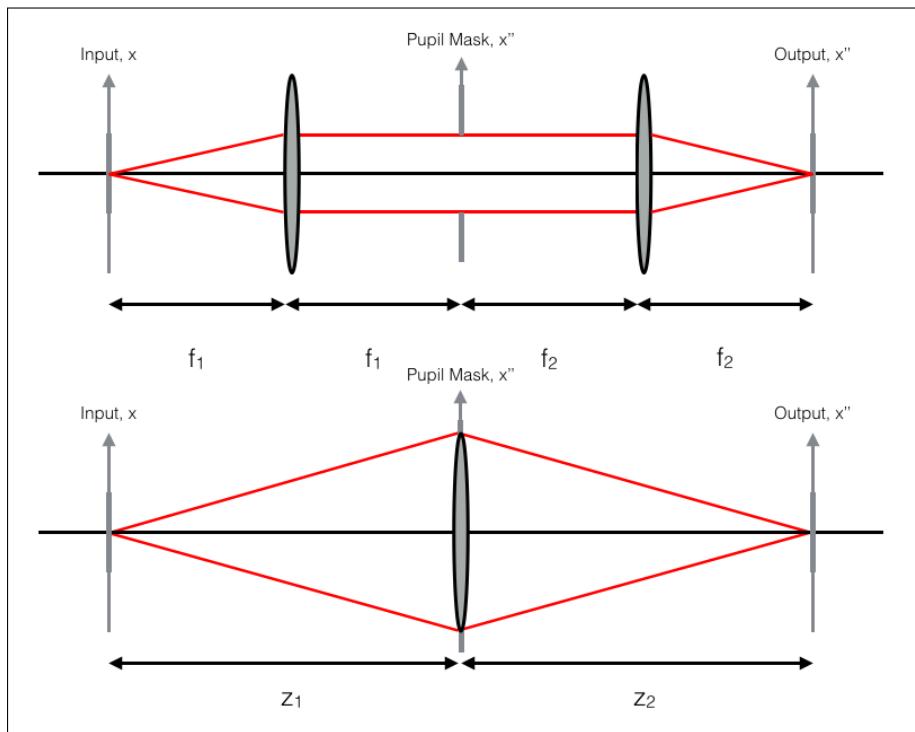


Figure 18.3: 4f system (top) and single lens (bottom) arrangement with a clear aperture

a virtual image) then z_1 must be larger than f_1 . The reason for this can be seen from the imaging equation, which is derived in section 3.1.

$$\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{f} \quad (18.4)$$

For a real image z_2 must be positive. It is therefore required for z_1 to be greater than f . If you have the same sized pupil mask with the same lenses, the single lens system will have a smaller numerical aperture.

The reason for numerical aperture being important is that it has an affect on the size of the point spread function, see section 3.3 for further explanation. If you consider a circular pupil, then the PSF of the system is the Fourier transform of the pupil function. Referring back to section 3.3, the Fourier transform of a circular pupil is a Jinc function.

$$h(x', y') = \frac{2J_1(r)}{r} \quad (18.5)$$

For both cases the shape of the PSF is the same but different in size. This arises because of a scaling factor which depends on the distance between the input and the lens. The same physical apertures, therefore, will give different sizes in the amplitude transfer function. The 4f amplitude transfer function will be proportional to $\frac{1}{\lambda f_1}$ and the single lens amplitude transfer function will be proportional to $\frac{1}{\lambda z_1}$. The larger amplitude transfer function will give a narrower PSF and therefore less

blur.[200] In the case of our system, we need to minimise the blur. Given the physical size of the aperture, the 4f is maximising the numerical aperture and thus producing a sharper image.

Secondly, imaging with the 4f is preferable as it produces a flat phase wavefront. As indicated in subsection 3.2.3, the general geometry (when there is some input a distance d from a lens) produces the 2D Fourier transform of input multiplied by a quadratic phase factor.

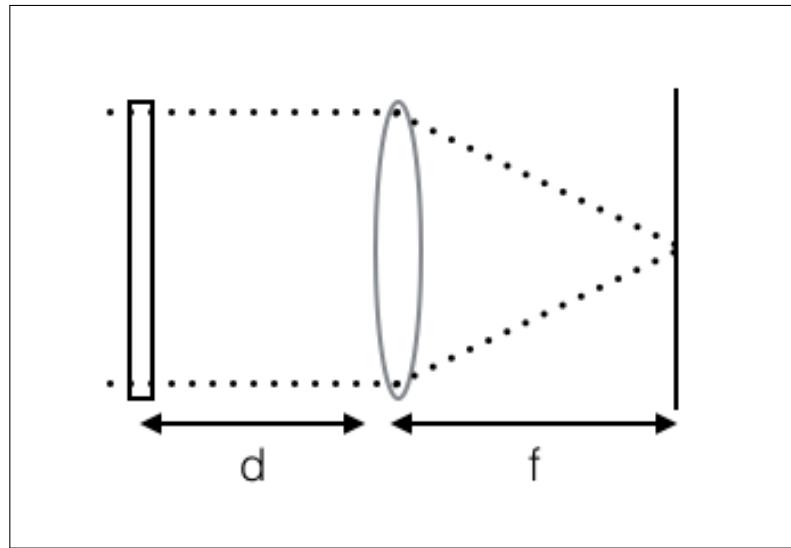


Figure 18.4: Geometry for performing the Fourier transform on a positive lens when the input is a distance d from the lens.

$$U_f(u, v) = \frac{-j}{\lambda f} \exp \left[\frac{jk}{2f} \left(1 - \frac{d}{f} \right) (u^2 + v^2) \right] F_0 \left(\frac{u}{\lambda f}, \frac{v}{\lambda f} \right) \quad (18.6)$$

Note that when $f = d$ the phase curvature disappears leaving the exact Fourier transform multiplied by a constant. As this distance d is always set to equal f within the 4f system, it will always produce flat phase imaging — regardless of whether it is magnifying or demagnifying.

$$U_f(u, v) = \frac{-j}{\lambda f} F_0 \left(\frac{u}{\lambda f}, \frac{v}{\lambda f} \right) \quad (18.7)$$

18.4 4f Implementation

Within our optical system, there are two 4f imaging systems, which are highlighted in Figure 18.5. The first is **L7** and **L8**, which images the output of the **SLM1** onto the galvonometer scanning mirrors **X1** and **Y1**. As well as producing the output of the SLM in a new plane, this system also de-magnifies the output. The beam waist at the SLM is 5mm. To ensure the galvonometers have maximum effect on the field view of the objective, a demagnification of 2.5 was chosen, giving a beam waist of 2mm. Therefore **L7** has a focal length of 500mm and **L8** equals 200mm.

The second 4f system is composed of lenses **L12** and **L13**, which images the output of galvome-

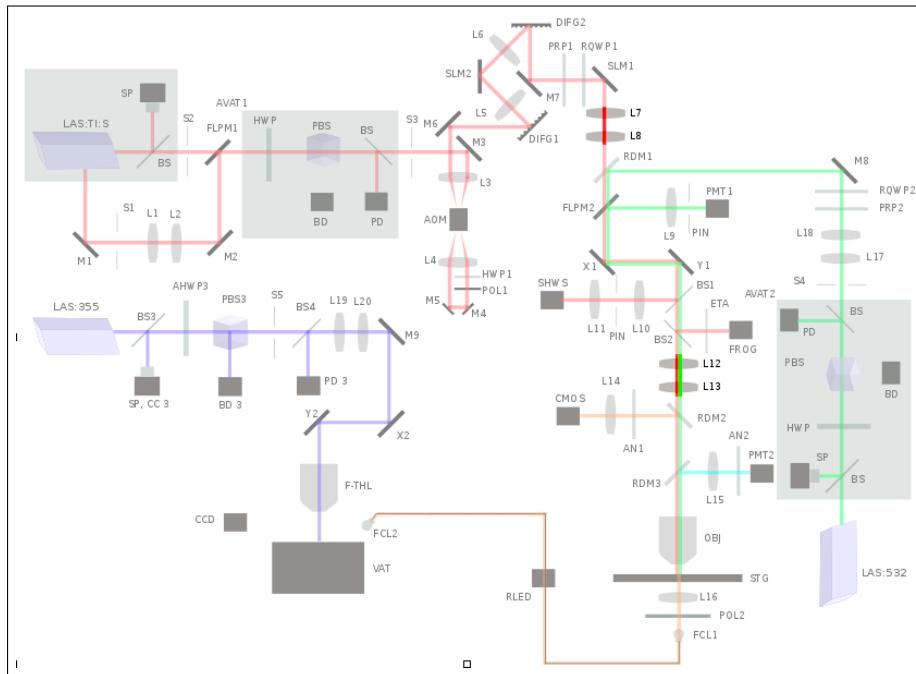


Figure 18.5: Optical Layout - The first 4f system is lenses **L7** and **L8**. The second 4f system is lenses **L12** and **L13**

ters **X1** and **Y1** onto the objective lens **OBJ**. In order to fill the objective, a magnification of 5 was chosen to increase the beam waist from $2mm$ to $10mm$. Therefore **L12** has a focal length of $150mm$ and **L13** has a focal length of $750mm$.

18.5 Summary

The 4f provides a flat phase method for imaging the output of the SLM onto the galvonometers and from the galvonometers onto the objective. Further it provides a method of magnifying and demagnifying the beam waist to ensure the galvonometers have a maximum field of view of the objective.

19 Aberrations

Aberration identification and correction is essential for various reasons, such as minimal dispersive delay, multiple foci generation at image planes, as well as beam control. Aberrations usually involve a spatio-temporal distortion of the beam around the focal region, and can result in poor machining quality, as well as poor resolution. Aberrations can usually be classified either as monochromatic aberrations, which are independent of incident optical frequency, such as spherical aberrations, astigmatism, and image distortion, or chromatic aberrations, where aberrations are dependent on incident optical frequency, such as chromatic focal shift. This chapter outlines the different methodologies used to detect and preserve the quality of the beam and the image plane through the optical system.

19.1 Achromatic Lens Design

Achromatic lenses (or doublets) are compound lenses. The generic doublet is a compound lens composed of two lenses (and hence 3 – 4 independent surfaces). The first (crown) lens is a convex lens made of a light crown glass material. In contrast, the second (flint) lens is a concave lens made of a thicker, dispersive flint glass, as shown in Figure 19.1.

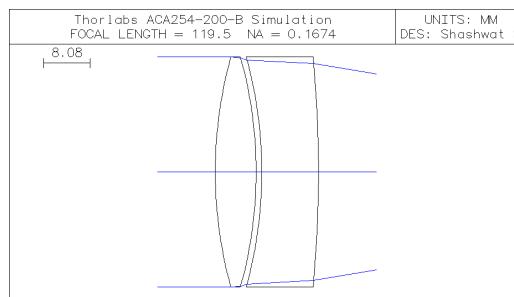


Figure 19.1: Lens Plan of a Designed Achromatic Doublet

Doublets have been used to minimize chromatic aberrations, or focal shifts due to the transmission spectrum. This effect is noticeable in ultrashort pulses due to the relatively large bandwidth of light used to generate the pulse waveform, as described in section 19.2.

19.1.1 Implementation and Supplier Choice

The chromatic focal shift over a set of wavelengths is caused by the non-linear dependence of the refractive index of a medium on wavelength. This inherently empirical relation is modelled over the regions of interest using the general Sellmeier equation [201]:

$$n^2(\lambda) = 1 + \sum_i \frac{A_i \lambda^2}{\lambda^2 - B_i} \quad (19.1)$$

where A_i, B_i are Sellmeier coefficients, empirically determined for $i = 3$. Based on Seidel aberration theory [202], adding air spacing between component lenses allows an extra surface redundancy

compared to a cemented doublet, which is utilized to eliminate other optical aberrations. This type of doublet, known as the Fraunhofer doublet, is manufactured by Thorlabs and selected for use to mitigate monochromatic aberrations. An example of a simulated ray trace analysis (RTA) between air-spaced and cemented doublets for focal length $f = 200$ mm provided by Thorlabs, is shown in Figure 19.2, Figure 19.3. The simulated RTA allows a comparison of the spatial distortion of focal regions in mm, as a function of the incident wavelength in μm .

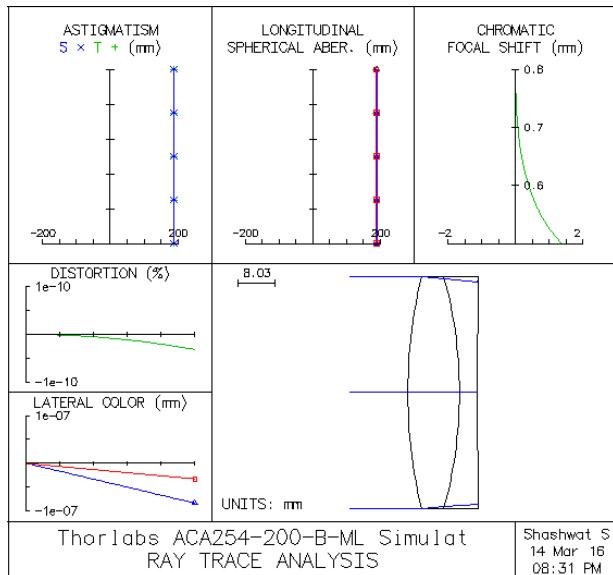


Figure 19.2: Cemented Doublet RTA

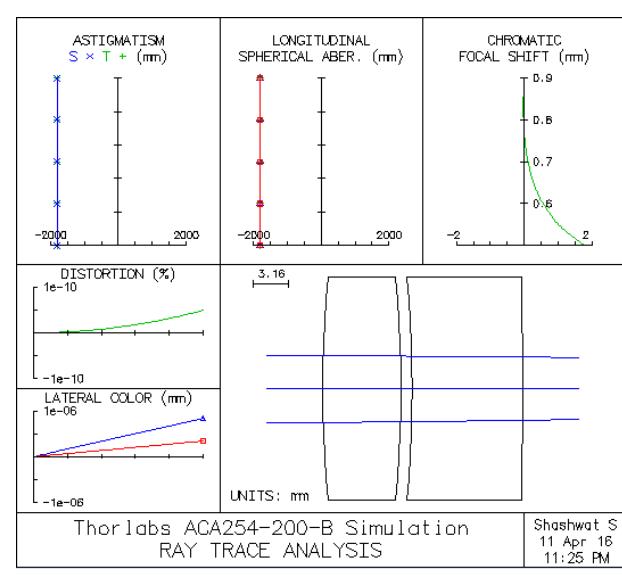


Figure 19.3: Air-spaced Doublet RTA

The comparison of Figure 19.2, Figure 19.3, prioritizes the extinction of various aberrations. The air-spaced doublets experience less lateral and axial chromatic focal shift across a large wavelength range, with negligible focal shift at 500 nm as compared to cemented doublets, with about 1.5 mm focal shift at 500 nm. Overall, both types of doublets account for uniformity in spherical aberrations and astigmatism, while distortion is minimized, at less than 5×10^{-11} . Hence, the system was designed using air-spaced doublets where possible. Unfortunately, the focal range of catalogued air-spaced doublets is limited; where a higher focal length was required, such as the magnification for the two 4F systems, cemented doublets with higher focal lengths were used. Although higher focal lengths usually lead to greater distortions for singlet lenses, the doublets chosen minimize the possibility of aberrations, as illustrated earlier.

Component price was also a key criterion in the selection process; suppliers such as Edmund Optics provide air-spaced doublets priced at USD 1095 while Thorlabs's doublets are priced at USD 475. Based on the cost for 20 such lenses and the unavailability of adequate anti-reflective coatings in Edmund Optics, Thorlabs was chosen as the primary lens supplier.

19.1.2 Lens Quality Control

All lenses were tested through simulations in Optics Software for Layout and Optimization (OSLO) to characterize performance. An example of the tests done is presented for a selected lens choice. The quality inspection for the lenses in the system setup was completed using key parameters such as the RTA to check for aberration extinction, and the modulation transfer function (MTF). An example of an RTA has already been presented in Figure 19.3.

The second measure of quality control used was the modulation transfer function (MTF). The MTF [203] defines the frequency response of the optical system to sinusoids of varying frequencies, which characterize the performance of the beam at sharp spatial changes in intensity, for example multiple foci, as explained in chapter 16. The MTF is mathematically defined for a particular frequency as $\max(\Delta I)/\max(\sum I)$. The MTF is a measure of the contrast that the lens maintains for sharp changes in spatial intensity; a well-behaved diffraction limited MTF implies that a high contrast ratio is maintained. Furthermore, the MTF is related to the point spread function (PSF), according to [204], as $MTF = FT[PSF]$, where the PSF is explained further in section 3.3. The MTF of Thorlabs ACA254-200-B is then shown in Figure 19.4.

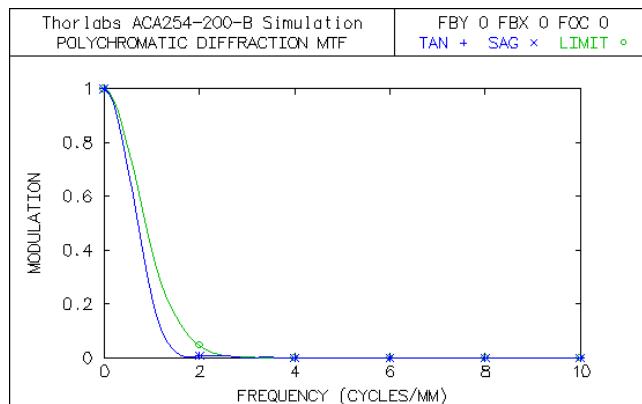


Figure 19.4: Polychromatic MTF Plot for ACA254-200-B

From Figure 19.4, the diffraction limit MTF in green illustrates the best possible MTF obtainable from the lens. The MTF plot shows that the Thorlabs doublet performs close to the diffraction limit. This MTF plot was then compared with cemented doublets, as well as singlet lenses, to maintain the performance limit.

19.1.3 Material Properties

A variety of crown and flint glasses can be determined and compared using a glass catalog, for example by SCHOTT, which compares each glass on a refractive index-Abbe Number plot. The Abbe Number (V_e) is defined as the following:

$$V_e = \frac{n_e - 1}{n_{F'} - n_{C'}} \quad (19.2)$$

where 'e' represents the centre wavelength of the lens, n represents the refractive index, where 'F', 'C' represent the respective Fraunhofer (H spectrum emission) lines at wavelengths 479.99 nm and 643.85 nm, respectively. V_e can then be used as a measure of the reciprocal of dispersion [205], and the respective materials can be used as an extra design parameter.

As light passes through a lens, there is a component that is either reflected or absorbed at each lens surface, due to total internal reflection at a surface. To reduce back-reflections for a specific wavelength range, the end surface of doublets were coated with an anti-reflective (AR) coating with ranges of 350 – 700 nm (A), 650 – 1050 nm (B). The B coating was used for lenses in the path of the Ti:sapphire laser, i.e. L1 – 11, and the A coating was selected for lenses L14 – 18. For lenses L19, 20, separate doublets with UV fused silica were chosen. Two lenses were requested without a coating for the 4F lens system **L12,13**, to accommodate for the optical system path, shown in Figure 4.1. The path was shared by the 532 nm diode-pumped solid state (DPSS) laser, and the 800 nm Ti:sapphire laser; however, neither AR coating fits the path specifications. The transmission specification of the lens is then determined mainly by the AR coating, as shown in Figure 19.5, Figure 19.6, provided by Thorlabs. As shown in Figure 19.5, an uncoated borosilicate crown lens (NBK7) transmits > 90%

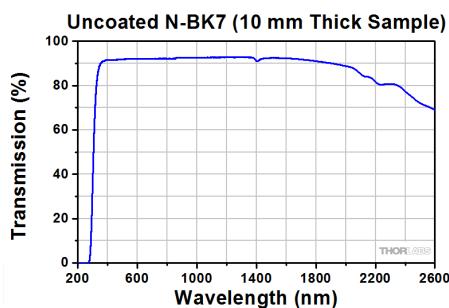


Figure 19.5: Uncoated Lens Transmission

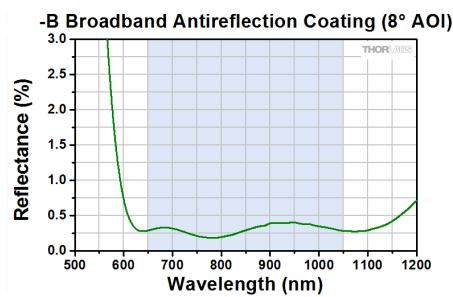


Figure 19.6: -B AR Coating Graph

light at both 532 nm and 800 nm. Similarly, the flint lens will also be able to transmit light at both wavelengths, provided there is no AR coating. With the specified customization, the -B coated Thorlabs doublet (label #ACA-254-200B) was then compared with an Edmund Optics cemented doublet, label #49-332-INK. The Thorlabs doublet specified above has been chosen as the most common doublet used in the system. The Edmund Optics lens is coated with a coating ranging from 400 – 1000 nm, and the edges of the lens were inked to avoid distortion. Then, Figure 19.3, Figure 19.4 were compared with Figure 19.7, Figure 19.8 to choose the type of lens.

The performance limits of the customized ACA254-200-B are the same as those specified in Figure 19.3, Figure 19.4. In comparison, Edmund Optics' 49-332 performs worse on chromatic aberra-

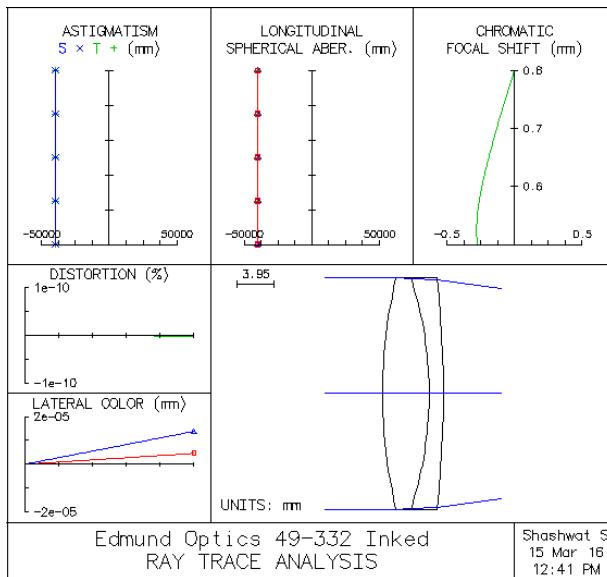


Figure 19.7: Edmund Optics 48-332 RTA

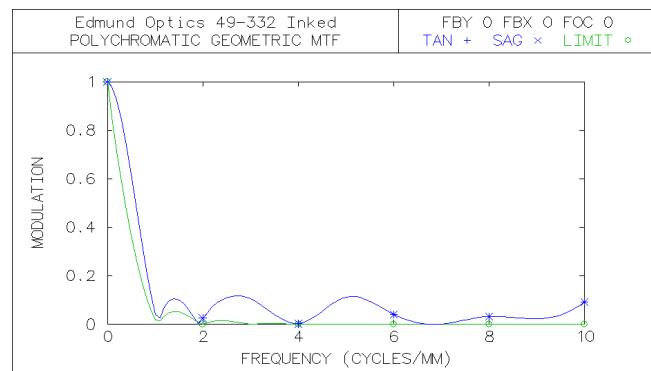


Figure 19.8: Edmund Optics 49-332 RTA

tion correction (as illustrated in the top right of Figure 19.7), and the MTF of 49-332 is marginally more dispersive, compared to the MTF of ACA254-200-B. These two reasons, as well as the benefits of selecting one lens supplier, mean that the customized Thorlabs lenses were also used for the anomalous case of **L12,13**, allowing the system to merge the STED and confocal microscopy functionality with the fabrication setup.

19.1.4 Summary

Achromatic doublets are versatile and well-known optical setups that reduce spatial aberrations for both continuous wave and pulsed lasers. Given the range of capabilities expected from the system, the achromatic doublets were chosen with specific performance metrics to reduce all spatial aberrations. However, despite this treatment of optical aberrations, doublets introduce further aberrations relating to the temporal model of the beam, which must be rectified by an alternative setup.

19.2 Ultrashort Pulse Maintenance

Beam propagation analysis shows that temporal pulse spreading is also a major aberration for fabrication purposes. Since both laser micro-machining and two-photon polymerization require specific pulse widths, it was important to ensure that the pulse width at the sample was regulated to a specified value.

A pulse with pulse width τ is formed by the superposition of sinusoids of a finite set of frequencies. The amplitude and phase of these sinusoids determines the shape and the width of the pulse, and so the time-bandwidth product (TBP) of a particular pulse is constant for a specific pulse shape. Given that the Ti:sapphire pulse shape is Gaussian (from the Coherent data sheet), then $TBP = 0.441 = \tau \Delta f$ [206]. Since $\Delta\lambda = \lambda^2 \Delta f / c$, the maximum spectral bandwidth $\Delta\lambda = 20$ nm for a minimum

pulse width $\tau = 50$ fs at full width, half maximum (FWHM). Since the spectral bandwidth is at just one order of magnitude less than the wavelength, pulse dispersion must be maintained using various techniques.

19.2.1 Group Velocity Dispersion and Chirp

The dispersion of a pulse occurs in media due to the varying group velocities of EM waves within the spectral bandwidth, a direct effect of chromatic aberrations. In the system, the dispersive flint lens is the primary cause of uncontrolled pulse dispersion, which must be rectified. The dispersion is quantified by using the rates of changes of phase φ with angular frequency ω . However, since $\Delta\varphi$ represents the wave-number k , the differentials of k produce (19.3), (19.4).

$$k' = \frac{\partial k}{\partial \omega} = \frac{1}{v_g} \quad k'' = \frac{\partial^2 k}{\partial \omega^2} = \frac{\lambda^3}{2\pi c^2} \frac{d^2 n}{d\lambda^2} \quad (19.3, 19.4)$$

(19.4) is also known as the group velocity dispersion (GVD) of the specific medium. The GVD helps quantify the degree of chromatic dispersion that a pulse experiences through the medium. The change in pulse width [207] and amplitude [208] are related as the following:

$$\frac{\tau_d}{\tau_0} = \sqrt{1 + \left(4 \ln 2 \frac{k'' z}{\tau_0^2}\right)^2} \quad \frac{I_d}{I_0} = \frac{\tau_0}{\sqrt[4]{\tau_0^4 + (4 \ln 2 k'' z)}} \quad (19.5, 19.6)$$

where the change in amplitude depends on the propagation distance z . To combine these expressions into a model for dispersion, the chromatic aberration effect acts as a convolution to the temporal profile of the pulse, implying that in the frequency (Fourier) domain, these effects can be shown to affect the pulse as shown in (19.8) [208].

$$I_{orig}(t) = I_0 \exp\left(-2 \frac{t^2}{\tau_0^2}\right) \quad I_{disp}(z, t) = I_d \exp\left(-2 \frac{(t - k' z)^2}{\tau_d^2}\right) \quad (19.7, 19.8)$$

where τ_0 represents the original pulse width. The dispersed pulse from (19.8) depends on the GVD of the medium. The GVD of three different materials in the system has been presented in Table 19.1.

Medium	k'	GVD at 800 nm (fs ² /mm)
N-BK7	5.1×10^{-9}	44.65
N-K5	5.2×10^{-9}	51.08
N-SF57	6.5×10^{-9}	219.09
Air	3.3×10^{-9}	0.02

Table 19.1: GVD of Different Media [209], [210]

Through the system, each pulse propagates through N-K5 (crown glass), N-SF57 (flint glass), and air. Since the Ti:sapphire pulse propagates through a maximum of 10 lenses at any point of time,

the primary source of dispersion occurs due to the achromatic lenses and the GVD of the N-SF57 glass. Alongside pulse spreading, the pulse also experiences a degree of attenuation, as the pulse energy is absorbed by the lenses. The degree of attenuation also depends on the pulse width, as shown in (19.6), Figure 19.10.

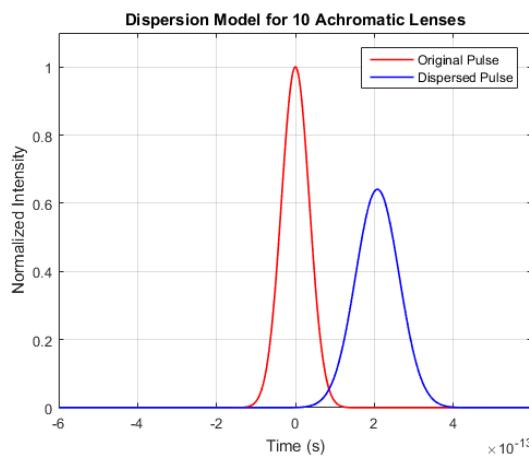


Figure 19.9: Pulse Model of 100 fs to 150 fs

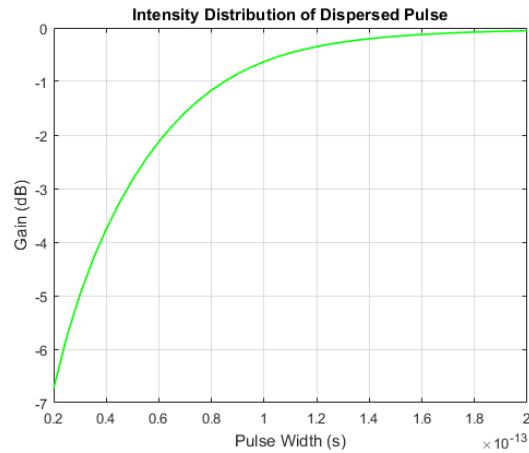


Figure 19.10: Dispersive Intensity Attenuation Simulation

From Figure 19.10, the intensity attenuation from the laser source to the objective can grow to 3 dB ($0.7I_{orig}$). In contrast, the pulse spreading over 10 achromats of about 17 mm thickness implies that, for low τ_0 , the minimum pulse width τ_d through the system is governed by the GVD term in (19.5). Without the pulse compressor, the dispersion effects limited the minimum pulse width to 150 fs, which does not satisfy the industry standard of 100 fs.

To achieve this industry standard, a pulse shaper, with specific information provided in section 20.8, was placed in the path, such that the compressed and compensated pulse from diffraction grating DIFG2, is dispersed by only 4 achromats. Then, the comparison of the pulse spreading has been presented in Figure 19.11, Figure 19.12, where the minimum pulse width achievable by the system is set to 94 fs.

From Figure 19.12, the pulse shaper further compresses the pulse to 50 fs, the minimum pulse width at the laser output. This provides system capability for laser micromachining and TPP at less than 100 fs. Bias chirp is often used in systems to compensate for pulse spreading, and hence the use of the pulse shaper to compensate for the pulse width is a useful method to account for the large pulse spreading that occurs through the system due to the achromatic lenses.

19.2.2 Spatio-Temporal Coupling

Apart from temporal shifting and spreading, the Ti:sapphire beam pulses also experience various spatio-temporal coupling (STC) effects through the system. An STC is described as any effect where temporal and spatial intensity components are not independent. The system experiences two differ-

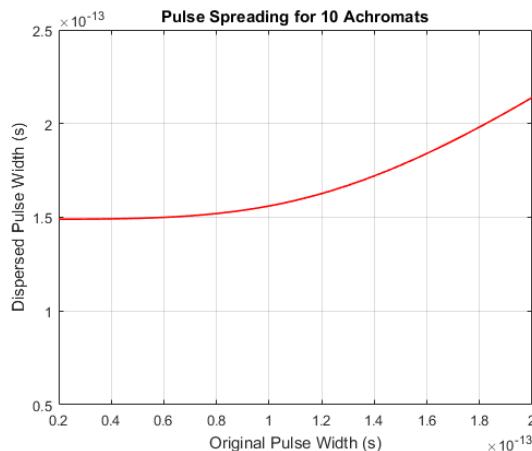


Figure 19.11: Pulse Spreading from Laser Source

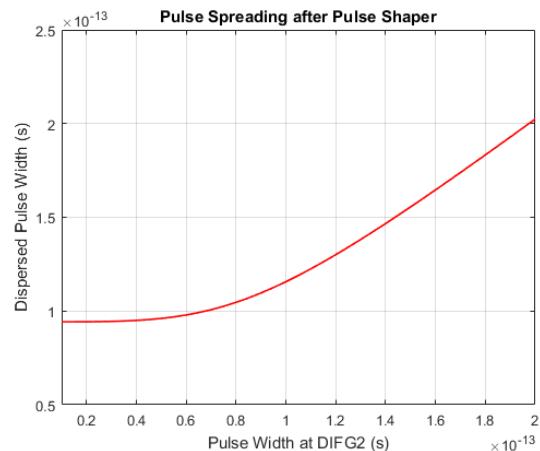


Figure 19.12: Pulse Spreading after Pulse Shaper

ent types of STCs [211], as explained below.

1. Angular Dispersion. The spatial separation of spectral components, which is utilized in the pulse shaper setup using the diffraction gratings to create a large separation, while the second diffraction grating (DIFG2) aims to combine all of the spectral components together to compress (or expand) the pulse. This is a key functionality of **DIFG1,2** for controlled pulse shaping, as illustrated in Figure 19.13.
2. Pulse Front Tilt. The spatial deformation of a pulse such that certain spectral components are shifted spatially, deforming the pulse. An example of unintended pulse shaping that might occur in this system is shown in Figure 19.14. In this system, pulse front tilt will be detected and compensated for by adaptive optics, as explained later.

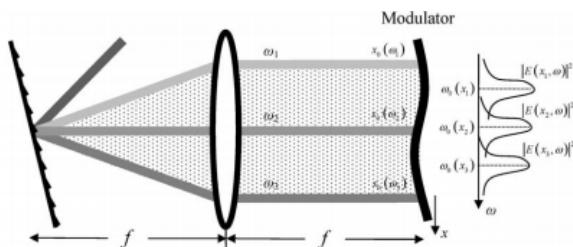


Figure 19.13: Example of Spatial Chirp [212]

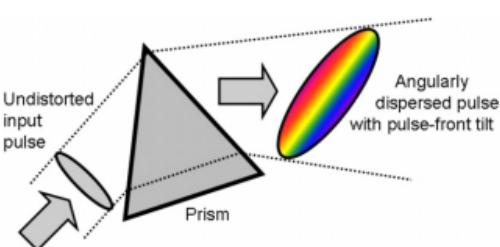


Figure 19.14: Example of Pulse Front Tilt [211]

Spatial chirp usually occurs due to misalignment of dispersive components, such as diffraction gratings, explored further in section 20.8. Unintended spatial chirp also occurs due to a beam offset from dichroic mirrors. Then, mathematically, an unchirped pulse is represented by (3.28), such that r and t are entirely independent. However, a dispersed pulse follows a mathematical relationship similar to (19.8), where

$$I(r, t) \neq I(r)I(t) \quad (19.9)$$

Since the expression produced in Equation 19.8 provides an intensity coupling, [211] provides a detail of the determinable STC characteristics for this system, given in Table 19.2.

Parameters	Coupling Type
(r, t)	Pulse-front tilt
(r, ω)	Spatial chirp
(k, ω)	Angular dispersion
(k, t)	Time versus Angle

Table 19.2: STC Classification for Intensity Coupled Function

The primary STCs in this system include pulse-front tilt and spatial chirp. One of the main illustrations of STC include a plot of pulse delay, as expressed by the coupled equation in (19.8), which is not separable into t , ω , or r . Since the primary source of STCs involve misalignment of pulse compressors, an angle misalignment of the pulse shaper was modelled in MATLAB in Figure 19.15 by plotting a 2-D intensity variation with pulse width and position from the peak of the pulse, based on (19.8).

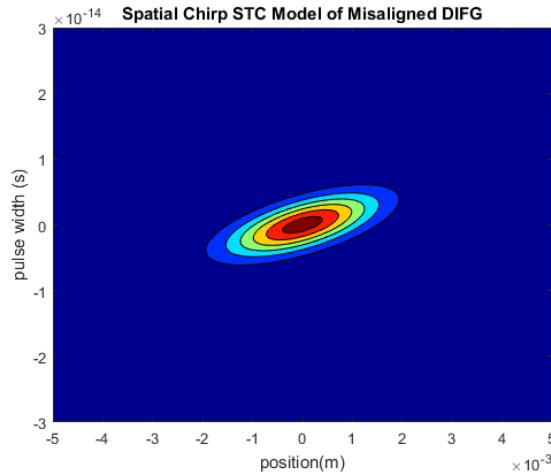


Figure 19.15: STC Model of Spatial Chirp

Given that the Gaussian pulse spatial and temporal intensity profiles are uncoupled (from (3.28)), the variance of the 2-D Gaussian should be independent. However, given the possibility of misalignment, from Figure 19.15, the coupling causes a change in pulse width related to the propagation position, as Figure 19.15 shows a dependence between the pulse width and the positional variation of the beam, a clear indication of an STC. Then, for this reason, there has been an inclusion of a an

ultrashort-pulse measurement device for the user to measure the temporal quality of the beam.

19.2.3 Pulse Control System Setup

Identification of STCs is key to developing a closed-loop model for maintaining the temporal distribution of the pulse through the system and achieving the required pulse from the entire system. For this system, a version of a frequency-resolved optical gating (FROG) was chosen for this system.

19.2.3.1 Theory

In 1997, Trebino proposed a novel method [213] known as a frequency-resolved optical gating (FROG) for determining the electric field variation of an ultrashort pulse. FROG utilizes a non-linear optical effect known as single harmonic generation (SHG) to generate an auto-correlation of the input pulse at specific time delays. The spectrometer then produces an intensity spectrogram instead of a conventional intensity-based mapping to provide a robust method to determine ultrashort pulse STCs accurately.

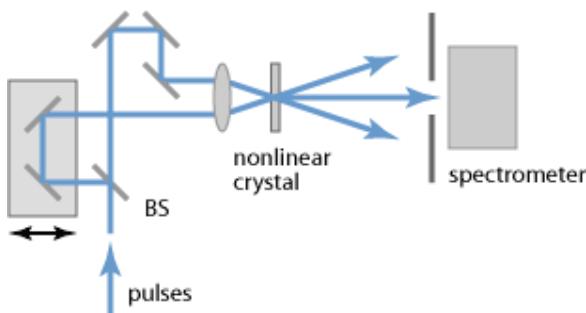


Figure 19.16: Implementation of SHG FROG [214]

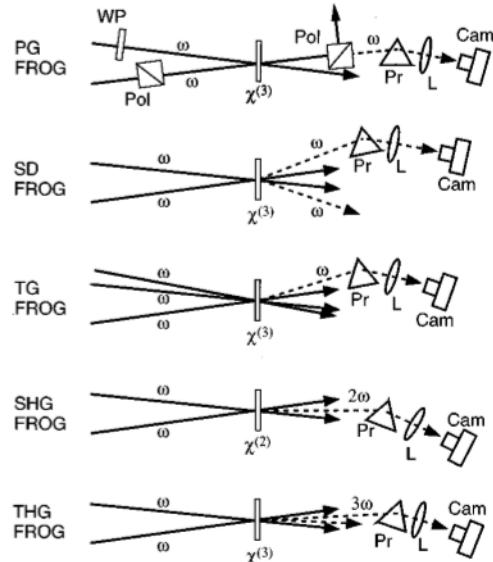


Figure 19.17: Other FROG Implementations [215]

From Figure 19.16, the beam splitter creates specific delays for which the pulse auto-correlation can be determined. The intensity spectrogram is then determined by the spectrometer from the electric field variation of the two pulses from (19.10) [215].

$$I_{FROG}^{SHG} (\omega, \tau) = \left| \int_{-\infty}^{\infty} E(t) E(t - \tau) \exp(-i\omega t) dt \right|^2 \quad (19.10)$$

where $I_{FROG}^{SHG} (\omega, \tau)$ represents the intensity spectrogram, with the SHG gating technique represented by a simple autocorrelation, as $E_{shg}(t) = E(t) * E(t)$. Several implementations of FROG (illustrated in Figure 19.17), such as polarization gating FROG (PG FROG) involved a polarization gating with an extra half-wave plate and several polarizers with a gating function $E_{pg}(t) = E(t) * |E(t)|^2 = E(t) * I(t)$.

While PG-FROG provides an intuitive spectrogram trace such as Figure 19.15, PG requires a higher minimum energy input of 1 pJ, compared to SHG FROG, which requires 10 nJ. Thus, implementing a non-linear SHG crystal provides the system with higher sensitivity, which can be provided by the beam sampler, allowing for real-time determination of misalignment.

A common alternative arrangement considered for ultrashort pulse measurement (UPM) is the spectral-phase interferometry for direct electric-field reconstruction (SPIDER), which incorporates dispersive delay and sum frequency generation (SFG) to delay the pulses. While SPIDER provides a straightforward method of measuring phase and electric field, measurements are susceptible to misinterpretation, as often the primary measurement is of the component of the pulse with the shortest component wave width, not of the pulse width itself. This was described as a coherent artifact by Trebino [216].

A disadvantage [214] is that the wave-front and the arrival time of the pulses cannot be determined, as the FROG treats the beam as spatially uniform. To tackle these shortcomings, the Shack-Hartmann wavefront sensor (**SHWS**) provides a straightforward method of determining the phase of the pulse and the spatial profile of the pulse. The SHWS does this using an array of micro-lenses; this is explained later in section 19.4.

19.2.3.2 System Implementation

In order to determine the pulse condition and the non-linear crystal performance, the group velocity mismatch (GVM) term is introduced as the maximum difference in group delay k' across the pulse bandwidth. The condition for the compatibility of the SHG crystal is that $TBP \ll GVM/GVD$. This condition is usually true for relatively transform-limited Gaussian pulses, as the ratio GVM/GVD is of the order 10^{16} , much greater than the $TBP = 0.441$.

An important issue with FROG is the direction-of-time ambiguity (DOTA) [217] that exists due to the auto-correlation function. An implication of DOTA is that the spectrogram trace is not unique; the same trace can be repeated by its mirror image, implying that any FROG trace is symmetric with respect to delay. In order to deal with this, Truong et al. [217] placed an etalon before the entrance of the FROG, which adds a second pulse of lower intensity but same shape, to distinguish the two pulses.

An etalon acts as a resonator, which allows for controlled release of the pulse, depending on the thickness and reflectivity of the surfaces of the etalon. By controlling the reflectivity of the surfaces, the amplitude of the second pulse is controlled by:

$$E_{meas}(t) = E(t) + \epsilon E(t - \tau) + \epsilon^2 E(t - 2\tau) + \dots \quad (19.11)$$

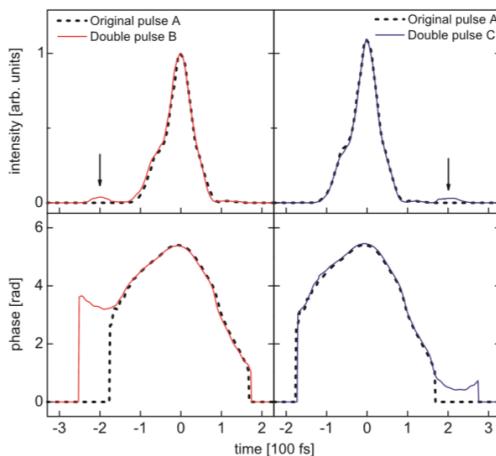


Figure 19.18: DOTA Elimination [217]

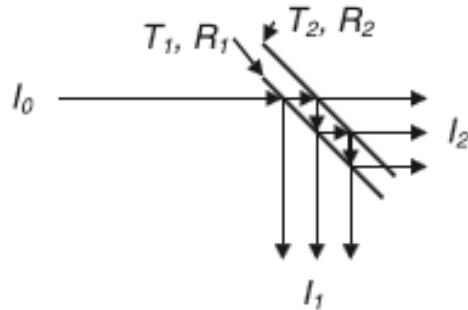


Figure 19.19: Etalon Schematic [217]

The etalon then generates an attenuated pulse train based on ϵ , or the Fresnel reflectance from Figure 19.19. Trebino's studies indicate a performance-guaranteed level of $\epsilon = R_1 = 0.5$ and back reflection of 10% was experimentally sound, since ϵ^2 and further orders can be ignored, and the FROG trace conducts an autocorrelation with a single delayed pulse. However, Zeek et al. note that the resonator effect for a thick etalon causes pulse distortion [218] as thick etalons have a thickness larger than the coherence length of the laser, which reduces the effect of constructive interference on the laser (and the interferometric nature of the etalon), causing the pulse shape to get distorted.

The coherence length for a laser is defined as the length for which the effects of interference on the laser can be predicted; this is key for the glass thickness as the multiple-reflection of the pulse occurs due to the interference of the beam, caused by the interferometric nature of the etalon. The coherence length for the Ti:sapphire pulse is given by $L = \lambda^2 / \Delta\lambda$, where L represents the coherence length. The difference between a glass with thickness $20 \mu\text{m}$ and an air-spaced etalon is that the latter is tougher to manufacture, and the studies cited illustrate that most air-spaced etalons, with spacing of $60 \mu\text{m}$ still cause distortion due to tolerances and thickness.

Instead, SLS Optics was chosen to provide a glass plate with thickness of $20 \mu\text{m}$. At the glass, the laser experiences Fresnel reflection with respect to a p-polarized beam, which is covered by the following equations. Instead of following the 50% - 10% ratio implemented by Zeek, an alternative was considered.

$$T_1 = \frac{4n_g \cos \theta_i \sqrt{1 - \left(\frac{\sin \theta_i}{n_g}\right)^2}}{n_g \cos \theta_i + \sqrt{1 - \left(\frac{\sin \theta_i}{n_g}\right)^2}} \quad (19.12) \quad R_2 = \left(\frac{\cos \theta_i - n_g \sqrt{1 - (n_g \sin \theta_i)^2}}{\cos \theta_i + n_g \sqrt{1 - (n_g \sin \theta_i)^2}} \right)^2 \quad (19.13)$$

where T_1 represents the transmittance at the first incident surface of the etalon, R_2 represents the reflectance at the second surface, n_g represents the refractive index of the glass, and θ_i represents

the angle of incidence. The Fresnel reflectance equation presented in (19.2.3.2) was optimized with respect to θ_i , such that the back reflectance of the etalon was 10%.

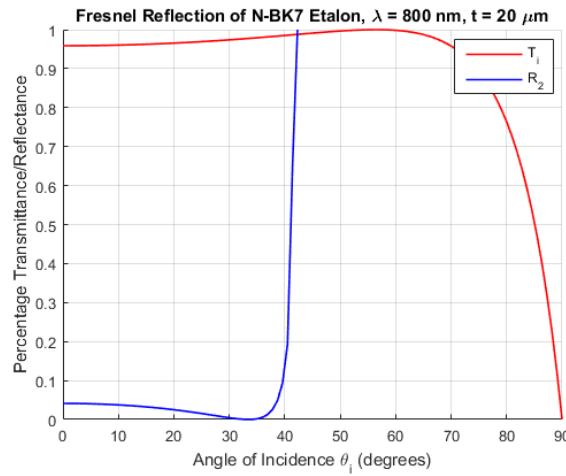


Figure 19.20: Fresnel Reflectance/Transmittance Model for N-BK7 Etalon

From Figure 19.20, ideally the etalon would be placed perpendicular to the beam itself, with a reflectance of under 10%. To bolster the reflectance, an AR coating corresponding to the tuning range of the laser was also requested. The transmittance of the first surface T_i was then 95%; this transmittance ensured the delayed pulse would form an identical pulse with a relative amplitude of 0.05, which is enough for detection of the delayed pulse, as explained in Figure 19.18, given a minimum pulse energy of 200 nJ. The purpose of the ultra-thin glass is to create a reflection of the incident pulses with a finite delay, determined by the value $2k't$, where t represents the glass thickness, and k' can be read from Table 19.1 for N-BK7. This benefits FROG, as the custom interferometer will cause a pulse delay of 200 fs, enough for FROG to detect the delayed pulse, as illustrated by Truong et al. [217].

19.2.4 Summary

The maintenance of ultrashort pulses poses a substantial challenge to minimize aberrations and improve system functionality across all processes. A closed-loop model to observe and modify ultrashort pulses was determined by modelling the possible aberrations and making layout decisions to improve and observe the pulse quality. In this setup, a combination of an etalon with a FROG setup ensured complete control over pulse measurement. The discussion will now move onto the importance of point spread function, and the effect of aberrations on the PSF and its relevance to image planes.

19.3 Effects on the Point Spread Function

The point spread function (referred to as the PSF onwards) is the 3 dimensional diffraction pattern seen when an infinitely small point source is emitted in a specimen and transmitted to the image plane through a lens. A further explanation of the fundamentals of PSF can be read in section 3.3. Monochromatic aberrations such as spherical, astigmatism and defocus manipulate the PSF from the ideal Airy disk. An understanding of how it is manipulated enabled us to incorporate adaptive optics in the system. For adaptive optics, the PSF is measured and then the conjugate of the aberration is produced on the SLM - thus the systems optics becomes aberration free. This section outlines the generalised pupil function and models the effects of common aberrations on the PSF using MATLAB.

19.3.1 Generalised Pupil Function

During the analysis in section 3.3, it was assumed the system was diffraction limited - so the PSF consists of the Fraunhofer diffraction pattern of the exit pupil. This fact allows aberrations to be included into the previous analysis. When a waveform error exists, it is imagined that a perfect spherical wave illuminates the exit pupil, however a phase plate has been positioned in the aperture, such that the wavefront leaving the pupil is deformed.[2]. The phase error is represented by $kW(x, y)$, where W is the effective path length error and $k = 2\pi/\lambda$. The complex amplitude transmittance $\mathcal{P}(x, y)$ is given by:

$$\mathcal{P}(x, y) = P(x, y) \exp[jkW(x, y)] \quad (19.14)$$

where $P(x, y)$ is the pupil function.

The complex amplitude transmittance $\mathcal{P}(x, y)$ is referred to as the generalised pupil function. The PSF of an aberrated system is the Fraunhofer diffraction pattern of the aperture with amplitude transmittance $\mathcal{P}(x, y)$.

The aberration function W can be defined using the geometry in Figure 19.21. In a system that is free of aberrations, the exit pupil is filled with a perfect spherical wave that converges to an ideal image point. A Gaussian reference sphere is a spherical surface which is centred on the ideal imaging point and passes through the point where the optical axis intersects the exit pupil. The aberration function $W(x, y)$ is defined with respect to this Gaussian reference sphere. $W(x, y)$ is the accumulated path length error by a ray as it passes from the Gaussian reference sphere to the actual wavefront.[2]

In conclusion, the sole effect of aberrations is to introduce phase distortions to the system. These phase distortions can have a severe effect on the accuracy of the imaging system.

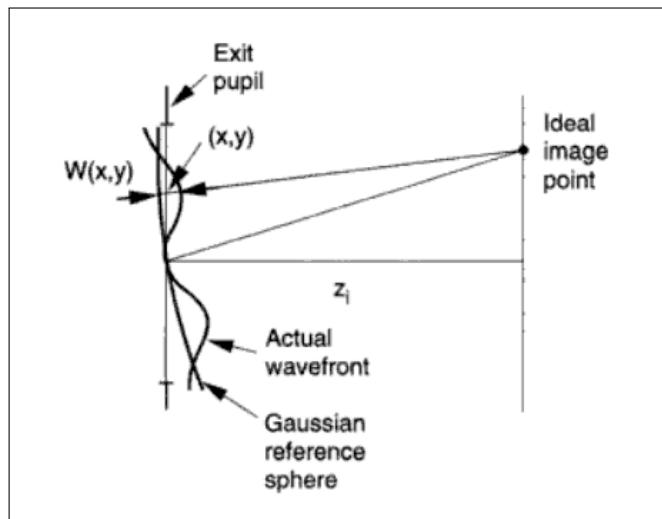


Figure 19.21: Geometry for the aberration function W [2]

19.3.2 Matlab Model of Aberrated PSF

In order to demonstrate how aberrations shape the PSF, a MATLAB simulation was used. Taking the MATLAB model of the ideal PSF in section 19.3 and replacing the pupil function with the generalised pupil function instead, aberrations could be added to the PSF. The generalised pupil function was implemented like so:

```

W = RMS_SA * sqrt(5) * (6*R_norm.^4 - 6*R_norm.^2 + 1); %Spherical Aberration wavefront
W(R_norm>1) = 0;
E = exp(1i*2*pi*W); %Complex amplitude
E(R_norm>1) = 0; %Impose aperture size
figure; imagesc(angle(E)/(2*pi)); colorbar; title('Wavefront Phase (waves)');

```

The first line defines the effective path-length error $W(x,y)$, this uses the Zernike polynomials that are explored further in section 19.4. The complex amplitude E is calculated and then the aperture is imposed. The variable E is thus the complex amplitude transmittance.

Figure 19.22 shows that an aberration free optical system produces a symmetrical Airy disk and that no phase has been introduced to the wavefront. In the case of a spherical aberration, the wavefront leaving the lens is no longer spherical with a centre position on the ideal image point. The wavefront is distorted due to the phase introduced in Figure 19.26 and so the PSF departs from the ideal behaviour. It can be observed that the PSF has an asymmetrical distribution where the intensity ratio between the central peak and the surround rings is shifted, with the latter becoming more prominent. This manipulation of the PSF will affect the system's image quality and fidelity, which will in turn affect the interpretation and measurement of what we are imaging. Furthermore, spherical aberrations will affect the efficiency of our confocal microscope as its optics are reducing the image

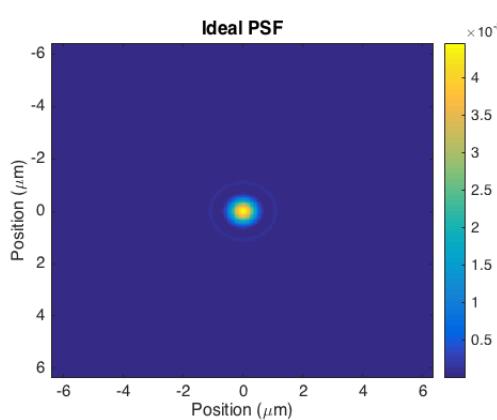


Figure 19.22: Aberration free PSF

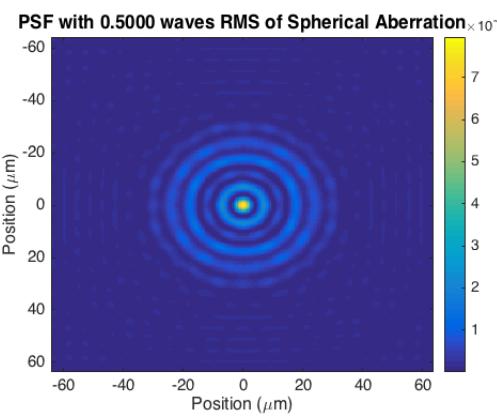


Figure 19.23: Spherical aberration

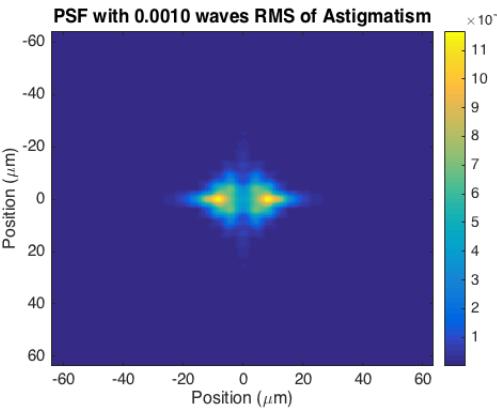


Figure 19.24: Astigmatism aberration

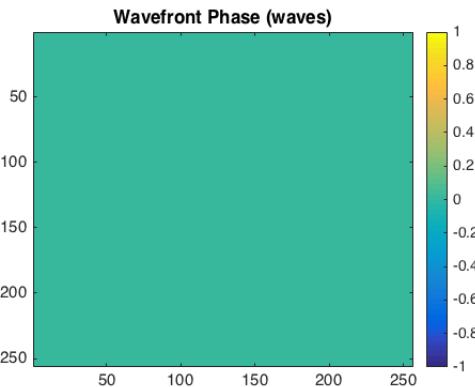


Figure 19.25: Phase plot for an aberration free PSF

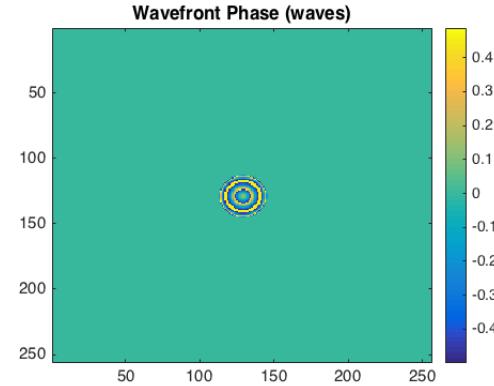


Figure 19.26: Phase plot for a spherical aberration

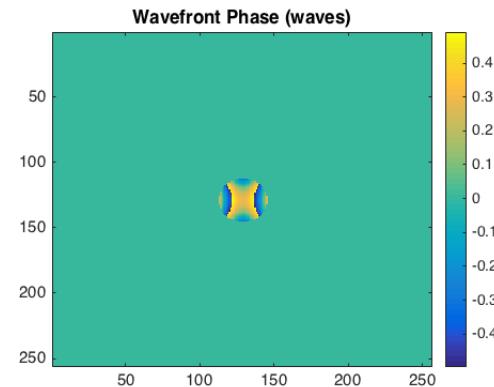


Figure 19.27: Phase plot for an Astigmatism

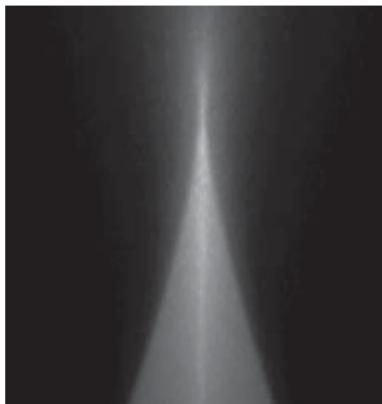
Figure 19.28: Plots of the PSF and phase applied, for an aberration free system, spherical aberration and an astigmatism

to obtain a diffraction limited point.[4]

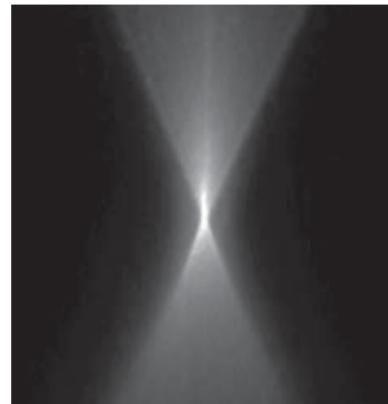
In conclusion, the PSF is important because it relates to how well we can focus light and this entire system design revolves around being able to focus light to the best of our ability. That is also why aberration elimination is so important. In this section we have seen that applying phases to the ideal PSF will deform it. This understanding of how the PSF is manipulated by phases led us to develop an adaptive optics system.

19.4 Adaptive Optics

Adaptive Optics is a modern technique that utilises feedback to correct for aberrations that occur during the propagation of light through an optical system. As the laser passes through lenses and other optical components, the wavefronts of the light can become distorted, resulting in decreased fabrication and image quality. Problems caused by spatial aberrations include reduced contrast, defocusing and other image distortions. Zernike Polynomials (discussed and generated in Section 19.4.1) provide a basis to characterise common optical aberration modes, each distinguished by their effects on the focal spot (PSF). For example, some significant spatial aberrations can cause the focal point to become elongated (as shown in Figure 19.29 from [219]), which reduces the resolution of both fabrication and imaging.



(a) Aberrated focal spot



(b) Focal spot with aberration correction

Figure 19.29: Comparison of a focused beam (viewed from the side) with spatial aberration (left) and after the aberration has been corrected using an SLM (right), from [219]

19.4.1 Modelling wavefront aberrations

A common method of modelling optical aberrations is to use Zernike polynomials, $Z_j(r, \theta)$. These polynomials are functions of r and θ in a polar coordinate plane and useful because they form an orthogonal basis set over the unit circle. This is advantageous since the pupil function of the microscopic objective (where we would usually model the wavefront aberration) lens is a circle. Each Zernike polynomial corresponds to a mode of optical aberration (some MATLAB generated common modes are shown in Table 19.3) and more complex optical aberrations can be generated by summing weighted basis modes, due to orthogonality.

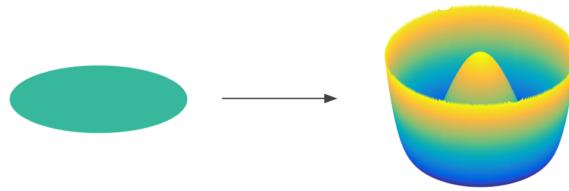
19.4.2 Correcting wavefront aberrations

Wavefront aberrations are caused by differences in optical path length across the profile of the laser beam throughout the system. In addition to this, different sections of the laser profile will also experience unequal propagation retardation due to variations in refractive index. Velocity of light in a

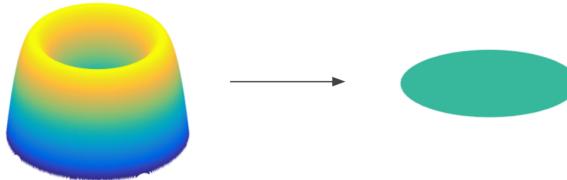
Optical Aberration	Zernike Index	Wavefront Profile	Description and Effect on Image
Piston	Z_1		Constant phase offset across the beam. This has no effect on image resolution since there are no relative phase changes, so the focal shape is unchanged.
Tip	Z_2		Linear gradient of phase, corresponding to a lateral movement of the focus but no focus shape distortion since the wavefront has no curvature. No effect on image resolution.
Defocus	Z_4		Optical axis change in focus location, but no focus shape distortion. No effect on image resolution, but the images will be out-of-focus.
Astigmatism	Z_5		Phase profile causes the perpendicular rays to converge to different focal points. Causes a reduction in image resolution due to the changed focus shape.
Coma	Z_8		Phase profile leads to a distortion of focal point characterised by a cometary shape. Reduced image resolution due to the changed focus shape.
Spherical	Z_{11}		Phase profile leads to off-axis rays converging in a shorter distance than on-axis rays, resulting in an optical-axis elongation of the focus point. Reduced image resolution due to changed focus shape.

Table 19.3: Table of some common types of optical aberrations, with the corresponding normalised Zernike polynomial representation plotted on the unit circle and a description of their effects. The Zernike indices are the Noll indices [220]. The shading of the wavefront profiles is interpolated between yellow and blue, corresponding to maximum positive phase shift and maximum negative phase shift, respectively.

medium with refractive index μ is $v = \frac{c}{\mu}$, where c is the velocity of light in a vacuum. This localised delay in wavefront is similar in effect to a longer path length.



(a) Example ideal wavefront profile before propagation (left) and an example measured wavefront after propagation (right).



(b) Example pre-aberrated wavefront with the conjugate of the propagation aberration (left) and the ideal wavefront profile after propagation, due to the correction

Figure 19.30: A MATLAB generated example representation of the wavefront correction from adaptive optics. Units are normalised and arbitrary, and only relative position of wavefront is of interest (i.e. Piston mode is considered ideal). Subfigure 19.30a shows a wavefront profile change due to propagation through the optical system with no correction and Subfigure 19.30b shows how, through pre-aberration of the input wavefront to the conjugate of the measured aberration, the wavefront after propagation is (in the ideal case) without aberration.

Predicting, characterising and quantifying these aberrations becomes a problem with a large number of degrees of freedom, especially in our heavily customisable optical system. Aberrations may vary depending on the laser type, pulse energy, pulse shape, presence of removable optical components, type of microscopic objective, specimen, illumination, and misalignments of the beam. Since the optical system is so complex, an analytic approach to calculating or approximating the wavefront aberrations would be cumbersome and extremely difficult.

Adaptive optics provides the system with a way to detect and measure wavefront aberrations, then compensate for these in an effort to keep the wavefront ideally flat. Direct-sensing adaptive optics uses these measurements of the wavefront reflected from the sample to form a control system with closed-loop feedback. The correction is performed by modulating the conjugate of the computed phase aberration to the beam in an attempt to cancel the the aberration; the idea is that once pre-aberrated, the aberration that occurs during propagation acts to further delay the sections of the beam that have been delayed least by the pre-aberration (since the conjugate aberrations are equal and opposite). This can be seen more clearly in Figure 19.30.

19.4.3 Shack-Hartmann Wavefront Sensor (SHWS)

The Shack-Hartmann Wavefront Sensor (SHWS) is a device that can measure the gradient of an incoming wavefront. It is constructed (as shown in Figure 19.31) as a CCD or CMOS camera in front of a micro-lens array (MLA) that receives a laser beam through a collimating lens.

When a flat phase wavefront is incident on the collimating lens, the propagating light is parallel to the plane of each micro-lens in the array and so generates a grid of equally spaced foci at the camera. These focal points form the undistorted reference points of wavefront measurement. When an aberrated wavefront is incident on the array, the propagation of light will not be parallel to the plane of the lenses since the wavefront has a non-zero gradient, therefore the focal spots at the input to the camera will be moved from their reference points. To visualise this see Figure 19.32.

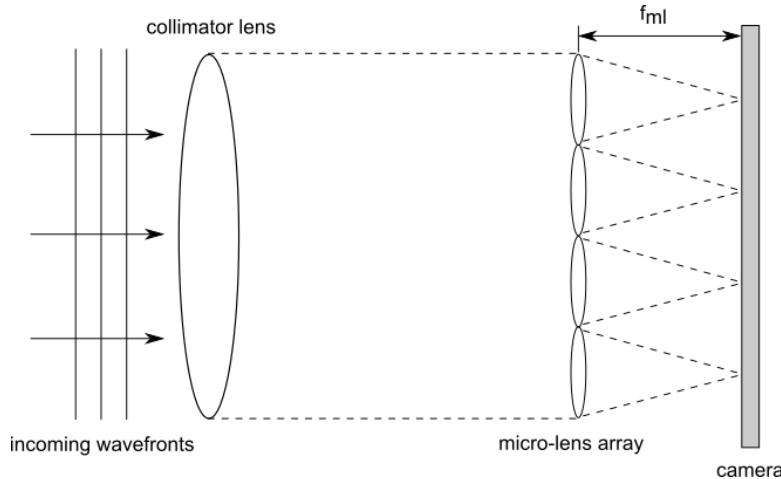
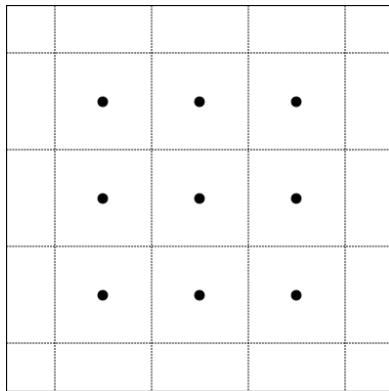
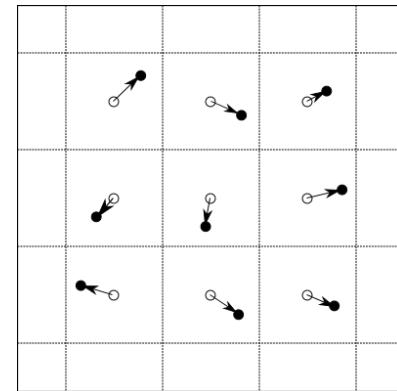


Figure 19.31: Simple diagram of the Shack-Hartmann Wavefront Sensor, f_{ml} is the focal length of the lenses in the micro-lens array



(a) Reference SHWS focal spots at camera sensor, for a flat phase wavefront



(b) Moved SHWS focal spots at camera sensor, due to wavefront aberrations

Figure 19.32: Comparison between the focal spot positions on the SHWS camera sensor for a flat phase wavefront and aberrated wavefront. The arrows show the vectors of $(\delta x, \delta y)$, the relative movement of the spots compared to the reference spots.

Then, using the geometry in Figure 19.33 provided by [221] we can see that the linearised infinitesimal gradient of the incoming wavefront $W(x, y)$ is

$$\frac{\Delta z}{\Delta y} = \tan \alpha = \frac{\delta y}{f_{ml}} \quad (19.15)$$

and similarly for the x direction. Since we have vectors $(\delta x, \delta y)$ for each pixel from the camera sensor

(see Figure 19.32) and f_{ml} is defined by the SHWS, we can use this expression to approximate the partial derivatives to the wavefront with respect to x and y :

$$\frac{\partial}{\partial x} W(x, y) = \frac{\delta x}{f_{ml}} \quad \frac{\partial}{\partial y} W(x, y) = \frac{\delta y}{f_{ml}} \quad (19.16)$$

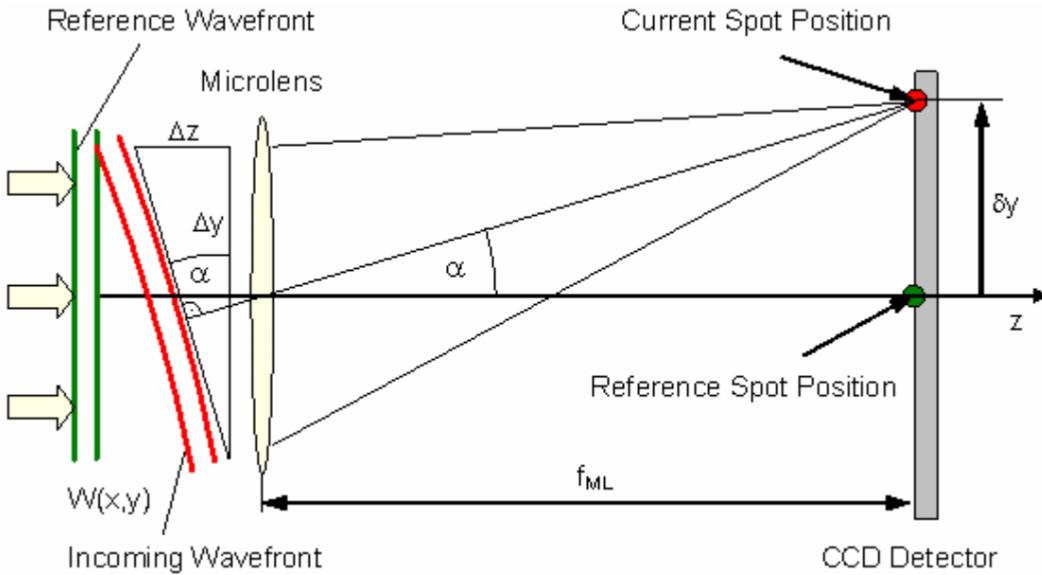


Figure 19.33: Geometry showing the relation between infinitesimal wavefront gradient and focal spot displacement, from [221]

Then, all that is left to compute an approximation to the wavefront $W(x, y)$ is to perform a two dimensional integration. $W(x, y)$ is then the profile of the aberration in units of length. Since it is difficult to change the optical path length for sections of the light, this profile is converted to a phase profile so that we can modulate the phase of the beam to correct the aberrations using the SLM. To compute the phase profile of the aberrated wavefront, we use the relation

$$\phi(x, y) = \frac{2\pi W(x, y)}{\lambda}. \quad (19.17)$$

ThorLabs produce SHWS available to buy off-the-shelf. Two main categories of SHWS they offer are those with CCD or CMOS sensors, with maximum frame-rate 15 Hz and 1120 Hz respectively. Our optical system uses high frequency laser repetition rates and short pulses and therefore a higher frame-rate is advantageous. In addition to this, the wavelength of light modulated by **SLM1** is 800 nm, for which the ThorLabs CMOS sensors have greater response compared with CCDs [222]. ThorLabs also offer various MLA pitch sizes, which affects the spatial resolution and wavefront accuracy. The 150 μm option provides higher resolution but lower wavefront accuracy than the 300 μm option. Since **SLM1** has a high resolution with which to display the correcting wavefronts, the 150 μm pitch size

MLA was chosen. Therefore the SHWS utilised is the ThorLabs WFS20-7AR, with CMOS sensor, 150 μm MLA and 400 nm to 900 nm working wavelength range.

19.4.4 Optical Layout

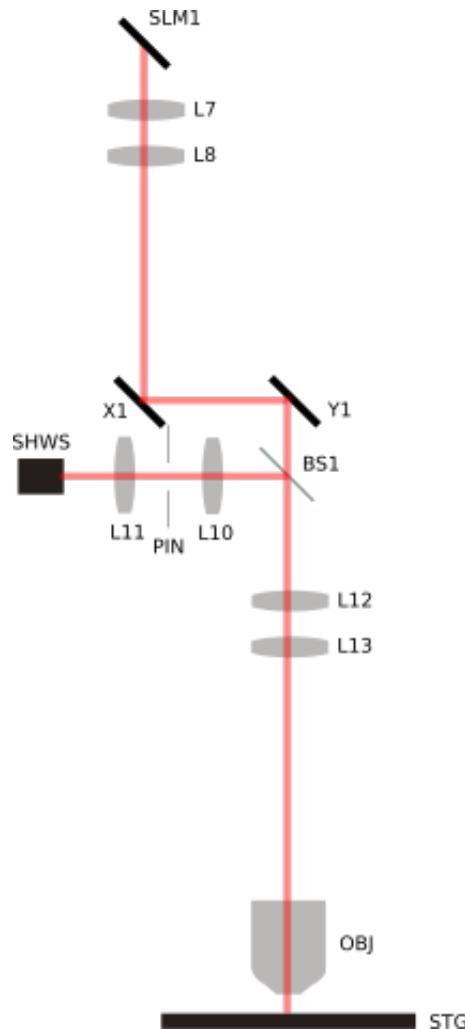


Figure 19.34: The designed optical path for Adaptive Optics

Using the SHWS and computation of the phase profile of the wavefront aberration described in 19.4.3, we have designed a direct wavefront sensing adaptive optics control system, which can be seen in Figure 19.34.

The ThorLabs WFS20-7AR **SHWS** receives the 800 nm Ti:Sapphire laser beam after reflection from the specimen, to measure the wavefront gradient. The 4f imaging system (defined in Section 18) comprised of lenses **L10** and **L11** are designed to magnify the beam to fill the SHWS input aperture and to ensure the phase distribution at the input of the SHWS corresponds to the phase at the entrance to the microscopic objective **OBJ** after reflection from the specimen on the stage, **STG**. This is where the phase profile is desired to be ideally flat. Because the beam waist is fully defined at the input focal plane of lens **L12** as 2 mm, the position of the input focal plane to the 4f arrangement is located one **L12** focal length from **L12** so that reflected light will be imaged from

the **L12**, **L13** arrangement to the input of this new 4f arrangement. The input aperture radius of the SHWS is 2.5 mm, so we require a magnification of $\frac{5}{4}$. This is achieved by designing **L10** to be focal length 100 mm and **L11** to be focal length 125 mm.

The use of the pinhole **PIN** (introduced by [223]) is used to filter out-of-plane reflected light from the specimen, using a similar principle to that of a confocal microscope. For this reason, the optimal pinhole diameter is designed in the same way as a confocal pinhole (explained in Section 13.1.4.3) and is calculated to be 35 μm .

The SHWS is connected to a PC which can perform the computation of the phase profile, $\phi(x, y)$ of the aberrated wavefront. This calculation is used to generate the conjugate phase profile, $\phi^*(x, y)$, which is modulated onto the laser beam by **SLM1** to create the pre-aberration discussed in Section 19.4.2.

19.4.5 Adaptive Optics Conclusion

With this optical design, direct-sensing adaptive aberration correction can be used for Laser Micro-machining, Two-Photon Polymerisation and Two-Photon Microscopy (including STED microscopy and imaging) since these processes all use the Ti:Sapphire laser that reflects from **SLM1**. To include direct-sensing adaptive aberration correction for Confocal Microscopy, Booth [224] explains how aberration correction would be needed in both the illumination path and detection path (after reflection from the specimen). Therefore, either another SLM would have to be added to the optical layout or **SLM1** would have to be moved into the optical path in both directions. In practice, if aberration correction was necessary for confocal microscopy, the user could apply summations of Zernike polynomial optical aberration modes to **SLM1** in an iterative optimisation procedure in order to decrease the aberrations (this could only be applied using the Ti:Sapphire laser, because **SLM1** is in the beam path). This is indirect-sensing adaptive optics and is described in [224]. Direct-sensing is more desirable since it is self-adapting whereas indirect-sensing requires significant user supervision, but results for both types will vary between processes and depend on the specimen and its geometry. It is anticipated that the Two-Photon processes will utilise adaptive aberration correction better since the out-of-focus light arriving at the SHWS is minimised.

20 Beam Control

20.1 Generic Implementation

The system utilizes an industry-standard power control setup, as shown in Figure 20.1.

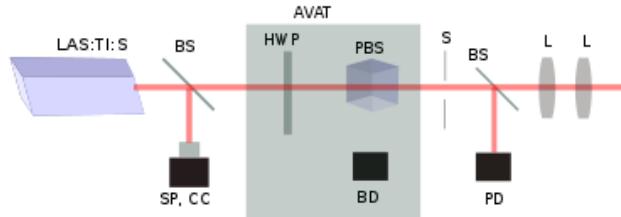


Figure 20.1: System Laser Power Control Setup

The laser is first sampled using a beam sampler (BS) to experimentally determine the spectral composition of the beam. This procedure is key to determining the pulse bandwidth for the Ti:sapphire laser **LAS:TI:S**, which can be used to verify the pulse width, as well as the linewidth of the continuous-wave (CW) DPSS lasers **LAS:355**, **LAS:532**. The system also utilizes a polarization-dependent optical attenuator, which is implemented using a rotating half-wave plate (HWP) and a polarizing beam splitter (PBS) to vary the intensity of the laser. The half-wave plate works to alter the polarization of the laser output (which is horizontally polarized), as explained through Figure 20.2.

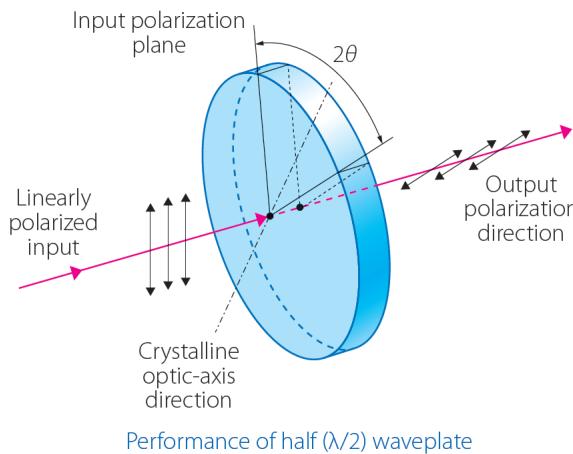


Figure 20.2: Half-wave Plate Functionality [225]

As explained by Figure 20.2, each half-wave plate has a fast axis and a slow axis, as explained in subsection 14.1.2. The component of the wave's polarization vector (indicated in red) parallel to the slow axis is retarded by π radians, which causes the change in the polarization vector, indicated by the right diagram. Then, in order to control the output polarization vector, the angle of the rotating half-wave plate φ is varied such that the amplitude of the horizontal component of the polarization

vector (not parallel to the fast/slow axes) varies according to Malus' Law:

$$I(\varphi) = I_0 \cos^2(2\varphi) \quad (20.1)$$

The assumption made in using Malus' Law is that the laser output is horizontal, given by the Coherent data sheet. Then to control the power, this beam is passed through a polarizing beam splitter, which allows the component parallel to the horizontal axis (I_x) to pass, while deflecting the rest of the beam to a beam dump. Then, a variation of φ from 0 rad to $\pi/4$ rad is enough to vary the polarization-horizontal component of the beam I_x from 0 to the maximum laser output. For the **LAS:TI:S** and **LAS:532**, this setup is automated entirely using a Newport automated VA toolkit (AVAT), while **LAS:355** is run using a simple rotating half wave plate setup **AHWP3**, with separate beam dump **BD3** and polarizing beam splitter **PBS3**.

This setup is then followed by a mechanical shutter **S5**, to stop unintended beam propagation. This method is preferred to switching the laser on/off as lasers require warm-up time up to 1 hr, and large voltage changes can damage the equipment. Instead the shutter can be shut to stop the beam propagating for hours at a time.

After the shutter, a known percentage of the beam is then sampled to a power detector **PD**, and the beam is spatially expanded using a telescopic arrangement to match the specific requirements of each system. Further details about component selection and beam control decisions have been explained below.

20.2 Choice of Lasers

The system requires three types of lasers: a Ti:sapphire pulsed and continuous wave laser with chirped amplified output (**LAS:TI:S**), a 532 nm CW DPSS laser (**LAS:532**), and a 355 CW DPSS laser for complete functionality(**LAS:355**).

20.2.1 Ti:sapphire Amplifier System

The system requires a Ti:sapphire seed laser at a centre wavelength of 800 nm, an oscillator to increase the repetition rate, and an amplifier output for chirped pulse amplification (CPA). In such a case, it was preferred to purchase an integrated setup, rather than separate components. For a pulsed laser, the key characteristics that were prioritized include the following list:

- Energy per pulse/Power output. This is important for all optical techniques in the system, to determine the system capability in terms of range of materials and resolution.
- Minimum pulse width. The minimum pulse width specified determines the quality of the beam at the output. It is important to note that the minimum pulse width achievable still remains at 95

fs due to dispersion, as explained in section 19.2.

- M^2 . Also known as the beam quality factor, this is a measure of the flatness of the spatial profile of the beam. A high M^2 implies a highly elliptical beam. Ideally, the beam profile of the objective would be circular, giving an ideal M^2 factor of 1.
- Price. The specifications are provided in Table 20.1.

Criteria	Coherent Libra	Thorlabs Octavius	Coherent Legend Duo Elite
Energy-per-pulse (mJ)	4	2×10^{-6}	5
M^2	1.3	1.3	1.35
Pulse Width (fs)	50	8	130
Price (USD '000)	200	170	250

Table 20.1: Amplifier System Specification Comparison

Then, the following components were compared using a multi-criteria analysis (MCA) in Table 20.2, to yield the Coherent Libra as the best option for the system specifications and capabilities.

Criteria	Weight	Coherent Libra		Thorlabs Octavius		Legend Duo Elite	
		Rating (5)	Score	Rating (5)	Score	Rating (5)	Score
Energy-per-pulse	4	4	16	2	8	5	20
M^2	4	4	16	4	16	3	12
Pulse Width	3	4	12	5	15	3	9
Price	3	3	12	4	12	2	6
Total		56		51		47	

Table 20.2: Multi-Criteria Analysis (MCA) for Amplifier System

Given the ratings based on a percentage point scale against the requirements, the Coherent Libra with Vitara T seed laser oscillator was chosen for the Ti:sapphire laser system. However, the optional seed laser (unamplified) output is at an angle to the main setup, for which a separate setup was used, as illustrated in Figure 20.3

The maximum power at the objective, however, is much lower than the power at the laser output, due to the absorption and reflection of intensity due to lenses and other optical components. By multiplying the transmissivity loss of 10 doublets - which is about 3% per doublet, as well as assuming

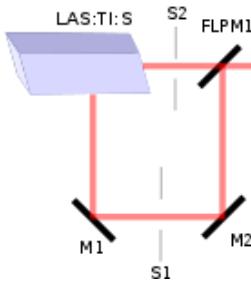


Figure 20.3: Mirrors used to redirect Vitara T HP output

a power loss of 30% due to the pulse-shaping diffraction grating setup, the real total power transmission from the laser output to the objective, is about 53% of the original output. This implies that the maximum pulse energy through to the objective will only be 0.76 mJ.

Then, after choosing the Coherent Libra + Vitara T HP amplifier system, the DPSS lasers were chosen on a similar MCA.

20.2.2 DPSS Laser Selections

The multi-criteria analysis were repeated for the two continuous wave (CW), diode-pumped solid state (DPSS) lasers, with centre wavelengths at 355 nm and 532 nm. The main criteria for these lasers include the laser power output required for STED and stereolithography, which was 5 W and 0.25 W, respectively. For this, the two DPSS lasers selected were selected from Coherent. For the 532 nm centre wavelength, the Coherent Genesis 532 CX STM (OEM version), with a power output of 5 W, was selected. For the 355 nm centre wavelength, the Coherent Genesis 355 CX STM (OEM version), with a power output of 0.25 W, was selected.

20.3 Variable Attenuation

As explained in section 20.1, the optical attenuator works by altering the polarization state of the incident light using a rotating half-wave plate, and deflecting a known portion of the beam using a polarizing beam splitter cube, into a beam dump. This setup was automated by Newport for both continuous and pulsed lasers with centre wavelengths greater than 400 nm, and hence this setup was chosen for the Ti:sapphire amplifier system and the 532 nm DPSS laser.

The main advantage of the Newport automated polarization setup is that the automation facilitates the debugging process and condenses the system risk by transferring it onto the manufacturer (Newport). While Newport does not provide actual specifications of the power constraints of the components and the absorption performance of the beam dump, most commercial half-wave plates can withstand powers of about 5 W cm^{-2} , and the beam dump has a rejection coefficient of about 10^{-4} , both of which satisfy the system capability. Thorlabs also provides multi-order half-wave plates, which involve retardations of integral multiples of $n\lambda/2$; however, a zero-order plate ($n = 1$) is sufficient for

this purpose and is generally considered more stable than multi order plates. Another advantage of using an automated VA lies in the software integration. Unfortunately, Newport does not provide an automated VA toolkit for a wavelength of 355 nm. In such a case, the variable attenuator was positioned in a similar fashion as section 20.1.

20.3.1 Alternatives Considered

One alternative to using a polarization-based variable attenuator is to use a neutral density filter wheel, or a graduated neutral density filter, as shown in Figure 20.4, Figure 20.5. A neutral density (ND) filter reduces the intensity of light at all wavelengths by either reflecting or absorbing the rejected light, and hence these could also be used in place of a half-wave plate and polarizing beam splitter.

The two main disadvantages with using either of the ND filter arrangements include component replacement/manufacture, discretization of levels, as well as system safety considerations. Firstly, the manufacturing of a graduated filter wheel would require customization of attenuation, as well as custom build, which would be tough to replace in case of manufacturing or operational defects. Furthermore, graduated filters do not have a setup to completely extinguish the beam, which a polarization-based setup allows.



Figure 20.4: Graduated ND Filter [226] Figure 20.5: Manual Replaceable ND Filter Wheel [227]



Secondly, no ND filter is perfectly continuous, as each level of attenuation would be generated by altering the level of coating. ND filters work by either absorption or reflection, which implies that the control over attenuation would be extremely difficult due to discretization of power levels, as well as the lack of a complete attenuation ability. Finally, both of these methods would pose a safety risk, as the ND filter could be damaged due to high power or heat absorption, since the power rejection at any point could reach about $1 - 2 \text{W cm}^{-2}$, which can cause material breakdown over prolonged exposure of an optical coating, thus destroying the filter performance. Alternatively, the random reflections from a reflective ND filter with the same power specifications could damage both system components, as well as have an effect on the sample by either heating up the sample or causing beam distortions [228].

Then, for these two reasons, a polarization-based setup with a controlled beam rejection system

was preferred over an ND filter setup with a filter wheel, or a graduated ND filter.

20.3.2 Laser-Induced Damage

One of the main considerations from a beam control perspective is the possibility of laser-induced damage in the form of field breakdown and component damage. Field breakdown occurs in any medium when the electric field density (or fluence) is strong enough to cause an electrical discharge, or a sharp reduction in medium resistance. Field breakdown is a danger to users, as well as system components, and is mainly caused by concentrated beams and high voltage power lines. The main components that require large voltages are the laser systems, which are positioned close to each other to minimize the length of the power rails. However, field breakdown due to the beam required more attention, since the safety considerations span the entire system.

The electric field that causes field breakdown $|\mathcal{E}|_{peak} \propto n$, the refractive index of the medium, indicating that the danger of field breakdown lies primarily in the free space propagation of high-fluence beams. Then, the following expressions hold for all Gaussian beams with a FWHM pulse width [6], where the subscript p indicates peak quantities relating to the individual pulse, rather than an overall quantity, and REP RATE indicates the repetition rate of the pulse.

$$P_{peak} = E_{peak} \times \frac{0.94}{\tau_p} = \frac{P_{avg}}{\text{REP RATE}} \times \frac{0.94}{\tau_p} \quad I = \frac{cn\epsilon_0}{2} |\mathcal{E}|^2 \quad (20.2, 20.3)$$

Given these quantities, combined with (19.7), the two equations that can describe $|\mathcal{E}|_{peak}(r)$ in terms of values from laser data sheets are in (20.4).

$$|\mathcal{E}|_{peak}(r) = \frac{1}{W(r)} \sqrt{\frac{2E_p}{cn\epsilon_0 \pi \times \text{REP RATE}}} \quad (20.4)$$

With these expressions, the maximum power levels for the beam waists at laser outputs through the system - ranging from 2 – 10 mm, depending on the output - were determined.

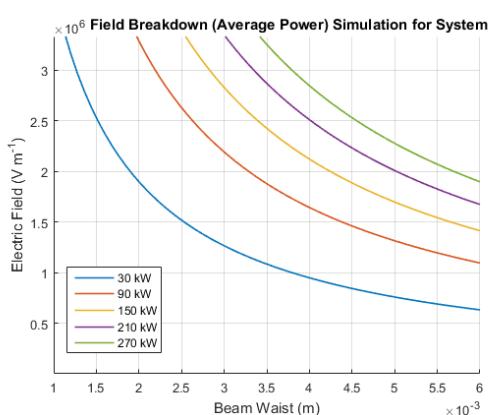


Figure 20.6: Field Breakdown Simulation

However, the maximum beam waist to land at the galvanometers, as specified in chapter 22, is 4 mm, given an error margin of 1 mm. Then, the electric field variation with pulse power P_{avg} was determined for a beam waist range $W(r)$ specified, as shown in Figure 20.6.

Through the system, the highest fluence occurs at focal regions in the 4F systems in lenses **L7,8, L12,13**. Assuming a safety factor of 1.2, the minimum beam waist to maintain a maximum field of $2.5 \times 10^6 \text{ V m}^{-1}$ is $10 \mu\text{m}$. Then the average pulse power

allowable for a Gaussian beam through the entire setup to avoid breakdown is 1.50 W, with a pulse energy of 0.15 mJ, enough for ablation techniques, as explained in chapter 6. Unfortunately, for applications involving multiple foci generation, the maximum power is lower than 0.75 W, as explained in chapter 16. **LAS:Ti:S** was also chosen with a maximum power of 0.15 mJ, accordingly. In order to determine the beam waist in the focal region, the Gaussian relation between beam waists before ω_{01} and after the lens ω_{02} and the focal length f is given as

$$\omega_{01}\omega_{02} = \frac{f\lambda}{n\pi} \quad (20.5)$$

This relation [229] indicates that for any lens, $\omega_{02} \propto \omega_{01}^{-1}$, a reflection of a high focal length to entrance aperture ratio, or a high f-number. To maintain a high f-number, the beam waist was maintained at about 1 – 3 mm, and **L7** was chosen to have a focal length of 0.5 m, since the minimum beam waist at **SLM2** is 5 mm, which is too high an aperture size to have a “slow” lens, with a large f-number.

A secondary concern through the system is the component damage that could occur in the form of either thermal or optical damage. This was easier to mitigate, as all components selected through the system were chosen with a minimum laser-induced-damage-threshold (LIDT) fluence of 5 J cm^{-2} , which exceeds the maximum system fluence at a component. The maximum fluence is then 0.03 J cm^{-2} when considering the power losses through the system (as explained in subsection 20.2.1), and is low enough to avoid any component damage. To conclude, any laser-induced damage has been mitigated by maximizing the f-numbers of lenses, as well as constraining the beam waist to accommodate for breakdown in focal regions, while also selecting components with a high LIDT to ensure safety considerations and component resilience has been justified.

20.4 Shutter

A requirement to speed up fabrication techniques in the system is the presence of a high-speed shutter. The shutter is useful in organizing samples where the fabrication requirements have arrays of similar components, for example in conjunction with multiple foci generation. In such cases, this shutter enables the system to generate such array structures by attenuating the laser for a short period, allowing the galvanometers to adjust the beam quickly, reducing the sample fabrication time significantly.

20.4.1 Considerations and Specifications

Unfortunately, the maximum shutter rate for mechanical shutters ranges to about 25 Hz at a burst, while fabrication methods require rates of orders of MHz. While mechanical shutters are useful to

ensure laser safety, optical shutters with a higher shutter rate can also be used, while compromising on attenuation. Most optical shutters are implemented using acousto-optic modulators (AOMs) or electro-optic modulators (EOMs), such as Pockels cells. The former are preferred in optical setups, since EOMs usually have a high power requirement, which could require high voltage rails to run through the system.

20.4.2 Implementation

Single-pass AOM arrangements have been implemented as optical shutters. An AOM works using acousto-optic diffraction of the pulse using an acoustic frequency ν_{RF} , representing a shift of about 80 fm for a centre wavelength of 800 nm, with a deflection $\theta_m = \arcsin(\lambda_{opt}/\lambda_{aco}) \simeq 10^{-6}$ rad. This is negligible both in terms of wavelength and deflection; however this 1st-order diffracted beam will affect the pulse width and profile. Instead, [230] proposed a novel double-pass arrangement to cancel out any angular dispersion and frequency shift. This is feasible as the second pass of the laser through the AOM yields a negative 1st-order diffraction, thus canceling out the initial 1st-order diffraction caused by the first pass Figure 20.8. The acoustic frequency signal is then provided by a TTL logic setup controlled by the PC, based on the tool path generation algorithm. One such example has been implemented in the system, as shown in Figure 20.7.

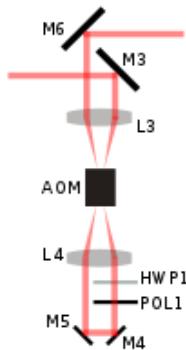


Figure 20.7: Optical Shutter Implemented

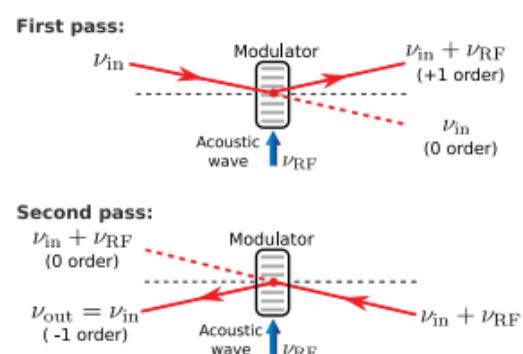


Figure 20.8: Zero Frequency Shift Illustration [230]

The key specification for this optical shutter include the shutter rate and the attenuation. [230] indicates the average shutter time is about 25 ns, a shutter rate of 40 MHz. The attenuation specified is about 46 dB, which yields a maximum attenuated energy of 3.8 nJ at maximum power. While this is not enough to ensure safety considerations are met (and hence the mechanical shutters are placed in the system), the attenuation is enough to avoid extraneous machining, and hence this arrangement has been utilized in this system to decrease fabrication timings.

20.5 Beam Redirection

One of the key considerations when working with complex, integrated optical systems is the need to redirect beam(s) of different wavelengths to various points in the system layout for different purposes.

Some of the ways in which light can be separated into different components based on wavelength include AOMs, EOMs, as well as dichroic filters. Optical modulators can easily function as tunable filters, as the first-order diffracted beam deflection varies based on the wavelength of the beam section 20.4. However, optical modulators are a complex and expensive way of separating the beam, as dichroic mirrors allow for controlled redirection of beams as required for the imaging devices, and are a cheaper alternative to beam redirection.

Dichroic mirrors (DM) work by transmitting a particular band of wavelength (known as the pass-band), while reflecting another wavelength band, all centred around a cut-off wavelength, with a large drop-off in transmittance/reflectance. If a dichroic mirror reflects shorter wavelengths, it is known as a longpass DM; the converse of which is a shortpass DM. **RDM1,2** were chosen to reflect 532 nm, 635 nm respectively and pass 800 nm, and hence a longpass DM was chosen with a cutoff at 650 nm, which satisfies the requirements for the input to the PMT and CMOS camera. **RDM3** must reflect wavelength of about 400 nm, and pass 800 nm, as the reflected 400 nm light from the sample must be redirected to **PMT2**. For this, a lower cut-off wavelength of 525 nm was selected in order to allow for stability in the transmittance/reflectance.

20.6 Fibre-coupled LED

Fibre-coupled light is used in optical communications and fabrication setups, as it provides a pragmatic method of redirection without complex physical arrangements such as mirrors. While fibre-coupling is often used for coherent light sources such as continuous wave (and sometimes pulsed) lasers, this was not implemented.

Firstly, fibre-coupling was not implemented for CW lasers because the LIDT fluence threshold for most fibers is too low. A polarization-maintaining fibre in Thorlabs has a diameter of $3 \mu\text{m}$ and a LIDT threshold intensity of 250 kW cm^{-2} for a CW source. However, the maximum intensity for **LAS:355** is 884 kW cm^{-2} , which would burn the fibre; the maximum power (and hence the intensity) of **LAS:532** is even higher. Secondly, fibre-coupling causes high dispersion for pulsed lasers, even in a single-mode fibre, due to the core of the fibre having a higher GVD compared to air (see Table 19.1), and hence the dispersion and pulse broadening models for lens GVDs developed in section 19.2 would be extended, causing severe pulse broadening.

Instead, the LED light sources used for stereolithography imaging and live transmissive illumina-

tion were chosen to be coupled, due to the low power - about 14 mW, intensity 49.5 kW cm^{-2} - and ease of beam redirection, which allows the butt-coupled LED to reach the de-coupling lens in the best position possible. The wavelength of the LED sources was set at 625 nm to avoid any possibility of extraneous fabrication or curing of the nearby sample due to energy absorption. A graded-index (GRIN) coupling lens was considered, since the graded index of the lens assures that light at different spatial points travels the same distance (by Snell's Law) and hence maintains coherence. However, for an incoherent LED, this was not required, and a simple fibre-collimating lens was placed at the end of the fibre opening to collimate the light to ensure that the collimated LED output is used for transmissive imaging.

20.7 Beam Expander/Condenser Design and Implementation

Beam expanders are a relatively old component, with various varieties of beam expanders used in different laboratories. The different configurations considered in the design of the system include simple telescoping arrangements, such as Galilean (concave-convex) and Keplerian (convex-convex) arrangement, a three-lens variable expander system using a relay lens, as well as prism beam expanders. Out of all the choices, however, the Galilean telescopic lens arrangement was chosen for all three beam expanders.

20.7.1 Design Alternatives

The main reason for considering a prism beam expander is because a prism beam expander offers the opportunity for variable beam expansion, when implemented using 2 prisms using the formula [231] given as

$$M = \left(1 + \frac{A^2}{2} \left(2K \frac{N_1}{N_2} + (N_1 - KN_2)^2 - 1 - K^2\right) - i_1 A \left(\frac{N_1^2 - 1}{N_1} - K \frac{N_2^2 - 1}{N_2}\right)\right) \quad (20.6)$$

where A is the transverse prism angle, K relates the transverse angle of the two prisms by $B + KA = 0$, i_1 represents the incident angle and N_1, N_2 represent the refractive indices of the individual prisms. While these prism beam expanders provide a useful method for controlled beam expansion, the main disadvantages include the actuation complexity of aligning the prisms to avoid any anomalous angular dispersion, and the astigmatic correction required. The astigmatic correction would be required as anamorphic prism pairs are used in industry to correct astigmatism in beams by expanding in a single lateral direction, depending on the orientation of the beam relative to the prism. Hence in this case, 4 prisms would be required, which would significantly increase the system complexity and error detection for a simple beam expander setup. The system requirement for beam waists

through the system indicate about 10 – 12 mm at the SLM for the Ti:sapphire laser, about 4 mm at the galvanometers for both **LAS:532** and **LAS:355**.

20.7.2 System Implementation

In terms of layout considerations, the beam expander for the Ti:sapphire was placed at the seed laser output, to magnify the beam waist from 1 mm to 5 mm, ensuring that the beam waist through the variable attenuator to **SLM2** remains about 5 mm. The amplified output already has a beam waist of 5 mm, and hence the magnification of **L1,2** will be 5.67, which is set by the focal length ratio f_2 / f_1 . The closest implementation is set to 5.67, yielding a beam waist of about 5.67 mm; this is not a problem as the extra portion of the beam will not be reflected off the SLM, and is of sufficiently low power to have no effect to system operation. The beam is then demagnified at the 4F setup of lenses **L7,8**. In order for the beam to fit the galvanometers' maximum beam waist specification of 5 mm, as explained in chapter 22. To do this, a demagnification of about 2.5 was chosen, to ensure that the variation of the beam with respect to the galvanometers had maximum effect on the field of view of the objective. This lead to the beam waist of 2 mm at the galvanometers. In such a case, the ratio of f_8 / f_7 was set to 2.5, yielding a beam waist of 2 mm. Finally, in order for the beam to fill the entrance aperture of the objective, the magnification of **L12,13** was chosen to be 5, as the beam waist at the objective would be equal to 10 mm, enough for the galvanometers to adjust the angle of the incident light.

20.7.3 Summary

Throughout this optical setup, various beam control techniques have been utilized to control the three different laser choices, as explained in this section. All the component selections as well as layout decisions chosen were based on integrating the process capabilities of the system to ensure that technical specifications of fabrication time, materials, and imaging requirements are met through various primary components such as a high-speed shutter with rate above 40 MHz, a femto-second pulsed laser, and relevant secondary components. Having considered the major aspects of beam control through the system, the final component of the setup involves the spatial and temporal beam shaping, which will be discussed in the following section to round off the beam control of this optical system.

20.8 Pulse Shaping

20.8.1 Motivation

Pulse shaping is the manipulation of an ultrashort laser pulses' temporal profile, thus providing the capability to produce a user defined wave form. Over the past 20 years, pulse shaping methods have had a rapid development and now provide a method to have extraordinary control of ultrashort pulses. The key motivation for the pulse shaper is to correct for pulse stretching through the optics due to the varying group velocities with the spectral bandwidth of a pulse. Two photon polymerisation, laser micro machining and STED are designed for 100fs pulses in order to produce the desired resolution of the system. Therefore the pulse shaper is required to be able to correct for the stretching as the pulses propagate through the system, so that at the objective the pulses are the required 100fs. STED also requires 20 – 30ps depletion pulse to suppress emissions around the focus and the narrow the point spread function, thus the pulse shaper must also accommodate for this.

This section will start by describing the basics of pulse shaping — based upon a linear time invariant filter and introduce the theory of diffraction gratings. It will then go on to provide an implementation of the pulse shaper, including the specification of selected components and instrument control.

20.8.2 Pulse Shaping Basics

The pulse shaper used in the system is based upon a linear, time invariant filter — a concept well known in electrical engineering which is commonly used to process electrical signals ranging from low frequencies (audio and below) to high frequencies (microwave).[232] Here the linear filter is used to produce specifically shaped optical waveforms on the order of femtoseconds. As shown in Figure 20.9, linear filtering can be described in either the time or frequency domain. In the time domain, the output of the filter $e_{\text{out}}(t)$ is given by the convolution of $e_{\text{in}}(t)$, the input pulse, and $h(t)$ the input response function.

$$e_{\text{out}}(t) = e_{\text{in}}(t) \star h(t) = \int e_{\text{in}}(t')h(t - t') dt' \quad (20.7)$$

If the input is a delta function, then the output is simply $h(t)$. For a sufficiently small input, the task of generating a specific output is, therefore, to generate a linear filter with the desired impulse response. As mentioned the filter can also be defined in the frequency domain, characterised by its frequency response $H(\omega)$. The output $E_{\text{out}}(\omega)$ is the product of the input signal $E_{\text{in}}(\omega)$ and $H(\omega)$. Where $e_{\text{in}}(t)$, $e_{\text{out}}(t)$, $h(t)$ and $E_{\text{in}}(\omega)$, $E_{\text{out}}(\omega)$, and $H(\omega)$ are Fourier transform pairs.

$$E_{\text{out}}(\omega) = H(\omega)E_{\text{in}}(\omega) \quad (20.8)$$

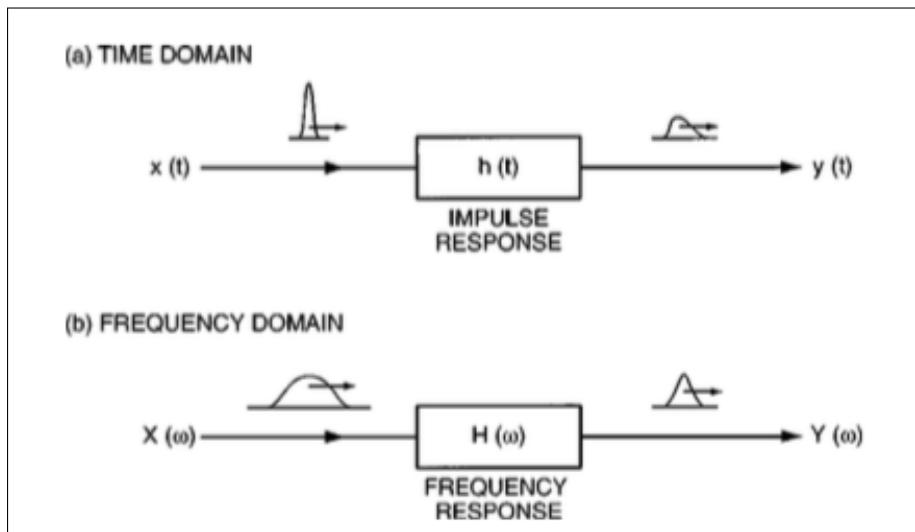


Figure 20.9: Pulse shaping by linear filtering. (a) Time domain view. (b) Frequency domain view.[232]

For a delta function input, $E_{\text{in}}(t)$ is 1, and so the output $E_{\text{out}}(t)$ is equal to frequency response of the filter, as in the time domain.

20.8.3 Diffraction Gratings

Diffraction gratings enable the optical system to separate light of different wavelengths with high resolution. As such, they enable the pulse shaper to stretch and compress the ultrashort pulses pre and post processing. Manufacturers produce a wide range of designs: transmissive, reflective, ruled and holographic, which are designed on extensive research, in order to produce dispersion free gratings. An understanding of the grating equation and the aforementioned types of grating is required to ensure the correct choice is made to meet our systems requirements.

When monochromatic light is incident on a grating surface, it is diffracted in discrete directions. This diffracted light then recombines to form a new set of wavefronts. The grating is, therefore, useful due to this unique set of discrete angles. The grating leads to constructive interference as the diffracted light from each groove is in phase with every other groove. These discrete angles are illustrated in Figure 20.10 and are determined by the grating equation:

$$\sin \theta_i + \sin \theta_d = \frac{n\lambda}{d} \quad (20.9)$$

Where λ is the wavelength of the light incident on the grating at an angle of incidence θ_i to the normal, d is the groove density, n is the order of diffraction and finally θ_d is the angle of diffraction. Diffraction gratings in certain configurations concentrate a larger proportion of the energy incident on them into a specific diffraction order. This phenomenon is called blazing. The Littrow configuration is a special geometry in which the diffraction and incidence angles are identical. This means the diffracted beam is back reflected onto the incident beam. This configuration is where the diffraction

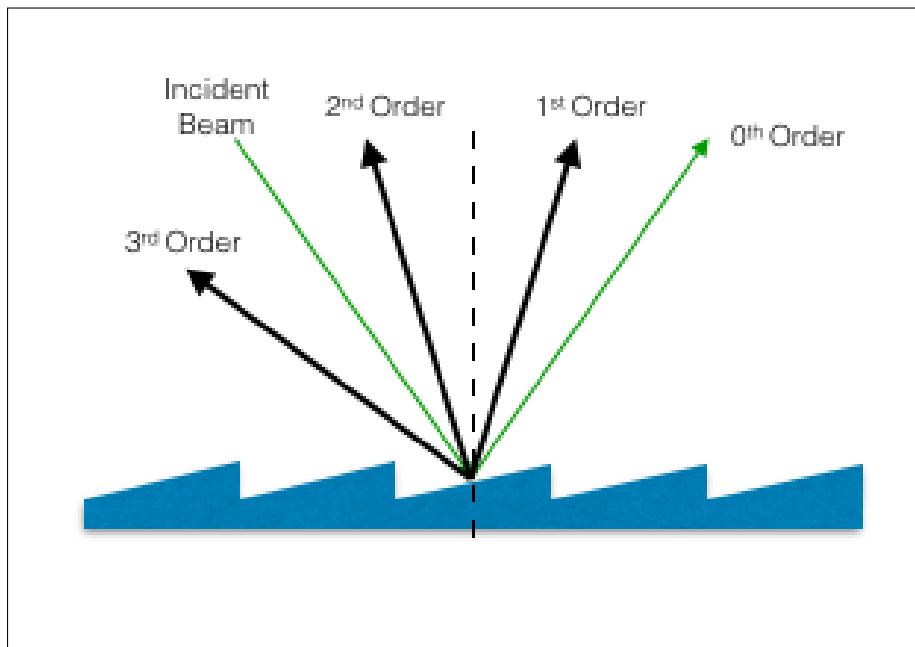


Figure 20.10: Geometry for diffraction gratings

grating is operating at its maximum efficiency.[233]

The spatial arrangement of the pulse shaper requires a reflection grating which consists of a grating superimposed on a reflective surface. Diffraction gratings are manufactured either by burnishing grooves with a diamond tool (ruled gratings) or holographically using the interference fringes generated at the intersection of two laser beams (holographic gratings). Ruled gratings are very efficient and are the best solution when working with low groove densities and in IR above $1.2\mu\text{m}$. Whereas holographic gratings are the best solution anytime the groove density is required to be above 1200g/mm and for use in the near UV, VIS, and NIR spectral regions.[234] Selection of the individual gratings depends upon the focal length of the lenses and the laser source.

20.8.4 Implementation

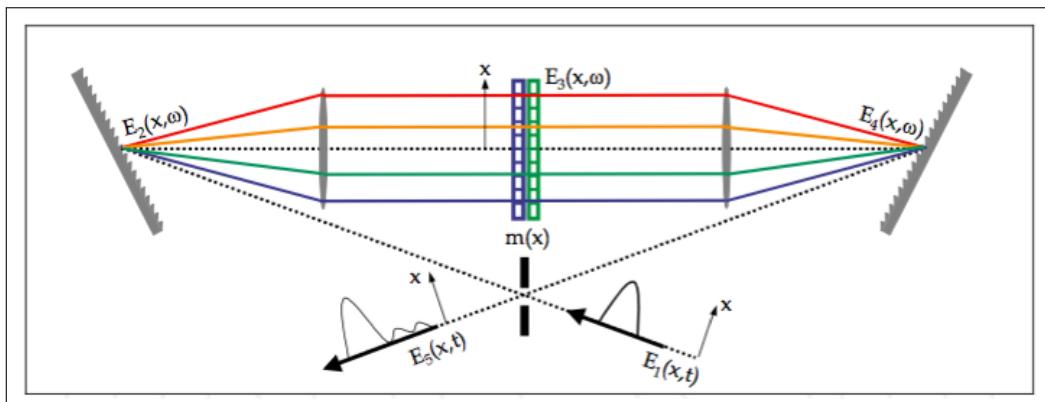


Figure 20.11: Geometrical layout for the zero compression pulse shaper[235]

Figure 20.11, shows the implementation of the pulse shaper, consisting of a pair of diffraction

gratings and lenses, and a SLM. The shaper is configured in a “zero dispersion pulse compressor” arrangement. The configuration — using the first diffraction grating — disperses the ultra short pulses at the input and then the first lens performs a Fourier on this angular dispersion converting it to a spatial separation at the back focal plane. The SLM is placed in this plane in order to manipulate the spatially dispersed Fourier components.

For this configuration to work it is required (in the absence of the SLM) that the input and output pulses are identical. Therefore the grating and lens configuration must be dispersion free. This can be achieved if the lenses are in a unit magnification telescope configuration: with the gratings positioned at the outer focal planes of the telescope. The system, thus performs two Fourier transforms upon the input and so remains unchanged. This condition depends upon various approximations: that the lenses are thin and aberration free; that chromatic dispersion through the system is small; and finally, that the gratings have a flat spectral response.[232]

The pulse shaper is designed for a specific laser source and SLM. The laser source has an energy per pulse of 4mJ at a centre frequency of 800nm . The SLM is the CRi D640-VN model, the selection of which is discussed in section 20.9. It has 640 pixels each with a width of $100\mu\text{m}$; an inter pixel width of $2\mu\text{m}$; and a pixel height of 5mm ; giving an overall dimension of $64 \times 5\text{mm}$.[236] The width of the SLM is the main criterion for the pulse shaper. It is desirable that a large bandwidth of the pulse is used so that we have an efficient modulation. However, a smaller bandwidth increases resolution - as each pixel modulates a smaller part of the most important sections of the spectrum.[235] It was decided that the cut-off criterion to be used was any part of the spectrum that contributes intensities of 2-3% of the maximum would be included. For our laser that is a range of 780nm to 820nm .

Having chosen the frequency range, we were positioned to select the grating density. In general, a higher groove density leads to a higher diffraction efficiency. For our centre wavelength, there are only a few suitable gratings. The decision was to use the Thorlabs’ GH13-18V - a holographic grating with a density of $1800/\text{mm}$, as this provided the highest dispersion efficiency for our required wavelengths. On the overall optical layout **DIFG1** and **DIFG2** will therefore be the aforementioned gratings. See Figure 20.12 for the efficiency curve.

The lens focal length selection depends upon the angle of incidence, spectral width, and SLM width. Figure 20.13 shows how the different design parameters are related geometrically. The angle of diffraction of the smallest incident wavelength is:

$$\beta_l - \beta_m = \arctan\left(\frac{w_l}{f}\right) \quad (20.10)$$

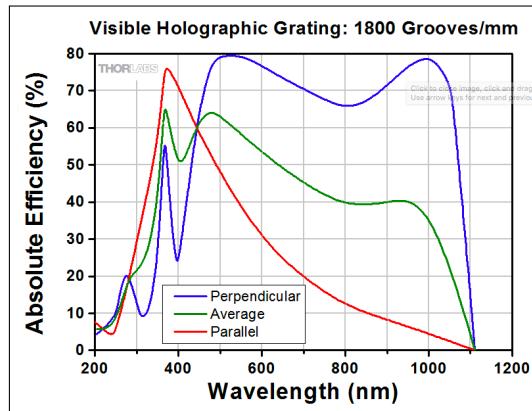


Figure 20.12: Grating Efficiency curve for Thorlabs' GH13-18V[237]

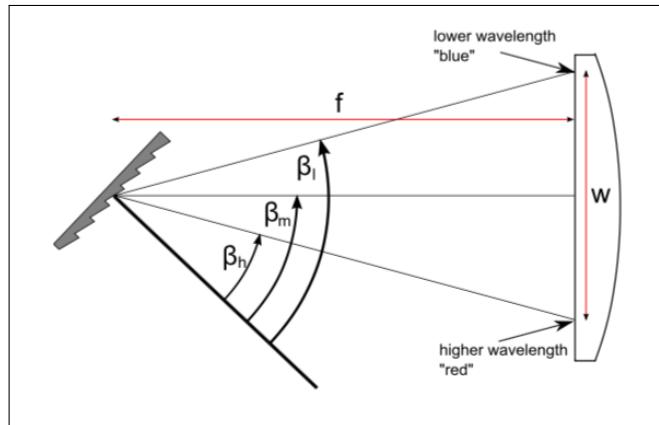


Figure 20.13: Geometry of the pulse shaper design parameters[235]

Ideally a near Littrow configuration is used in order to maximise efficiency. MATLAB was used in order to work through different focal lengths and input angles to see which combinations would produce the required spectral width and give a near Littrow configuration. It was decided to use lenses of focal length 200mm and an input angle of 29.5° . This gives the lower wavelength cutoff a diffracted angle of 64.6° and the upper wavelength cut-off an angle of diffraction of 82.8° - thus producing a spectral width 64mm and filling the SLM. On the overall optical layout **L5** and **L6** will therefore be Air-Spaced Doublets with a focal length of 200mm .

At the Fourier plane the beam width should not exceed the width of a single pixel on the SLM. The pixel width is 5mm . The beam width in the Fourier plane can be calculated using the equation

Component	Part Number	Technical data
Diffraction Grating	Thorlabs GH13-18V	Holographic reflective, 12.7m^2 , grating density = 1800
Lens	Thorlabs ACA-254-200-B	Air-Spaced Doublet, focal length $f = 200\text{mm}$
SLM	CRi D640 - VN	Transmissive 640 pixels

Table 20.3: Components table for the pulse shaper

below.

$$W'_0 = \frac{\lambda}{W\pi} f \quad (20.11)$$

[235] Using $\lambda = 780\text{mm}$, $W_0 = 1\text{mm}$ and $f = 200\text{mm}$ gives a beam width of $52\mu\text{m}$ which is well within the pixel width.

Finally the damage threshold of the SLM must be considered. The average energy of a laser pulse is about $4\mu\text{J}$. Focused at a height of 64mm and spread to a width of $50\mu\text{m}$, gives a fluency of $12.5\mu\text{J}/\text{cm}^2$ which is far below the SLM's damage threshold, and can therefore be used without damaging it.

20.8.5 Control

The control strategy for the pulse shaper will be to open loop. This requires knowledge of the input pulse and careful calibration of the gratings and lenses. The user will provide the desired output waveform and this transfer function will be programmed into the SLM. If there is additional linear distortion between the input and output (for example, due to a phase aberration) the pulse shaper can be programmed so that the transfer function compensates for this.

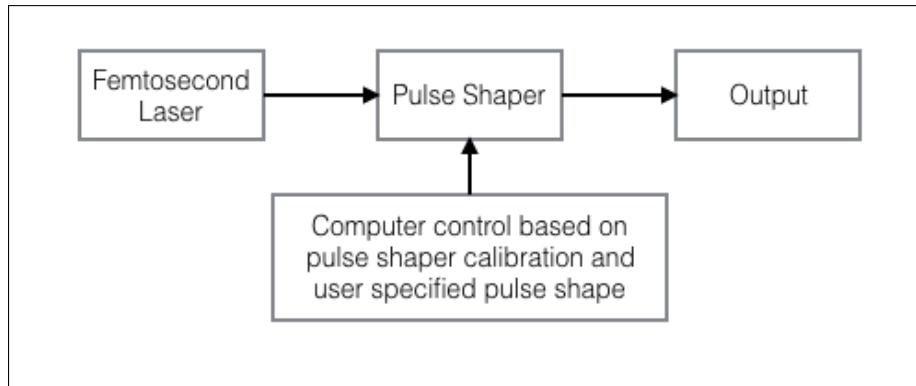


Figure 20.14: Pulse Shaper Control block diagram

20.8.6 Comparison of Dispersive Components

The choice of the linear SLM and diffraction gratings was based on a comparison of various pulse shaping techniques utilized today. The implementation chosen was carefully chosen to extract the maximum utility for the system specifications.

Dispersive components are used in the pulse shaper to spatially separate the constituent waves in the pulse bandwidth for further modulation. The implementations considered in this system include diffraction gratings, prisms, and dielectric mirrors.

In order to maximize the spectral separation of the pulse, the spatial deviation of individual component waves should be as large as possible. This is useful when the entire range of the bandwidth falls on the linear SLM, which can efficiently generate specific pulses by modulating the spectral components. The two most dispersive components considered include the diffraction grating and the prism. Given the deviation angle for a specific wavelength, the deviation for pulse bandwidth $\Delta\theta_m$ was calculated for a pulse bandwidth of 30 nm, using the equations below.

$$\Delta\theta_{DG} = \frac{m\Delta\lambda}{d\sqrt{1 - \left(\frac{m\lambda}{d} - \sin\theta_i\right)^2}} \quad \Delta\theta_{PR} = \frac{\partial n}{\partial \lambda} \alpha \Delta\lambda \quad (20.12, 20.13)$$

where m represents the diffraction order, d represents the inter-slit distance, θ_{DG} represents the incidence angle at the diffraction grating, θ_{PR} represents the incidence angle at the prism, and α represents the apex angle of the prism. Since the grating offers greater spatial variation for a large spectrum as $\Delta\theta_{DG} > \Delta\theta_{PR}$ for the wavelength range, the spatial variation allows finer separation of individual spectral components, improving the performance of the grating pair in comparison to the prism. Hence, grating pairs are hence more useful in compressing (and expanding) pulses than prism pairs.

In comparison, dielectric mirrors fall short in terms of dispersive capability. In practice, multiple reflections off dielectric mirrors are required for a sufficient pulse shortening effect; hence gratings were selected as the primary dispersive component. The only disadvantage of grating pairs is that the power attenuation can often be as high as 30%, a major consideration in the safety specifications considered earlier. However, since the system incorporates a high-power amplified laser, the system provides ample power for micromachining, while providing functionality for custom pulse shaping.

20.9 Pulse shaping using an SLM

The main role of the pulse shaper is to modify the pulse length and profile the beam of the laser passing through the system and used for fabrication through laser micromachining, 2photon polymerization and STED. SLM based on liquid crystal technology is one of the options considered and is seen to be the better one to serve as spatially patterned masks. It is the component that manipulates the fourier transform of the beam before it again recombines to be sent forward to the system, modifying the pulse amplitude and phase as needed by each fabrication technique. This is controlled by a computer electronically and the software for that is usually provided with the SLM by the company providing it. The electro-optic properties of the liquid crystals used in SLM are given in chapter 15 in addition to a detailed explanation of how an SLM works. Here, we discuss in brief the relevant functioning of an SLM to modulate phase and amplitude in pulse shaping, look at the advantages of a liquid crystal based SLM and make a choice of SLM based on key characteristics of relevance to our system.

The light needs to split up into its spectral components before the spatial light modulator can be used for spectral shaping. This is done in the system through diffraction gratings. In the setup, the left and right half are a mirror image of each other and the light then recombines before exiting this section of the setup. Figure 20.15 shows an example of the splitting and recombination with different frequencies of light. Depending on the frequency the colours split up, at different angles, converting the spectral difference into a spatial difference. [160] Each colour is focussed on different pixels of the SLM, which means each colour is separately modulated. This again is a 4f system, the principles of which have been explained in chapter 18.

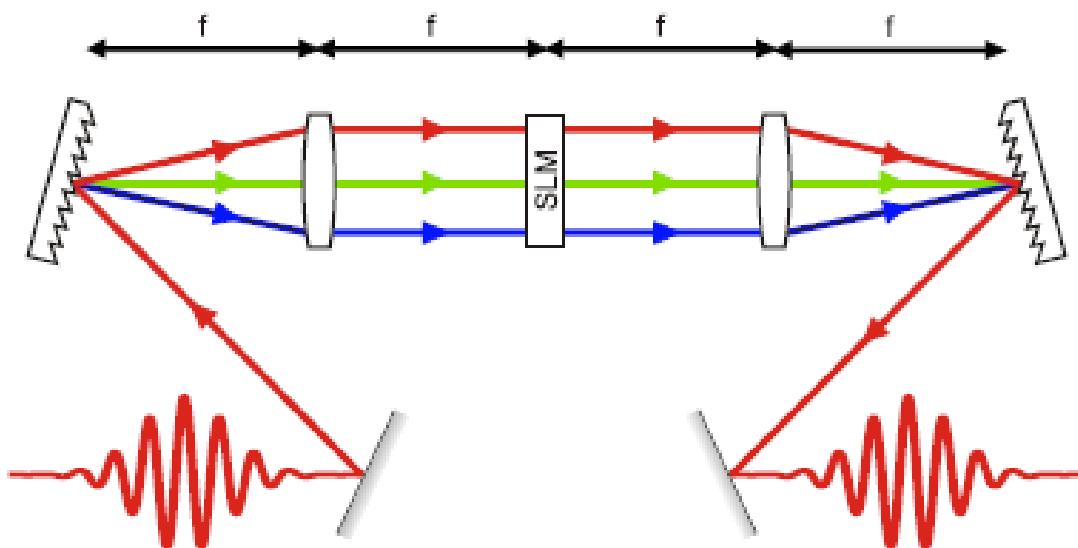


Figure 20.15: Example of setup for pulse shaping [160]

SLM based on two broad technologies are used for pulse shaping, which requires phase modulation of the light; liquid crystals and deformable mirrors. Deformable mirrors change the length of the physical path of the light. This is made possible by the surface of the mirror being deformed, so that the distance travelled by different parts of the spectrum can be changed. [160] They are microscopic hinged mirrors that either tilt towards or away from the light source to reflect or block the light respectively.

SLM technology is more centred around liquid crystals and their many applications owing to their property of birefringence. Birefringence is the ability of a material to change the refractive index depending on the polarization of light and its direction of propagation. The SLM is composed of a nematic liquid crystal layer between two glass layers for transmissive, and a glass layer and silicon coating for reflective set up of the SLM. The molecules of the liquid crystal realign on application of an external electric field such that the direction of long axes of the molecular tilts, causing a change in the refractive index (birefringence). In order to apply an external field, the surface of the glass has a thin film of tin oxide painted, and is composed of a number of pixels (separate electrodes) allowing individual and more accurate control. [238] A minimum phase control of 2π is needed to be able to phase modulate freely. For achieving pulse position modulation, Weiner has experimentally found the relation for a phase versus voltage calibration to be related by the transfer function of,

$$T(V) = \sin^2 \frac{(\phi(V)) - (\phi_x)}{2} \quad (20.14)$$

Applying a voltage across the liquid crystal causes the molecules to tilt from original position. This gives a non-linear variation of phase with voltage[239], and would be the process for phase modulation. The retardance caused by this is given by, [239]

$$\tau = 2\pi(n(\theta) - n_o) \frac{d}{\lambda} \quad (20.15)$$

$$\tau_{max} = 2\pi(n_e - n_o) \frac{d}{\lambda} \quad (20.16)$$

$$\frac{1}{n^2(\theta(V))} = \frac{\cos^2(n(\theta(V)))}{n_e^2} + \frac{\sin^2 n(\theta(V))}{n_o^2} \quad (20.17)$$

Where τ is the retardance, λ is the centre wavelength, d is the thickness of the liquid crystal, n_e is the refractive index along the extraordinary axis and n_o is the refractive index along the ordinary axis. Retardance is the representation of the integrated effect of birefringence on a beam.

Amplitude variation on the other hand, requires the light to be first polarized before entering the SLM, which is fairly true in our system as the light from the TiS laser is polarized and is further

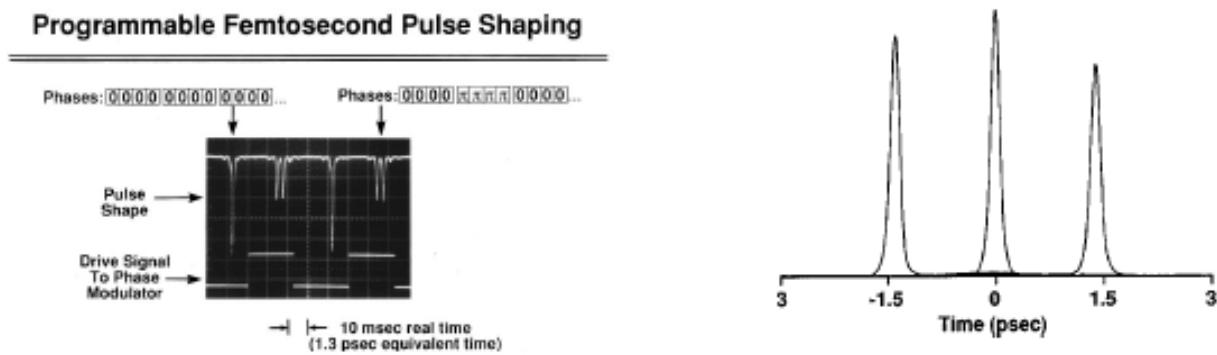


Figure 20.16: Phase change leading to shaping of pulse accordingly [238]

also passed through the polarising beam splitter. The variable attenuation setup hence gives the horizontally polarised light. This is important because unpolarised light wouldn't be allowed to pass through a nematic liquid crystal medium as the molecules would realign with the electric field on application of voltage. This can be understood better in Figure 20.17. The SLM can be used either in reflective or transmissive mode, the reflective one forcing the beam to go through the mask twice and any effect on it would be doubled, implying higher phase modulation rates being allowed.[239] We will be using a transmissive SLM instead to meet other more important benefits achieved by arranging the system as shown in subsection 20.8.2 Amplitude modulation can be done using a linear SLM as each frequency of light after the splitting is oversampled, hence multiple pixels are modulating each of the focussed frequencies.[7]

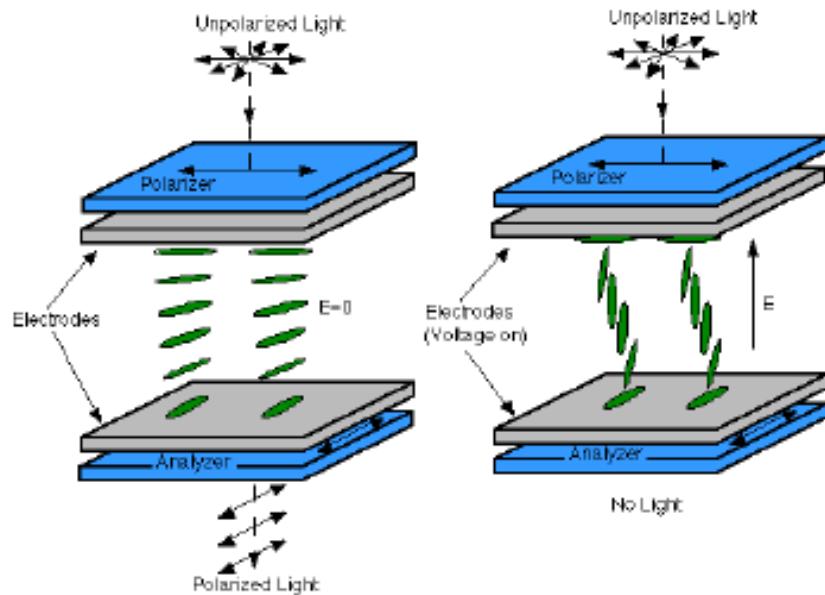


Figure 20.17: Voltage through SLM restricts the passing of unpolarised light [240]

Although using a linear SLM and only having direct control over the phase has been used for pulse shaping most frequently, Frunker and Silberg [173] found that using a 2D SLM capable of both amplitude and phase modulation independently would help to remove errors in accuracy due

to alignment and the gap between the two masks introducing further optics. However, a linear SLM would be beneficial in terms of response time and the ability to produce complex spectral phases and ultrafast amplitude modulation. Further, there would be an added cost to using a 2 dimensional SLM. Key advantages of using an LC based hence SLM over other dynamic modulators would be [238]:

1. Analogous phase control of each separate pixel as compared to other systems in which only binary phase modulation is possible
2. Programmable control of the pulse shape on a millisecond timescale
3. Faster response time

In choosing an SLM, not a lot of options are available in the market for linear SLMs anyway and Cambridge Research and Instrumentation (CRi) has usually emerged to be the best choice, due to better accuracy and involvement in research. Our constraint would be to find a linear SLM which is transmissive, has 2 masks and has operating wavelength of 800nm. With CRi there are two options, either a 128 pixel or 640 pixel SLM [22]. An increased number of pixels would give higher resolution and better accuracy which is crucial in our system, hence it would be better to choose the one with 640 pixels. Further to choose between a VN or NM model, the response time is 35ms for a VN model which is half that of NM. Our choice hence corresponds to CRi D640 – VN model. The distance between the two mask is 1.03mm while the inter-pixel gap is 2 μm which are both fairly low and help in achieving accuracy by minimising cross talk. The SLM is addressed by 12 bits and it achieves a good transmission efficiency of 88%. The SLM is packaged as shown in Figure 20.18 and the transmission set up can be understood through Figure 20.19



Figure 20.18: CRi SLM-640 [241]

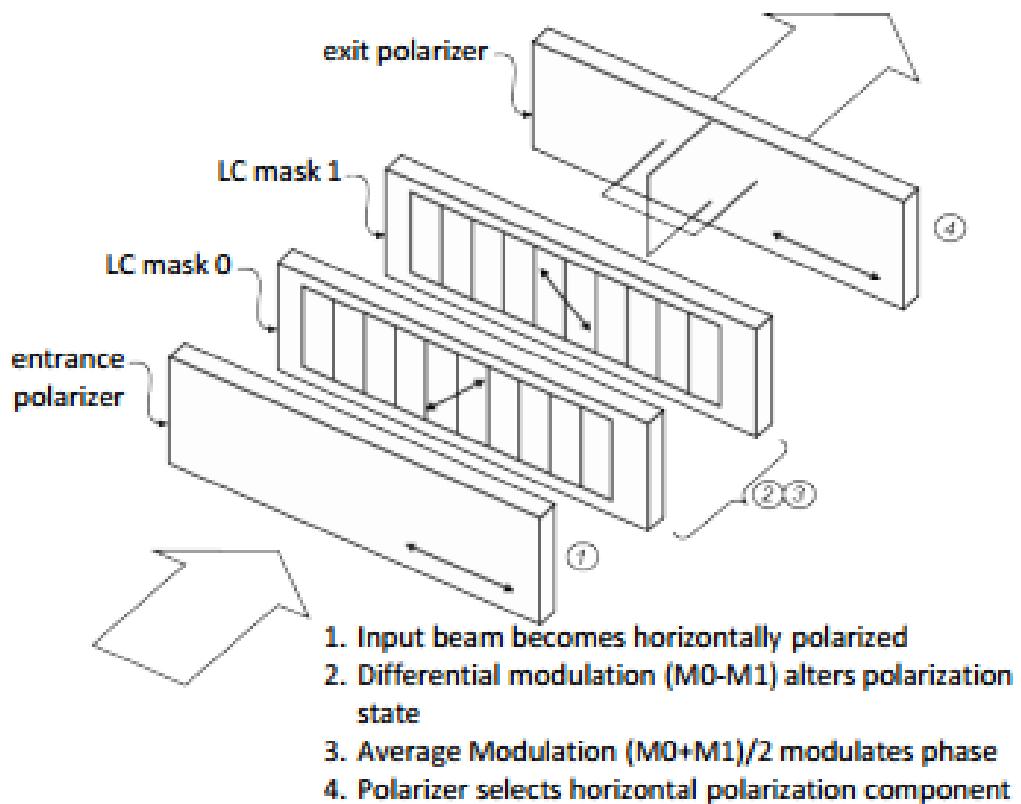


Figure 20.19: Transmissive 2 mask set-up of SLM[241]

20.9.1 Conclusion

An optimal pulse shaper design was presented which counteracts the stretching of the system, therefore ensuring that two photon polymerisation and laser micro machining have the required 100fs pulse thus improving fabrication and imaging resolution. The pulse shaper is also able to stretch the 80MHz laser output so that it can be considered quasi-continuous and produce the depletion pulse for STED. The system is based on the common 4f zero dispersion arrangement. The SLM allows our users the capability to implement complex waveforms. In conclusion, the addition of the pulse shaper provides our system with greater manufacturing capabilities through a robust, repeatable and user friendly system.

Part V

Mechanical Design

21 Overall Mechanical Design

21.1 Introduction

For technical fields such as optomechanics, and indeed many others, the practicing Engineer frequently calls upon past experience and knowledge to inform their design decisions. They accrue a wealth of understanding that allows for a certain grasp of what 'feels right', when to apply a proven and well documented design pattern, and ultimately, to be able to comprehend when a solution makes sense. Such expertise is not assumed here. The material covered in the following section will provide a clear and concise overview of the physics at play underpinning the optomechanical design process; from the forces brought about through different mounting solutions, to the resulting degradation in optical performance. A rigorous derivation of the equations used throughout will not be provided, but left to the reader to seek out at their own discretion.

21.2 Optical Table

Sources of vibrations can range from vehicular traffic from nearby streets to the sway of the very building housing the fabrication unit, with a myriad of sources in between. For an application requiring such high precision, and working at such fine a scale, it is of the utmost importance to be able to not only isolate the working section of the system from exterior sources of vibrations, but also effectively remove any vibrations created upon the working section itself, such as the movement of an active stage, or the rotary motion of a cooling fan. As such, the motivation of the next section is to cover the vibrational theory required to understand the design and requirements of an optical table, and also provide the justification of the final choice of table for the system.

21.2.1 Vibrational Theory

Sources of vibrations can be broadly lumped into three categories [242]

1. Seismic
2. Acoustic
3. Direct contact

Seismic vibrations drive the motion of the ground relative to the unit, acoustic vibrations concern the changes in pressure of the air around the unit, and direct contact sources are those caused by motion of mechanical elements inside of the unit. Sources, as previously mentioned are numerous and those shown in table 21.1 provides but a small sample of the typical sources expected to be prevalent during the operation of the fabrication unit.

Source	Frequency Range (Hz)	Amplitude (in)
Air Compressors	4 - 20	10^{-2}
Foot Traffic	1 - 5	10^{-5}
Vehicular Traffic	5 - 100	10^{-2}
Building Motion	4 - 15	10^{-1}
Elevator Motion	0 - 40	10^{-3}

Table 21.1: Typical sources of vibrations and their amplitudes [242]

The table is by no means exhaustive and the entries are meant only to provide estimates of the frequency ranges and amplitudes expected, quantifying such complex behavior is difficult, and in the absence of the specific environment that will house the fabrication unit, downright impossible. Nonetheless, the estimates given above can help to outline the typical range of frequencies that the unit is expected to be exposed to, and thus those that it must be shielded from the most.

Compliance offers a useful metric by which to assess the performance of a system in response to a vibrational forcing function and is defined as:

$$\text{compliance} = \frac{\text{amplitude of excitation}}{\text{amplitude of forcing}} \quad (21.1)$$

Although optical tables are complex mechanical systems, with several modes of vibration that can couple together to produce complex frequency responses and resonance curves, intuition about their operation and design can be gained by consideration of the simple mass spring damper as depicted in figure 21.1.

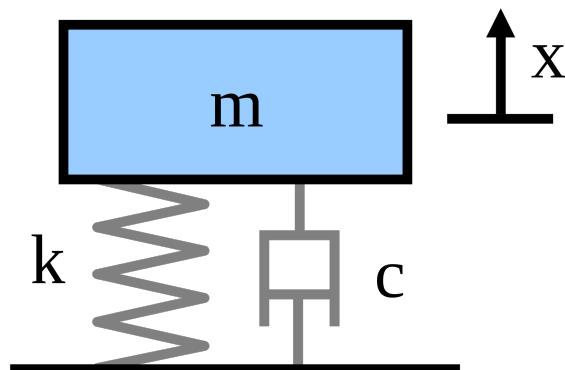


Figure 21.1: Simple Mass-Spring-Damper system [243]

When subjected to the forcing function:

$$F = F_0 \sin(\omega t) \quad (21.2)$$

It can be shown the system is governed by the following differential equation:

$$m\ddot{x} + c\dot{x} + kx = F_0 \sin(\omega t) \quad (21.3)$$

Which admits the solution, given in terms of the systems compliance as:

$$\frac{x}{F} = \frac{1}{\sqrt{(k - m\omega^2)^2 + (c\omega)^2}} \quad (21.4)$$

Note, analysis is done with a simple sinusoidal forcing function because any arbitrarily complex forcing function can be decomposed into its constituent frequencies, and the response of the system built up through a superposition of resulting frequency responses. This can be seen to contain three significant terms in the denominator: a stiffness term, a mass term and a damping term, each of which dominates over a certain range of frequencies; stiffness at low frequencies, damping at resonance, and mass at high frequencies.

21.2.2 Resonance

Resonance is the sharp peak in the frequency response of a system, corresponding to the systems natural frequency. The action of resonance can be understand by considering the progression of a system from static deflection, to resonant frequency. As the forcing function increases in frequency the system begins to lag behind due to its finite momentum preventing it from being able to reverse its motion at the same instant as the forcing reverses its. This lag increases, and reaches a point where the maximum of the forcing function coincides with the maximum acceleration of the forced system, at this frequency given by:

$$f_n = \sqrt{\frac{k}{m}} \quad (21.5)$$

A large amplification of the input function can be achieved. It is extremely important the optical table is designed such that its natural frequency is shifted out of the range of the typical background vibrations. This is generally considered 0 – 150Hz, as can be reasoned from table 21.1. Care needs taken to ensure advances in stiffness aren't counteracted by corresponding increases in mass to remove any net shift in natural frequency.

21.2.3 Table Design

The ideal table top requires both a high natural frequency and to be itself well damped. Natural frequency is dictated by both stiffness and mass, as given by equation 21.5. Thus sufficient stiffness must be provided with minimal use of mass. In order to understand the design, the stress distribution throughout a beam section must be understood, as demonstrated by figure 21.2.

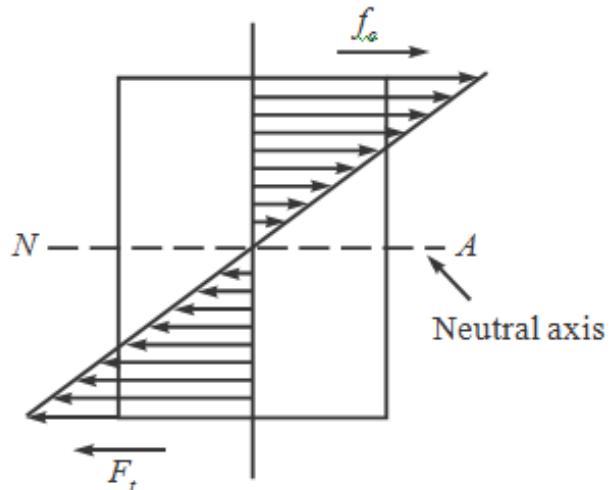


Figure 21.2: Stress distribution in a beam segment [244]

From the diagram, it is apparent a mass efficient design would be to use a hollow shell, as the vast majority of the stress in the solid structure is located at the extremities and therefore little is gained from a solid interior. However, for a large table, there would be a tendency to sag in the middle and there would be very little damping provided by the structure, as the only source is due to frictional effects at grain boundaries [245].

The most common solution to this issue is to use a honeycomb mesh, this can provide flexural rigidity in several axis, the importance of which can be seen by inspection of the the first few vibrational modes of a flat bed, as shown in figure 21.3

Providing stiffness along only a few directions would result in a table design susceptible to some of the modes not compensated for, and thus wouldn't provide a flat mounting surface. Providing stiffness in too many adds excess mass reducing the natural frequency of the table, the honeycomb mesh provides a balance, ensuring sufficient rigidity whilst increasing the mass of the table only slightly. The mesh also provides the table with extra damping, allowing for removal of vibrations created upon the tabletop.

21.2.4 Vibrational Isolation

The previous section was concerned with the table tops ability to handle internally generated vibrations, this section focuses on isolation of the table top from the vibrations of the ground. This is known

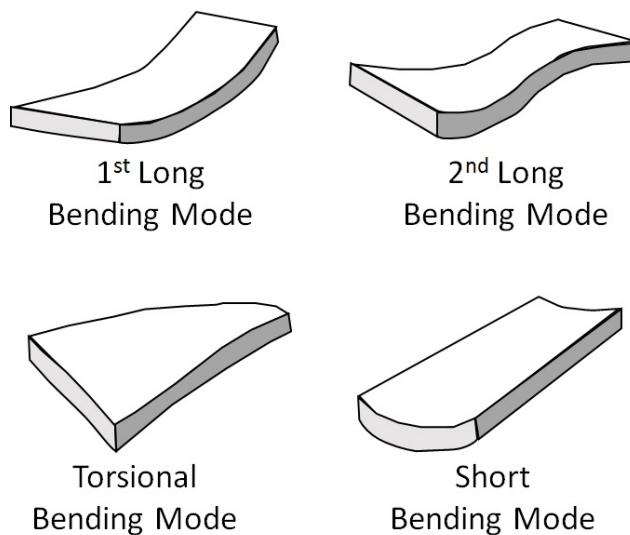


Figure 21.3: Bending modes for optical tabletop [242]

as seismic mounting. This involves mounting the tabletop on a spring system in order to decouple it from the movement of the ground below, an air cushion acts as a spring and an orifice plate as a damper. This isolation is also why the table in itself must be well damped, as the vibrations cannot be passed from the tabletop to the ground below, and thus must be compensated for in the design of the table top.

21.2.5 Table Choice

Several vendors provide optical tables, ultimately it was decided to adopt a **T1530QK** from Thorlabs. The bundle consists of a 1.5m by 3 m by 310mm tabletop and a set of active isolator table legs. The active legs require a constant source of compressed air, this will be provided by a **PTA513** air compressor, also sold by Thorlabs. The table costs £8,674.22 and the air compressor £441, totaling £9115.22.

21.3 Mounting Optics

Mounting an optical element can be done, broadly speaking, and perhaps rather coarsely too, in one of two ways; bonding it, or clamping it [246]. Bonding employs the use of an adhesive to secure optics in place with a plethora of specialized products on the market for optomechanical use. Clamping utilizes a retaining surface and application of a preloading force to provide a more temporary solution. Both methodologies offer up their own respective benefits and draw backs, the ripples of which propagate far throughout the rest of the design procedure, such that the first decision to be made was which one to adopt.

21.3.1 Adhesive Bonding

Bonding optical elements provides a permanent and rigid mounting solution. It alleviates the necessity of the preloading force, the effects of which are examined throughout the next section, but suffice it to say the resultant reduction of stress throughout the optic is beneficial to optical performance.

In order to assess the viability of optical adhesives, analysis of the forces experienced throughout the adhesive layer is required. The main source of stress throughout said layer is due to differences in coefficients of thermal expansion between the two materials being bonded together. Figure 21.4 below demonstrates the how the interaction gives rise to stress, the differences in expansion coefficients causes a warping of the adhesive layer resulting in a stress given by: [247]

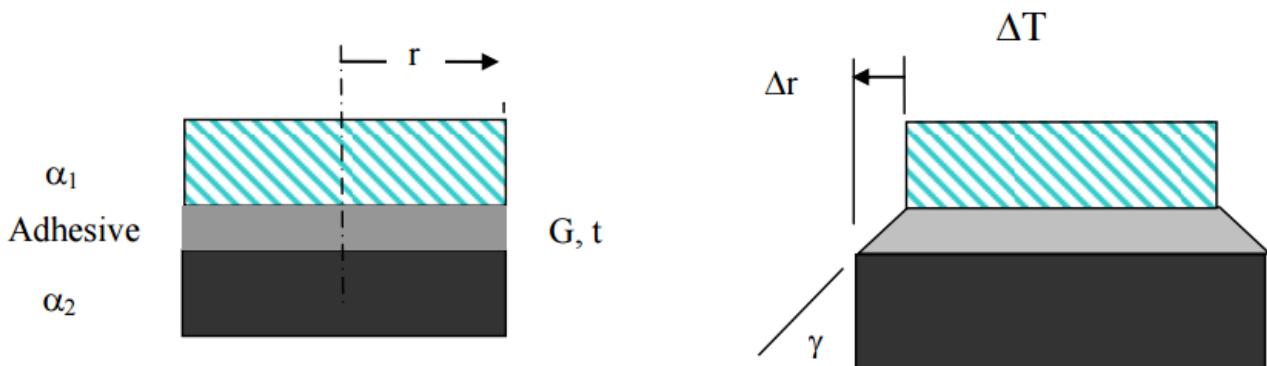


Figure 21.4: Stress induced due to differences in CTE values between bonded surfaces [248]

$$\tau_{max} = \frac{Gr}{t}(\alpha_1 - \alpha_2)\Delta T \quad (21.6)$$

Where τ_{max} is the maximum shear stress experienced, G the shear modulus of the adhesive, r is half of the bond length, t is the bond length, the α values the coefficients of thermal expansion for the two bonded materials and ΔT the change in temperature. This equation provides a conservative estimate of the resulting maximum stress, several factors can effect the accuracy of this approximation such as bond size, adhesive thickness and adhesive shear modulus, under such conditions a more accurate equation is given by [247]. For the purpose at hand however, the approximation is sufficient.

To relate to the specific scenario at hand, an example bond is taken between N-BK7, $\alpha = 7.1 \times 10^{-6} \text{K}^{-1}$ and aluminum $\alpha = 23.1 \times 10^{-6} \text{K}^{-1}$. A bond width of 3mm is assumed and thickness 0.5mm and take as an approximate worse case scenario of temperature change as 50 K. For **Norland NOA60** with $G = 350 \text{Mpa}$ we get a resultant $\tau_{max} = 1.68 \text{Mpa}$, we can thus be assured of safe operation. It is also worth noting that adhesives are rated for a specific working temperature range, for our example adhesive this is -125 to 150 degrees Celsius, exceeding this range would indicate

a serious operational malfunction and therefore we can rest assured that in such a scenario, it is perhaps fair to say, adhesive failure is of low importance.

The use of adhesives comes with a rather large set of unwanted consequences which shall be listed here. Adhesives quite often require a lengthy curing time, and for many the action of UV light is required, this would both extend the assembly time and in some scenarios necessitate the use of rather complex jigs and other such equipment. The finality provided by the adhesion causes any servicing or repair of components to become arduous and more costly. Finally, and more technically, in a process known as outgassing [247] adhesives release gaseous particles, these particles can collect on the surfaces of optics damaging coatings and degrading performance. This effect is most pronounced at high temperatures, or vacuum conditions, but also prevalent under the conditions we would expect to find during normal operation.

21.3.2 Clamping

Clamping an optical element requires the use of a retaining surface (the surface of which it is pressing against) and a preloading force (the provider of said pressing). This preloading force is the source of the main draw back of this technique, and as such, its effects shall now be analyzed.

Considering the scenario depicted by figure 21.5, that is, the use of a sharp edge mount. The maximum compressive stress experienced in the lens is given by [247]

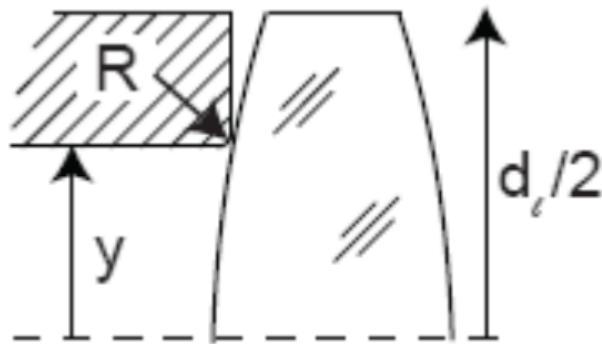


Figure 21.5: Lens mounted with sharp edge contact [247]

$$\sigma_a = 0.4 \left(\frac{F}{2\pi y} \frac{E}{R} \right)^{\frac{1}{2}} \quad (21.7)$$

Where σ_a is the maximum compressive stress, F is the applied force, y the distance from the center of the lens to the point contact and R the radius of curvature of the corner. In order to evaluate this, we require an appropriate value of the preloading force, this is given by

$$F = ma_g \quad (21.8)$$

Where m is the mass of the optic, and a_g the acceleration factor, i.e. the maximum expected acceleration normal to the retaining surface. Throughout shipping and handling a typical value of $\alpha_G = 15$ is to be expected, then a safety factor of 2 is applied before computing the force. To allow for calculation, an example optic is taken to be an N-BK7 $\phi = 1''$ lens, with Youngs modulus $E = 82$ GPa, and approximate $R = 0.05$ mm. Using a simple cylindrical approximation to the dimensions, the weight can be calculated as ≈ 5 g. This results in a maximum stress value of $\sigma_z \approx 65$ MPa.

To evaluate the acceptability of this stress, understanding of its effect on the mechanical stability of the optic, and the impact on optical performance is needed. First to be handled is resulting mechanical impact. Due to the catastrophic nature with which glass typically fails, it is crucial to ensure that there is no such risk imparted upon the optic by the induced stress. To convert this applied stress into a probability of failure, invocation of Weibull statistics is required. Weibull statistics provides a relation between applied stress and a resultant probability of failure, as given by [247]

$$P_f = 1 - \exp \left(- \left(\frac{\sigma}{\sigma_0} \right)^m \right) \quad (21.9)$$

Where P_f is the probability of failure, σ the applied stress, σ_0 the characteristic stress and m the Weibull modulus. For N-BK7, we have $m = 30.4$ and $\sigma_0 = 70.6$ MPa, this results in $P_f \approx 0$ and thus safety under the action of the applied preloading force is assured.

In order to evaluate the effect on optical performance, first an appreciation of the underlying physics that causing the degradation must be gained. Birefringence [249] is a property of optical materials that results in differing indexes of refraction for light of different polarization and direction of travel, it is prevalent naturally in materials with asymmetric crystal lattices but can also be induced through stress. Quantification of the effects of stress birefringence is done in terms of optical path difference per unit path length, most typically in nm/cm. The resulting difference in optical path length is given by [247]

$$OPD = K_s \sigma t \quad (21.10)$$

Where K_s is the stress optic coefficient, σ the applied stress and t the thickness in the direction of travel. For N-BK7 we have $K_s = 2.77 \times 10^{-12}$ /Pa, which results in $OPD = 5.4 \times 10^{-6}$ nm/cm which can safely be ignored.

21.3.3 Choice of Solution

In order to fully evaluate the impact of either choice, an MCA was performed as shown in table 21.2, the highest weight was assigned to the overall impact on optical performance and the least significance given to ease of dis-assembly. Ultimately, clamping came out ahead due to the simplicity it affords, making it a much easier method to handle throughout all stages of product life.

Criteria	Weight	Clamping		Bonding	
		Rating	Score	Rating	Score
Ease of assembly	4	5	20	3	12
Ease of dis-assembly	3	5	15	1	3
Optical impact	5	3	15	2	10
Mechanical impact	3	2	6	5	15
Total			56		40

Table 21.2: MCA used to evaluate the preferred mounting solution

21.4 Seeking a Clamping based Solution

Two main methods were considered in the design of a clamping based mounting arrangement; use of a cage system, or integrating the mounts directly into the units housing, both are evaluated in this next section.

21.4.1 Cage System

Employing an optical cage system throughout the fabrication unit allows for the use of an established product line of building blocks from which to construct the unit. Many of the optics used in the system, from simple lenses to complex PMTs are sourced from Thorlabs, Thorlabs provide a number of cage systems, and provide support for said cage systems integrated directly into their products. As such, many of the components of the system are readily compatible, and exploiting this compatibility allows for efficiency in design. The cage also allows for components to be directly mounted to the table top which helps reduce the vibrations the optic receives due to the natural damping provided.

21.4.2 Custom Mounting

The use of an integrated method allows for a more bespoke design potentially reducing the size of the final unit. However, integrating all of the mounts into the housing of the unit would require manufacture of complex, bespoke parts to an extremely high degree of accuracy. The design of housing with simple optical mounts would be possible, however, layering in complexity such as the ability to adjust positioning to account for manufacturing imperfections quickly complicates the design.

A bespoke solution would allow for more modularity in the design which helps with isolation of various sub systems, however in coupling together all of the optical mounts into a single discrete section of the housing the system loses resilience to damage. To exemplify, damage to any area of the housing containing the optics could result in the need to replace the section of housing entirely, disturbing all of the other optics, with a cage system, the modularity in the cage systems itself allows for small sections to be replaced fairly easily.

21.4.3 Choice of Solution

Evaluation of the choice was quantified and an MCA was performed as shown in table ??, the highest weight was assigned to the modularity provided by the design, this was due to the fact a more modular design would help break the complex system up into simpler constituent parts. The lowest weight was assigned to size, it was included as a parameter due to its relevance, however its impact scaled so due to overall system size not being one of the core specifications aimed for. The optical cage was preferred, ultimately making up for what it lacked in modularity, in simplicity of implementation, and efficient use of native support offered by many of the chosen optical components. The choice was made to adopt the 30mm cage system supplied by Thorlabs due to it offering the largest amount of compatibility, both with optics supplied by Thorlabs, and optics sourced from other vendors.

Criteria	Weight	Cage System		Integrated Design	
		Rating	Score	Rating	Score
Complexity	3	5	15	3	9
Modularity	5	2	10	5	25
Size	2	2	4	5	10
Efficiency	4	5	20	1	4
Total			49		48

Table 21.3: MCA used to evaluate the preferred mounting solution

21.4.4 Custom Cage Components

As previously stated, support for an optical cage throughout the system was plentiful in the chosen optics, plentiful but not universal. In such a case where no native support was available, or in some such situations where the offered solution was cumbersome, custom mounts were designed to allow for direct integration into the cage system. The parts were designed to be as simple as possible, all incorporating a retaining surface and rear support providing a preloading force to acquire the mount. Examples of the some of the mounts incorporated into the final design are shown in both figure 21.6 and figure 21.7.

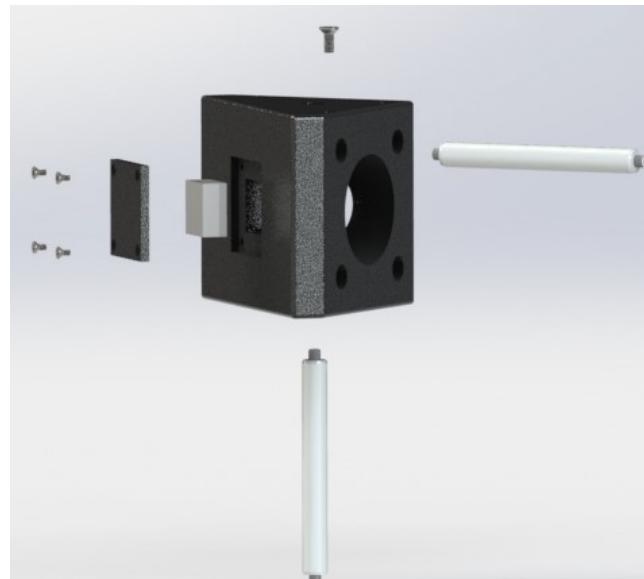


Figure 21.6: Corner mount for diffraction grating

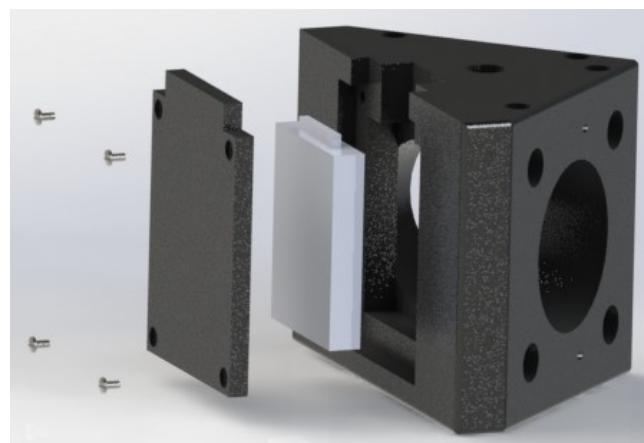


Figure 21.7: Corner mount for SLM

All of the custom mounts are made from black anodized aluminum. In the anodizing process, aluminum is coated with a layer of aluminum oxide forming a hard outer shell. Use of this material provides us with strong and light weight components, and the hard surface acts as an ideal retainer [250].

21.5 Final Design

Coverage the design of the housing will be done in a later section, as it requires treatment of the finite element method beforehand, but so as to round off this section, an overview and conclusion on the final mechanical design of the optical part of the system, sans housing, shall be provided.

21.5.1 Final Cage System

The cage system was required to be extensive in order to accommodate every aspect of the optical design, and thus many components were used throughout its construction. The constituent elements are illustrated in table 21.4

Component	Quantity	Individual Cost	Total Cost
Cage Mounts			
CP11/M	3	£11.17	33.51
CP02T/M	4	£11.76	47.04
CP06/M	12	£12.31	147.72
KCB1C/M	35	£102.86	£3600.10
C6W	7	£44.98	£314.86
B3C/M	7	£17.57	£122.99
B1C/M	7	£13.23	£92.61
B5C1/M	7	£22.05	£154.35
GCM012/M	2	£132.30	£264.60
GHS003/M	2	£14.70	£29.40
Cage Rods			
ER025	16	£3.63	£58.08
ER05	16	£3.63	£58.08
ER1	24	£3.63	£87.12
ER1.5	12	£4.19	£50.28
ER2	8	£4.37	£34.96
ER4	48	£5.10	£244.80
ER8	10	£8.41	£84.10
ER10	8	£9.14	£73.12
ER12	40	£12.08	£483.20
ER24	24	£21.27	£510.48
Total			£6491.40

Table 21.4: Table of cage components

21.5.2 SolidWorks Model

A model was built from scratch in SolidWorks, the ability to visualize the final design is extremely important but perhaps more important still was the role the model played in the design process. For such a complex problem with such an open ended solution, being able to interact with a fully three dimensional realization of the fabrication throughout all stages of design is an extremely powerful tool for both iteration and validation of design decisions. It allows for evaluation of various geometries quickly where hand calculations would be cumbersome, and provides instant verification on the viability of certain choices. It is hard to cover the intricacies of the design process, as many decisions have several equally valid solutions, and many decisions can end up carrying very little weight. Instead, the the final system shall be presented pictorially, and then an attempt to encapsulate all of the salient design information and ramifications in a final conclusion made. The final fabrication unit is shown from two alternate angles in figures 21.8 and 21.9. The housing is absent as previously stated, as too are the supporting cage rods, the cage rods were suppressed in order to aid clarity of the images. Of final note is the use of 'dummy' parts where CAD files were not supplied by the manufacturer, all such parts were recreated with accurate dimensions but left as generic white boxes. In one case, that of the objective turret, neither CAD file nor dimensions were made available and thus, in lieu of sheer guess work, the turret was omitted from the model entirely.

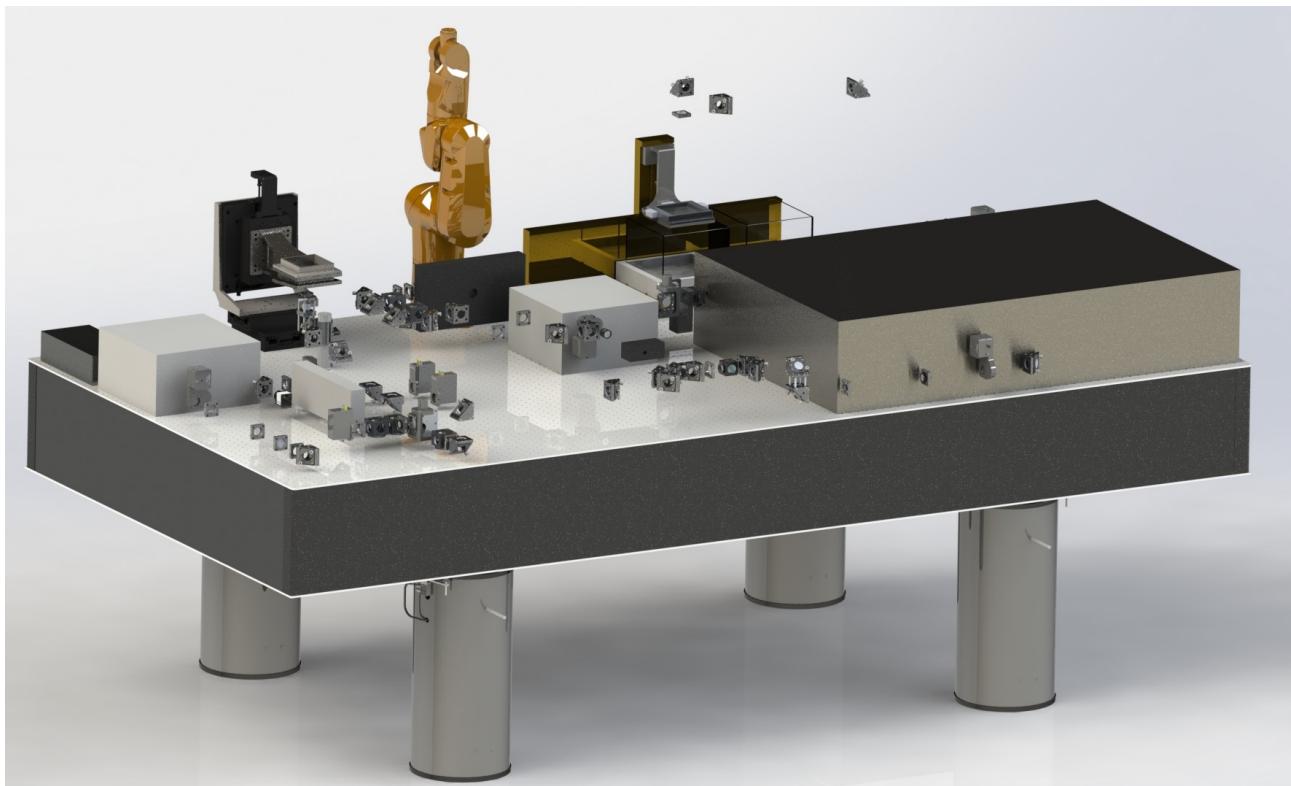


Figure 21.8: Mechanical Layout

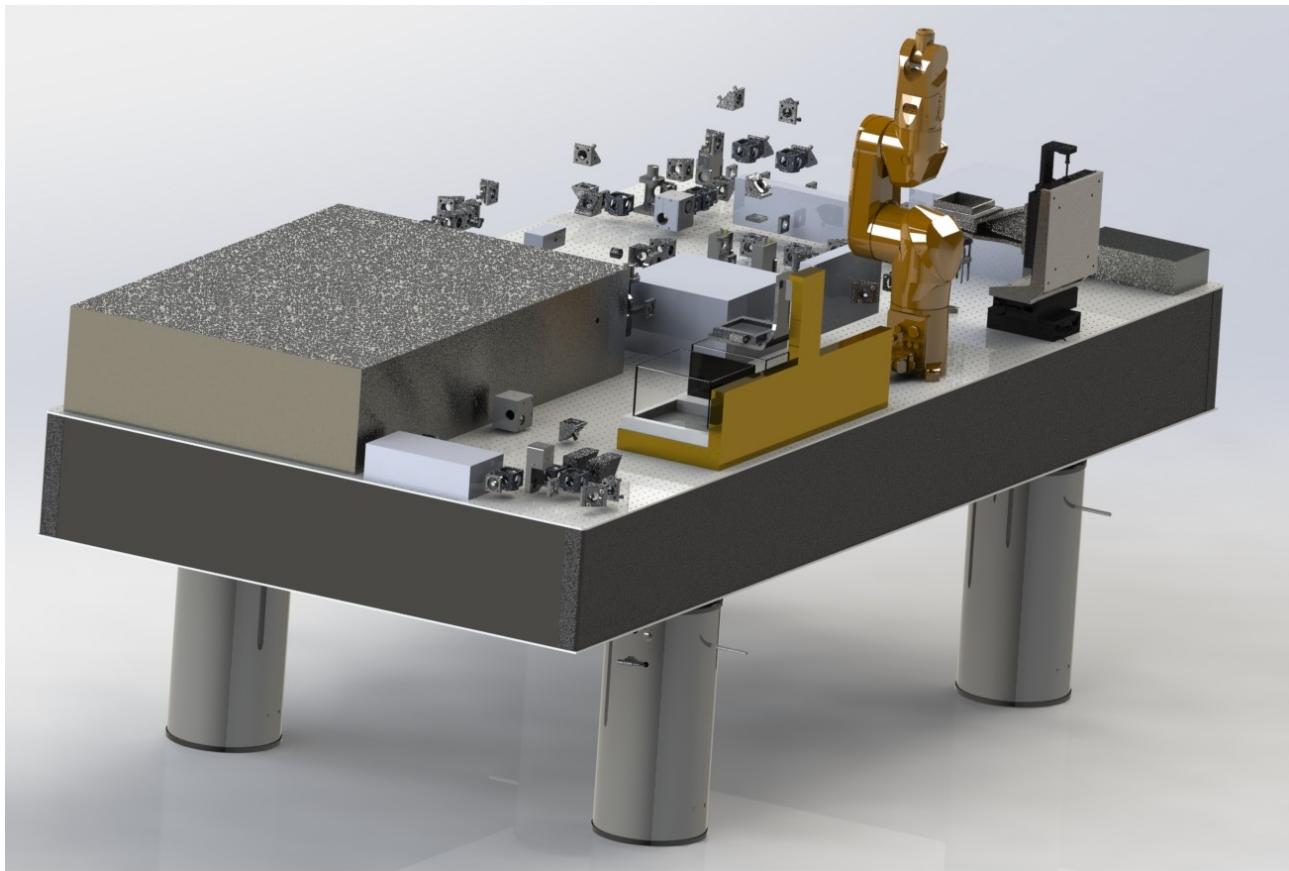


Figure 21.9: Mechanical Layout - alternative view

21.5.3 Design in Review

The final design provides a successful realization of the underlying optical layout. We provide secure, adjustable mounts for all of the optical elements, and allow for relatively easy (if not slightly time consuming) construction of the entire fabrication unit. The unit is built such that individual parts are readily accessible, and replacement requires only small disturbance to the larger system. The optics are mounted directly to the tabletop to provide as much shielding for vibrations as we could reasonably achieve and we facilitate the complete automation of our fabrication processes.

Moving on from just merely praising the design, dedication of the following paragraph is an evaluation from a more critical angle. The nature of design, is iterative, this was just the case in the design of the optical layout of the system, large changes were made throughout the project, processes dropped and then brought back upon further inspection. This was all in aid of reaching the final, rather impressive, system. The nature of the mechanical design complimented this process. The cage system provided a large amount of flexibility and freedom with which to place the optics, and the table provided a large canvas upon which to do this. A design was very rarely 'locked in', and upon simple changes to the layout, such as the removal of lenses to relocation of a shutter, could all easily accommodated for. However, therein lies the problem. Allowing the cage layout adapt to

changes over time, the resulting design is rather dispersed over the table, and separate subsystems become somewhat rigidly coupled together. The control hardware is located in a discrete module, as will be discussed later, but from what is on show in the two images of the mechanical layout above, one could reasonably identify three distinct subsystems; the optics, the lasers, and the staging areas, from a layout point of view however, one might not. In the final design the three are interleaved with one another, meaning changes in one subsystem can lead to issues arising in another. If the three were entirely decoupled, and some simple optics used to bridge them together, the final design would be more accommodating to large changes in one of the subsystems. In striving to design the system to allow for 'micro' changes, the ability to cope with 'macro' ones becomes hindered. Design of a single housing 'shelf' with which to mount all of the optics would have allowed for much better modularity, however designing in such a way would have meant small, and perhaps worthy to note, more frequent changes, would have been much harder to accommodate. In a battle between two evils, a small solace can be taken from the lesser coming out the victor, although this is ultimately down ones own discretion, the justifications given throughout the design section would provide evidence that this was indeed so.

21.6 The Finite Element Method

21.6.1 Introduction

The finite element method provides a robust framework with which to solve complex boundary value problems to partial differential equations [251]. Its power lies in the reuse-ability and generality of the procedure, and direct software integration in many CAD packages makes it an indispensable tool for the design process. It can allow for quick evaluations of designs, as well as much in depth analysis from complex simulations. The method was utilized in the design of the system and as such, the next section is devoted to covering the underlying theory.

21.6.2 Theory

Prior to the discovery of the finite element method, one such methodology for finding approximate solutions to complex partial differential equations was the Rayleigh-Ritz method [252], in which the form of solution is assumed, or perhaps more appropriately, reasoned with some intuition of the underlying physics, and then the solution found under said assumptions. This provides an adequate formulation for simple problems, and in such a case if the real solution exists in the subset of solutions encapsulated by the assumed one, it can be shown that this method does in fact provide the correct solution. However, in seeking high order solutions, the method becomes quickly unstable and as such quickly detracts from its usefulness. There is also no framework in place with which to handle arbitrarily complex geometries. All is not lost though, a clever re-purposing, as demonstrated later,

underpins part of the finite element method.

The method involves initially discretizing a given continuum, into a set of finite elements. Many potential elements exist, for one, two and three dimensional problems, some providing a more general purpose role, and others allowing for better representation of complex geometries. Figure 21.10 shows some such elements.

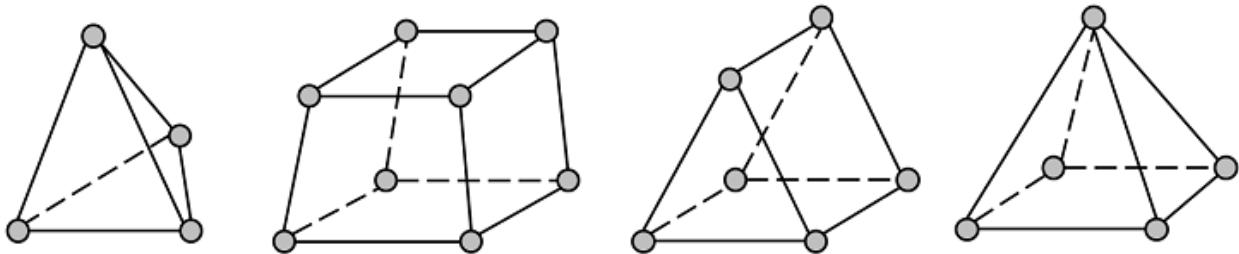


Figure 21.10: 3D finite elements [253]

The meshing process will be covered shortly, but for now, analysis is continued under the assumption of a readily computed mesh. In order to allow for handling of elements stretched and contorted so as to represent a given geometry, a mapping is performed to move analysis from the physical domain of the problem, into a natural co-ordinate system for a given element, for example, into natural tetrahedral coordinates. Such a mapping is depicted in figure 21.11.

The labeled nodes will too be discussed shortly, but for now the aim of the diagram is to depict the transformation of the unsuitable topology of the elements acquired in the real domain, to the well ordered ones offered in the natural coordinate systems. The method through which this is done is the use of shape functions, in order to understand their action, first the a fore mentioned nodes must be explained.

These nodes represent the points at which the solution will ultimately be calculated, and the results at these nodes used to interpolate the solution at any arbitrary location on the geometry. The functions used to map the real geometry into the ideal one can be shown to also be the functions required to perform the interpolation of nodal values, as such, the restrictions on the shape functions can be asserted. At each node there must be a function that is valued exactly one, and all others must be zero. This restriction links the number of nodes that can be used to the order of the polynomials used in the shape functions. In the natural coordinate system for a three dimensional element, the axis are expressed along three of the edges of element and the vertexes chosen to be at the most convenient locations, for example, the tetrahedral element is expressed in terms of ζ , η and ξ , with the vertexes located at $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. The simplest of elements use each vertex

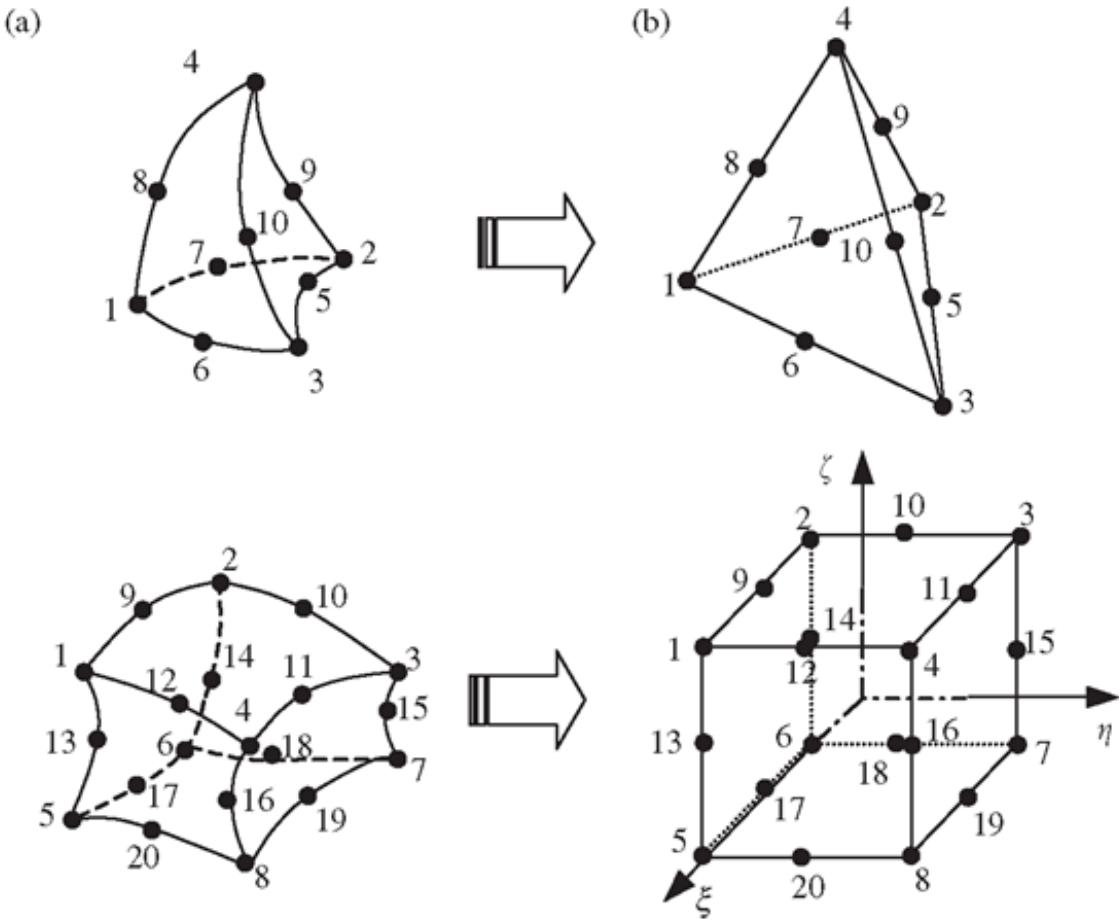


Figure 21.11: Mapping elements in the physical domain, into a natural coordinate system [254]

as a node and expresses all of the shape functions as linear functions of the coordinate system. To continue the example of the tetrahedral element, four such shape functions are:

$$N(\zeta, \eta, \xi) = 1 - \zeta - \eta - \xi \quad (21.11)$$

$$N(\zeta, \eta, \xi) = \zeta \quad (21.12)$$

$$N(\zeta, \eta, \xi) = \eta \quad (21.13)$$

$$N(\zeta, \eta, \xi) = \xi \quad (21.14)$$

$$(21.15)$$

Note that using linear functions restricts the element to nodes only located at the vertexes due to the requirement of the functions evaluating to precisely zero at all other nodes, to alleviate this problem higher order polynomials can be used, allowing for several nodes along the edges. In the case demonstrated above, the differential relation between strain and the coordinate system means

the linear functions yield constant strains throughout the element. The resulting form of the strain as a consequence of a given element can be used in conjunction with prior knowledge as to the form of the desired stress distribution to aid in the choice of most appropriate element for a given problem.

It is in the shape functions that the resurgence of the Rayleigh-Ritz method can be seen. Instead of making the approximation across the entire geometry, it is instead made on smaller scale, across each of the individual elements. This piece-wise formulation allows for complex stress distributions to be represented by assembling vast amounts of simplex ones, thus the risk of instability of Rayleigh-Ritz approximations is averted. Truly a cake both had, and eaten.

After meshing and formulation the task at hand is to now solve the given set of equations. For the case of solving for a stress distribution, the problem comes down to energy minimization and virtual work. Energy minimization is a concept used in many fields of engineering and is a process by which a solution is found by formulating the energy of a system and calculating the state the system must be in to minimize its associated energy. Virtual work states that for a body in equilibrium the work done by external forces moving through a set of displacements is equal to the integral over the body of the product of stresses and strains compatible with the given displacements [251].

$$\Sigma \text{External Forces} \times \text{Displacements} = \int_v \text{Stresses} \times \text{Strains} \quad (21.16)$$

The ultimate goal here is to arrive at the simple, and easy to handle stiffness formulation:

$$\mathbf{F} = \mathbf{KU} \quad (21.17)$$

Having formulated the stiffness matrix for the given problem and supplied the set of forces, manipulating somewhat to handle reaction forces and known displacements, a simple matrix inversion can be performed to acquire the set of displacements and the set of displacements to acquire the nodal stresses and strains. A plethora of exotic methods exist for inverting matrixies, and for the sparse and typically banded matrixies common to these problems, efficient inversion is readily available.

21.6.3 Meshing

Throughout the above analysis, the ability to mesh a given geometry was assumed, in this section, treatment is given to cover methods employed to acquire such a meshing. The complexity of the algorithms employed is great, and literacy documenting such algorithms scarce. Thus the section will attempt only to present a qualitative outline of the most popular meshing algorithm.

The grid-based approach to meshing has become popular due to the methods ability to produce extremely uniform internal elements. In this method, an infinite template grid is produced, and

overlaid onto the given topology. Points on the grid lying outside of the boundary of the object are removed, and a resultant zigzag border is left [255]. The points making up the zigzag are then relocated onto the boundary of the object to provide a finished mesh. A slight alteration of this method first creates a zigzag lying entirely inside of a the object, and then a separate triangulation algorithm used to complete the interior mesh.

A question that is raised here, and in many other meshing algorithms is that of how to convert a set of nodes into a mesh of triangles, i.e. what is the best triangulation for a set of nodes? Generally considered to be the best solution is Delaunay triangulation. Here, the mesh is chosen so as to maximize the sum of the smallest angles of the triangles, the aim being to minimize the use of thin elements whenever possible [255].

21.6.4 Refinement

There are two types methods currently used to acquire refinements to the finite element method [256]

1. **p - method**

2. **h - method**

In the h-method, the characteristic length of the element, h , is reduced and in the p-method the order of polynomial used in the shape function, hence the p , is increased. Figure 21.12 demonstrates the two effects.

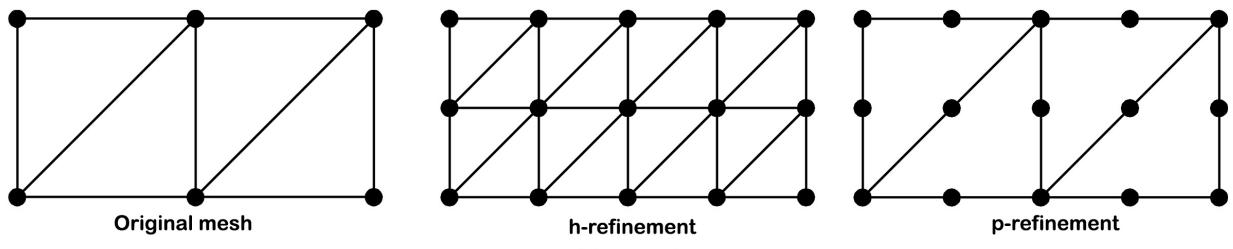


Figure 1.

Figure 21.12: h and p methods of refinement [256]

The h-method refines the mesh used during the calculations, finer meshes allowing for better representation of a given geometry along with more flexibility with which to model changes in the stress field. The p-method gives each individual element more freedom with which represent the stress field, for example, increasing from a second to a third order polynomial allows for the stresses throughout an element to vary in a quadratic manner as apposed to a linear one. Both methods of refinement come at the cost of extra computational power, so intelligent choices of refinement is key. As mentioned previously, a priori knowledge as to the form of the distribution can help make this choice. If the expected stress field were to vary quadratically, and elements in use that only

provide linear element wise stresses, a vast improvement can be made through p-refinement to better represent the given problem. In such a situation, employing h-refinement to acquire a finer mesh would prove imprudent. Conversely, if a low order polynomial can sufficiently represent a given distribution, h-refinement would be more apt.

Both methods of refinement provide a useful tool, that is, the ability to check a solution for convergence. It is important to remember that the finite element method provides only an approximation, and as such it is crucial to verify the validity of a given approximation if meaningful design can be achieved as a result. If in applying a refinement, the solution drastically changes it is apparent that previous approximation was too crude, the method of refining and checking a newly computed solution against a previously computed one can allow for the identification of convergence.

21.6.5 Examples in Design

Instead of covering each and every example of the use of the finite element throughout the design, instead this section shall conclude with perhaps the most interesting of them all, and further examples shall be given in a latter section of the report, covering the design of one of the custom staging areas.

In designing a product, all aspects of its life must be accounted for, from the details of its construction to its ultimate decommission. Even those stages that are only transient in nature are important to be given proper due diligence. For the vast majority of the life cycle of a fabrication unit, the unit is stationary. From a structural point of view, material can thus be saved throughout the design of the units housing as it is required to experience little shock. The period in which we would expect the largest amount of disturbance is that during which the fabrication unit is being delivered. As previously mentioned, although only short relative to the life of the product, it is vital no damage is caused throughout the duration. Although the nature of such a period is random, the finite element method provides us with a means to tackle such a stochastic process. Integrated directly into the SolidWorks finite element suite, is the ability to perform random vibrational analysis. Several forms of input are catered for, and in this case a power spectral density curve was used. Power spectral density is a measure of power as a function of frequency, it provides a distribution of the power throughout the spectrum of frequencies of a given signal. The power spectral density curve used is shown in figure 21.13.

In order to efficiently evaluate the stresses experienced the analysis was performed only on the more massive parts of the structure, and in such cases complex bodies were replaced with mass stand-ins. The assumption here being that the complex parts being replaced would have been validated during their specific design process, and any parts brought from a manufacturer validated by said manufacturer. The design of the housing shall be covered in the proceeding section, but so as

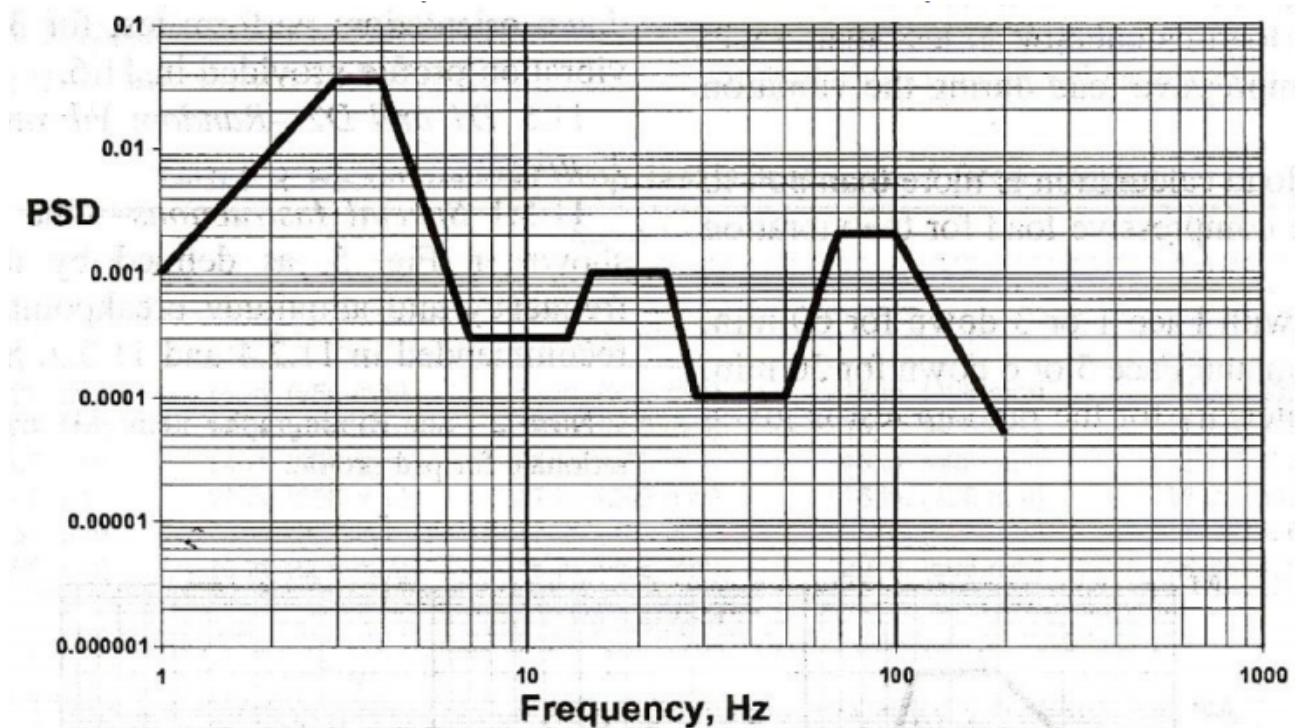


Figure 21.13: Power spectral density of a delivery truck [247]

round out the treatment of the finite element method a portion of its design is being exemplified here. The result of meshing the section of the housing to be analyzed are shown in figure 21.14.

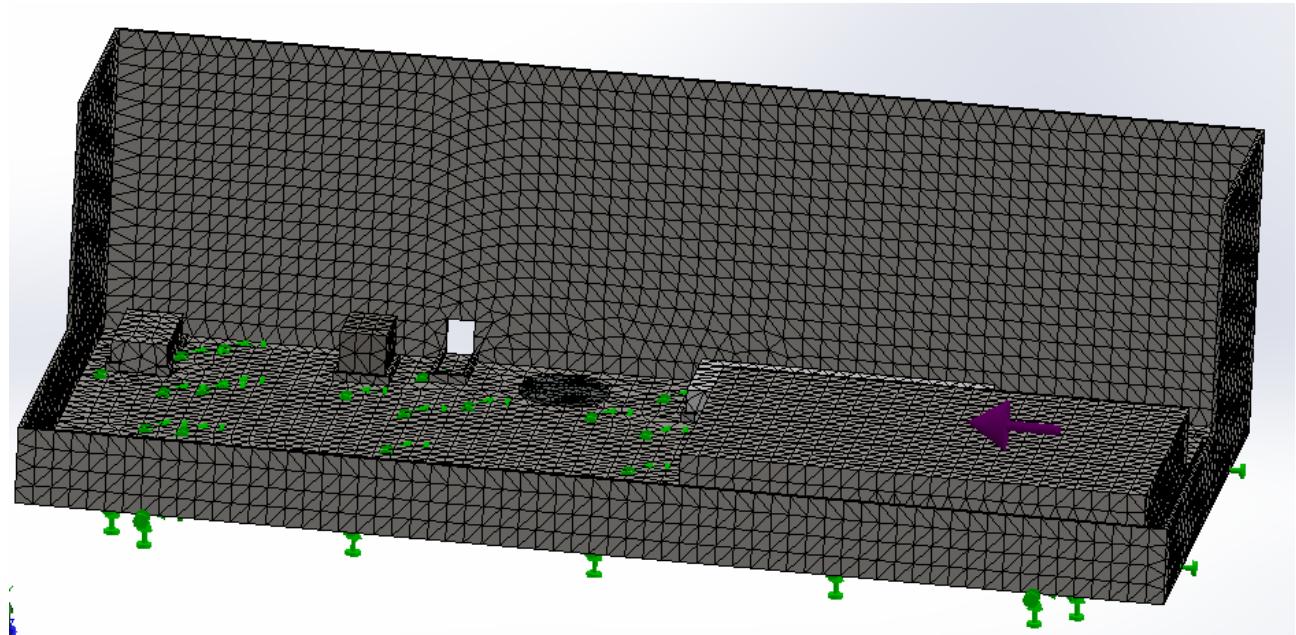


Figure 21.14: Resulting finite element meshing

And after running the simulation, the stress distribution given in figure 21.15 was acquired

The results show areas of concentrated stress, but all well within a margin of safety, and thus confidence can be gained in the structural integrity of the overall structure of the housing, and indeed the mounts used to secure the largest of components.

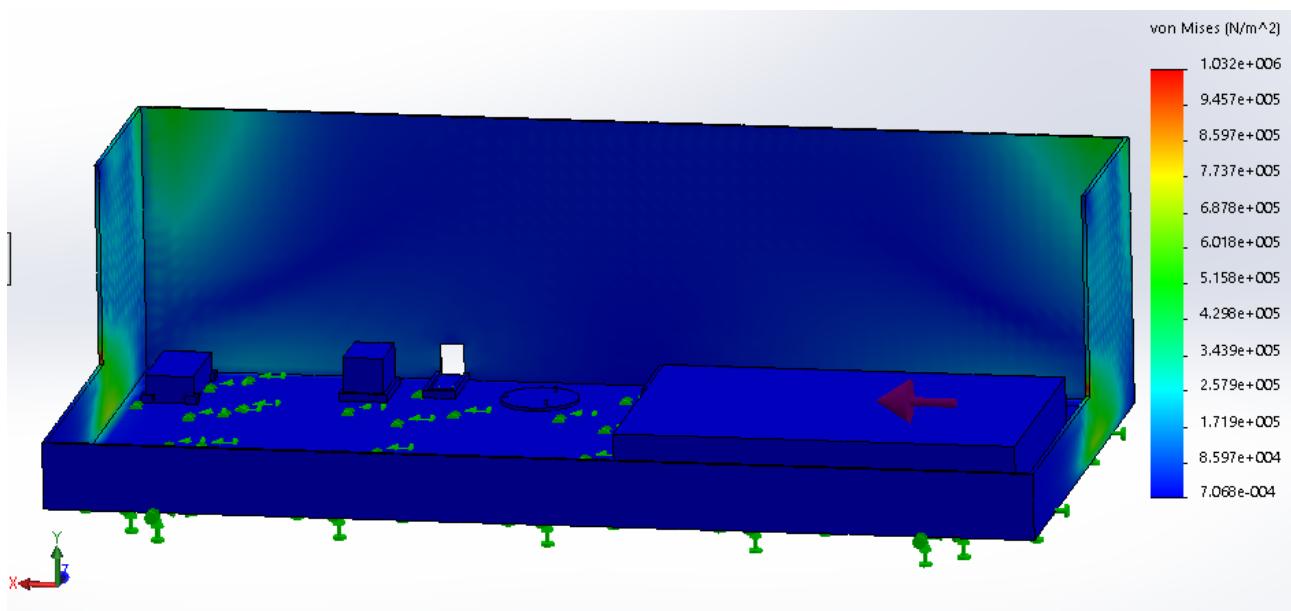


Figure 21.15: Resulting finite element distribution

21.7 Housing Design

In the design of the housing, some core tenants were laid out. The integration of an optical cage running throughout the system was, for one of several reasons, to facilitate ease of access and replacement of certain optics. As such, the housing was required to allow complete access to the interior optics. Furthermore, the as previously discussed, the mechanical layout of the system resulted in the coupling together of several subsystems, so as to remove the coupling of yet another subsystem, the control deck was decided to be an entirely separate unit.

For the main section of the housing, that is, the section containing the optics, stages, lasers and robotic arms was designed with a simple, 'fold-out' approach. This is demonstrated in figures 21.16 and 21.17. As can be seen, the main body opens so as to provide unrestricted access to the interior, and electronically controlled locking mechanism is in place so as to ensure access is only permitted at a time when it is safe to do so. The control deck is not shown in the images, but simply slots on top of the lower half of the system, it to is completely removable. The housing is made from stainless steel so as to provide a sturdy, tarnish resistant finish. Holes in the base aligned with the mounting holes in the table allow for the optics to be secured, and the larger structures are provided with inbuilt, simple corner restraining mounts. The verification of the performance of the housing was done as shown in the earlier section on the finite element method.

To handle temperature control throughout the system, thermostatically controlled cooling fans were integrated. Finite element analysis readily provides a method with which to simulate and analyze heat distribution throughout a structure. Although wherever possible, CAD files supplied by

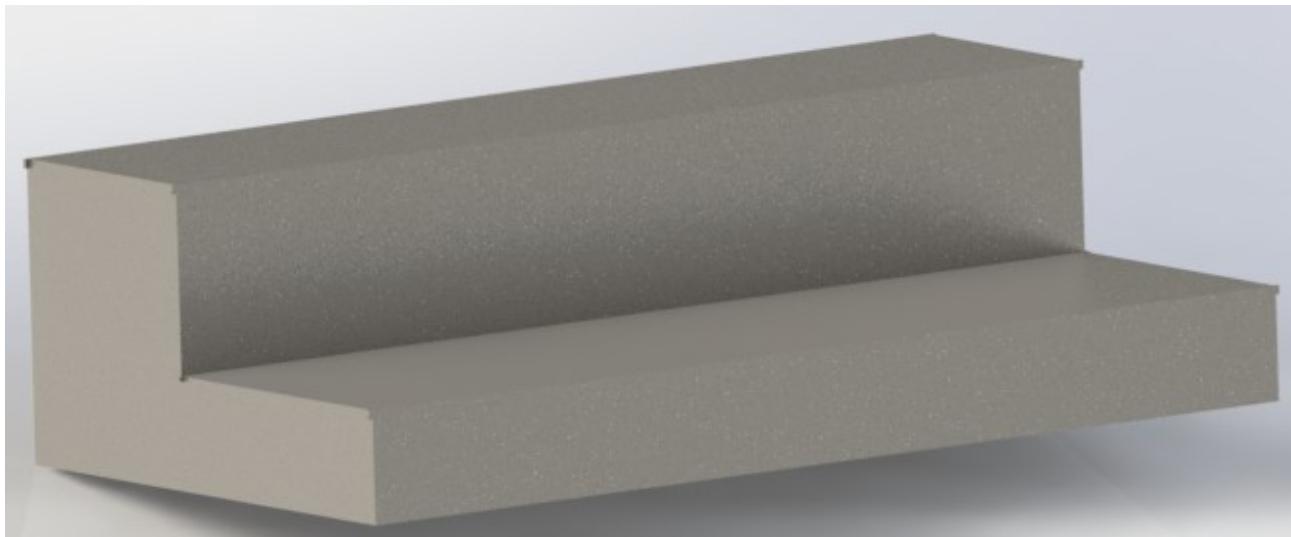


Figure 21.16: Housing in closed form

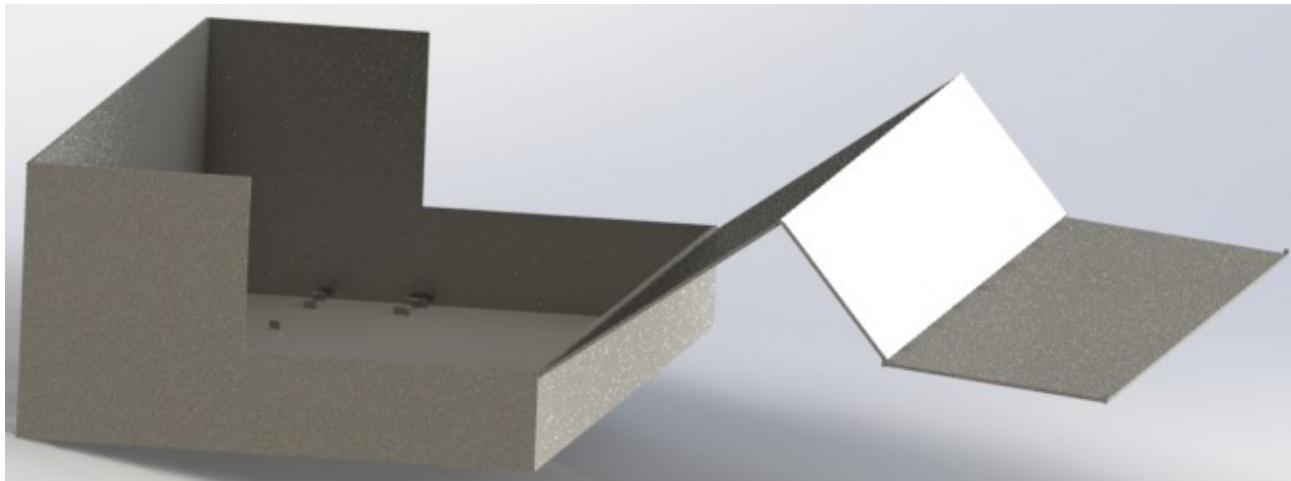


Figure 21.17: Housing in open form

manufacturers of the components used throughout the system were implemented in the model, many of the more complex models were not provided. As such, stand-ins were made with as accurate as possible replications of the specified dimensions. The materials used were often not specified by the manufacturer, and in many cases the complex inner workings, such as the insides of the lasers were not modeled. All such components were also those which produce the heat throughout the system, thus making any attempt at full temperature simulation extremely crude. The crudeness of the approximation in combination with the inexpensive nature of cooling fans meant a conservative approach was taken and bank of cooling fans added. A row of eight **5915PC-20T-B30-AM0** ac axial cooling fans provided by Minebea were used, the fans combined with thermostats and integrated controllers allow for complete temperature control of the internal environment.

The different fabrication techniques offered by the system require an assortment of different lenses to be used for their own respective manufacture. In order to accommodate for this with-

out necessitating human interaction, a mechanism is needed to automate the switching of lenses. Several such devices were considered, and ultimately it was chosen to integrate a motorized linear objective changer with controller as supplied by Edmund Optics. The model selected supports up to six objective lenses and costs £4780.

22 Automated Stage and Galvanometers

The ability to image, and fabricate in the specification volume of $8cm^3$, with incremental motion of $< 100nm$, accurately and reliably is crucial. The specification volume of $8cm^3$ was chosen, as this is far greater to existing products in the market, that can combine stereolithography fabrication, with the high resolution TPP fabrication, this is one the key unique selling points, which ensures the entire product is highly competitive in its market. Confocal imaging, TPP and laser micro machining require high scanning speeds ($10mm/s$) across the sample. To achieve high scanning speeds, over a large sample area of $8cm \times 8cm$, prior literature [1], [257] reveals incorporating both dual-axis scanning galvo scanning mirrors (X, Y) (low inertia, thus high speed), with a movable stage (much higher inertia, thus low speed) (X, Y, Z). This section outlines the design of the galvo scanning mirror, and custom stage design.

22.1 Galvo scanning mirrors

Prior literature [258], shows TPP, confocal and laser micro machining using Galvo scanning mirrors, for a beam deflection scanning system. Although there are other options such as piezoelectric and magnetostrictive actuators. Below is a table contrasting the laser beam scanning systems.

Criteria	Galvo mirror	Piezoelectric mirror	Magnetostrictive mirror
Manufactures	Many	Many	Few
Max scan angle	$\pm 20^\circ$	$\pm 1^\circ$	$\pm 0.1^\circ$
Resolution	$8 \times 10^{-4}^\circ$	$6 \times 10^{-6}^\circ$	$2 \times 10^{-4}^\circ$
Max Oscillating Frequency	$1kHz$	$2kHz$	$1kHz$
Price (Including controllers)	£1,786	£1,888	£2,000

Table 22.1: Comparison of laser beam scanning systems

Evaluating the laser beam scanning system, the fabrication and imaging systems require a high speed, high resolution scanning system, which can cover the entire field of view. Magnetostrictive mirror actuators seem too experimental, to incorporate in a reliable design, whereas piezoelectric show high resolution, and speed, however they only operate over a too smaller scan angle. Therefore galvo mirrors were chosen, due to their large scan angle to cover the entire field of view, coupled with their more than sufficient resolution and speed, additionally they are a mature technology, that are supplied from many manufacturers.

Galvo (galvanometer) mirror scanner, is a sensitive limited rotation DC motor. When a small current flows in the rotor coil, in the presence of the stator magnetic field, this leads to a offset force (i.e. a torque) applied to the rotor (arising from the Lorentz force). The torque acts to rotate the shaft,

against the restoring spring force. The shaft is coupled to the mirror, which allows rotation of the mirror. Galvo mirrors are current controlled, and commonly controlled with closed loop feedback.

ThorLabs Large Beam Diameter Dual-axis scanning galvo system **GVS012/M** (£2,440 including power supply), is selected for the design. Some of its specifications are listed in table 22.1 (as the galvo scanning mirror), are more than sufficient for the fabrication and imaging processes. Additionally it was chosen for its large galvo mirrors to support a large beam diameter of $10mm$, which is important to ensure the incoming beam underfills the mirror, and to prevent damage to the mirrors (As using a large beam width for the same energy, reduces the light intensity contacting the mirror). Protected silver mirrors were chosen, as they are optimized to reflect a wavelength range of $500nm - 2.0\mu m$, which encompasses all wavelengths from the processes that pass through the galvo mirrors.

22.1.1 Control

The dual-axis Galvo Scanning mirror system chosen comes with a analog control electronics board for each galvo, that uses a calibrated PID (proportional integral derivative) controller to control the current through each galvo, using a capacitive encoder to provide rotation feedback. Additionally a Thorlabs Galvo System Linear Power Supply would be implemented, to source the driver boards with $\pm 15VDC, 3A$. The mirror rotation is changed, by applying an input voltage (1V per 1°) to the driver boards.

The angular position, of the mirror galvos, requires high speed synchronization with the high speed shutter, and stage to operate the fabrication processes correctly. This is achieved with the NI-DAQ LabView controller, **NI USB-6353**, connected to the system PC. In Labview the galvo mirrors, can be controlled as a resonant, raster scanning method, or can be dynamically controlled using the open source library *ReplicatorG*, that takes Gcode as input [259]. The system would be calibrated using error mapping, to account for optical aberrations that cause unwanted beam deflection from the manufacturing tolerances of the lenses in the beam path, and from aberrations discussed in 7.2.1.4.

22.1.2 Scanning Area

A key operating value is the scanning area in the sample, as this controls the maximum scanning angle, max voltage, scanning frequency, scanning speed, lateral resolution and the input digital samples per second needed. Ideally the scanning area should be maximized, as it has a significantly faster scanning speed compared to the high inertia stage. Constraining factors on the scanning area are, galvo maximum mirror angle rotation ($\pm 20^\circ$), objective lens, optical aberrations and the maximum angle that can be passed through the two 4F lenses (**L12, L13**) that still fills the back aperture of the objective. The galvo design presented earlier in section 7.2.2, was designed for large angles,

whereas the imaging and TPP fabrication design utilizes small angles, which allow focusing of a scanned beam onto a regular microscope objective for small scan areas.

A optical model of the sample scanning, illustrating one axis of the galvo scanning mirrors, seen in figure 22.1. This model assumes, the objective lens can be modeled as a thin lens, where the incident beam angle on the back aperture is the same on the output of the objective. The thin lens, is an oversimplification for the real objective lens, however this is justified for illustrating design choices, and theoretical capabilities of the scanning area. The model is only valid for small angles ($2\theta \leq 10^\circ$ [260]), and assumes that the beam correctly fills the back aperture of the objective, from passing through the 4F lenses, which only holds true for $2\theta \leq 5.2^\circ$ calculated from the 4F lens setup.

To steer the beam to x (mm) position off center, the mirror must rotate by $\theta = \frac{1}{2}\tan^{-1}\left(\frac{x}{f_o}\right)$. Hence a rotation angle of the mirror corresponds to a beam sample displacement of $x = f_o\tan(2\theta)$ (mm), to maximize this, as 2θ is maximum at 5.2° , the objective focal length f_o can be increased.

Presently the objective chosen for TPP, Laser micro machining and imaging is the objective lens **MPLAPON 100XO** ($f_o = 1.8\text{mm}$). This corresponds to a theoretical maximum scanning sample area of $327\mu\text{m} \times 327\mu\text{m}$, however in practice this is constrained by the sample Field of View. The FOV diameter $d = \frac{\text{fieldnumber}}{\text{magnification}} = \frac{26.5\text{mm}}{100} = 265\mu\text{m}$, which corresponds to an internal scanning area square, of length $L = \frac{265\mu\text{m}}{\sqrt{2}} = 187\mu\text{m}$. To increase the scanning area, another objective **UPLSAPO 20XO** was included (section 17.10.6), to complement the main objective, with a $f_o = 9\text{mm}$, which makes the scanning area, an impressive $937\mu\text{m} \times 937\mu\text{m}$ (which is again constrained by the FOV).

The maximum angle requested by the controller would be $\pm\theta = \frac{5.2^\circ}{2} = 2.6^\circ$, which corresponds to an input voltage of $\pm 2.6\text{V}$, which is within the input voltage range for the Ni-DAQ driver (± 10). The maximum frequency of oscillation is calculated to be 895Hz , for the maximum deflection of $\pm\theta = 2.6^\circ$, using data from the galvo data sheet. The individual laser points that can be reliably scanned, for maximum θ deflection are, $\text{LaserPoints} = \frac{2\theta}{\theta_{\text{resolution}}} = \frac{5.2^\circ}{0.0008^\circ} = 6500$. The frequency to scan a whole image of the 2D scanning area is $= \frac{1\text{DScanFrequency}}{\text{LaserPoints}} = \frac{895}{6500} = 0.138\text{Hz}$. This is a reasonable scanning image frequency for the imaging techniques such as confocal microscopy, TPM and STED. Below is a table showing the calculated specifications for each objective.

This design allows fabrication and imaging in two modes, a low resolution 144nm , large scanning area $937\mu\text{m}^2$ mode, and a high resolution 28.8nm , small scanning area $187\mu\text{m}^2$ mode. This mode

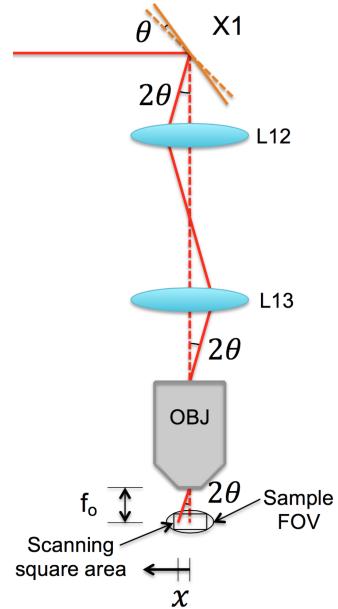


Figure 22.1: Galvo model

Criteria	UPLS APO 100XO	XSLPLN25XGMP
X, Y Scanning area	$187\mu m \times 187\mu m$	$937\mu m \times 937\mu m$
X Max scanning speed (mm/s)	167	838
X Lateral resolution (nm)	28.8	144

Table 22.2: Galvo mirror system specifications, using both objectives

change, is enabled by switching the objective using the motorized turret system. Often depending on the process, either a large scanning area, or a high resolution is preferred. These scanning areas are small compared to the large scanning areas of the stereo lithography fabrication process, $\approx 80mm^2$, however the design aim for the main stage galvo scanning system was to design for high sub micron resolution.

The NI Data Acquisition controller board, has 16 bit outputs, at a voltage range of $\pm 5V$, which output a maximum of 1.25×10^6 samples per second, the necessary input digital samples per second needed per galvo is $\frac{2V_{MaxInput}}{DACResolution} = \frac{2 \times 2.6}{10/2^{16}} = 3.41 \times 10^4$ samples per second, which is well within the DAC output sampling limit.

22.2 Automated Stage

The stage physically moves the sample around in three dimensions, it is specified to move the sample in a volume of $8cm^3$, with a incremental motion $< 100nm$. It must also hold the sample accurately level, and mitigate the effects of vibration, mechanical loading and thermal expansion. The stage requires the use of transmissive illumination for the imaging methods, and for live imaging of the TPP fabrication, therefore the stage needs to allow a light source to illuminate the sample from underneath.

To achieve a reliable fabrication volume of $8cm^3$, a safety factor of 1.25 is included, to mitigate the effects of the objective contacting the cassette resin containing walls, and provides room for the sample to be lowered from the objective to be moved over to the automatic cassette transfer system. Thus a translation volume of $8cm^3 \times 1.25 = 10cm^3$ will be designed for. There are no stage systems that provide the requirements listed, therefore a custom stage was designed using discrete linear actuators, discrete parts and custom components.

22.2.1 Linear Stages

A linear stage restricts movement to a single axis, comprising of a movable platform and a fixed base. The platform and the base are connected via a guide that constrains the motion to the desired axis. The linear stage is moved by a linear drive system, such as a linear motor, ball screw or a piezo drive [1], [261].

Table 22.3 was constructed below comparing the available guide types. [262] [263] [264] [265].

Criteria	Air Bearings	Linear-motion Bearings	Crossed Roller Bearings	Recirculating Ball Bearings
Travel Range	Unlimited	< 250mm	< 250mm	Unlimited
Stiffness (N/m)	$10M$	$17.9M$	$30M$	$30M$
Load Capacity (kg)	170	100	300	200
Motion sensitivity ($gram$)	0.45	2.88	1.50	2.20

Table 22.3: Comparison of guide types

In a nanopositioning linear stage, the stiffness and motion sensitivity are most important, to reduce deflection. Analyzing table 22.3, it is clear that either air bearings or crossed roller bearings guide types are suitable, with crossed roller bearings preferred for their high load capability.

A table was constructed below comparing the available linear drive systems

Criteria	Ironless Linear Motor	Ball Screw with Stepper Motor	Piezo Drive
Minimum Incremental Motion	$1nm$ [266]	$50nm$ [267]	$0.5nm$ [268]
Travel	Unlimited	Unlimited	$100\mu m$
Backlash/Hysteresis	None	Low	Medium
Speed Stability	High	Medium	Medium

Table 22.4: Comparison linear drive systems

For the nanopositioning linear stages, the minimum incremental motion and the travel are crucial for the design. Research from table 22.4, shows although piezo drives have the best minimum incremental motion (Reliable motion), they suffer from small travel ranges, and hysteresis (due to the piezoelectric effect [269]). The best drive method, for the problem is to use an Ironless Linear motor, which can span the required translation length of $10cm$, whilst having adequate minimum incremental motion, with negligible backlash and good speed stability when translating linearly.

Analyzing all crossed roller bearing, Ironless Linear motor, $10cm$ travel linear stages, the Newport **XMS100**, with the optional high load crossed roller bearings was chosen, for the X, Y stages. It has a travel of $10cm$, with a minimum incremental motion of $1nm$, however can only support a normal load of $300N$. PID Feedback is provided with a optical linear encoder, which also incorporates absolute home position and limit signals for the linear stage. To achieve X, Y directions, the stages are stacked on top of each other connected with M6 screws, with a granite base plate **GB100** (for good mounting surface flatness), as recommended in their manual [270].

For the vertical Z stage, Newport's **XMS100V** was chosen, which is a **XMS100** mounted vertically

on a plate, with a low friction air cylinder to counteract the weight force of the moving platform and any load. Chosen for its 10cm travel range, with a minimum incremental motion of 50nm and can support loads up to 100N along its axis. This is connected to the lower stages using a Newport L bracket fitting. The two air cylinder valves are controlled from the supporting control electronics, and the high pressure air source, is from the Thorlabs **PTA513** air compressor, which is also used for the optical table.

To maximize the linear stages operating lifetime, they should be operated with minimized loads, with the loads acting close to the central axis of the stage, as to avoid off center moments, which leads to unequal loading of the precise bearings. This means the center of mass of any assemblies joined to the stage the lower X, Y stages should align with their central axis.

22.2.1.1 Control

The two **XMS100**, and **XMS100V** stages, will be each controlled with their recommended Newport **XPS-DRV02** driver cards, which slot into a Newport **XPS-Q4** 4-axis ethernet controller. This control assembly performs closed loop PID control for each stage, using the optical encoder feedback, to pulse width modulate the input voltage signal to each 3-phase ironless brushless motor. The control assembly correctly powers the stages, and synchronizes them with a small on-board CPU.

This is controlled over ethernet, to the main operating CPU in the product, which is interfaced with Newport drivers, that read the gcode output from the software suite, to control the stage. The Newport drivers will be configured to use Sgamma trajectory profiles, which provide smooth sigmoid acceleration profiles, thus the velocity and position profiles are smooth. This avoids the common trapezoidal velocity profiles, between points that induce abrupt acceleration changes, causing unwanted large oscillations in the system.

22.2.2 Transmissive Sample Illumination

The stage incorporates transmissive sample illumination. To ensure constant illumination of the sample when imaging from a fixed objective, the transmissive light source must remain fixed, directly underneath the transmissive sample. Therefore the stage must not occlude the illumination beam, and must allow the illumination light to illuminate the sample from underneath on the glass surface of the cassette.

The transmissive light source, realized in CAD in figure 22.2, is constructed using Thorlabs optical cage system standardized parts, which allow the optical components to be fixed precisely along their common optical axis. The incoming fiber optic, carrying the red illumination light, is connected to the lower **FCL1** Thorlabs collimator, which then passes through the polarizing filter **POL1**, and the condenser lens **L16**, which then shines the red polarized light onto the underneath of the sample,

which is imaged by the objective.

As the light source is fixed to the optical breadboard table, it must be positioned far enough away from the stage, such when the X movable platform, is fully extended by 5cm in the x axis, it will not contact or collide. Incorporating a safety factor of 1.4, hence the edge of the light source is $1.4 \times 5\text{cm} = 7\text{cm}$ away from the X movable platform edge. This leaves a 2cm gap between the light source and the stage when at the maximum X translation, for any additional cabling.

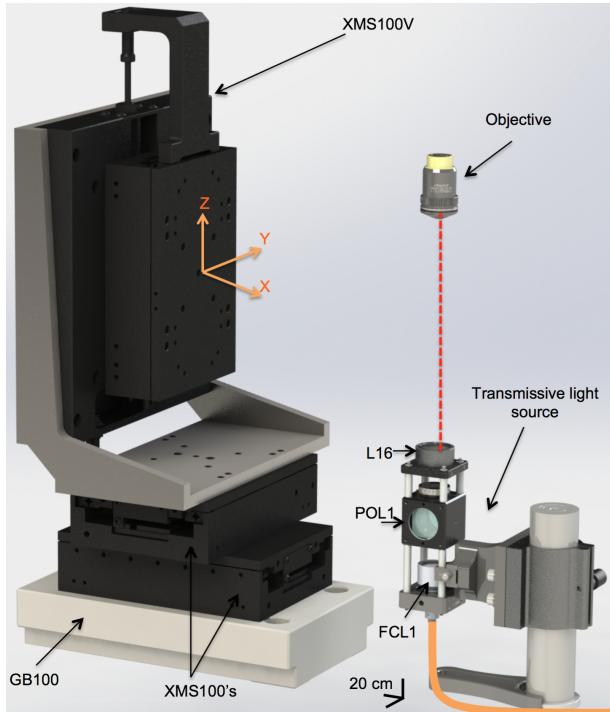


Figure 22.2: Stage layout of standardized components

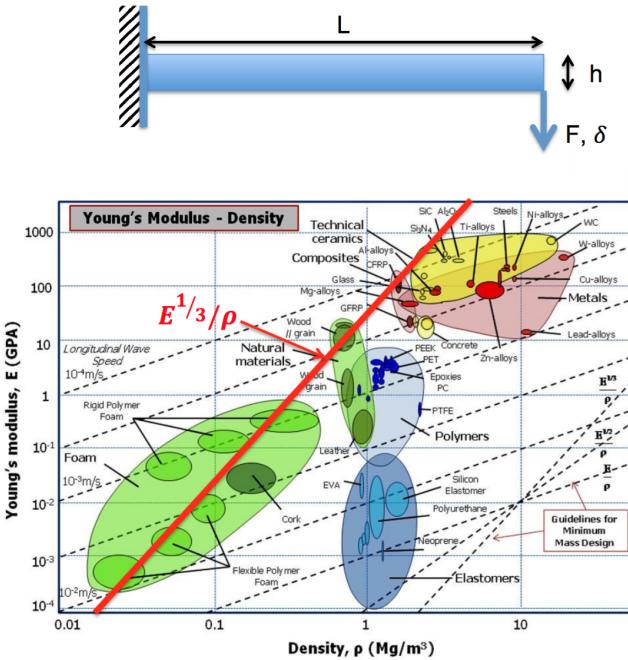


Figure 22.3: Simple Cantilever approximation, and Ashby Material selection chart [271]

22.2.3 Custom Cantilever Design

To position the sample so it is in contact with the illumination beam path, which is fixed away from the moving linear stages, a custom cantilever component was designed. As the entire stage is nanometer precise, to ensure the fabrication and imaging processes, can reliably meet nanometer resolutions, the cantilever must produce minimal tip deflection, to ensure no sample tilt is introduced. Additionally the cantilever must have its weight minimized, to reduce the loading on the lower linear stages and to prevent a substantial shift in the center of mass off the center axis of the linear stages.

The cantilever must hold the cassette mechanism (consisting of the housing clamps and cassette), which the sample rests on, a blank cassette mechanism weighs 0.523kg , incorporating a safety factor of 1.4 for dynamic sample weight, a uniform load can be assumed at the end of the cantilever of $1.4 \times 0.523\text{kg} = 0.732\text{kg}$. Hence the cantilever is designed to support a maximum uniform end load of $0.732 \times g = 7.18\text{N}$.

22.2.3.1 Cantilever Material Selection

Approximating the cantilever as a simple (uniform cross section) cantilever fixed into a wall (welded to a plate that is bolted onto the vertical stage), as seen in figure 22.3, with a length L , height of h , width w , and a force F applied at the tip, causing a tip deflection δ . Here L , and F are fixed as described above, and w is also fixed to accommodate the transmissive light hole. The free variables being h , and the material, with the objective to minimize tip deflection δ .

Here the second moment of area is $I = \frac{wh^3}{12}$ [272], with a tip deflection of, $\delta = \frac{FL^2}{2EI}$, assuming simple beam bending theory assumptions [273] [274], where E is the Young's modulus of the material. Substituting I into the above equation, tip deflection becomes $\delta = \frac{4FL^2}{Ewh^3}$. Since the beam, has a mass $m = \rho whL$, this can simultaneously combined with the δ equation substituting out h , as it is a free variable. Thus:

$$\delta = \left(\frac{4FL^6w^2}{m^3} \right) \left(\frac{\rho^3}{E} \right) \quad (22.1)$$

Therefore to minimize tip deflection δ , the material parameters (right bracket above) must be minimized, by maximizing $\frac{E^{1/3}}{\rho}$, the merit index. Taking a constant deflection, such that a constant $k = \frac{E^{1/3}}{\rho}$. Taking logs yields, $\log(E) = 3\log(\rho) + 3\log(k)$, an equation of a line on a log-log plot. By using the gradient of the line, sliding it upwards on a $\log(E)$ vs $\log(\rho)$ plot, to maximize the merit index, which maximizes E and minimizes ρ . This was performed in figure 22.3, illustrated as the red superimposed line. Suitable materials that maximize the merit index are the ceramic Silicon Carbide and composite carbon fiber reinforced polymers (CFRP). A subset of these materials are evaluated in Finite Element Analysis (FEA) simulations, against each other with the final component geometry to determine the most appropriate material for minimizing the tip deflection δ .

22.2.3.2 Cantilever Geometry Optimization

Cantilever geometry was designed starting with the basic cantilever model described above, illustrated in figure 22.3. An iterative design approach was taken, using FEA simulations (with the material Silicon Carbide) as feedback for geometry changes that minimized the tip deflection of the component under the loading of the sample and cassette assembly. Initially incorporating a modified I cross section design, due to its optimal stiffness (from its high second moment of area $I = \iint_R y^2 dA$) for bending in the vertical axis.

After 7 design iterations, the design seen in figure 22.4 was reached. The cross sectional cuts, illustrate the initial I section (with a thick web to reduce horizontal deflections, and to increase the horizontal natural frequency mode), which initially has a fixed beam width, constrained by the mounting screw holes for the vertical stage. The beam width gradually increases, causing majority of the cross sectional area to be farthest from the neutral axis, increasing the second moment of area, thus in-

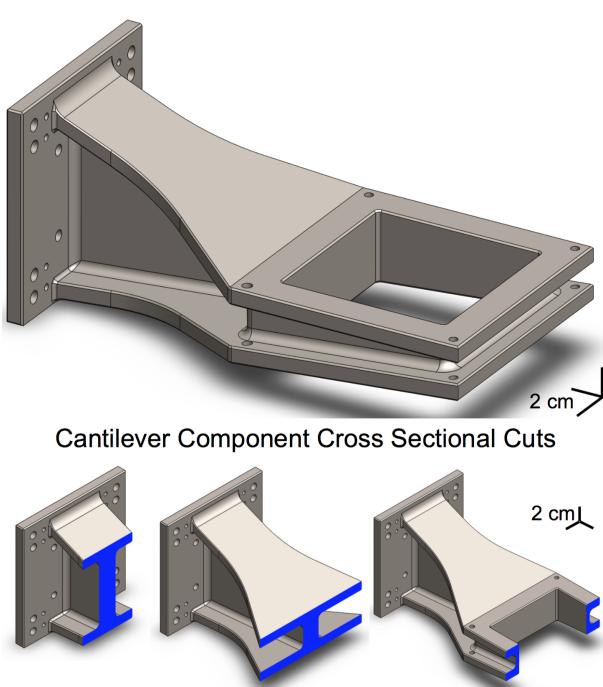


Figure 22.4: Cantilever Geometry, with cross sectional cuts

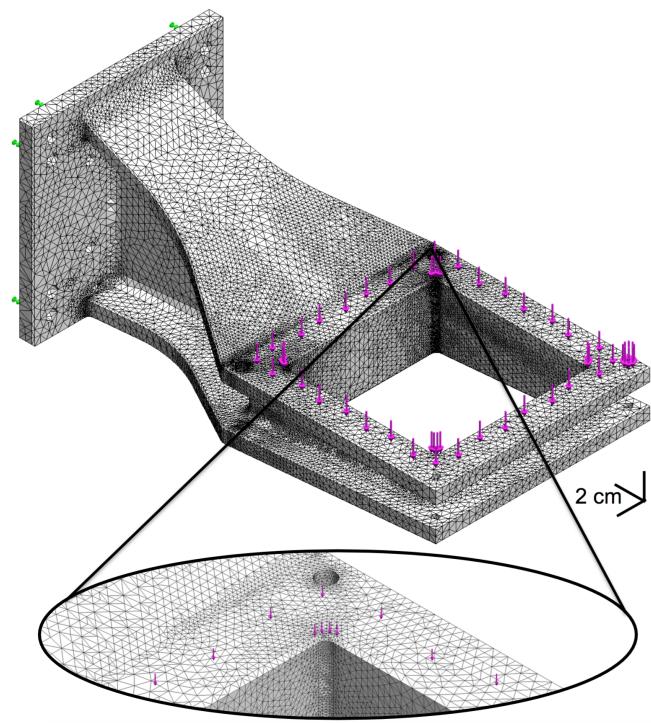


Figure 22.5: Component adaptive mesh, with close up mesh inspection

creasing local stiffness. The width of the beam is increased along its length, until it reaches the fixed width to appropriately contain the transmissive light hole, which is $100\text{cm} \times 100\text{cm}$. A split I beam section was incorporated around the transmissive light hole, along with the end of the cantilever. This custom cantilever geometry minimized the volume, hence the mass of the component. The mounting plate was of fixed dimensions, therefore the beam height initially started out at its largest acceptable value, reducing along the length of the beam. The beam height was reduced to a fixed dimension at the end of the cantilever to allow for the optical light source and objective to be mounted in close proximity without coming in contact with the cantilever component.

22.2.3.3 Cantilever FEA

Finite element analysis simulation, in Solidworks was carried to test various materials on the final geometry, and assist in the geometry design. In the structural mechanics simulation, the component is modeled as finite tetrahedral elements (mesh), which are simulated using the virtual work principle, to produce stress, strain and displacement results for the component.

The finite element analysis method was used, first the material for the component was selected. The fixed fixtures were defined, here the mounting plate to the vertical stage surface, is assumed fixed, since it will contain 20 screws to securely attach the component. Secondly the external load of 7.18N was applied distributed over the cassette base clamp contact surface.

The mesh for the simulations, were generated using the h-adaptive meshing method, which iteratively simulates the component, identifying areas where a higher mesh is needed (around stress

concentrations), to then refine (using smaller finite tetrahedral elements) the mesh locally at those points. This was performed using a fine mesh to start with, and iterating using a h-adaptive meshing 3 times. The adaptive meshing method offers superior solution convergence, utilizing computational resources efficiently. The generated mesh can be seen in figure 22.5, which shows a close up of mesh inspection. The mesh inspection correctly shows a suitable mesh has been generated, that correctly uses a finer mesh around the small fillet geometry of the transmissive hole in the cantilever (where stress concentrations are likely to occur).

The simulations were carried out initially for aiding the geometry design, then finally in selecting the choice of material, to minimize tip deflection. Potential materials were chosen and simulated, of which the results are tabulated below.

Material	Weight of component (kg)	Tip deflection δ (nm)
CFRP (Hexcel AS4C 3K)	1.40	164
Silicon Carbide Ceramic	2.45	91
Aluminum Alloy (7079 Alloy)	2.13	528

Table 22.5: Cantilever component material testing results

Analysis of the results in Table 22.5, shows clearly the best material for reduced tip deflection is a Silicon Carbide Ceramic, however it does come at the cost of increased mass to the component, however this is acceptable as long as it is within the working limits of the stages. The simulated tip deflection results, are approximately accurate to within $\pm 5\%$ of the actual values, due to the simple simulation performed. Additionally the aerospace grade aluminum alloy (7079 alloy), was tested to compare as a base line material for comparison. One of the many FEA simulation results can be seen in figure 22.6.

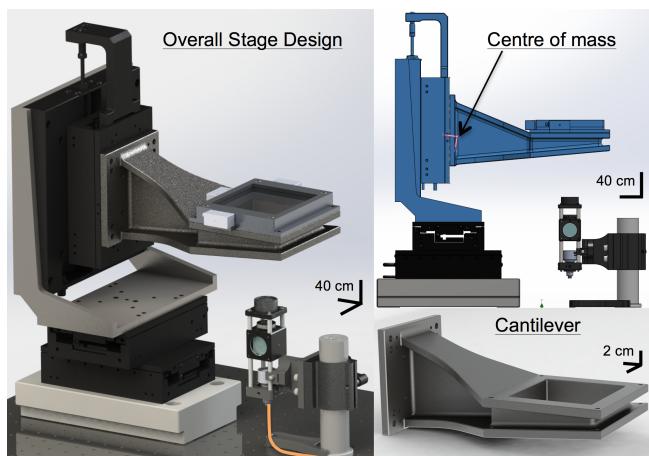
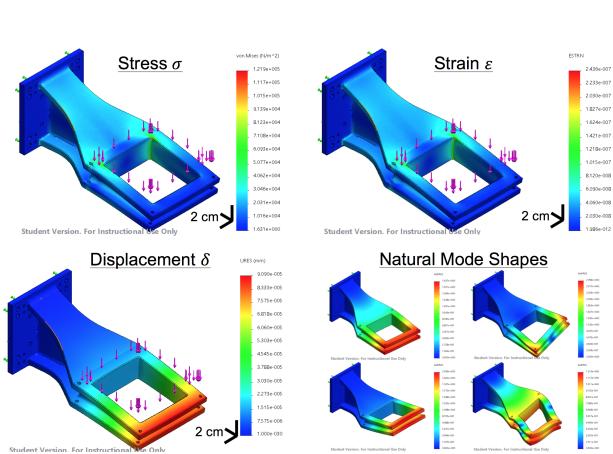


Figure 22.6: Silicon Carbide Ceramic FEA results Figure 22.7: Overall stage design, and center of mass diagram

Selecting the ceramic material Silicon Carbide, further simulations were carried out to determine the natural frequencies and natural mode shapes using a similar finite element method. The first few mode shapes can be seen in figure 22.6. The natural frequencies were 1.1kHz , 1.96kHz , 2.91kHz and higher multiples thereof. These are acceptable, as they are sufficiently high, and the stage will operate with very low speed dynamics, which will not drive any oscillations at these frequencies.

The cantilever component, can be custom manufactured from the supplier “Silicon Carbide Products” [275], for less than £1,000, constructed from sintering SiC grains to form the ceramic.

22.2.4 Overall Stage Specifications

Firstly the overall stage dynamics are affected by the weight of the assemblies that each linear stage has to support, here the vertical Z stage supports a force of 24.6N , the Y stage supports 127N and the X stage supports 162N . All the loading forces are well within the loading capacity for each linear stage. Each linear stages, will travel a small displacement before reaching their maximum operating speed, when traversing across to a new point in space. The worst case values were calculated using a linear acceleration model, for the given moving platform mass, and the stages constant force it can apply (16N for all stages).

Stage axis	Travel (mm)	Minimum Incremental Motion (nm)	V_{max} Operating Speed (mm/s)	Travel (mm) to reach V_{max}
X	100	1	300	46.4
Y	100	1	300	36.5
Z	100	50	300	9

Table 22.6: Overall stage dynamics

Evident from the table above, that the lower stages will have slower dynamics, due to their increased inertia, meaning that for the given constant force, hence constant acceleration, it takes a greater displacement to reach the V_{max} operating speed. The fastest stage, was designed to be the Z stage, as this is required for fast focusing during imaging, and layer stepping during fabrication. These low speeds are beneficial, as they are far away from any resonant natural frequencies of the mechanical components in the assembly. The final stage system can be seen in situ, in figure 22.7.

The final stage assembly, comprising of the cassette base clamp and the custom cantilever component, the center of mass for the assembly on the Y stage was checked, and it was found to be 10mm off the center X, Y central axis, see figure 22.7. This off center moment is within the operating limits, of both supporting X, Y stages. The total cost of all stage components included is £25,868, which is reasonable for its impressive capabilities, and is a small percentage of the total system cost.

22.2.5 Calibration

For ideal operation the sample should be held level to within $\pm 100\text{nm}$, in this design, the cantilever has a worst case tip deflection of 91nm , over its entire length. However when this is spread across the scanning area, taking for worst case the largest scanning area, of length 937mm , the Z error is only 0.852nm which is well within acceptable limits. Furthermore error mapping would be implemented, which would consist of calibrating known spatial positions in space, which would remove any remaining errors from manufacturing tolerances of standardized parts, to produce the most accurate and precise stage system.

22.3 Summary

The combined automated stage, with the galvo beam scanning system meets the specification to deliver a nano-positioning system with incremental motion $< 100\text{nm}$ in each dimension, which can robustly fabricate and image in the specification volume of 8cm^3 . In a typical TPP fabrication application, the desired $3D$ structure to fabricate could be as large as a few cm in size, which would be sliced into $2D$ layers, each layer further subdivided into $2D$ square grid pattern. Thus allowing the high beam scanning spatial resolutions (28.8nm) inside each galvo scanning square ($187\mu\text{m}^2$). This original design allows nanometre precision and resolution, over large volume sizes, which is key for a multi purpose imaging and fabrication stage and beam scanning system.

22.4 Further Development

22.4.1 Stage

The basic FEA simulations for the cantilever, did not account for dynamic movement of the sample, and or cyclic loading fatigue, caused by prolonged and constant use. These could be accounted for in specialized FEA software such as ANSYS.

22.4.2 Galvo scanning mirrors

Optical ray tracing software could be used, to model the non-trivial case of off-axis light tracing through the objective and into the sample and or the photoresist. Similarly analysis could be performed for the illuminating light source, using non-sequential ray tracing. Both these effects could be simulated in specialized optic ray tracing software such as Zemax.

23 Sample Logistics

23.1 Specification

In order to integrate Stereolithography into the build process, a means for transferring samples from the main stage to that of the Stereolithography resin vat was required. However, not only must it be able to transfer the object but also position it accurately and in a repeatable manner too so as the fabrication can continue uninterrupted without the need for additional time consuming calibration processes. The method must also have no impact on the capabilities of the fabrication and imaging techniques.

23.2 Cassette

Due to the potentially minute scale of the structures being fabricated, it is clear that the structures cannot be moved around directly. Hence a cassette system was devised to sit atop each of the 2 stages. This would be a standard part for all processes and would hold the build platform upon which the samples would reside from when they enter the machine to when they leave (rendered as glass in the centre of the above figure).

The cassette would have suitable means for its location and orientation to be determined accurately in space allowing any change of its position compared to its last machining run to be accounted for by adjustment to the stage positioning.

The design can be seen in figure 23.1 and outlined below are some of the key aspects in the design process.

23.2.1 Build Area

The build area of the cassette is the surface upon which the samples are stored while being machined or observed. They come in two main derivatives. The first is a glass panel and the other being a vacuum chuck either of which can be implemented in the cassette system. Being optically transparent the glass allows beams to approach and leave the sample from both above and below, critical for 'Bright-field Microscopy' and 'Polarised Light Microscopy' imaging modes that are implemented in this system. It should be noted however that the refraction created by the glass will not cause a problem, but will need to be accounted for in the optics of these processes.

The vacuum chuck solves the issue of the glass being unable to adhere to samples such as



Figure 23.1: Cassette 3D CAD Render

semiconductor wafers without clamping, however being opaque it limits the processes that can be implemented while it is in use. The porous nature of the ceramic also makes it unsuitable for very small samples hence it can only be used for macroscopic items with microscopic features making it ideal for wave-guide writing in glass substrates and machining of semiconductor wafers.

Due to the clear advantage of being able to implement both systems, the cassette has been designed so that with the same external shape and interface either build surface can be used. This allows the user to choose which build surface is most appropriate. For example the machining of optical wavguides in a glass sample with laser micro-machining may use the vacuum chuck to keep the sample stationary while stereo-lithography will build upon the glass panel version of the cassette.

23.3 Cassette I.D. Positioning and Calibration

The cassettes require a system to attach them to the stage, and the design in figure 23.2 was the final version with 3 optical encoders positioned atop a polished flat mounting surface. The cassette can be seen mounted in figure 23.3. A variety of aspects were considered in the design of this clamp they key points of which are outlined below.

23.3.1 Repeatable Positioning

As the cassettes cannot be located with nanometre accuracy, a method is required to account for any error in the cassettes positioning after being relocated. This can be broken into 2 stages. 1. Identifying the positioning error and 2. correcting that error.

23.3.1.1 Identifying the error

The simple way to identify location is to place markers or use a fiducial which can be tracked in space. Ordinarily, this is an optical process with cameras observing an area where a clear pattern exists. Computer vision software then determines the location of the pattern on the camera's image plane. More advanced systems can track the size and distortion of the fiducial and hence also identify a measure of distance from the camera and angle to it. However, these processes cannot achieve the sub-wavelength levels of accuracy that we need. The solution to this problem is to use high resolution encoders.

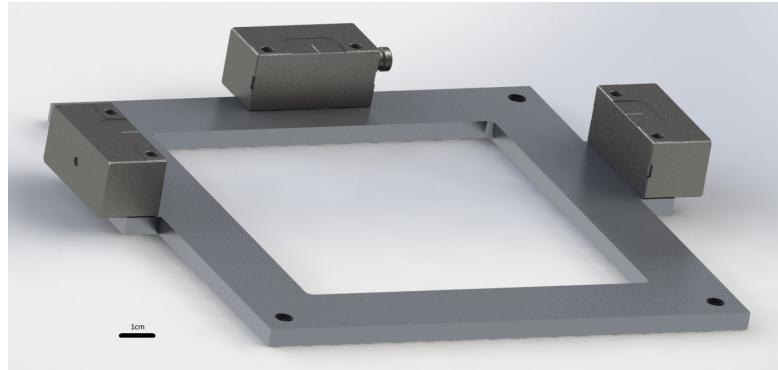


Figure 23.2: Cassette-Stage Interface

23.3.1.2 Encoders

Encoders come in linear or rotary measure and operate on a variety of principles including: optical, magnetic, and interferometric laser. They can also give incremental or absolute readout of position.

For our uses, we need an absolute readout of position so as to be able to determine the location of the cassette without zeroing movement. We also want the encoders to be linear in operation and sufficiently compact to not impede the function of the fabrication and imaging systems.

The only encoder found to satisfy the requirements on both sub micron resolution and also absolute readout without zeroing motion is the 'Resolute' encoder manufactured by 'Renishaw PLC'. This is an optical encoder that reads a $30\mu\text{m}$ pitch coded track, much like a bar-code, to attain a coarse location before using a number of confidential methods to interpolating the fine position by way of phase comparison. Ultimately leading to a 1nm resolution.

The implementation of the encoders can be thought of as restraining a rigid body. Each encoder can measure and hence determine a single dimensional measure, so they behave like higher pairs in Gruebler's equation 23.1 where F is the total number of degrees of freedom in the model, n is the number of links (including the reference frame), l is the number of lower pairs and h is the number of higher pairs. [276] It should be noted that this equation applies to planar mechanisms hence the cassette must be constrained to a plane of motion before applying this rule. We have satisfied this requirement however by clamping the cassette to the surface of the stage and hence restraining its motion to that in the plane of the stage.

$$F = 3(n - 1) - 2l - h \quad (23.1)$$

Without any encoders, but with the cassette remaining on the stage plane, there are no higher or lower pairs, and two links exist (the cassette and the stage) hence:

$$F = 3(2 - 1) - 2 \times 0 - 0 = 3 \quad (23.2)$$

So three DoF (Degrees of Freedom) remain; as would be expected of an unconstrained planar mechanism (x and y movement over the plane and rotation about an axis perpendicular to the plane). Adding three encoders, however, provides three higher pairs:

$$F = 3(2 - 1) - 2 \times 0 - 3 = 0 \quad (23.3)$$

So the cassette now has zero degrees of freedom relative to the stage and its position in space is

defined in its entirety. This allows our control system to 'know' the location of the cassette (and by extension build volume) with respect to the objective lens, within the bounds of error set out by the encoders measuring the error and the stage accuracy.

The inclusion of three encoders is not sufficient to determine the position of the cassette, the encoders must be able to measure motion in the unconstrained degrees of freedom. The ultimate layout can be seen in figure 23.2. In these positions both the 'X' and 'Y' position can be tracked to 1nm and by comparing the encoder readings the rotation of the cassette can also be identified. Some basic trigonometry giving the angular resolution as being 4.093×10^{-7} degrees or 1.473 milliarcsecond.

23.3.1.3 Calibration

Encoders require two components to operate: the 'Read-head' and the 'track' or 'scale'. These are attached to the two objects that you wish to determine the relative position of. In this case, the scale is attached to the cassette and the read-head is attached to the stage. However, we cannot position either of these components with perfect nm accuracy, hence they must be calibrated before first use.

The calibration process is effectively a reverse of the ordinary position checking process. However this time the position relative to the objective lens is identified by probing for the edges of the build volume with the imaging capabilities of the system, a number of points can be identified along each edge of the cassette aperture and the stage coordinates recorded for each point. This will allow a line to be fit to each edge and a virtual build volume created. Combining this with the encoder readout for that particular state will associate the build volume positioning with the encoder values which can be then used in the opposite direction. Furthermore due to variations between each cassette they will all require calibration which will result in different calibration details for each cassette.

23.3.1.4 Clamping

With the cassette on the stage, it will be experiencing forces due to the acceleration of the stage, this could cause the cassette to slide over the stage or even lift off of it entirely, hence this must be considered and accounted for, potentially with some form of clamping mechanism to prevent any movement.

From the stage specification the maximum acceleration possible is 5m/s^2 in the z direction and the same in each of the x and y directions resulting in 5m/s^2 vertically and a combined $5\sqrt{2}\text{m/s}^2$ laterally. The worst case scenario is for the stage accelerating downwards while also accelerating laterally in both directions at maximum rate. When combined with gravity the resulting response force of the base supporting the stage is given by equation 23.4 (mass taken from solidworks analysis as 0.36kg):

$$F = ma = 0.36 \times (9.8 - 5) = 1.73N \quad (23.4)$$

The other vertical force to consider is air resistance upon the build platform as it moves. This force can be estimated by equation 23.5. Where ρ is the density of air, v is the velocity of the object through the air, C_D is the drag coefficient and A is the cross sectional area.

$$F = \frac{1}{2}\rho v^2 C_D A \quad (23.5)$$

As a worst case scenario v must be taken at its maximum. The stage has 350mm of travel, if accelerating at $5m/s^2$ the maximum reachable speed can be found by equation 23.6 where v_f is the final velocity, v_o is the initial velocity, a and d are the acceleration and displacement respectively.

$$v_f^2 = v_o^2 + 2ad \quad (23.6)$$

Substituting in values results in a maximum velocity of $3.5m/s$ returning this to equation 23.5 and using $C_D = 1.28$ (for a flat plate perpendicular to the flow) results in $0.096N$ of force. Hence the force between the cassette and stage is 1.63. Hence the cassette will not lift off of the stage, however it may still slide. For the cassette to remain stationary the accelerative force upon it must not exceed the force of static friction between it and the stage. This is given by equation 23.7 with C_f being the coefficient of friction chosen as 1.15 for clean dry aluminium on aluminium and R is the force between the stage and cassette calculated above.

$$F = C_f R = 1.15 \times 1.63 = 1.87N \quad (23.7)$$

Under maximum lateral acceleration the force reaches $0.36 \times 5\sqrt{2} = 2.55N$ this exceeds the threshold for slippage hence something must be done to secure the cassette when accelerating horizontally.

The use of physical pins or clamps poses the issue that they are imperfect, the base cannot be machined to nm accuracy and even if it could be then wear and damage would change that. So slippage may still occur, however not at the macroscopic scale. The solution to this is to increase the force holding the cassette to the stage, this in turn increases the critical threshold for slippage and hence, the accelerative forces will not exceed the value any more. The method however must be lightweight so as not to burden the stage and it must also be possible to disengage it in order not to overload the robotic arm used for moving the cassette detailed in section 23.5.

The solution is to use electromagnets with ferrite studs in the cassette. This is a switchable, lightweight and compact system. It can introduce additional newtons of force on demand, more than enough to prevent slippage.

23.3.2 Cassette Identification

As each cassette is unique and requires calibration, it is undesirable for the cassette to need calibrating before every use. Hence a means of uniquely identifying each cassette enables the system to write the cassette calibration details to a database which can be called upon during every load thereafter in order to attain the relevant offsets for that specific cassette.

23.3.2.1 Types of I.D.

The simplest of ID methods is a barcode, cheap and well known, a barcode encodes data with a series of narrow and wide light and dark stripes. These stripes are read by a laser scanner (sometimes a camera) which can detect the reflection of the laser light on the light bars. Narrow bars encode a '0' and wide bars a '1' hence a simple digital code now exists which can encode any digital sequence. It is limited however by the resolution and range of the scanner, this limits the amount of data that can be stored, however the system is very cheap.

A 'QR code' is a form of two dimensional barcode (others being 'Aztec Code' 'Micro QR Code' 'Data Matrix' and 'Microsoft Tag') [277] however this time instead of encoding the data in the width of the alternating colour stripes, the data is determined by the colour of the pixel. Otherwise the systems are very similar, the difference being that QR codes can encode more data and are read by a camera as opposed to a laser scanner.

RFID (Radio Frequency IDentification) is another form of identification that encompasses a wide range of ID methods. Markers can be either powered or passive, and systems can operate on a wide range of frequencies enabling read distances to be anywhere from close contact, up to 100's of metres. This versatility with range removes the restrictions on alignment and aiming of the reader to the identification at the two methods above require, however it is also more expensive and difficult to implement as an RFID tag would have to be attached to the cassette somewhere whereas an optical system can just be printed onto the cassette directly.

23.3.2.2 Comparison

Table 23.1 shows a multi-criterion analysis between the tree major identification modes. RFID technology has come up short against the two optical methods due to the system not requiring the benefits that it brings, along with its cost and difficulty of integration. Meanwhile the optical methods are cheap and satisfy requirements to just hold an ID number, all other information can be stored on a database referenced by said I.D.

-	-	Barcode		QR Code		RFID	
Comparison	Weight	Points	Score	Points	Score	Points	Score
I.D. Size	3	2	6	3	9	1	3
Reader size	4	2	8	2	8	1	4
Data Capacity	1	1	1	3	3	2	2
Reliability	5	2	10	2	10	3	15
Read Speed	2	2	4	2	4	2	4
System Cost	3	2	6	2	6	1	3
Total	-	-	35	-	40	-	31

Table 23.1: Cassette Identification MCA

23.4 Storage

An advantage of having uniquely identifiable cassettes that can be loaded automatically is that they can also be stored automatically. This enables the system to work in a manufacturing line environment optimally. This fabrication facility may be preceded by a cassette and sample preparation station. It may also be succeeded by a post-processing and cassette cleaning station. If no storage were to be included in the system, then the failure of any one process would halt the entire production line. For example, if the post-processing station were to fail then the fabrication facility would be unable to offload its current job and hence will have to wait. Unable to take on a new cassette the preparation station will also be halted. Introducing storage into the above scenario results in the fabrication facility continuing to operate if post processing fails as it offloads the completed cassettes to storage. Similarly, if the preparation station fails then it should have filled up the storage for the incoming jobs and hence the fabrication facility can continue until the issues are solved. This decouples the actions and operation of the system from the rest of the assembly line enabling greater 'up time' and hence higher manufacturing throughput.

A scenario for seeing the benefits of this is in 24 hour unattended operation. With no storage, when the machine technicians go home at the end of the day there are no cassettes being loaded or offloaded and hence the system can only complete the job it is currently running. However, with the storage, the 'In' storage can be filled and the 'Out' storage emptied before the end of the day. At which point the fabrication facility can run overnight processing the jobs in its stockpile.

23.5 Transport

23.5.1 Specification

As specified in section 24.1 a means must be implemented to transport the cassettes between stages and storage with sufficient accuracy to be placed within the $\pm 0.15mm$ tolerant read distance of the encoders after which they can determine the exact location down to $1nm$ in each direction. This must also be performed by a system that does not affect the operation of the optics, hence vibration on the table must be avoided or fabrication paused while it moves.

23.5.2 Result

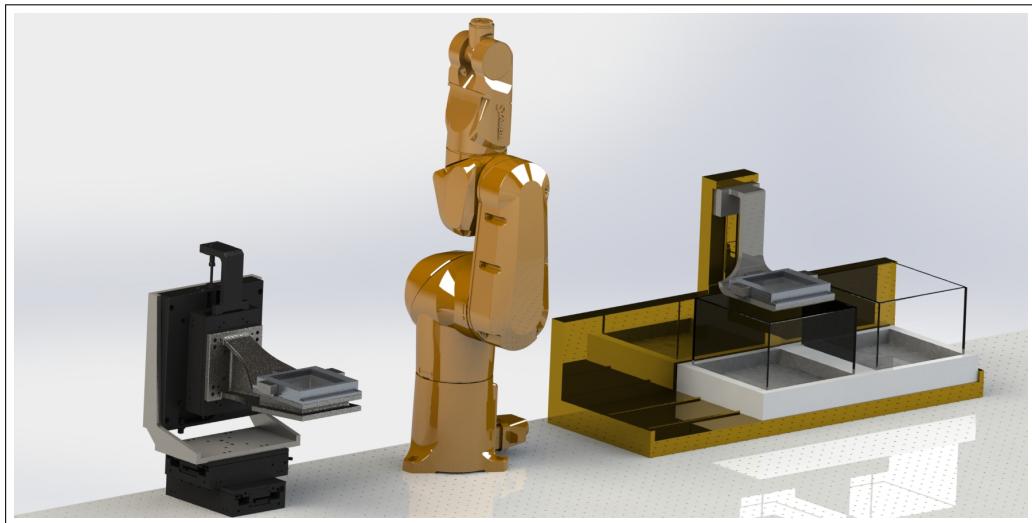


Figure 23.3: Both stages with the robotic arm

The solution is to use a robotic arm, such as the 'Staubli TX2-40' seen in the render of figure 23.3. This is a 6 degree of freedom robotic arm with nominal load of $1.7kg$, and can reach $51.5cm$ from its base, this is sufficient to be able to reach both work points as well as a storage location and an input/output tray. It also has a $\pm 0.02mm$ repeatability, sufficient accuracy to allow the cassette to be placed within the read envelope of the encoders.

23.6 System Summary

Due to the requirement for fabrication samples to be moved between work points with sufficient accuracy a system has been devised that by way of interchangeable stages or 'cassettes' the samples can be relocated by a robotic arm, no matter how small they are. This has been enabled by instead of having to be positioned with nm accuracy, which is very expensive and time consuming, they can be positioned with $0.1mm$ accuracy before the error in position is measured and counteracted by offsets and adjustments to the fabrication toolpath.

Part VI

Control

24 Software

24.1 Overview

The facility comes equipped with a robust, user-friendly software suite that will enable the end user to define the object to be fabricated as well as the fabrication parameters. All the user has to do is provide a 3D model file, define the desired size of the model, and select the desired fabrication material, the software will handle everything else.

The software program was designed from the ground up in both C++ and Matlab, with C++ being used for computationally expensive portions of the program, while the Matlab portion acts primarily as the GUI wrapper, but also handles 2D manipulation as Matlab provides a very robust and powerful toolbox for doing so.

In this chapter, I will discuss the process of designing the software, from defining the desired functionality, to the designing of the algorithms used, to testing examples of potential end use cases. Particular emphasis shall be put on the mathematical theory and performance of the algorithms.

24.1.1 Capabilities

The main function of the software is to translate a supplied 3D design file into control signals and tool paths for the lasers to take. It does this by first slicing the 3D model into distinct layers, each of these layers corresponding to a layer fabricated by either Stereolithography or Two-Photon Polymerisation. The end user can view each of the layers in the GUI to ensure that the data from the model file was read correctly. Note this layer data can also be used for the Live Imaging Feedback as described in Section 8.4.1. Then the program runs a resolution separation algorithm on each layer to separate the high resolution portions of the layer, to be fabricated by Two-Photon Polymerisation or STED Lithography, from the lower resolution portions of the layer, to be fabricated by Stereolithography. The user will also be able to view the results of this algorithm in the GUI to ensure that no errors have occurred. Then the software will run tool path generation on each of the layers, resulting in a set of lines defining the movement of the laser on the surface of the material. This information, which can be viewed by the user, will be fed into the input of the control system for fabrication. An example of the user's perspective of the software is shown below in Figure 24.1.

24.2 STL File Format

The object to be fabricated will be provided by the user in the form of an STL (Standard Tessellation Language) file. This file format is an industry standard file format for storing a 3D CAD model. The file only provides the surface geometry of the object; no texture, color, or material data is provided in the file. Furthermore, the file provides no units for scale, this means that the object can be scaled up

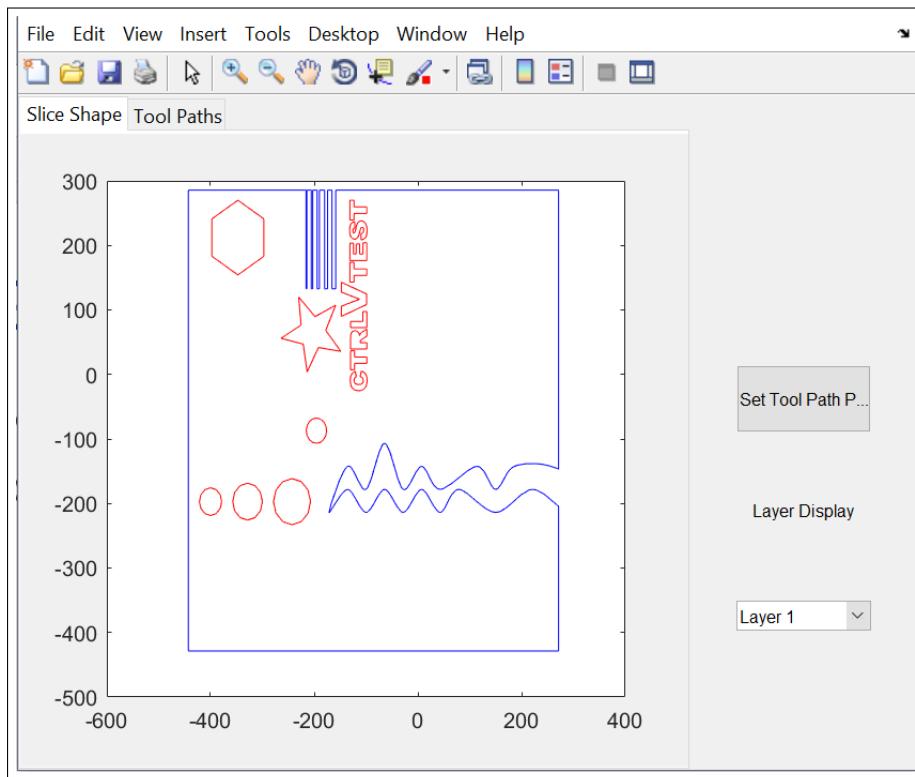


Figure 24.1: User end GUI: layer and tool path data, options for system parameters

or down to any size that the user desires (provided it fits within feature resolution and build volume parameters). The surface geometry is described with an array of triangles (tessellation), each triangle has 3 vertices and a normal vector associated with it. An important thing to note is that all the triangle's vertices must be vertices of another triangle as shown in Figure 24.2. The binary file format can be seen in Figure 24.3.

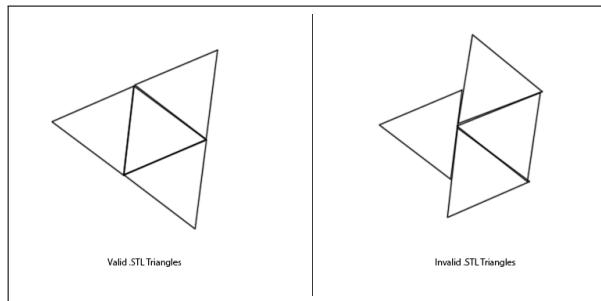


Figure 24.2: Example of valid vs invalid triangle formation for STL format.

24.2.1 Data Structures for STL File

The choice of C++ for the reading of the STL file was deliberate in that C++ provides many native tools for reading binary files into data structures. Three classes were defined to store and manipulate the data: v3 (3D vector), Triangle, and TriangleMesh. The TriangleMesh class is simply a class to contain a vector of triangles and parameters of the model (dimensions and number of triangles). Simple vector operations were implemented in v3 to allow for easier adding, subtracting, dot product-ing,

```

UINT8[80] - Header
UINT32 - Number of triangles

foreach triangle
REAL32[3] - Normal vector
REAL32[3] - Vertex 1
REAL32[3] - Vertex 2
REAL32[3] - Vertex 3
UINT16 - Attribute byte count
end

```

Figure 24.3: Structure of the binary STL file.

and cross product-ing later in the program.

24.2.2 Parsing of STL File

The reading of the file is done simply by reading the STL file as input stream and reading sequential blocks of bytes corresponding to the structure shown in Figure 24.3. A simple flow chart for this process is shown in Figure 24.4. Care must be taken to ensure that the file location provided to the program is valid otherwise the program will fail. In the event that an incorrect file path is provided to the program, the program will return an error message and halt execution.

24.3 Layer Slicing

The layer slicing algorithm is, at its essence, an implementation of vector algebra in 3D space. The goal is to have, for each layer, a set of points defining the intersection of the TriangleMesh (defining the 3D model) with the slicing plane. For simplicities sake, the slicing plane was chosen to be in the z-direction.

24.3.1 Theory

First, there is a very simple test to determine whether there will be an intersection or not. This test is merely to check whether the height (z-coordinate) of the slicing plane is between the maximum and minimum (inclusive) z-coordinates of the triangle. This is shown in Figure 24.5. If this condition is satisfied, then we are guaranteed some sort of intersection. There are distinct cases of intersection, as shown in Figure 24.5.

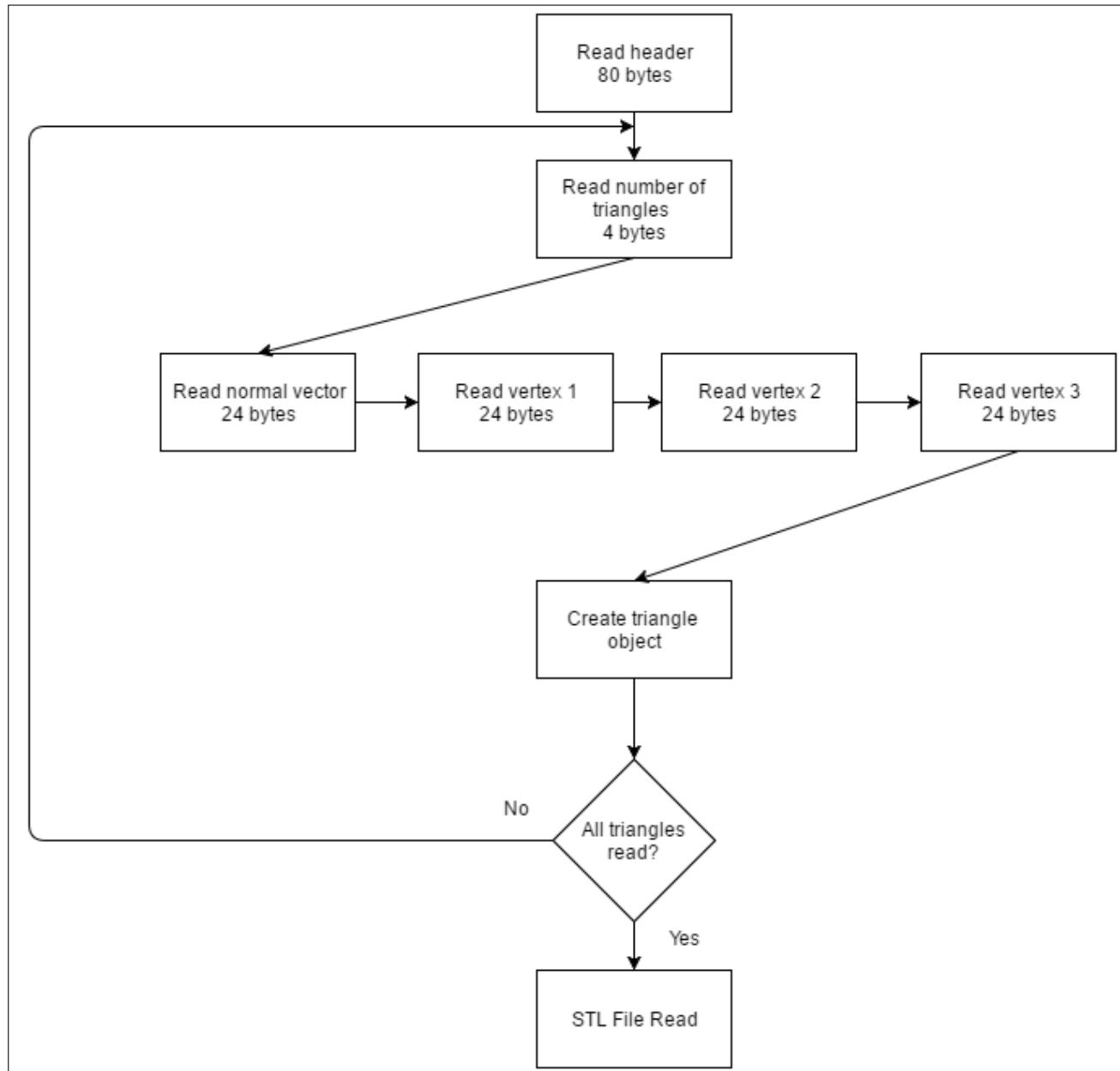


Figure 24.4: Flowchart for parsing STL files.

Case Number	Description
1	'Normal' Intersection
2	Single Vertex Intersection (Edge)
3	Single Vertex Intersection (Middle)
4	Two Vertex Intersection
5	Three Vertex Intersection (In Plane)

Each of these cases will be handled differently. By returning to the definition of an STL file (Figure 24.2), we can safely ignore cases 2 and 5. For case 2, the point of intersection will be picked up by another triangle's intersection, so to ensure we don't have double records of intersection we will ignore this case. For case 5, this triangle can be defined as part of the interior surface of a layer; we

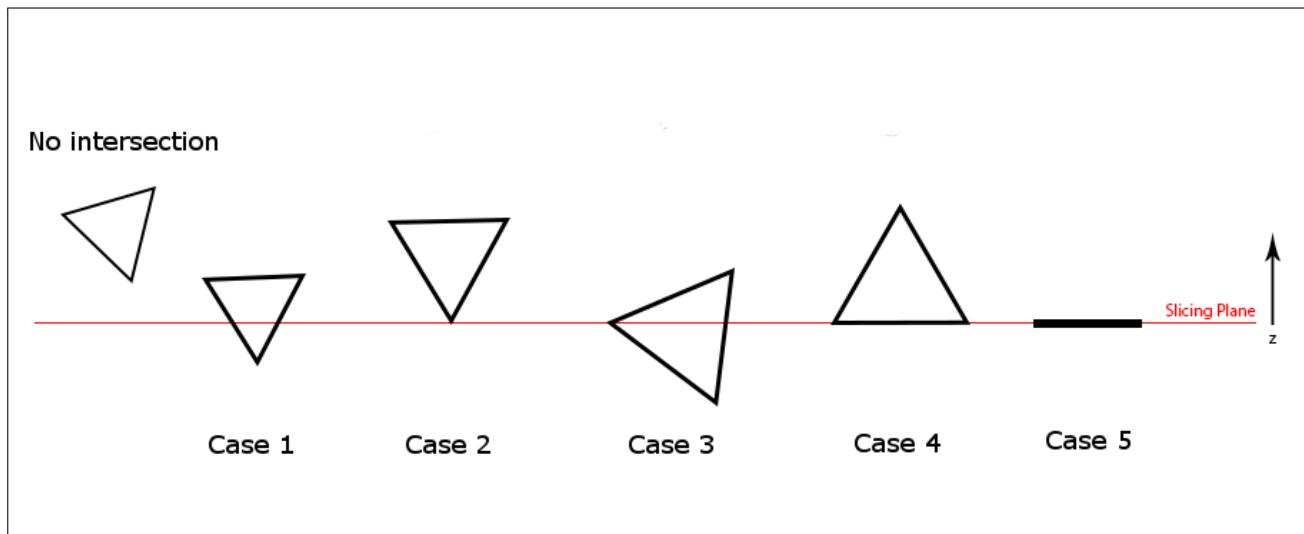


Figure 24.5: Different cases of intersection of triangle and slicing plane.

are only interested in the layer boundaries defining the change from exterior to interior, so we can ignore case 5 too.

This brings the discussion to cases 1,3 and 4. We will start with the simplest one, case 4. As shown in Figure 24.5, the two points intersecting the plane are simply the two vertices with the same z-coordinate as the plane.

Next, we shall consider case 1. It is evident that the points of intersection can be described by the intersection of a line with a plane. This is a simple vector algebra problem to solve. If we define the points of interest as shown in Figure 24.6.

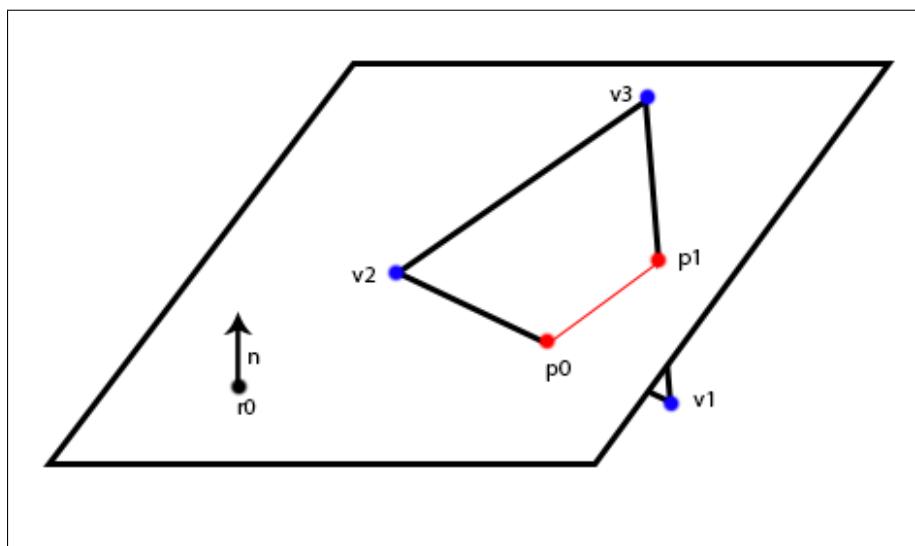


Figure 24.6: Case 1 intersection.

The two line segments intersecting the plane are given by, where d_{ij} defines the direction between

v_i and v_j and λ moves along the line segment:

$$p_0 = v_1 + \lambda d_{21} \quad (24.1)$$

$$p_1 = v_1 + \lambda d_{31} \quad (24.2)$$

We can also define the plane in the traditional way, where r_0 is an arbitrary point on the plane and n is the normal vector:

$$(r - r_0) \cdot n = 0 \quad (24.3)$$

From Figure 24.6 it can be seen that p_0 and p_1 must lie in the plane, hence (for p_0):

$$(p_0 - r_0) \cdot n = 0 \quad (24.4)$$

$$(v_1 + \lambda_0 d_{21} - r_0) \cdot n = 0$$

The solution to this equation for λ_0 is given by: $\lambda_0 = \frac{(r_0 - v_1) \cdot n}{d_{21} \cdot n}$. A similar result can be found for p_1 .

Case 3 is a mix of both case 1 and 4. One of the points of intersection is given simply by the vertex, whereas the other point is given by the solution to the intersection of a line and a plane as given above.

In review, the points of intersections for the given cases above:

Case Number	Points of Intersection
1	$p_0 = v_1 + \frac{(r_0 - v_1) \cdot n}{d_{21} \cdot n} d_{21}$, $p_1 = v_1 + \frac{(r_0 - v_1) \cdot n}{d_{31} \cdot n} d_{31}$
2	Ignore
3	$p_0 = v_2$, $p_1 = v_1 + \frac{(r_0 - v_1) \cdot n}{d_{31} \cdot n} d_{31}$
4	$p_0 = v_2$, $p_1 = v_3$
5	Ignore

24.3.2 Algorithm for Slicing

Now that we have a methodology for slicing the layers, we must implement this into code. The full algorithm can be seen in Figure 24.7. It is an implementation of the theory stated above, first checking the boundaries on the triangle against the layer height, then checking which case it is and appropriately applying the vector algebra to find intersections. The algorithm is $O(n)$ in number of layers and $O(n)$ in number of triangles.

Note that a new type of data class has been introduced to store the intersection segments: Line-Segment. The class is a simple way of storing a line segment in the form of two points, a magnitude, and direction.

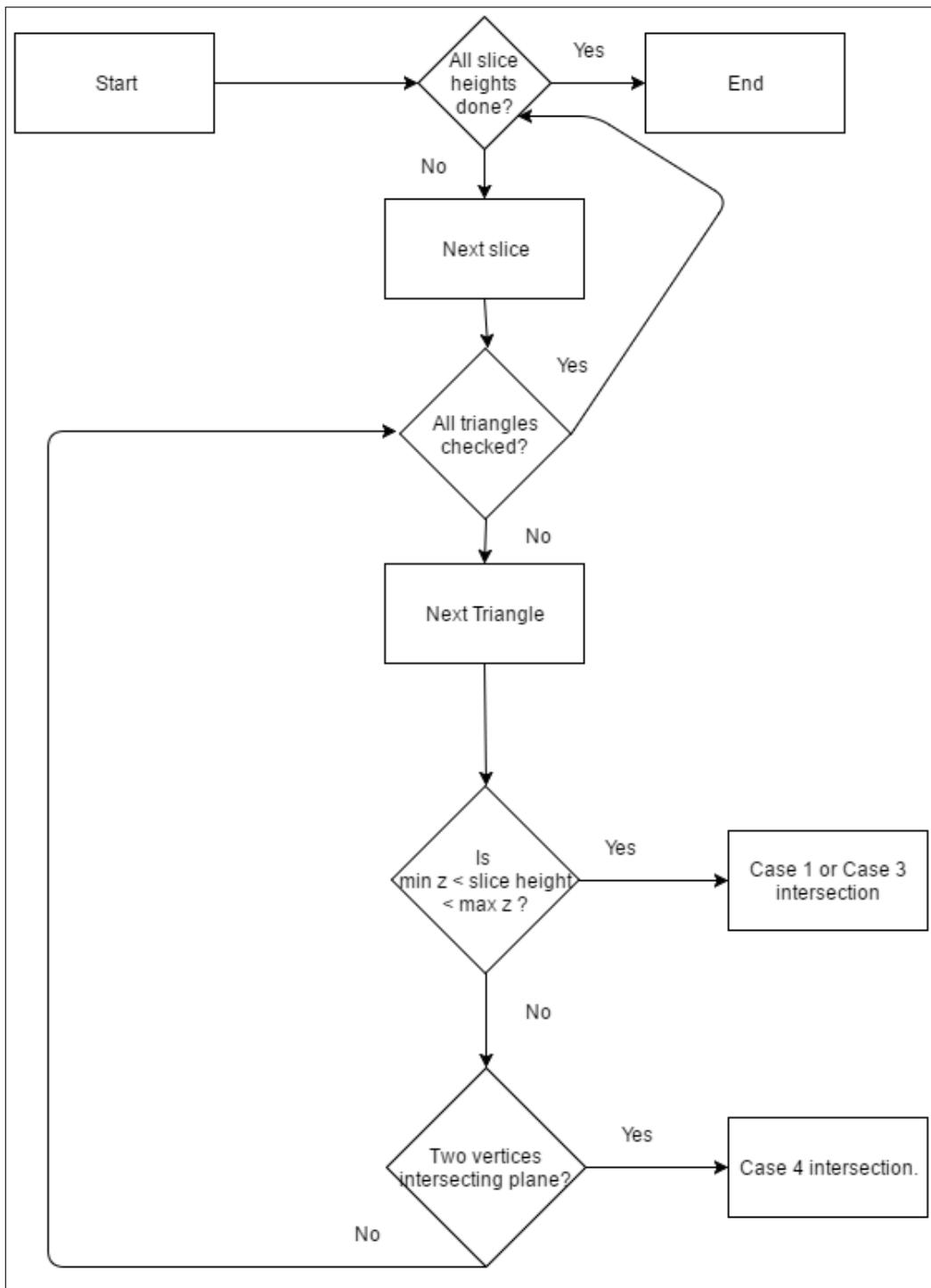


Figure 24.7: Slicing Algorithm Flowchart.

24.4 Contours

Now the data for the layer slices are stored in unordered arrays of line segments. While this is functional, it makes more sense organizationally and for further use in tool paths to reorganize this data in an ordered arrays of points defining the vertices of the 'contours'. It makes sense to classify each contour as an inner or outer contour as shown in Figure 24.8. Outer contours define the outer boundary of material whereas inner contours define the inner boundary of material.

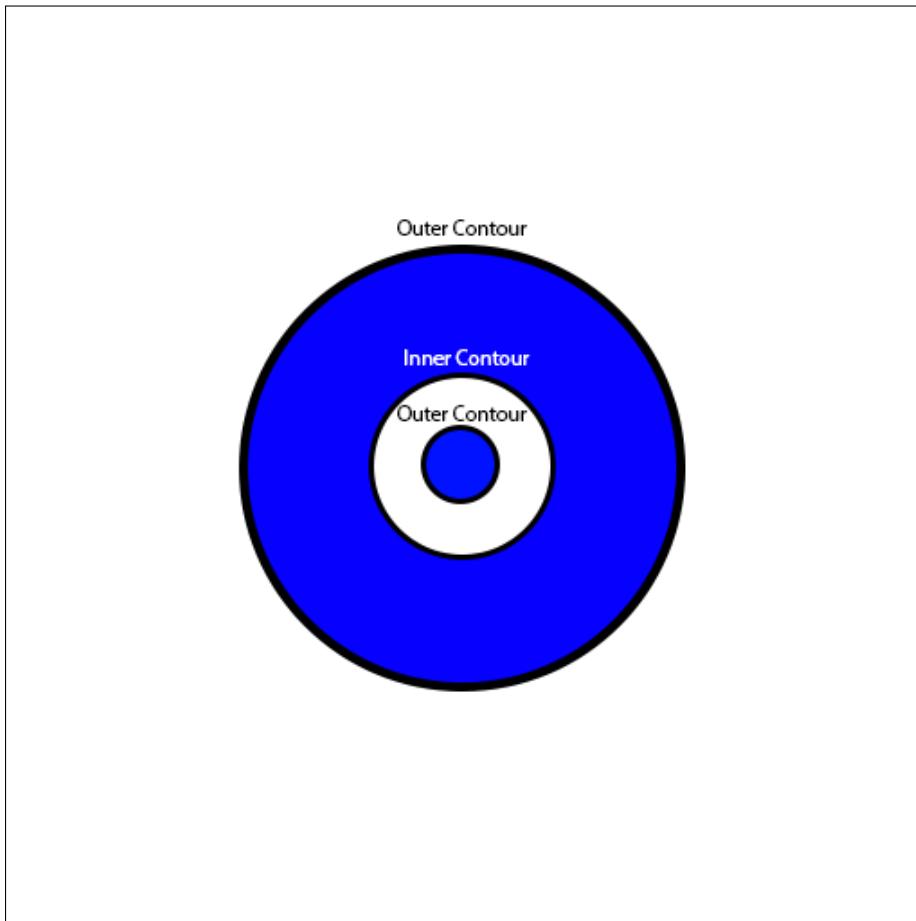


Figure 24.8: Example of inner vs outer contours.

First, a Contour class must be defined. In defining the class a decision had to be made for the data structure to hold the point data.

24.4.1 Linked List vs. Vector

The selection of data structure came down to two different types: the built-in vector object provided by the Standard Library of C++ or a linked list implementation. The merits and disadvantages of both types will be discussed here.

Beginning with the Standard Library implementation of vector(which is implemented in C++ as an array), we will examine first allocation. For maximum efficiency, a vector object should have a pre-allocated size, that is, the size should be defined in the object definition, eg - `std::vector<v3> vector(size)`. While it is possible to allocate an empty vector and increment the size every time you append a data point to it, this results in inefficiencies since the entire vector array may need to be moved to a different physical memory location to accommodate the larger memory size. Since we do not know the size of the contour beforehand, this constant appending of elements may result in program slowdowns. However, a vector implementation is much simpler and easier and does not deal with memory pointers, which if mishandled, can lead to memory leaks. Another advantage of vectors is that accessing individual elements have a constant, quick access time due to the way vectors are

implemented.

Linked lists are a more complicated data structure that require some explanation. A linked list is a data structure in which ordered items are linked via memory pointers. Each 'list item' has two components: a data component and a pointer to the next element. Typically the end element of a list has a null pointer to signify the end of the list and the first item has a null data component to signify the beginning of the list. A visual aid of this data structure is shown in Figure 24.9. Again, we will begin with allocation. Linked lists have no performance impact on allocation, in fact, you cannot pre-allocate the size, only dynamic allocation is possible. The list elements do not have to physically be in a continuous location like a vector since you merely point to the memory location of the next element. A disadvantage of linked lists is accessing individual elements by index, to access the n_{th} element of a linked list, you must cycle through $n - 1$ list elements to reach that element. This results in minor program slowdown. Also, since a linked list deals with pointers, care must be taken to manage the memory correctly, or memory leaks may occur.

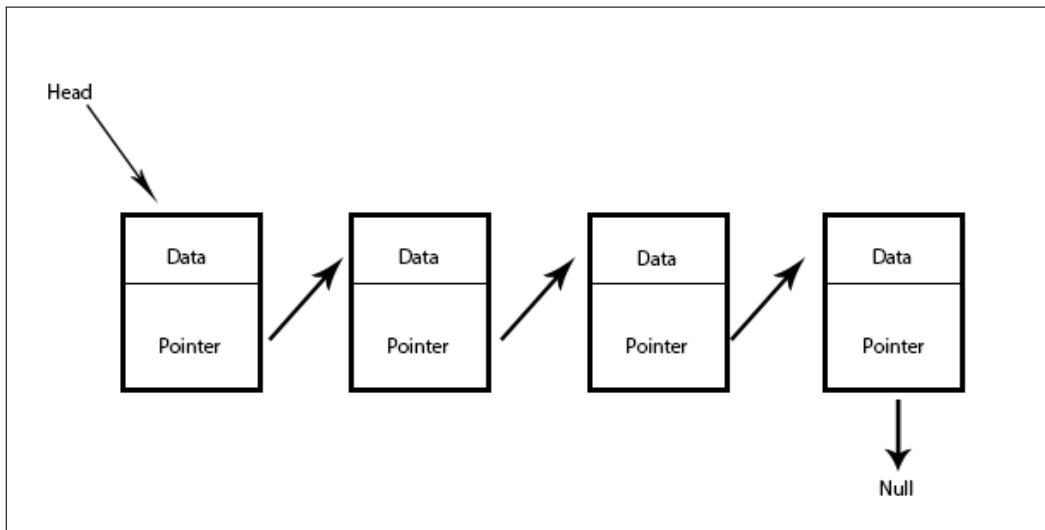


Figure 24.9: Data Structure of a Linked List.

Keeping these differences in mind, I selected a linked list implementation for the Contour class for two reasons. A linked list imparts an inherent order to the memory elements, similar to how a point connects to the next in a contour, a list element points to the next one. Also, the contours will be dynamically allocated, so a linked list is much more preferable in terms of performance. The performance hit in accessing the n_{th} element is unlikely to occur often since we will likely be cycling through the list rather than picking out a random element.

The implementation of the linked list was slightly different than the typical one in that the final element points to the first element, 'closing the loop'. This change required a list element to hold two properties: the point (v3) and a end value (boolean). The end value would be false on every element

except for the last one to signify the end of the loop.

24.4.2 Creating the Contour Objects

The next problem to tackle was converting the array of line segments into a set of contours defining the layer. The heart of the problem is connecting vectors head to tail until a closed loop is formed. In practice, this problem is a bit more complex than first meets the eye. Note that a set of line segments may contain any number of contours.

The simplest way of doing this would be to take the first line segment, assign one point as head and one as tail, then cycle through the remaining vector searching for a point that matches the head and appending the other point in the segment to the list and calling that one the head. A flow chart of this algorithm is shown in Figure 24.10. This algorithm would return the set of line segments that were not part of the contour to be ran through the algorithm again. This algorithm would be run until the set returned has no elements, then and only then would we have the full contour set that describes the layer. This algorithm would be $O(n)$ in the size of the original set of line segments best case and $O(n^2)$ in worst case.

However this method has one crucial fault that is a result of floating point numbers.

24.4.3 Floating Point Inaccuracies

The algorithm in Figure 24.10 relies on matching points in different line segments, however, since the (x,y,z) data are stored as floats, exact matching is impossible due to general inaccuracies in floating point numbers. For example, in floating point arithmetic, $10 \cdot 0.1 = 1$ may not be true.

The first solution to this problem that comes to mind is to say that the floats are equal if they are within a certain magnitude of error of each other. However, finding an appropriate threshold that universally worked was problematic. If the threshold was too large, like 10^{-4} then erroneous points are picked up, particularly in places whether the contour curves sharply. This effect occurs since to simulate sharp curves, there will be lots of very, very small line segments making up the curve. The way the algorithm works is that the first point that satisfies the condition is taken. And on the other end of the scale, if the threshold is too small, like 10^{-7} then sometimes there will be no match for the point.

So to try and combat this, a sliding threshold was implemented. If one of the coordinates (x,y) is accurate to a very tight threshold, then the threshold is relaxed for the other coordinate. However, this implementation ran into the same problems as above. The algorithm for creating the contour had to be fundamentally changed.

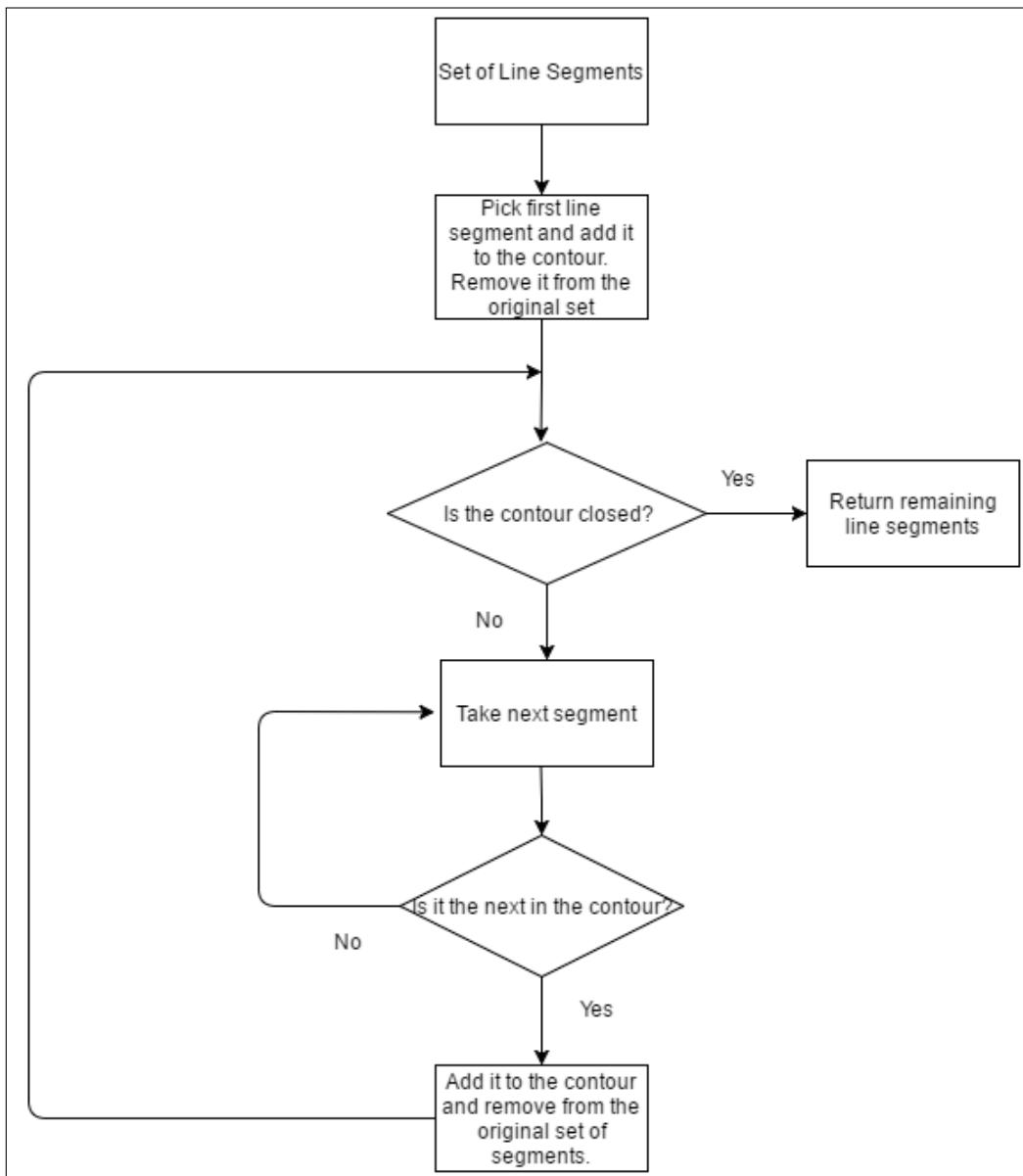


Figure 24.10: Old Contour Generation Algorithm.

24.4.4 The New Contour Creation Algorithm

To ensure accurate contour creation, the algorithm had to become far more selective. Instead of stopping when a point that falls within the threshold is found, we make it a minimization problem. The algorithm now searches for the point with the minimum distance from the 'head' point. This results in a much more accurate, but slower algorithm. We are now guaranteed $O(n^2)$ speed with the size of the line segment data set input. A flow chart of this algorithm is shown below in 24.11. The sacrifice in speed was absolutely necessary for the increased accuracy.

24.4.5 Classification of Contours

For purposes that will be discussed later in Section 24.5, we need to classify each contour as an inner or outer contour. The goal is to identify contours that sit fully within one another, an example of this is

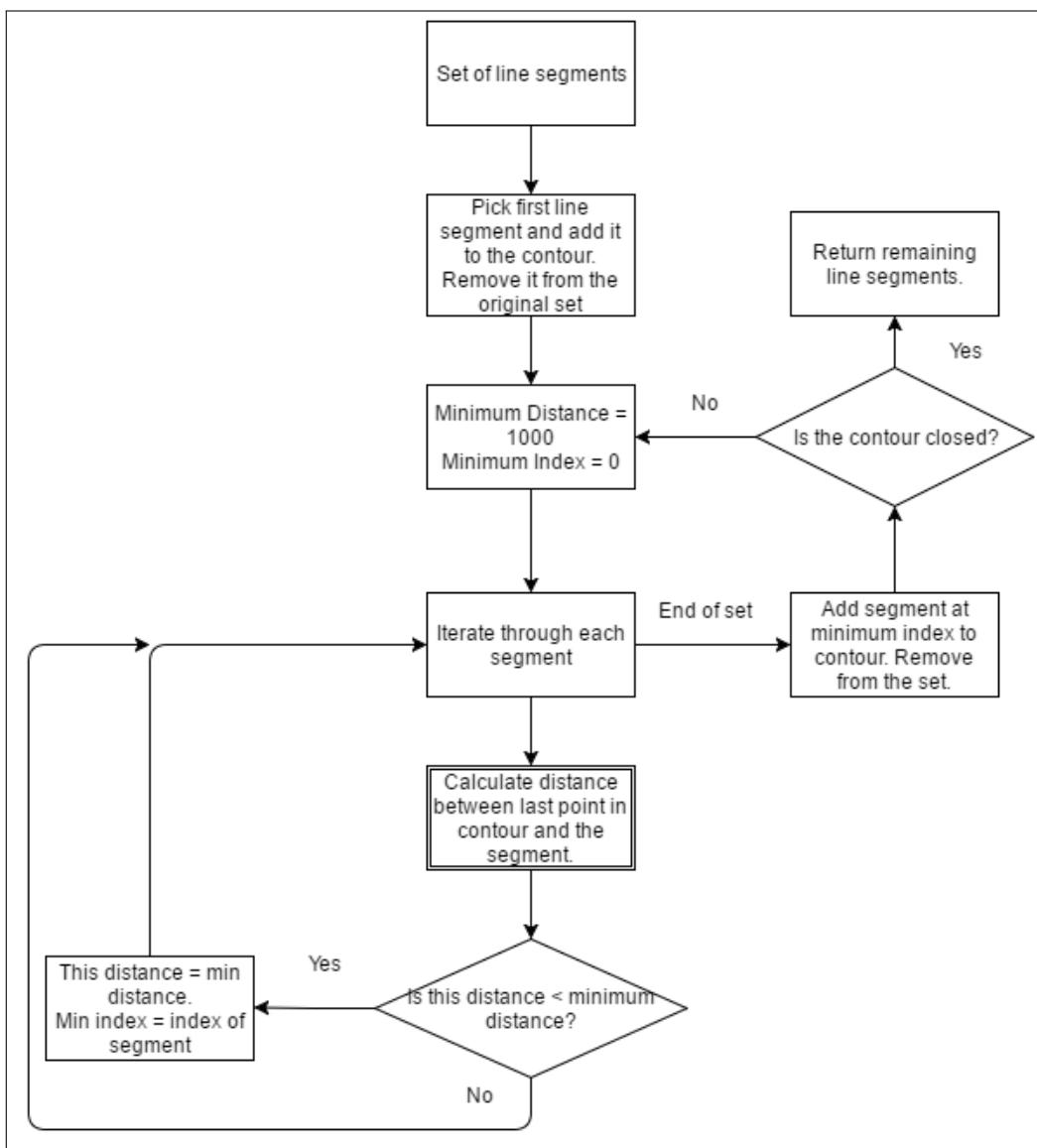


Figure 24.11: New Contour Generation Algorithm

shown in Figure 24.8. Two methods for classification and the rationales for the one implemented are discussed below.

First, a technique, that is common to both methods, for determining if a contour is definitely not inside another will be discussed here. If we consider each contour within a bounding box, we can quickly eliminate some contours from being within one another if the bounding boxes are not within one another or intersect. The contraposition of this statement is not guaranteed to be true. This is illustrated in Figure 24.12. This is a simple way of reducing computation effort, thus increasing performance.

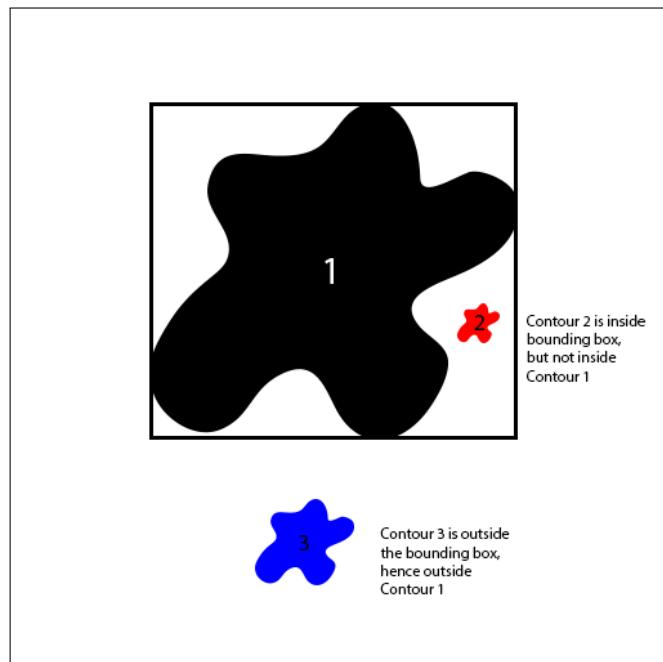


Figure 24.12: Application of bounding box exclusion.

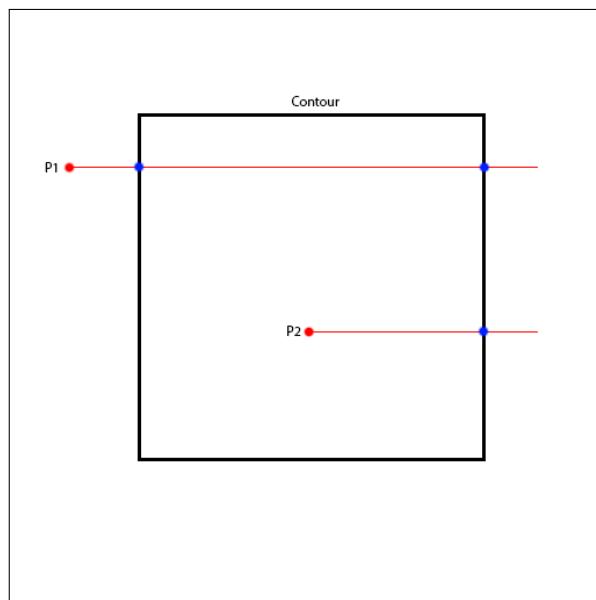


Figure 24.13: Example of Jordan Curve Theorem. P1 is outside the contour and intersects it twice. P2 is inside the contour and intersects it once.

The first method that will be discussed will be the application of the Jordan Curve Theorem with ray casting. The Jordan Curve Theorem states that a Jordan curve divides a plane into an exterior and interior region, and any continuous path that connects from one region to another must intersect the curve somewhere. From this we can extrapolate that if a point is within a curve, a ray casted in any direction will intersect the curve an odd number of times, whereas if a point is outside the curve, a ray casted will intersect the curve an even amount of times. This is shown in Figure 24.13. Using this principle we can determine if a point, and hence the contour that this point belongs to, falls within another contour. However, this method has certain challenges associated with ray casting. Due to the way of storing the points, we would need to check intersection of the ray with every single consecutive pair of points in the contour, this is very computationally expensive.

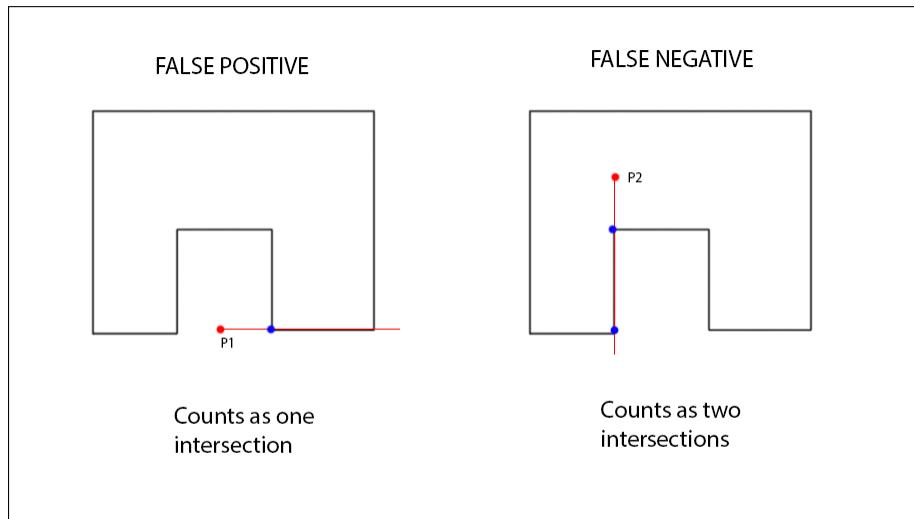


Figure 24.14: Example of problematic corner cases. P1 is outside but registers as inside. P2 is inside but registers as outside.

Also, corner cases are particularly difficult to resolve, what happens if the ray is collinear with the curve? If we count a collinear intersection as two intersections, then we may register false negatives, as shown in Figure 24.14. If we count a collinear intersection as one intersection, then we may register false positives as shown in Figure 24.14.

The second method relies on the principle of summation of angles of a polygon. If a point is on the inside of a closed curve (polygon), then the summation of all the angles created by the lines from the point to consecutive vertices is nonzero. However, if the point is on the outside of the curve, then the summation of all the angles will be zero! This effect is illustrated in Figure 24.15. We can exploit this to determine if a point, and thus a curve lies within another. Similar to the method making use of Jordan Curve Theorem, we need to calculate the angles of every single pair of points in the contour, this is computationally expensive. However, the main reason for the implementation of the angle method is the simplicity and elegance of the algorithm involved. The algorithm to implement

this is much simpler and quicker to implement than the Jordan Curve method.

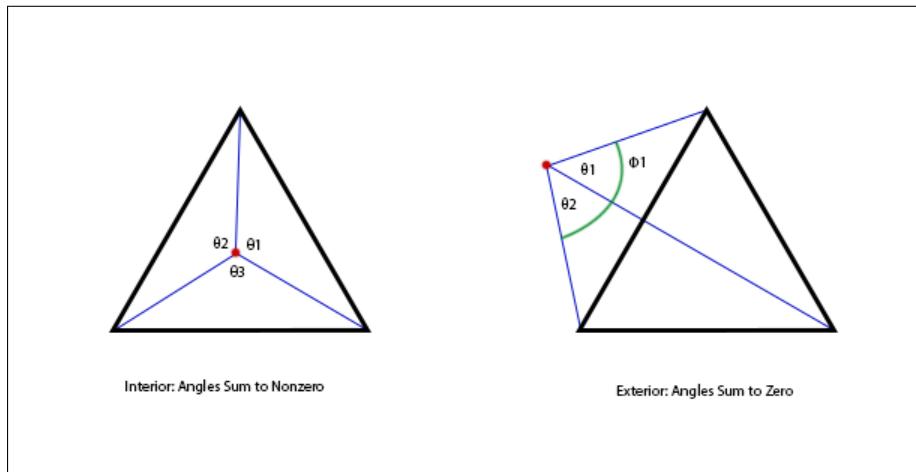


Figure 24.15: Example of summation of angles technique.

For each contour, we first assume that it is an outer contour. We then check every other contour to see if the first one was inside of them; if it is, we toggle the inner/outer state. This algorithm has performance $O(n^2)$ with the number of contours in a layer.

24.4.6 Results

An example output of the layer slicing component of the program is shown in Figure 24.16 and Figure 24.17. Two models were chosen for testing the layer slicing program, a test object file designed to test a 3D printer's capabilities and an R2-D2 model file. The test object file is made of 6746 triangles and the R2-D2 object is made of 101384 triangles. This allows us to test the scalability of our program as well as test performance as the size of the model increases.

A 3D view of the full sliced R2-D2 model is shown in Figure 24.19 to show the correctness of the layer slicing.

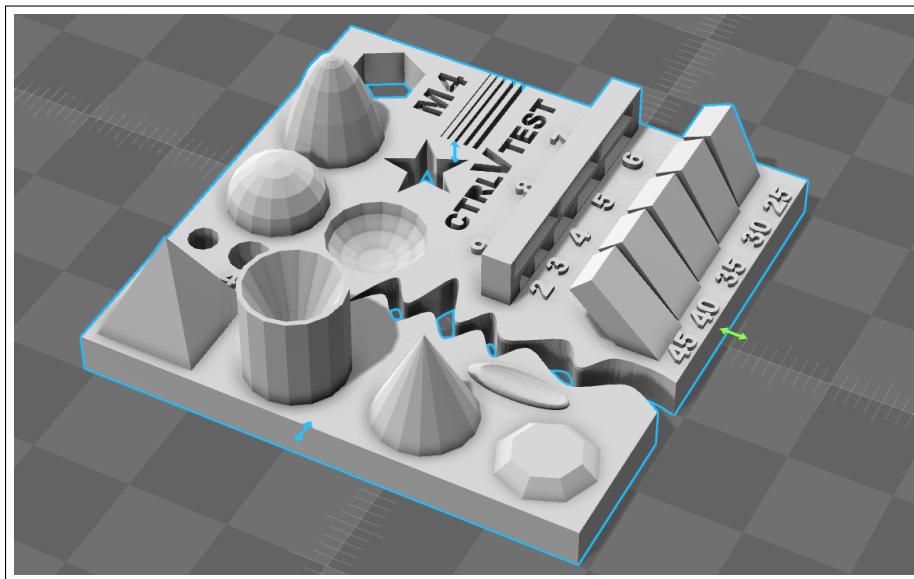


Figure 24.16: The 3D object, Test Object

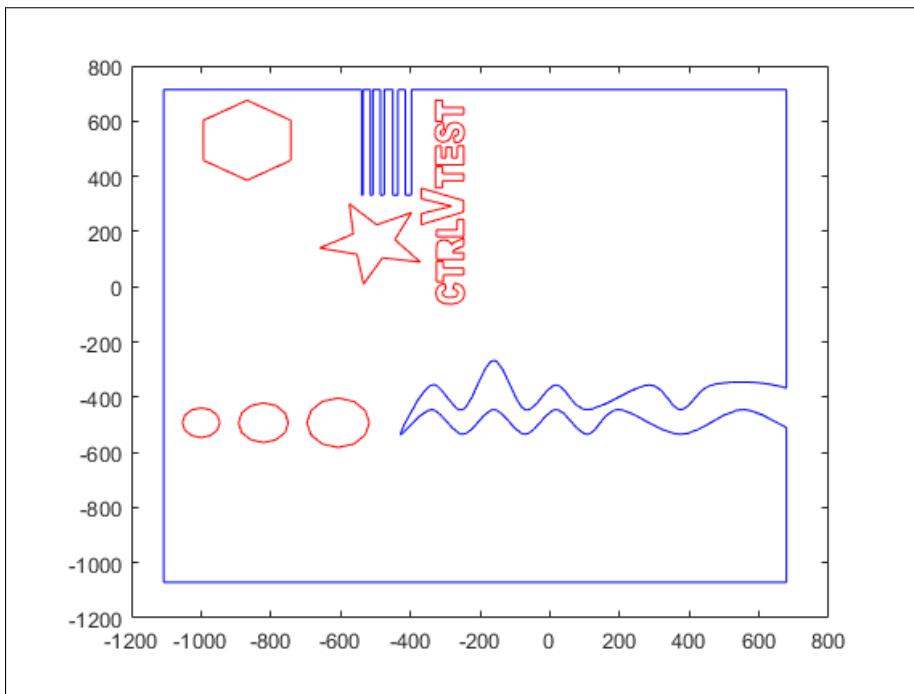


Figure 24.17: Example of layer slice of the Test Object. This is the first layer. Units are in micrometers

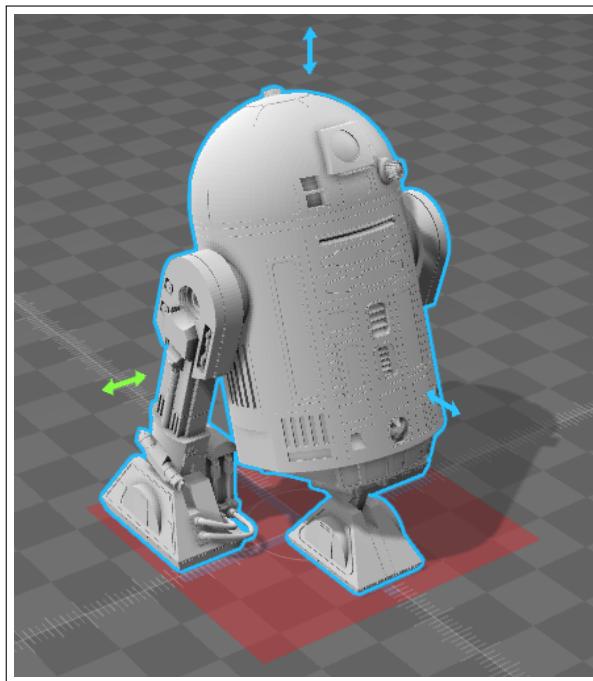


Figure 24.18: The 3D object: R2-D2

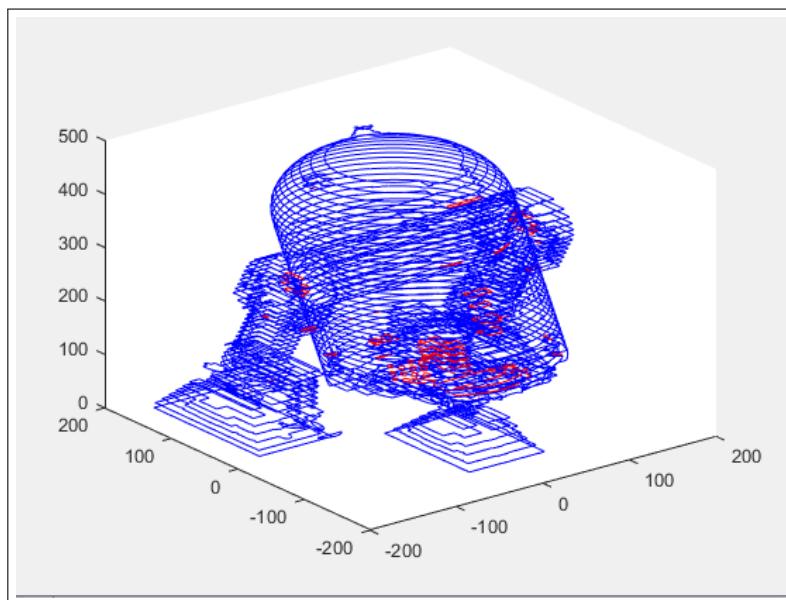


Figure 24.19: A 3D visual of the sliced R2-D2 (50 layers). Units are in micrometers

24.4.7 Performance

There are two parameters to test performance over: number of layers and size of the model. In light of this, a test was organized to test the running time of the layer slicing program over multiple models and number of layers. The results of the tests are shown in the table below. Note, that the tests were conducted on a machine with an Intel i5-6200U.

Number of Layers	Test Object(s)	R2-D2 Model (s)
25	6.57	18.07
50	12.81	36.77
75	19.15	53.28
100	25.72	70.78
125	32.61	87.85
150	38.6227	105.88

An interesting thing to note is that the average time per layer is relatively constant in a model, approximately 0.255 seconds for the Test Object and approximately 0.715 seconds for the R2-D2 model. This shows that the layer slicing algorithm is linear in time with the number of layers. Another interesting thing to note is that the number of triangles appears to have a smaller effect than predicted. The R2-D2 model has 15.03 times the number of triangles as the Test Object, but takes on average about 2.8 times as long. This would suggest a log relationship between time and number of triangles. In summary we have a layer slicing and organizing algorithm that is $O(n)$ with number of layers and $O(\log(n))$ with number of triangles (related to model complexity).

24.5 Resolution Separation

From here on out, the decision was made to work in Matlab due to the robust and powerful built in tools Matlab has for displaying and manipulating 2D structures. This is a compromise as Matlab is a much slower language than C++ is, however, the time saved writing and debugging similar functions that already exist in Matlab is well worth the compromise on speed.

24.5.1 Purpose of Resolution Separation

Our facility has the capability to fabricate at three different resolutions, a stereolithography process that fabricates at 12 micrometers, a two-photon polymerisation process that fabricates at 100 nanometers, and a STED lithography process that fabricates at 50 nanometers. What this allows us to do is to fabricate large macro structures using stereolithography and then fabricate micro features using two-photon polymerisation or STED lithography. To accommodate this desired feature, we need some methodology to determine which portions of the model to fabricate with each method. This is the idea of resolution separation.

The aim of this module of the program is to take the layer data and separate this into two distinct regions: high resolution and low resolution. This process will be very computationally expensive, although the Matlab library simplifies the code significantly.

24.5.2 Theory

There are a number of ways of doing this. The central idea to the two possibilities that are presented below is the idea that all the high resolution regions will be around the edges of the contour. Since the only time we can have fine resolution is at the interface between material and no material, we need only concern ourselves with this behavior. All the material in the interior can be treated as low resolution.

From this the problem emerges of defining an interior region that can be sufficiently described as low resolution. Two possibilities presented themselves as solutions: maximal area inscribed rectangles and polygonal offsetting.

24.5.3 Maximal Area Inscribed Rectangles

One method for selecting interior regions would be to find the maximum axis-aligned rectangle that can be inscribed within the contours. A rectangle can easily be described as low resolution, hence it satisfies this requirement. This could be iteratively done such that the size of the rectangle is at least some predefined minimum. This would result in a series of rectangles defining the low res, interior region and everything else would be the high res region. An example of this is shown in Figure 24.23.

However, calculating a maximal area inscribed rectangle in a n-sided, possibly concave with holes polygon is exceptionally computationally expensive and challenging. A heuristic algorithm can be employed, similar to that described at the d3plus blog [278], however, it is still prohibitively expensive over the long run to employ such a method.

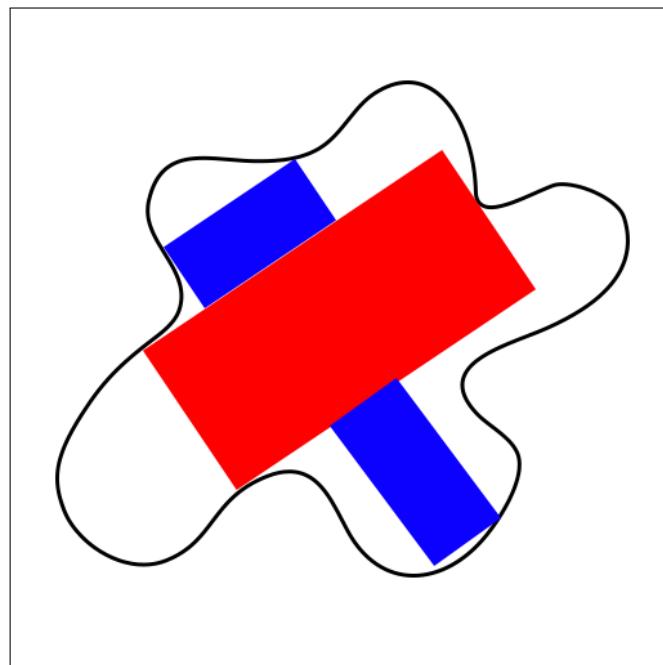


Figure 24.20: An example of the maximal rectangle method for resolution separation.

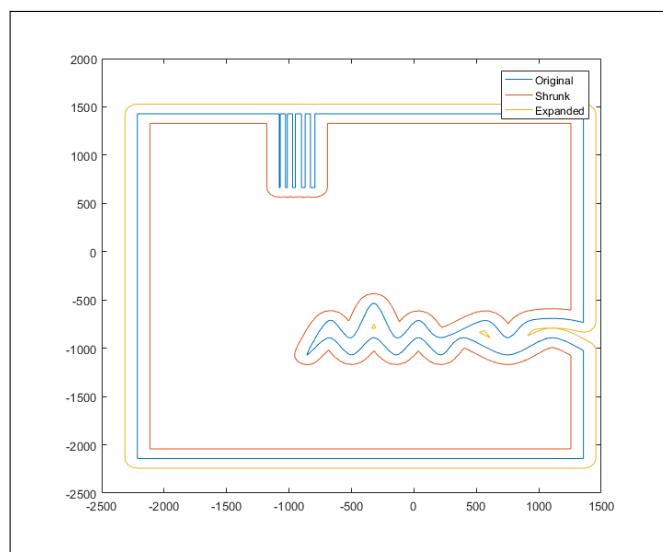


Figure 24.21: This is an example of the both shrinking and expanding a contour, for use in resolution separation. Units are in micrometers.

24.5.4 Polygonal Offsetting

Another method for selecting interior regions is to shrink the outer contours down by a set amount, the area between the original contour and the new shrunk contour is high resolution area and the area in the shrunk contour is low resolution. A similar thing can be done with inner contours, except we expand the contour out. An example of both of these is shown in Figure 24.21. This has a two fold effect, it separates the high resolution edges from the low resolution interior while the shrinking/-expanding action also smooths the interface between the high resolution and low resolution portions. This means that errors at the interface are less likely.

The principle behind shrinking/expanding contours is fairly simple in concept, difficult in implementation. To shrink a contour, each edge must be moved perpendicular to itself by the defined amount, then intersections of the new edges must be resolved (for shrinking) or creation of new vertices to close the loop (expansion). The latter operation is the real complexity in the process.

Thankfully, a Matlab library exists for exactly this purpose, `bufferm2` which can be found on the Mathworks File Exchange [279]. This library can both shrink or expand a contour by a set amount.

24.5.5 Algorithm Implementation

To implement polygonal offsetting, we merely need to iterate through each contour in a layer, either expanding or shrinking as needed. There is a small condition though, if the contours are smaller than a defined amount (in this case five times the linewidth of the low resolution process), then we call the contour high resolution and skip the shrinking/expansion. The algorithm flowchart is shown in Figure 24.22. The result is two sets of contours defining the high resolution regions and low resolution regions.

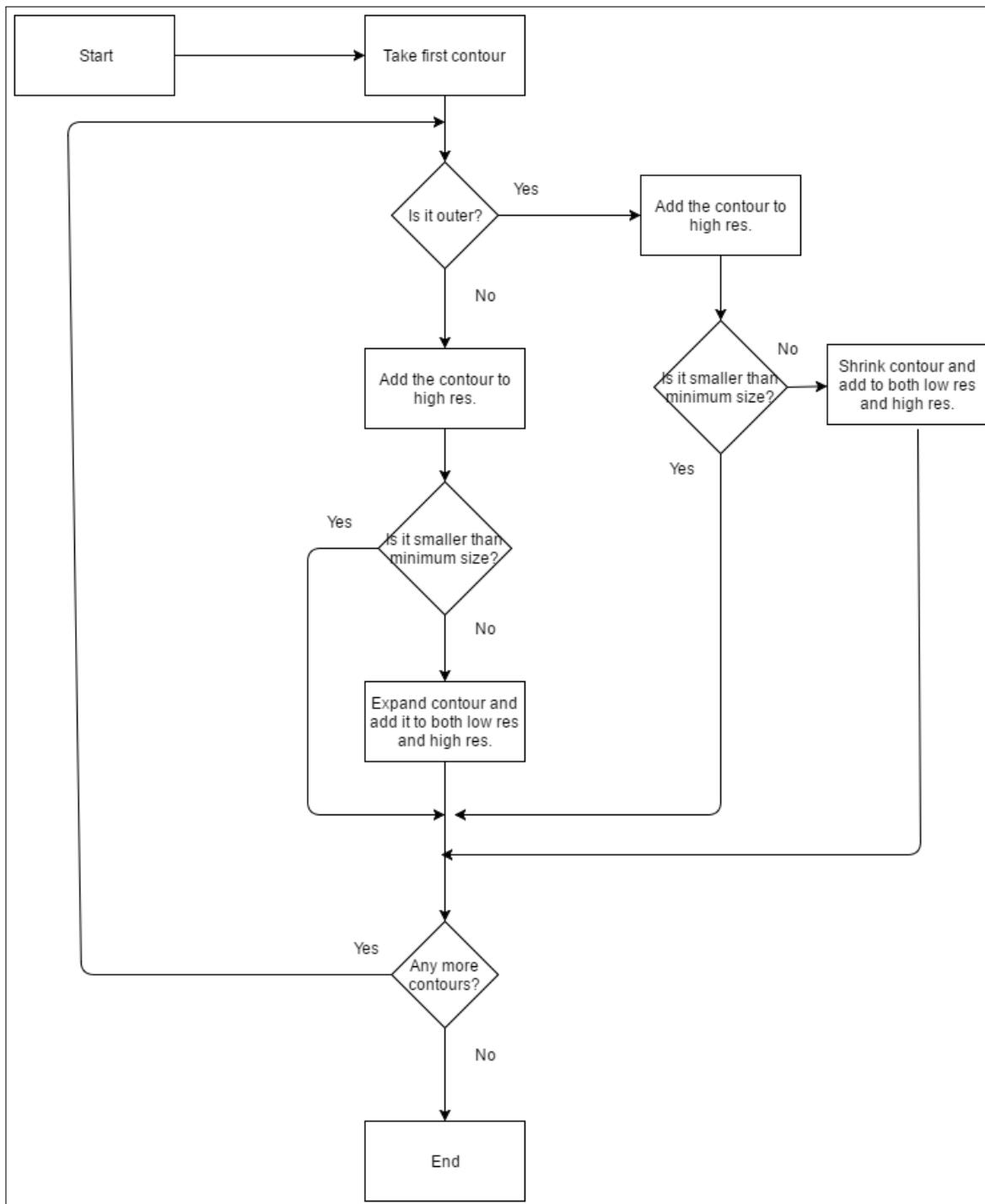


Figure 24.22: Resolution Separation algorithm flowchart.

24.5.6 Reorganising Data

At this point, the data structure for the layer is getting somewhat complex. There are separate structures for low resolution and high resolution with any number of contours in each. To condense this data before tool path generation, we can take advantage of Matlab's built in polygon operations, `polybool`, to reduce the data structure complexity to merely two contours, one for high resolution and one for low resolution.

We need to add all of the outer contours while subtracting the inner contours, however, nested

contours prove a problem if we just blindly do this. Depending on the order of operations we may lose data. To overcome this problem, we can be more careful with how we do the operations. We can first subtract all of the inner contours that are of a subset of each outer contour, then union all of the resulting contours. This preserves all of the relevant data.

24.5.7 Results and Performance

Some example output of the resolution separation data is shown in Figure 24.23. Although the method is not as clean as maximal area rectangles would theoretically be, it is sufficient for our purposes and is much faster.

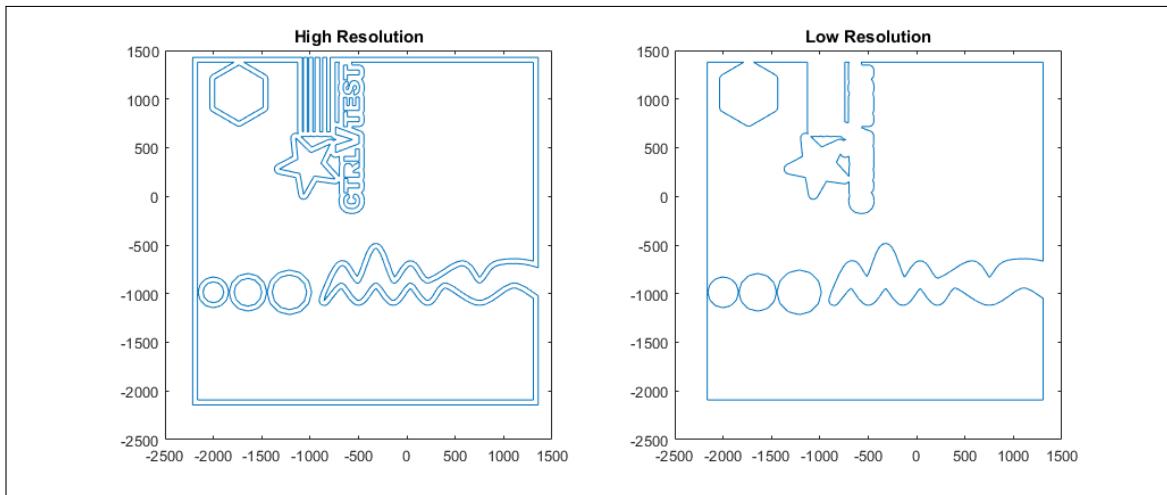


Figure 24.23: This is an example of the output of the resolution separation algorithm.

Tests were ran with Test Object (as number of triangles should have no impact at this stage) for dependency of this algorithm on number of layers and number of contours in each layer. The results of these tests are shown below in the table:

Number of Layers	Total Time (s)	Number of Contours
25	134.9	404
50	312.95	768
75	493.9	1099
100	655	1491

With this data we can see a linear relationship between the time taken and the number of contours and hence the number of layers. This is the expected performance of the algorithm since we loop over each layer once, hence, $O(n)$.

24.6 Tool Path Generation

The final step in the software is to generate tool paths for each layer. A tool path can be defined as the path the laser must travel on the surface of the material in order to create the desired layer.

24.6.1 Theory

At this point, we have prepared the data so that this step is as simple as possible; for both high res and low res we have one contour each defining the boundaries. We simply need to now convert these contours into tool paths. The simplest way of doing this is determining the intersection points between the contours and a set of lines that define the movement of the laser, interpolating between the intersection points gives the appropriate tool paths. This is demonstrated in Figure 24.24. We expect there to be an even number of intersection points due to the Jordan Curve Theorem as discussed in Section 24.4.5, this makes it trivial to interpolate between consecutive points.

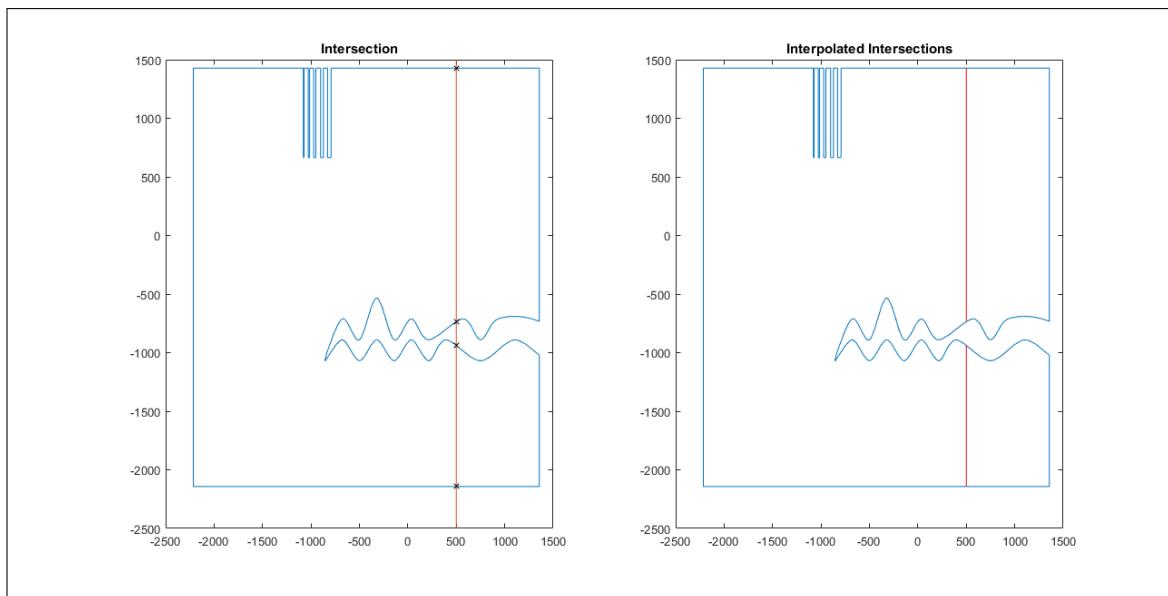


Figure 24.24: This is an example of how intersection points are interpolated. Units are in micrometers

24.6.2 Set of Intersecting Lines

We must first define a set of lines to test for intersection against. We know for example that the linewidth for the high res (two-photon polymerisation) is 100 nanometers and for the low res (stereolithography) is 12 micrometers. It follows that the set of lines should be parallel separated by the respective line width.

We want the tool paths to be able to be at any angle that the user defines. To do this we must come up with a generalized system of determining the equations of these lines. First, the possible area that the lines must cover can be determined from the bounding box of each layer. The maximum area that would need to be covered is given by the diagonal of the bounding box, since this is the maximum distance between two points on a rectangle. This gives us an easy way of determining the number of lines in the set. In the following equation n is defined as the number of lines, l_1 and l_2 are the lengths of the bounding box and w is the linewidth.

$$n = \frac{\sqrt{(l_1^2 + l_2^2)}}{w} \quad (24.5)$$

To generate each successive line, we merely need to move a linewidth's distance perpendicular to the direction of the line.

$$\Delta x = w \cdot \sin(\theta + 90) \quad (24.6)$$

$$\Delta y = w \cdot \cos(\theta + 90) \quad (24.7)$$

Now the final step is to generate the first line upon which we iterate. We want the line to be large enough such that as we move across the layer we stay fully intersected with the bounding box. First, if the angle is positive we want to start on the top left corner of the bounding box, whereas if the angle is negative we want to start on the bottom left of the bounding box.

To ensure the length of the line is appropriate, we know the maximum distance we can move in the direction of either x or y is given by (Given for x, for y replace sin with cos):

$$n \cdot \Delta x = \sin(\theta + 90) \cdot \sqrt{(l_1^2 + l_2^2)} \quad (24.8)$$

So if we extend the lines out by this distance we guarantee that we will always intersect with the layer as we move across this.

24.6.3 Algorithm Implementation

Matlab quite nicely provides a very simple to use method for determining intersections between polygons and lines, `polyxpoly`. This means the algorithm for tool paths becomes very simple, we merely iterate through each layer, generate a set of lines from both high res and low res and solving intersections for each respective contour. A flowchart representation of this algorithm is shown below in Figure 24.25.

24.6.4 Results and Performance

Some example output of the tool path generation algorithm is shown below in Figure 24.26. Note that high resolution data is shown in blue and low resolution is shown in red.

In analysing the performance of this algorithm we are mainly concerned with two things: how the algorithm scales with number of layers and how it scales with the overall size of the model (impacting number of lines). However, these are intrinsically related, as the number of layers is directly related to the size of the model. Hence, we can only test against number of layers. Again, the Test Object

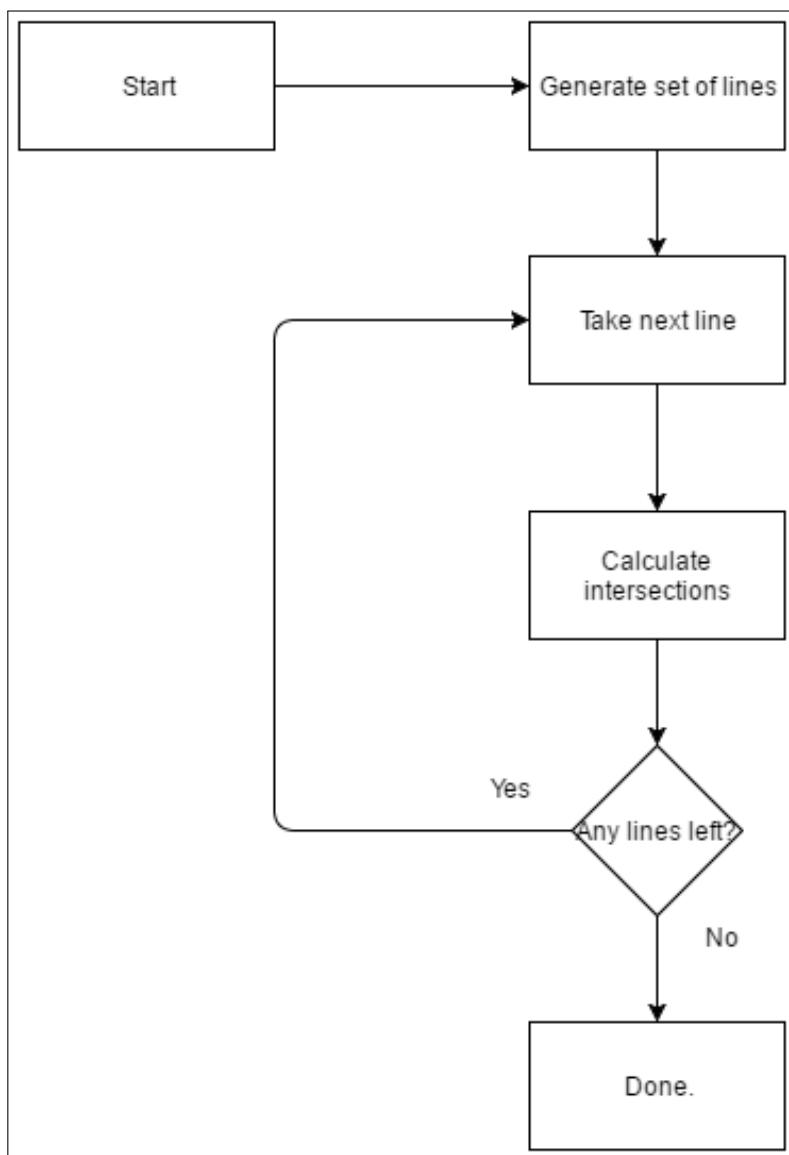


Figure 24.25: Flowchart for the tool path generation algorithm.

was used in this testing with the results shown in the table below:

Number of Layers	Time Taken (s)	Average Time/Layer (s)
25	604.9	24.19
50	3149.5	62.9
75	6889.0	91.9
100	12584	125.8

We can see that the total time taken increases significantly with the increase in the number of layers. However, there is a nonlinear effect due to the inherent relationship between the size of the model and the number of the layers. Since the size of the model is linked to the number of layers by the fixed layer height, as the number of layers grows, so does the size of the model. This means that there is a larger surface area to cover at each layer, hence, more tool paths to be generated. This is

the causes of the nonlinearity of the performance.

The lengthy times should also be taken with a grain of salt as the tool path generation was performed on a machine with an Intel i5-6200U processor, it is very likely that a machine running this software would have a much more powerful CPU to take advantage of. Furthermore, since the code is merely looping over every layer, performance can be further improved by taking advantage of threading or GPU acceleration to parallelize the process.

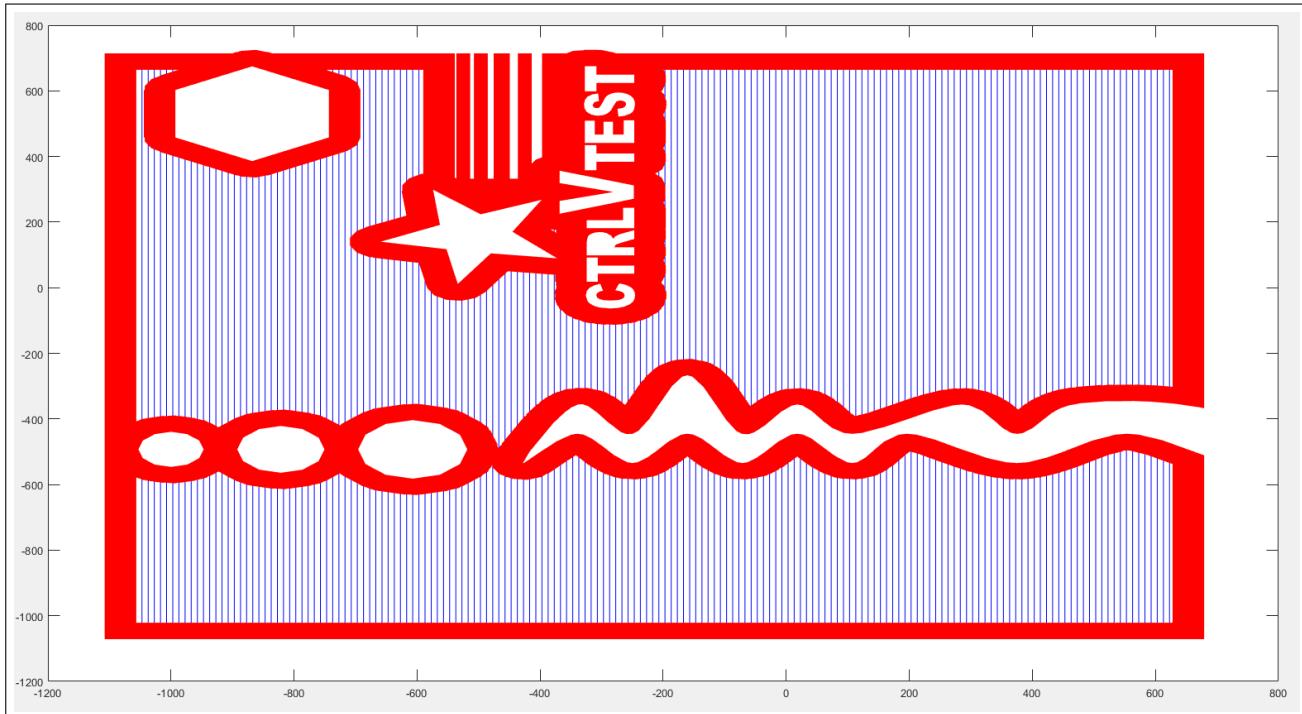


Figure 24.26: Example output of tool path algorithm. Note that the red area appears to be solid because the high resolution has a much smaller line width and hence larger density.

24.7 Conclusion

In this last chapter, a robust software suite has been described for use with our system. Its primary function is to convert 3D models into actionable tool paths for input to our control system. However, it also provides the user with a easy-to-use and powerful GUI for ease of use. The implementation of each section of the software has been described in detail and the performance quantified. The software really brings the multiple fabrication methods: stereolithography, two-photon polymerization, STED lithography, and laser micro-machining together into a cohesive package such that a user can take advantage of everything there is to offer in our system.

Part VII

Conclusion

25 Applications

The applications of the design facility are manifold. The application of 3D printing as such is growing at a phenomenal rate and our facility really advances the breadth of applications by being at the cutting edge. The facility can be used to develop structures as developed as shark skin, add lotus effect in materials etc. At Purdue University, a structure inspired by a crab claw has been 3D printed. Using piezoelectric materials in the bone, force can be produced to then enable the hand to grab a ball, for example, as shown in Figure 25.1. Using higher resolution fabrication even the most minuscule features can be brought in. Units produced by our fabrication system can be widely used

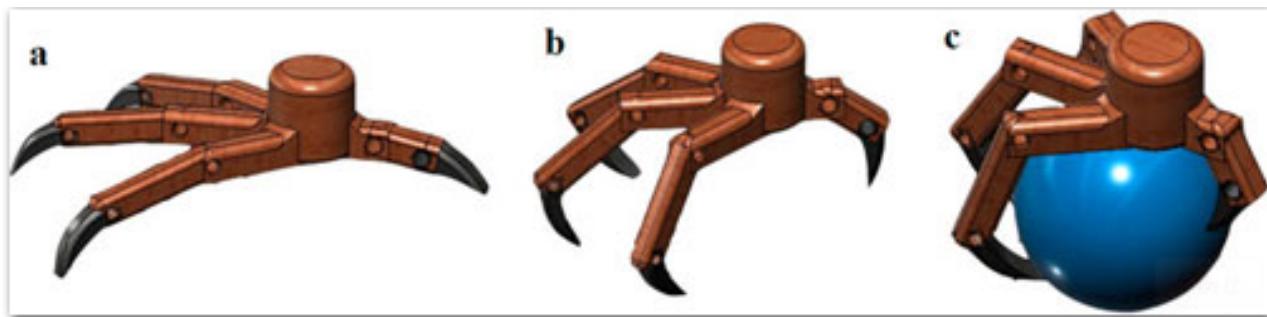


Figure 25.1: Design inspired by crab claw for piezoelectric structures [280]

in biology, for example, to estimate aberrations in human eyesight with high accuracy and fast time response, to visualize transparent biological specimens etc. [166] A key application with much scope in the future would be lab on a chip which is the study of biological samples.

Some industries which would benefit much from the capabilities of our system would be health-care, semiconductor electronics and automotive which utilise MEMs – micro electro mechanical systems.

MEMs came up as a type of technology of SLMs that could have been used in the fabrication set up, however, we could even fabricate MEMs from this system at a high level of accuracy. The use of MEMs is widespread, in insulin pumps, hearing aids, safety airbags systems in automobiles etc.

The ability to manufacture on a wide range of materials like polymers, glass, metals and ceramics lend credence to the wide applicability of the fabrication facility. These applications range from biophotonics, optical materials, integrated optics, functional nanomaterial surfaces and MEMs.

26 Business Case

The aim of the design project was to design an optical nano-fabrication facility that can be used for a variety of applications. Industries that would benefit include automotive, aerospace, healthcare and dental industry, and further many academic institutions could utilise for research purposes. Market research carried out by decision databases [241] reveals that application in healthcare & dental industry is expected to grow at the highest CAGR (compound annual growth rate) over the forecast period of 2011 and 2016.

Being able to produce an object at nano scale with a huge variety of materials with utmost care to accuracy poses its own challenges. This requires the use of cutting edge technology that the design project has used so as to push the limits of current technology. The business case put forward looks at the unique selling points of our facility that sets it apart, analyses the market through the lens of Michael Porter's five forces and then finally costs the product.

26.1 Unique selling points

A combination of 4 fabrication techniques: laser micro machining, stereolithography, two photon polymerisation and STED lithography, and 4 imaging techniques: confocal, bright field, 2 photon and STED microscopy, has been used. Each of these have their own advantages and extend the ability of our system, integrated together.

Laser micro machining, gives the added capability of reductive manufacturing that may be a more efficient way of fabrication depending on application. It has a much larger range of compatible materials than other fabrication processes and further the SLM adds the ability to create unique beam shapes for all sorts of extensions like multiple foci generation and bessel beams.

Stereolithography is the most traditional way of 3D manufacturing as it quite simple to set up. Including this technique in our design enables large building area and high fabrication speed. Two photon polymerization is a high resolution (100nm) version of stereolithography that enables much more detailed fabrication system. Having live feedback loop control lends further accuracy to the system which is crucial at nano scale manufacturing.

Having optical tweezing in our system adds the unique ability to interact and exert controlled forces on objects. This can be used for assembly in a variety of cases when there are several parts that need to be put together to build one structure. It is also useful to power micromachinery, sorting of cells and microfluidic control. Holographic optical tweezing (HOT) traps can be created to move and rotate several larger objects together.

Various imaging techniques employed enable the system to image in 3D, at high resolution and even real time. Bright field microscopy which has a simple set up can be used to image lower

resolution fabrication. Using polarized light increases the contrast of a transparent sample by a simple modification. Confocal microscopy increases the resolution and enable 3D image formation of both transparent and opaque material. Two photon microscopy the increases the resolution even further to open up the frontier of in-depth structure detection, while being just a reutilisation of the already existing two photon polymerization set up.

Really then, what pushes the limits in resolution of imaging is STED, which gives fabrication and imaging resolution at 50 nanometers. With only one STED imaging facility by Leica and no fabrication facilities incorporating STED in the market, it really sets our product apart for reaching such high resolutions. Being such a cutting edge technology adds to the longevity in the relevance of our system in the market.

The intelligent software allows the user to combine high and low resolution fabrication leading to optimised manufacture. As one of the key features of our facility is the ability to build one structure using different manufacturing and imaging techniques to give the most optimised manufacture. The ability to identify from the provided CAD, the best fabrication technique for the system further reduces the time taken to fabricate a unit. Further, the user can view the build plans and adjust parameters to their desire.

Striving for accuracy, the system includes set up for aberration correction through adaptive optics. Further, the speed of fabrication is increase by the ability to produce multiple foci and have simultaneous manufacture. At each step of the design process accuracy and response time were taken as key measures in making a decision. Using this combination of fabrication and imaging techniques increases the range of materials that can be employed by the facility, hence increasing the breadth of applications manifold.

An automated cassette system integrates the various manufacturing techniques, allowing for a single structure to be built by various techniques. The power of the system design lies in the ability to combine all these different processes and integrate it into a system that functions as a whole. Using the most efficient manufacturing technique, rather even a combination of them and imaging in real time and high resolution, sets the design apart from any competitors.

The range of materials that can be used is another unique selling point of the product borosilicate glass, silicon and metals, photoresist material like SU-8, and many more can be manufactured by the system.

26.2 Porter's Model

A strategic marketing plan would be needed to build partnerships with major research facilities and industry players that would benefit from our product. The capital investment is quite high but relatively

fair in terms of investments in similar levels of research. Further, formally doing an analysis based on porter's model affirms the positive business prospects of our fabrication facility as depicted in Figure 26.1. The five forces to consider would be bargaining power of buyers and suppliers, risk of substitutes and new entrants and internal jockeying of power in competitors. [281] The bargaining power of suppliers is noticeably key as many components of our system has off the shelf components and they are quite expensive. Hence, forming partnerships would be pivotal for keeping the process of manufacture smooth and efficient. There isn't a risk of substitutes for many years as the technology is quite far from maturity. Further the costs involved in substitution or entering the market are quite substantial. With a low number of companies present in the market, the competition isn't fierce but is rather built through reputation and relationships. We have 5 different kinds of manufacturing techniques combined with real time imaging and adaptive optics. Our product will not be competing on price, but rather on accuracy and breadth of applications.



Figure 26.1: Michael Porter's five forces analysis of industry

26.3 Cost

The overall cost of the design project would be just over £1.2 million which includes research and development costs, costs of testing the facility after building, to optimise the system and ensure all safety features. It would seem viable then to sell the product at £2 million.

To put things in perspective, a STED microscope in the market is priced at £640k, which is just

26.3. COST

one of the techniques of our system. Nanoscribe, our nearest competitor currently prices a facility called 'Photonic Professional GT' at £1.7 million which is only capable of manufacture through photo polymerisation and minimum thickness of 0.001 mm while providing the similar ability of real time imaging. Further, the majority of the cost of our system comes from the lasers bought off the shelf, adding up to £540k.

27 Conclusion

The design of an integrated facility has been discussed in the report, that has the ability to fabricate structures of sizes till 8x8x8cm with details down to 50nm. The project aimed to build a fabrication facility that fabricates structures at nano-scale using the properties of light. These include, using photo-initiated chemical reactions to polymerise materials, like in 2 photon polymerisation and stereo-lithography. Using high intensities of laser, the large blocks of material can be used in a reductive fashion to create structures as in laser micromachining. Optical tweezing, further uses the ability of light to exert force that although small, can move objects at nano scale. Further, the various imaging techniques enable real-time and high resolution imaging.

This involved, looking at the various fabrication techniques and making decisions to include techniques that helped us achieve speed, accuracy and scale in the structures that would be built using the facility. These decisions were based on the design specifications and the core capabilities that a technique added. Four fabrication techniques have been incorporated in our system: laser micromachining, 2 photon polymerization, stereo lithography and STED. They have further been integrated all together, taking advantage of each to optimise the system as a whole. Correspondingly, several imaging techniques are employed to ensure that the fabrication has been done as per the design required. There is a combination of confocal, 2 photon, real time and high resolution imaging to account for a variety of applications.

The optical system has been set up to have complete control of the light, tracking the intensities and pulse widths along the system. All optical components were also chosen to minimize aberrations and provide flexibility to the entire system. Further the Pulse shaper and spatial light modulator gives added control on the phase and amplitude of the light and allows aberration correction. Design of stage and galvanometers allows accurate control over the movement of light and the object being fabricated. Further, having developed an intelligent software optimises the fabrication process, and automates it just on the assignment of a CAD file. Adaptive optics has allowed for increased accuracy in the system and the challenge of speed is tackled by employing the right techniques suitable for the application through the software and parallelisation through the spatial light modulator.

The design project has hence achieved the design specifications set out. Each of the techniques have been integrated in one housing facility, making the system capable of accurate fabrication and imaging at higher speed. The facility developed is unparalleled in terms of the variety of techniques included and the breadth of applications that can be achieved at nano-scale.

Appendices

A Optical Component List

Ref No.	Component Description	Supplier	Component No.
LASERS			
LAS:TI:S	Pulsed 800 nm Ti:sapphire	Coherent, Inc.	Libra, Vitara-T, REVOLUTION-30 (10 kHz)
MECHANICAL SHUTTERS			
	1" Low Speed Optical Shutter		SH1/M
S1-5	Shutter Controller (x5)	Thorlabs	TSC001
	15 V PSU Shutter Controller		TPS008
BEAM SAMPLERS			
BS1,2	650 – 1050 nm AR Coating BS	Thorlabs	BSF-20B
BS3,4	UV Fused Silica Beam Samplers		BSF-20UV
MIRRORS			
M1-7	Ultrafast Low GVD, 700 – 930 nm	Newport Corp.	10B20UF.25
M8	2" Broadband Dielectric Mirror	Thorlabs	BB2-E02
RDM1,2	650 nm Longpass Dichroic Mirror	Edmund Optics	64-702
RDM3	525 nm Longpass Dichroic Mirror		64-634
FLPM1,2	Owis Flip Mirrors	Owis	46.060.38KN
VARIABLE ATTENUATION			
AVAT1	Automated VA Toolkit (Pulsed)	Newport	M-CCVA-TL-KT
AVAT2	Automated VA Toolkit (CW)		M-CCVA-L-KT
SP,CC3	CCD Spectrometer		CCS-200/M
	CCS Cosine Corrector for Fibre	Thorlabs	CCSB1
BD3	CW Beam Dump for 200 – 410 nm		BT600/M
PBS3	UV Polarizing Beam Splitter	Newport	05SC16PC.24
AHWP3	Mounted Rotating Half-Wave Plate	Thorlabs	AHWP05M-340

LED ILLUMINATION

RLED	625 nm (Red) FC LED, 14 mW		M625F2
	4 Channel LED Driver Setup	Thorlabs	DC4100
FCL1,2	Fibre-Collimating Lens, $f = 35.41\text{mm}$		F810SMA-635

LENSES. AR Coating: (UV) - 245 – 410 nm, (A) - 350 – 700 nm, (B) - 650 – 1050 nm

L1	N-BK7 Concave Lens (B), $f = 30\text{mm}$		LC1060-B-ML
L2-6,L8-11	Air-Spaced Doublet (B), $f = 200\text{ mm}$		ACA-254-200-B
L7	Cemented Doublet (B), $f = 500\text{mm}$		AC254-500-B-ML
L12	Air-Spaced Doublet, $f = 150\text{mm}$		ACA254-150-B
L13	Air-Spaced Doublet, $f = 750\text{mm}$	Thorlabs	AC508-750-B-ML
L14-16	Air-Spaced Doublet (A), $f = 30\text{mm}$		254-030-A
L17	N-BK7 Concave Lens (A), $f = 50\text{mm}$		LC1715-A-ML
L18	N-BK7 Convex Lens (A), $f = 200\text{mm}$		LA1708-A-ML
L19	UVFS Concave Lens (UV), $f = 50\text{mm}$		LC4357-UV
L20	UVFS Convex Lens (UV), $f = 200\text{mm}$		LA4102-UV-ML

OPTICAL MODULATORS

AOM	Acousto-Optic-200 MHz, 780 – 850 nm	Gooch&Housego	AOMO 3200-1113
SLM1	CRi Transmissive 1-D Spatial Light Modulator	Meadowlark Optics	CRI-SLM-640
SLM2	Pluto Reflective Spatial Light Modulator	Holoeye	PLUTO-NIR-015 (HR)

OBJECTIVES

OBJ1	Universal Oil	Olympus	MPLAPON 100XO
OBJ2	Universal Oil	Olympus	UPLSAPO 20XO
OBJ3	Universal Dry	Olympus	LUCPLFLN 40X
OBJ4	Polarized Light	Olympus	UPLFLN 100XOP
OBJ5	High Dry	Olympus	LCPLN 100XIR
OBJ6	Multiphoton Oil	Olympus	XSLPLN25XGMP
F-THL	F- θ Lens for Stereolithography	couriertronics	

IMAGING

CMOS	CMOS Camera for Live Imaging	pco.edge	4.2 cameralink
PMT1,2	Bialkali Amplified PMT, 280 – 630 nm	Thorlabs	PMM01
CCD	Color CCD Camera, Res: 1280×1024	Thorlabs	DCU224C

WAVEPLATES, POLARIZERS

HWP1	Zero-Order Half-Wave Plate, 808 nm		WPH05M-808
POL1	Mounted Circular Polarizer, 780 nm		VBC05-780/M
POL2	Mounted Circular Polarizer, 633 nm	Thorlabs	VC5-633/M
RQWP1	Zero-Order Quarter-Wave Plate, 808 nm		WPQ05M-808
RQWP2	Zero-Order Quarter-Wave Plate, 532 nm		WPQ05M-532
PRP1	Vortex Phase Plate	Holoor	V-L-209-M-Y-A
PRP2			V-L-209-Q-Y-A
AN1,2	Optical Spectrum Analyzer, 350 – 1100 nm	Thorlabs	OSA201

SENSORS

SHWS	Shack-Hartmann Wavefront Sensor	Thorlabs	WFS20-7AR
FROG	Frequency-Resolved Optical Gating Kit	Swamp Optics	8-50-USB
ETA	Glass-Based Fabry-Perot Etalon (20 μm)	SLS Optics	Custom-Made
PD3	Thermal Power Detector, 10 μW Resolution	Thorlabs	S302C

PULSE SHAPING EQUIPMENT

DIFG1,2	Reflective Diffraction Grating	Newport	10HG2000-475-1
SPM1,2	Spherical Mirrors, $f = 457.2$ mm	Edmund Optics	NT32-831
FM1,2	Folding Mirror, $\lambda/10$ flatness	Thorlabs	PDF10-03-M01

Bibliography

- [1] Tommaso Baldacchini. *Three-Dimensional Microfabrication Using Two-Photon Polymerization*. Elsevier, 2015.
- [2] J.W. Goodman. *Introduction to Fourier Optics*. 2nd. MaGraw-Hill, 1996.
- [3] Tony Wilson and Stephen Morris. *C14 Applied Optics*. 2015.
- [4] Rudi Rottenfusser, Erin Wilson, and Michael Davidson. *ZEISS Microscopy Online Campus / The Point Spread Function*. URL: <http://zeiss-campus.magnet.fsu.edu/articles/basics/psf.html> (visited on 04/12/2016).
- [5] *Super-Resolution Microscopy Tutorial*. URL: <http://advanced-microscopy.utah.edu/education/super-res/> (visited on 02/18/2016).
- [6] R. Paschotta. *Gaussian Beam / The Encyclopedia of Laser Physics and Technology*. 2008. URL: https://www.rp-photonics.com/gaussian_beams.html (visited on 03/09/2016).
- [7] 2016. URL: <http://bnonlinear.com/>.
- [8] Ruediger Grunwald and Martin Bock. "Programmable micro-optics for ultra short pulses". In: *SPIE* (2010).
- [9] Lonnie Lucas and Jim Zhang. *Femtosecond laser micromachining: A back-to-basics primer*. URL: <http://www.industrial-lasers.com/articles/print/volume-27/issue-04/features/fermosecond-laser-micromachining-a-back-to-basics-primer.html>.
- [10] B. C. Stuart et al. *Optical ablation by high-power short-pulse lasers*. Tech. rep. California: Lawrence Livermore National Laboratory. URL: <http://www.mfeit.net/physics/papers/josab1396.pdf>.
- [11] Chris B. Schaffer, André Brodeur, and Eric Mazur. *Laser-induced breakdown and damage in bulk transparent materials induced by tightly focused femtosecond laser pulses*. Tech. rep. Cambridge: Applied Sciences, Cambridge and Department of Physics, Harvard University. URL: http://mazur.harvard.edu/publications/Pub_282.pdf.
- [12] Shane M. Eaton, Giulio Cerullo, and Roberto Osellame. *Fundamentals of Femtosecond Laser Modification of Bulk Dielectrics*. Tech. rep. URL: http://www.springer.com/cda/content/document/cda_downloaddocument/9783642233654-c1.pdf?SGWID=0-0-45-1315251-p174192069.

BIBLIOGRAPHY

- [13] E. G. Gamaly et al. *Ablation of solids by femtosecond lasers: ablation mechanism and ablation thresholds for metals and dielectrics*. Tech. rep. Australian National University and P.N. Lebedev Physical Institute, Moscow, Russia. URL: <http://arxiv.org/ftp/physics/papers/0102/0102046.pdf>.
- [14] Susanne Pfalzner. *An Introduction to Inertial Confinement Fusion*. CRC Press, 2006.
- [15] URL: <https://www.miniphysics.com/number-density-of-free-electrons.html>.
- [16] W. Perrie et al. *Femtosecond laser micro-structuring of aluminium under helium*. Tech. rep. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.605.6074&rep=rep1&type=pdf>.
- [17] Chris B. Schaffer et al. *Micromachining bulk glass using femtosecond laser pulses with nano-joule energy*. Tech. rep. Harvard and Cambridge University. URL: http://mazur.harvard.edu/publications/Pub_264.pdf.
- [18] Dmitriy A. Dvoretskiy et al. *High-energy, sub-100 fs, all-fiber stretched-pulse mode-locked Er-doped ring laser with a highly-nonlinear resonator*. Tech. rep. Bauman Moscow State Technical University, Yaroslav-the-Wise Novgorod State University and Fiber Optics Research Center of the Russian Academy of Sciences. URL: https://www.osapublishing.org/view_article.cfm?gotourl=https%3A%2F%2Fwww.Eosapublishing%2Eorg%2FDirectPDFAccess%2FC8181378-97A4-8DD5-FAC797E997DD763F_333772%2Foe-23-26-33295%2Epdf%3Fd%3D1%26id%3D333772%26seq%3D0%26mobile%3Dno&org=.
- [19] Marti Duocastella and Craig B. Arnold. *Bessel and annular beams for materials processing*. Tech. rep. Princeton University. URL: <https://www.princeton.edu/~spikelab/papers/081.pdf>.
- [20] Kholoud Kandil. *Laser in Industry*. URL: <http://www.slideshare.net/kokoisawesame/laser-drilling-47644341>.
- [21] Sambit Mitra et al. *Micro-channel machining using high-energy femtosecond Bessel beams*. Tech. rep. Aix-Marseille University and Karlsruhe Institute of Technology. URL: http://www.lp3.univ-mrs.fr/IMG/pdf/IONS_Karlsruhe_2015_Abstract_final_.pdf.
- [22] Ronald D. Schaeffer. *Fundamentals of Laser Micromachining*. CRC Press, 2012.
- [23] *Femto to short pulse: new precision in industrial production*. URL: http://www.opli.net/magazine/lasers/2013/news/rofin_StarFemto_FX.aspx.

BIBLIOGRAPHY

- [24] Eleri Williams. *Experimental and Theoretical Investigations of Nanosecond Fibre Laser Micro-machining*. Tech. rep. University of Cardiff. URL: <http://orca.cf.ac.uk/72916/1/2015WilliamsEPPhD.pdf>.
- [25] Benay Sager and David W. Rosen. *ON THE USE OF ANGLED, DYNAMIC LASER BEAMS TO IMPROVE STEREOLITHOGRAPHY SURFACE FINISH*. Tech. rep. The Woodruff School of Mechanical Engineering in Georgia Institute of Technology, 2004.
- [26] *Stereolithography FAQ*. URL: [http://www.sci-proto.com/home/stereolithography+\(sla\)/stereolithography+faq/default.aspx](http://www.sci-proto.com/home/stereolithography+(sla)/stereolithography+faq/default.aspx).
- [27] *VAT Photopolymerisation*. URL: <http://www.lboro.ac.uk/research/amrg/about/the7categoriesofadditivevatphotopolymerisation/>.
- [28] Michael P. Lee et al. *Development of a 3D printer using scanning projection stereolithography*. Tech. rep. University of Glasgow, 2005.
- [29] *Uses for FEP Tubing*. URL: <http://www.fluorotherm.com/uses-for-fep-tubing/>.
- [30] Amr Khalil. *Optical Advances Speed Rapid Prototyping*. URL: <http://www.makepartsfast.com/2009/03/708/optical-advances-speed-rapid-prototyping/>.
- [31] *laser scanning lens theory*. URL: <http://specialoptics.com/download/10/>.
- [32] *WHICH IS THE PERFECT F-THETA LENS FOR MY HIGH-SPEED SCANNING APPLICATION?* URL: <http://motioncontrol.aerotech.com/2015/06/03/which-is-the-perfect-f-theta-lens-for-my-high-speed-scanning-application/>.
- [33] Daniel Malacara, Robert Celotta, and Judah Levine. *Geometrical and Instrumental Optics*. Academic Press, 1989.
- [34] *Laser scanners*. URL: <http://www.zamisel.com/SSpostavka2.html>.
- [35] Trevor L. Cass. *A short history of the Galvanometer*. URL: <http://www.museumoftechnology.org.uk/stories/galvos.html>.
- [36] *Galvanometer*. URL: <http://hyperphysics.phy-astr.gsu.edu/hbase/magnetic/galvan.html>.
- [37] Theodore Mirtchev, Robert Weeks, and Sergey Minko. *Optimizing the Feedback Control of Galvo Scanners for Laser Manufacturing Systems*. Tech. rep. Algonquin College, 2010.
- [38] *Aircraft Pitch: Simulink Controller Design*. URL: <http://ctms.engin.umich.edu/CTMS/index.php?example=AircraftPitch>.
- [39] Professor Fearings. URL: ProfessorFearing.

BIBLIOGRAPHY

- [40] Nano Scribe. *Micro Rapid Prototyping*. 2016. URL: <http://www.nanoscribe.de/en/applications/micro-rapid-prototyping/>.
- [41] Nano Scribe. *Photonics*. 2016. URL: <http://www.nanoscribe.de/en/applications/photonics/>.
- [42] University of Tampere. *Tailored bioabsorbable implants and scaffolds for biomedical and tissue engineering applications*. 2016. URL: <http://docplayer.net/1016265-Tailored-bioabsorbable-implants-and-scaffolds-for-biomedical-and-tissue-engineering-applications.html>.
- [43] Serge L. Li Hoi Foo-Gregory. "Analysis Of Two-photon Polymerization For The Development Of A Reconfigurable Micro-device Manufacturing System". PhD thesis. University of Michigan, 2012.
- [44] Shoji Maruo, Osamu Nakamura, and Satoshi Kawata. "Three-dimensional microfabrication with two-photon-absorbed photopolymerization". In: *Optical Society of America* (1997).
- [45] Göppert-Mayer. "Über Elementarakte mit zwei Quantensprüngen". In: *Ann. Phys.* : 401.273–294 (1931).
- [46] University of Iowa. *2 Photon Photo Lithography*. 2016. URL: <http://www.tuckerlaboratory.org/sites/tucker/files/2PP%20Figure.png>.
- [47] Intech. *Intech*. 2016. URL: <http://www.intechopen.com/source/html/45302/media/image7.png>.
- [48] Newport Corporation. *Three-Dimensional Microfabrication by Two-Photon Polymerization*. Newport Corporation.
- [49] Boris Chichkov. *Two-photon Polymerization*. 2016. URL: http://asdn.net/asdn/nanotools/two-photon_polymerization.php.
- [50] Daniel S. Correa et al. "Polymerization". In: ed. by Ailton De Souza Gomes. InTech, 2012. Chap. 15.
- [51] J. Serbin et al. "Femtosecond laser-induced two-photon polymerization of inorganic–organic hybrid materials for applications in photonics". In: *Optical Society of America* 28.5 (2003).
- [52] Florin Jipa et al. "Updates in Advanced Lithography". In: ed. by Sumio Hosaka. InTech, 2013. Chap. 3, Femtosecond Laser Lithography in Organic and Non-Organic Materials.
- [53] James (Ed.) Pawley. *Handbook of Biological Confocal Microscopy*. Springer, 2006.

BIBLIOGRAPHY

- [54] Edmund Optics. *Beam Expanders*. 2016. URL: <http://www.edmundoptics.com/resources/application-notes/lasers/beam-expanders/>.
- [55] Tommaso Baldacchini, Scott Snider, and Ruben Zadoyan. “Two-photon polymerization with variable repetition rate bursts of femtosecond laser pulses”. In: *OSA* 20.28 (2012).
- [56] W. J. Schwenger and J. M. Higbie. “High-speed acousto-optic shutter with no optical frequency shift”. In: *AIP* (2012).
- [57] A. M. Weiner. “Femtosecond pulse shaping using spatial light modulators”. In: *AIP* (2000).
- [58] Gaszton Vizsnyiczai, Lorand Kelemen, and Pal Ormos. “Holographic multi-focus 3D two-photon polymerization with real-time calculated holograms”. In: *OSA* 22.20 (2014).
- [59] Yihong Liu, David D. Nolte, and Laura J. Pyrak-Nolte. “Large-format fabrication by two-photon polymerization in SU-8”. In: *Applied Physics A* (2010).
- [60] Jesper Serbin, Aleksandr Ovsianikov, and Boris Chichkov. “Fabrication of woodpile structures by two-photon polymerization and investigation of their optical properties”. In: *OSA* (2004).
- [61] Math Works. *What Is Camera Calibration?* 2016. URL: <http://uk.mathworks.com/help/vision/ug/camera-calibration.html?refresh=true>.
- [62] Pavel Křížek, Ivan Raška, and Guy M. Hagen. “Flexible structured illumination microscope with a programmable illumination array”. In: *OSA* (2012).
- [63] Math Works. *Install and Use Computer Vision System Toolbox OpenCV Interface*. 2016. URL: <http://uk.mathworks.com/help/vision/ug/opencv-interface.html#bujcpsp>.
- [64] Tommaso Baldacchini and Ruben Zadoyan. “In situ and real time monitoring of two-photon polymerization using broadband coherent anti-Stokes Raman scattering microscopy”. In: *OSA* (2010).
- [65] Math Works. *Detecting a Cell Using Image Segmentation*. 2016. URL: <http://uk.mathworks.com/help/images/examples/detecting-a-cell-using-image-segmentation.html?refresh=true>.
- [66] Math Works. *edge*. 2016. URL: <http://uk.mathworks.com/help/images/ref/edge.html>.
- [67] Stephanie A. Pruzinsky and Paul V. Braun. “Fabrication and Characterization of Two-Photon Polymerized Features in Colloidal Crystals”. In: *Advanced Functional Materials* (2005).
- [68] scikit-learn developers. *Tutorial: Machine Learning for Astronomy with Scikit-learn*. 2016. URL: http://www.astroml.org/sklearn_tutorial/_images/plot_ML_flow_chart_1.png.
- [69] Phil Simon. *Too Big to Ignore*. Wiley, 2013.

BIBLIOGRAPHY

- [70] Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [71] Math Works. *Image Classification with Bag of Visual Words*. 2016. URL: <http://uk.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html>.
- [72] Math Works. *Statistics and Machine Learning Toolbox*. 2016. URL: <http://uk.mathworks.com/products/statistics/>.
- [73] Yin-Wen Chang et al. “Training and Testing Low-degree Polynomial Data Mappings via Linear SVM”. In: *Journal of Machine Learning Research* (2010).
- [74] Richard Wollhofen et al. “120 nm resolution and 55 nm structure size in STED-lithography”. In: *OSA* 21.9 (2013).
- [75] Matthew J. Lang and Steven M. Block. “Resource Letter: LBOT-1: Laser-based optical tweezers”. In: *American Journal of Physics* 71.3 (Mar. 1, 2003), pp. 201–215.
- [76] Joshua W. Shaevitz. *A Practical Guide to Optical Trapping*. Aug. 22, 2006.
- [77] Rob Phillips et al. *Physical Biology of the Cell*. 2 edition. London : New York, NY: Garland Science, Nov. 21, 2012. ISBN: 978-0-8153-4450-6.
- [78] Yasuhiro Harada and Toshimitsu Asakura. “Radiation forces on a dielectric sphere in the Rayleigh scattering regime”. In: *Optics Communications* 124.5-6 (Mar. 1996), pp. 529–541.
- [79] Adams Stratton Julius. *Electromagnetic Theory*. In collab. with Osmania University and Digital Library Of India. McGraw Hill Book Company, 1941.
- [80] Jennifer E. Curtis, Brian A. Koss, and David G. Grier. “Dynamic holographic optical tweezers”. In: *Optics Communications* 207.1 (2002), pp. 169–175.
- [81] Aymara Jurado and Justin Kiggins. “Polystyrene Bead Size Behavior in Optical Trapping”. In: (2010).
- [82] Yuichi Kozawa and Shunichi Sato. “Optical trapping of micrometer-sized dielectric particles by cylindrical vector beams”. In: *Optics express* 18.10 (2010), pp. 10828–10833.
- [83] Qingchuan Guo et al. “Comparison of in Situ and ex Situ Methods for Synthesis of Two-Photon Polymerization Polymer Nanocomposites”. In: *Polymers* 6.7 (July 14, 2014), pp. 2037–2050.
- [84] Timo A. Nieminen et al. “Optical tweezers computational toolbox”. In: *Journal of Optics A: Pure and Applied Optics* 9.8 (2007), S196.
- [85] David M. Carberry David C. Benito. “Fabrication of photonic crystal templates using holographic optical tweezers and adhesion via entropic attraction”. In: *Proceedings of SPIE - The International Society for Optical Engineering* 7038 (2008).

BIBLIOGRAPHY

- [86] A. Terray, J. Oakey, and D. W. M. Marr. "Fabrication of linear colloidal structures for microfluidic applications". In: *Applied Physics Letters* 81.9 (Aug. 26, 2002), pp. 1555–1557.
- [87] Péter Galajda and Pál Ormos. "Complex micromachines produced and driven by light". In: *Applied Physics Letters* 78.2 (Jan. 8, 2001), pp. 249–251.
- [88] *Image Gallery - Nanoscribe GmbH*. URL: <http://www.nanoscribe.de/en/image-gallery/?tag=Microfluidics> (visited on 04/18/2016).
- [89] Jonas Enger et al. "Optical tweezers applied to a microfluidic system". In: *Lab on a Chip* 4.3 (June 2004), pp. 196–200.
- [90] Emma Eriksson et al. "A microfluidic system in combination with optical tweezers for analyzing rapid and reversible cytological alterations in single cells upon environmental changes". In: *Lab on a Chip* 7.1 (Jan. 2007), pp. 71–76.
- [91] C.R. Deckard. *Method and apparatus for producing parts by selective sintering*. US Patent 4,863,538. 1989. URL: <https://www.google.com/patents/US4863538>.
- [92] unknown. *Selective Laser Sintering*. URL: <http://www.custompartnet.com/wu/selective-laser-sintering>.
- [93] Vinicius Tribuzi et al. "Birefringent microstructures fabricated by two-photon polymerization containing an azopolymer". In: *Optical Materials Express, Washington : Optical Society of America* 3.1 (2013), pp. 21–26.
- [94] Kotaro Obata et al. "High-aspect 3D two-photon polymerization structuring with widened objective working range (WOW-2PP)". In: *Science & Applications* (2013).
- [95] PCO AG. *PCO.edge 4.2 scientific CMOS camera*. URL: <http://www.pco.de/scmos-cameras/pcoedge-42/>.
- [96] Kenneth R. Spring and Michael W. Davidson. *Modulation Transfer Function*. URL: [ModulationTransferFunction.html](http://www.microscopyu.com/articles/modulationtransferfunction.html).
- [97] Philip C. Robinson and Michael W. Davidson. *Introduction to Polarized Light Microscopy*. URL: <http://www.microscopyu.com/articles/polarized/polarizedintro.html>.
- [98] Douglas B. Murphy, Kenneth R. Spring adn Thomas J. Fellers, and Michael W. Davidson. *Introduction to Optical Birefringence*. URL: <http://www.microscopyu.com/articles/polarized/birefringenceintro.html>.
- [99] *Using Darkfield Microscopy To Enhance Contrast: An Easy and Inexpensive Method*. 2016. URL: <http://public.wsu.edu/~omoto/papers/darkfield.html>.

BIBLIOGRAPHY

- [100] *Dark Field Viewing*. 2016. URL: <http://www.ruf.rice.edu/~bioslabs/methods/microscopy/dfield.html>.
- [101] *Nikon Phase Contrast Microscopy*. 2016. URL: <http://microscopyu.com/tutorials/java/phasecontrast/positivenegative/index.html>.
- [102] *Microscopy Enhanced by the Wave Characteristics of Light*. 2016. URL: <http://cnx.org/contents/DMKzyEh-02/Extended-Topic-Microscopy-Enhanced>.
- [103] Hellen C. Ishikawa-Ankerhold, Richard Ankerhold, and Gregor P. C. Drummen. “Advanced Fluorescence Microscopy Techniques—FRAP, FLIP, FLAP, FRET and FLIM”. In: *Molecules* 17.4 (2012), p. 4047.
- [104] *Handbook of Biological Confocal Microscopy*. Springer Science & Business Media, 1995.
- [105] Denis Semwogerere and Eric R. Weeks. “Confocal Microscopy”. In: *Encyclopedia of Biomaterials and Biomedical Engineering, Second Edition* (2014).
- [106] K.H. Drexhage. *Dyes and lasers*. Ed. by Fritz P. Schäfer. Springer-Verlag, 1990.
- [107] Derek K. Toomre, Matthias F. Langhorst, and Michael W. Davidson. *Introduction to Spinning Disk Confocal Microscopy*. URL: <http://zeiss-campus.magnet.fsu.edu/articles/spinningdisk/introduction.html>.
- [108] Michael W. Davidson. *Concepts in Digital Imaging Technology Photomultiplier Tubes*. URL: <https://micro.magnet.fsu.edu/primer/digitalimaging/concepts/photomultipliers.html>.
- [109] Kenneth R. Spring, Thomas J. Fellers, and Michael W. Davidson. *Signal-to-Noise Considerations*. URL: <http://www.olympusmicro.com/primer/techniques/confocal/signaltonoise.html>.
- [110] CJR Sheppard and DM Shotton. *Confocal Laser Scanning Microscopy*. Bios Sci. Pub. et al., 1997.
- [111] Matthew R Foreman et al. “Determination of the transfer function for optical surface topography measuring instruments—a review”. In: *Measurement Science and Technology* (2013).
- [112] Gu M. “Principles of Three-Dimensional Imaging in Confocal Microscopes”. In: *Singapore: World Scientific* (1996).
- [113] Sheppard C J R and Gu Mand Mao X G. “Three-dimensional coherent transfer function in a reflection mode confocal scanning microscope”. In: *Opt. Commun.* 81 (1991), pp. 281–284.

BIBLIOGRAPHY

- [114] Colin J. R. Sheppard. "Scanning Confocal Microscopy". In: *Encyclopedia of Optical Engineering* (2003).
- [115] CJR Sheppard and DM Shotton. *Confocal Laser Scanning Microscopy*. Bios Sci. Pub. and Springer, 1997.
- [116] Tom Vettenburg. *3DPSF 3DMTF widefield confocal*. URL: https://en.wikipedia.org/wiki/Optical_transfer_function#/media/File:3DPSF_3DMTF_widefield_confocal.png.
- [117] *ETH - CERL - 2-Photon Polymerization*. URL: <https://www1.ethz.ch/cartilage/research/BioFabrication/2-Photon/index> (visited on 03/25/2016).
- [118] ThermoFisher Scientific. *Fluorescent Probes for Two-Photon Microscopy*. URL: <https://www.thermofisher.com/uk/en/home/references/molecular-probes-the-handbook/technical-notes-and-product-highlights/fluorescent-probes-for-two-photon-microscopy.html>.
- [119] Richard K.P. Benninger and David W. Piston. "Two-Photon Excitation Microscopy for the Study of Living Cells and Tissues". In: *Curr Protoc Cell Biol.* (2013).
- [120] *Leica STED demo*. 2010. URL: <http://microscopy.duke.edu/news/STED.html>.
- [121] Remko RM Dijkstra. "Brightness characterization of single core and core/shell Yb3+, Er3+-doped NaYF4 upconversion nanoparticles". In: (2012).
- [122] *FSRA STED*. 2008. URL: <http://mochriellab.yale.edu/STED.html> (visited on 03/09/2016).
- [123] *Olympus Quarter Wave Plate*. 2016. URL: <http://www.olympusmicro.com/primer/techniques/polarized/quarterwavelengthplate.html>.
- [124] Volker Westphal and Stefan W Hell. "Nanoscale resolution in the focal plane of an optical microscope". In: *Physical review letters* 94.14 (2005), p. 143903.
- [125] Javad N Farahani, Matthew J Schibler, and Laurent A Bentolila. "Stimulated emission depletion (STED) microscopy: from theory to practice". In: *Microscopy: science, technology, applications and education* 2 (2010), p. 1539.
- [126] Katrin I Willig et al. "Nanoscale resolution in GFP-based microscopy." In: *Nature Methods* 3.9 (2006), pp. 721–723.
- [127] Emiliano Ronzitti, Benjamin Harke, and Alberto Diaspro. "Frequency dependent detection in a STED microscope using modulated excitation light". In: *Optics express* 21.1 (2013), pp. 210–219.

BIBLIOGRAPHY

- [128] Katrin I Willig et al. “STED microscopy with continuous wave beams”. In: *Nature methods* 4.11 (2007), pp. 915–918.
- [129] Giuseppe Vicidomini et al. “STED nanoscopy with time-gated detection: theoretical and experimental aspects”. In: *PLoS One* 8.1 (2013), e54421.
- [130] Marcel Leutenegger, Christian Eggeling, and Stefan W Hell. “Analytical description of STED microscopy performance”. In: *Optics express* 18.25 (2010), pp. 26417–26429.
- [131] Frederik Gorlitz et al. “A STED microscope designed for routine biomedical applications”. In: *Progress In Electromagnetics Research* 147 (2014), pp. 57–68.
- [132] Yong Wu et al. “Resonant Scanning with Large Field of View Reduces Photobleaching and Enhances Fluorescence Yield in STED Microscopy”. In: *Scientific reports* 5 (2015).
- [133] Gael Moneron et al. “Fast STED microscopy with continuous wave fiber lasers”. In: *Optics express* 18.2 (2010), pp. 1302–1309.
- [134] Richard Wollhofen et al. “120 nm resolution and 55 nm structure size in STED-lithography”. In: *Optics express* 21.9 (2013), pp. 10831–10840.
- [135] Silvia Arroyo-Camejo et al. “Stimulated emission depletion microscopy resolves individual nitrogen vacancy centers in diamond nanocrystals”. In: *ACS nano* 7.12 (2013), pp. 10912–10919.
- [136] *MR121 Specifications*. 2016. URL: <http://www.scbt.com/datasheet-211929-mr-121.html>.
- [137] *SigmaAldrich Nano-diamonds*. 2016. URL: <http://www.sigmadralich.com/catalog/product/aldrich/636428?lang=en®ion=GB>.
- [138] *Aberrior STAR specifications*. 2016. URL: <http://www.sigmadralich.com/catalog/product/sigma/95348?lang=en®ion=GB>.
- [139] Thomas A Klar, Richard Wollhofen, and Jaroslaw Jacak. “Sub-Abbe resolution: from STED microscopy to STED lithography”. In: *Physica Scripta* 2014.T162 (2014), p. 014049. URL: <http://stacks.iop.org/1402-4896/2014/i=T162/a=014049>.
- [140] Tomasso Baldacchini. *Three-Dimensional Microfabrication Using Two-Photon Polymerization*. Elsevier.
- [141] *STED Microscopy and Lithography*. 2016. URL: <http://www.jku.at/applphys/content/e81290/e279893>.
- [142] Travis J Gould et al. “Auto-aligning stimulated emission depletion microscope using adaptive optics”. In: *Optics letters* 38.11 (2013), pp. 1860–1862.

BIBLIOGRAPHY

- [143] Daniel Rönnlund. "Super resolution optical imaging–image analysis, multicolor development and biological applications". In: (2014).
- [144] Thomas A Klar, Richard Wollhoven, and Jaroslaw Jacak. "Sub-Abbe resolution: from STED microscopy to STED lithography". In: *Physica Scripta* 2014.T162 (2014), p. 014049.
- [145] Benjamin Harke and Matthias Reuss. "Sub-25 nm Resolution STED Microscopy". In: ().
- [146] Benjamin Harke et al. "Three-dimensional nanoscopy of colloidal crystals". In: *Nano letters* 8.5 (2008), pp. 1309–1313.
- [147] Andriy Chmyrov et al. "Nanoscopy with more than 100,000'doughnuts". In: *Nature methods* 10.8 (2013), pp. 737–740.
- [148] Stefan W Hell et al. "Measurement of the 4pi-confocal point spread function proves 75 nm axial resolution". In: *Applied Physics Letters* 64.11 (1994), pp. 1335–1337.
- [149] Stefan Hell and Ernst HK Stelzer. "Properties of a 4Pi confocal fluorescence microscope". In: *JOSA A* 9.12 (1992), pp. 2159–2166.
- [150] Martin Schrader et al. "4Pi-confocal imaging in fixed biological specimens". In: *Biophysical journal* 75.4 (1998), pp. 1659–1668.
- [151] Brynmor J Davis et al. "4Pi spectral self-interference microscopy". In: *JOSA A* 24.12 (2007), pp. 3762–3771.
- [152] by J. Demmerle, C. Eggeling, and L. Schermelleh. *A Short Introduction to Super-Resolution Microscopy*. URL: http://www.micron.ox.ac.uk/research/super_resolution_microscopy.php.
- [153] Lothar Schermelleh et al. "Subdiffraction Multicolor Imaging of the Nuclear Periphery with 3D Structured Illumination Microscopy". In: *Science* 320 (2008), pp. 1332–1336.
- [154] *Holoeye presentation*. 2016. URL: <http://www.lastek.co.kr/updata/f11804878071.pdf>.
- [155] Nathan Jay Jenness. "Three-dimensional Holographic Lithography and Manipulation Using a Spatial Light Modulator". Department of Mechanical Engineering and Materials Science Duke University, 2009.
- [156] 2016. URL: <http://www.ipms.fraunhofer.de/content/dam/ipms/en/documents/ao-development.pdf>.
- [157] Hällstig. "Nematic Liquid Crystal Spatial Light Modulators for Laser Beam Steering". Acta Universitatis Upsaliensis, 2004.

BIBLIOGRAPHY

- [158] Elroy Louis Pearson. "MEMS spatial light modulator for holographic displays". Massachusetts Institute of Technology, 1973.
- [159] MEMS. 2009. URL: <https://www.youtube.com/watch?v=CNmk-SeMOZI>.
- [160] Daan Sprünken. "A 2d Spatial Light Modulator for spatio-temporal shaping". Faculty of Science and Technology university of Twente, 2008.
- [161] Thomas Becker et al. "MEMS spatial light modulators with integrated electronics". In: *SPIE* (2003).
- [162] Jonathan Dunayevsky and Dan M. Marom. "MEMS Spatial Light Modulator for Phase and Amplitude Modulation of Spectrally Dispersed Light". In: *Journal of microelectromechanical systems* (2013).
- [163] Uzi Efron. *Spatial Light Modulator: Materials, Devices and Application*. United States of America: Marcel Dekker, 1999.
- [164] LCoS SLM. 2016. URL: https://www.hamamatsu.com/resources/pdf/ssd/e12_handbook_lcos_slm.pdf.
- [165] Wolfgang Osten and Nadya Reingand. *Optical Imaging and Metrology: Advanced Technologies*. first edition. Wiley-VCH Verlag GmbH & Co. KGaA., 2012.
- [166] Kujawinska Małgorzata, Porras-Aguilar Rosario, and Zaperty Weronika. "LCoS SPATIAL LIGHT MODULATORS AS ACTIVE PHASE ELEMENTS OF FULL-FIELD MEASUREMENT SYSTEMS AND SENSORS". In: *Metrology and Measurement Systems* 19.3 (2012).
- [167] Melia Bonomo. 2016. URL: http://laser.physics.sunysb.edu/~melia/SLM_intro.html.
- [168] Spatial Light Modulator. 2016. URL: <http://holoeye.com/>.
- [169] Denis Andrienko. *Introduction to liquid crystals*. International Max Planck Research School, 2006.
- [170] U. Efron, S. T. Wu, and T. D. Bates. "Nematic liquid crystals for spatial light modulators: recent studies". In: 3.2 (1986).
- [171] Michael G. Robinson, Jianmin Chen, and Gary D. Sharp. *Polarization Engineering for LCD Projection*. England: publisher, 2005.
- [172] Eliyahu Osherovich. "Numerical methods for phase retrieval". Israel Institute of Technology, 2012.
- [173] E. Frumker and Y. Silberberg. "Two-dimensional phase-only spatial light modulators for dynamic phase and amplitude pulse shaping". In: *Journal of Modern Optics* (2009).

BIBLIOGRAPHY

- [174] 2016. URL: <http://www.laser2000.co.uk/>.
- [175] Hidetomo Takahashi, Satoshi Hasegawa, and Yoshio Hayasaki. "Holographic femtosecond laser processing using optimal-rotation-angle method with compensation of spatial frequency response of liquid crystal spatial light modulator". In: *Appl. Opt.* 46.23 (2007), pp. 5917–5923. DOI: 10.1364/AO.46.005917. URL: <http://ao.osa.org/abstract.cfm?URI=ao-46-23-5917>.
- [176] Gaszton Vizsnyiczai, Lóránd Kelemen, and Pál Ormos. "Holographic multi-focus 3D two-photon polymerization with real-time calculated holograms". In: *Opt. Express* 22.20 (2014), pp. 24217–24223. DOI: 10.1364/OE.22.024217. URL: <http://www.opticsexpress.org/abstract.cfm?URI=oe-22-20-24217>.
- [177] Alexander Jesacher and Martin J. Booth. "Parallel direct laser writing in three dimensions with spatially dependent aberration correction". In: *Opt. Express* 18.20 (2010), pp. 21090–21099. DOI: 10.1364/OE.18.021090. URL: <http://www.opticsexpress.org/abstract.cfm?URI=oe-18-20-21090>.
- [178] J.W. Goodman. *Introduction to Fourier Optics*. 2nd. McGraw-Hill, 1996.
- [179] *PLUTO Phase Only Spatial Light Modulators (Manual)*. NIR-015. V 1.7. HOLOEYE Photonics AG. Jan. 2016.
- [180] R. W. Gerchberg and W. O. Saxton. "A practical algorithm for the determination of the phase from image and diffraction plane pictures". In: *Optik (Jena)* 35 (1972), pp. 237+.
- [181] Eric R. Dufresne et al. "Computer-generated holographic optical tweezer arrays". In: *Review of Scientific Instruments* 72.3 (2001).
- [182] Dustin Stuart, Oliver Barter, and Axel Kuhn. "Fast algorithms for generating binary holograms". In: *arXiv preprint arXiv:1409.1841* (2014).
- [183] Roberto Di Leonardo, Francesca Ianni, and Giancarlo Ruocco. "Computer generation of optimal holograms for optical trap arrays". In: *Opt. Express* 15.4 (2007), pp. 1913–1922. DOI: 10.1364/OE.15.001913. URL: <http://www.opticsexpress.org/abstract.cfm?URI=oe-15-4-1913>.
- [184] Michael A. Seldowitz, Jan P. Allebach, and Donald W. Sweeney. "Synthesis of digital holograms by direct binary search". In: *Appl. Opt.* 26.14 (1987), pp. 2788–2798. DOI: 10.1364/AO.26.002788. URL: <http://ao.osa.org/abstract.cfm?URI=ao-26-14-2788>.

BIBLIOGRAPHY

- [185] Andrew G. Kirk and Trevor J. Hall. "Design of binary computer generated holograms by simulated annealing: coding density and reconstruction error". In: *Optics Communications* 94.6 (1992), pp. 491 –496. ISSN: 0030-4018. DOI: [http://dx.doi.org/10.1016/0030-4018\(92\)90592-F](http://dx.doi.org/10.1016/0030-4018(92)90592-F). URL: <http://www.sciencedirect.com/science/article/pii/003040189290592F>.
- [186] Rainer G. Dorsch, Adolf W. Lohmann, and Stefan Sinzinger. "Fresnel ping-pong algorithm for two-plane computer-generated hologram display". In: *Appl. Opt.* 33.5 (1994), pp. 869–875. DOI: 10.1364/AO.33.000869. URL: <http://ao.osa.org/abstract.cfm?URI=ao-33-5-869>.
- [187] A.M. Howatson. *Engineering Tables and Data*. Springer Netherlands, 2012. ISBN: 9789401093149. URL: <https://books.google.co.uk/books?id=VpznCAAAQBAJ>.
- [188] Christopher Cullinan et al. "Computing performance benchmarks among cpu, gpu, and fpga". In: [www.wpi.edu/Pubs/E-project/Available/E-project-030212-123508/unrestricted/BenchmarkingFinal](http://www.wpi.edu/Pubs/E-project/Available/E-project-030212-123508/unrestricted/BenchmarkingFinal.pdf) (2013).
- [189] *Nikon MicroscopyU / Introduction to Microscope Objectives*. URL: <http://www.microscopyu.com/articles/optics/objectiveintro.html> (visited on 04/05/2016).
- [190] M. Françon. *Progress in Microscopy: International Series of Monographs on Pure and Applied Biology: Modern Trends in Physiological Sciences*. Elsevier, Sept. 11, 2013. ISBN: 978-1-4831-8459-3.
- [191] *Introduction to Reflective Objectives / Edmund Optics*. URL: <http://www.edmundoptics.co.uk/technical-resources-center/microscopy/introduction-to-reflective-objectives/> (visited on 02/04/2016).
- [192] *Quick Guide Microscopes*. URL: https://www.mitutoyo.co.jp/eng/products/menu/QuickGuide_Microscopes.pdf.
- [193] *Numerical aperture*. In: *Wikipedia, the free encyclopedia*. Feb. 22, 2016.
- [194] *Nikon MicroscopyU / Concepts and Formulas / Numerical Aperture*. URL: <http://www.microscopyu.com/articles/formulas/formulasna.html> (visited on 04/06/2016).
- [195] Michael W. Davidson and Mortimer Abramowitz. "Optical Microscopy". In: *Encyclopedia of Imaging Science and Technology*. John Wiley & Sons, Inc., 2002. ISBN: 978-0-471-44339-1.
- [196] Michael Bass et al. *Handbook of Optics, Third Edition*. 3rd ed. New York, NY, USA: McGraw-Hill, Inc., 2010. ISBN: 978-0-07-149889-0.

BIBLIOGRAPHY

- [197] *Objective Back focal plane diameter*. Physics Forums - The Fusion of Science and Community. URL: <https://www.physicsforums.com/threads/objective-back-focal-plane-diameter.312669/> (visited on 04/14/2016).
- [198] James Pawley, ed. *Handbook of Biological Confocal Microscopy*. 3rd edition. New York, NY: Springer, June 2, 2006. ISBN: 978-0-387-25921-5.
- [199] R Culp. *Telescope lens magnification*. 2012. URL: <http://www.rocketmime.com/astronomy/Telescope/Magnification.html> (visited on 04/15/2016).
- [200] *MIT Optics Course Lecture 23*. URL: http://ocw.mit.edu/courses/mechanical-engineering/2-71-optics-spring-2009/video-lectures/lecture-23-imaging-with-a-single-lens/MIT2_71S09_lec23.pdf (visited on 03/09/2016).
- [201] R. Paschotta. *Sellmeier Formula / The Encyclopedia of Laser Physics and Technology*. 2008. URL: https://www.rp-photonics.com/sellmeier_formula.html (visited on 03/13/2016).
- [202] Vladimir Sacek. *9.3 Designing Doublet Achromat / telescopeOptics.net*. 2006. URL: http://www.telescope-optics.net/designing_doublet_achromat.htm (visited on 03/13/2016).
- [203] G.D.Boreman. *Modulation Transfer Function / SPIE*. 2001. URL: http://spie.org/publications/optipedia-pages/press-content/tt52/tt52_131_modulation_transfer_function/ (visited on 03/14/2016).
- [204] Hamed Hamid Muhammad. *Medical Image Analysis / Digital Image Analysis*. 2008. URL: <http://medim.sth.kth.se/612872/F/F9.pdf> (visited on 03/14/2016).
- [205] E. Hecht. *Optics*. 1998.
- [206] Ihar Schatsinini. “Free Clusters and Free Molecules in Strong, Shaped Laser Fields”. Dissertation. Freien Universität Berlin, 2009.
- [207] R. Paschotta. *Chromatic Dispersion / The Encyclopedia of Laser Physics and Technology*. 2008. URL: https://www.rp-photonics.com/chromatic_dispersion.html (visited on 03/15/2016).
- [208] Sophocles J. Orfanidis. *Electromagnetic Waves and Antennas*. 2014.
- [209] *The Effect of Dispersion on Ultrashort Pulses - Newport*. 2016. URL: <http://www.newport.com/The-Effect-of-Dispersion-on-Ultrashort-Pulses/602091/1033/content.aspx> (visited on 03/15/2016).
- [210] Paul J Wrzesinski et al. “Group-velocity-dispersion measurements of atmospheric and combustion-related gases using an ultrabroadband-laser source”. In: *Optics Express*. 2011.

BIBLIOGRAPHY

- [211] Selcuk Akturk et al. "Spatio-temporal couplings in ultrashort laser pulses". In: *. 2010.
- [212] Xun Gu, Selcuk Akturk, and Rick Trebino. "Spatial Chirp in ultrafast optics". In: 2004.
- [213] Rick Trebino et al. "Measuring ultrashort laser pulses in the time-frequency domain using frequency-resolved optical gating". In: 1997.
- [214] *Frequency-resolved Optical Gating / The Encyclopedia of Laser Physics and Technology*. 2016. URL: https://www.rp-photonics.com/frequency_resolved_optical_gating.html (visited on 04/14/2016).
- [215] Patrick O'Shea et al. "Highly simplified device for ultrashort-pulse measurement". In: 2000.
- [216] Justin Ratner et al. "The coherent artifact in modern pulse measurements". In:
- [217] N X Truong, J Tiggesbäumker, and T Döppner. "Shaper-assisted removal of the direction-of-time ambiguity in second-harmonic generation frequency-resolved optical gating". In:
- [218] E. Zeek et al. "Simultaneous automatic calibration and direction-of-time removal in frequency-resolved optical gating". In:
- [219] Hamamatsu Photonics. *Opto-Semiconductor Handbook*. Available at https://www.hamamatsu.com/jp/en/hamamatsu/overview/bsd/solid_state_division/related_documents.html (30/11/2015).
- [220] Robert J. Noll. "Zernike polynomials and atmospheric turbulence*". In: *J. Opt. Soc. Am.* 66.3 (1976), pp. 207–211. DOI: 10.1364/JOSA.66.000207. URL: <http://www.osapublishing.org/abstract.cfm?URI=josa-66-3-207>.
- [221] ThorLabs. *Optical Wavefront Sensors (Shack-Hartmann) WFS Series Operation Manual*. Available at https://www.thorlabs.co.uk/newgroupage9.cfm?objectgroup_id=2946&pn=WFS150-5C (30/01/2016). 2015.
- [222] 2016. URL: http://www.thorlabs.co.uk/NewGroupPage9.cfm?ObjectGroup_ID=5287.
- [223] Saad A. Rahman and Martin J. Booth. "Direct wavefront sensing in adaptive optical microscopy using backscattered light". In: *Appl. Opt.* 52.22 (2013), pp. 5523–5532. DOI: 10.1364/AO.52.005523. URL: <http://ao.osa.org/abstract.cfm?URI=ao-52-22-5523>.
- [224] Martin J Booth. "Adaptive optical microscopy: the ongoing quest for a perfect image". In: *Light Sci Appl* 3 (Apr. 2014), e165–. URL: <http://dx.doi.org/10.1038/lsa.2014.46>.
- [225] *Crystalline Quartz Waveplates*. 2016. URL: http://www.altechna.com/product_details.php?id=877 (visited on 05/05/2016).

BIBLIOGRAPHY

- [226] *Round Continuously Variable Metallic Neutral Density Filters*. 2016. URL: https://www.thorlabs.com/NewGroupPage9.cfm?ObjectGroup_ID=1393 (visited on 03/25/2016).
- [227] *Manual Filter Wheel Mounts with Neutral Density Filters*. 2016. URL: https://www.thorlabs.com/newgroupage9.cfm?objectgroup_id=1444 (visited on 03/25/2016).
- [228] R. Paschotta. *Neutral Density Filters / The Encyclopedia of Laser Physics and Technology*. 2008. URL: https://www.rp-photonics.com/neutral_density_filters.html (visited on 03/25/2016).
- [229] Mark L. Kushner. *Computational Plasma Science and Engineering Group - ECE355*. 2001. URL: http://uigelz.eecs.umich.edu/classes/pub/ece355/handouts/example_gaussian_beam_focused_lens.pdf (visited on 03/28/2016).
- [230] W.J.Schwenger and J.M. Higbie. “High-speed acousto-optic shutter with no optical frequency shift”. In: 2012.
- [231] J.R.Priebe. “First-Order Equation for Prismatic Magnification”. In:
- [232] A. M. Weiner. “Femtosecond pulse shaping using spatial light modulators”. In: *Review of Scientific Instruments* 71.5 (2000), pp. 1929–1960. DOI: <http://dx.doi.org/10.1063/1.1150614>. URL: <http://scitation.aip.org/content/aip/journal/rsi/71/5/10.1063/1.1150614>.
- [233] *Thorlabs Grating Tutorial*. URL: <https://www.thorlabs.com/tutorials.cfm?tabID=0ca9a8bd-2332-48f8-b01a-7f8bf0c03d4e> (visited on 03/20/2016).
- [234] *Diffraction Gratings - Horiba*. URL: <http://www.horiba.com/uk/scientific/products/optics-tutorial/diffraction-gratings/> (visited on 03/20/2016).
- [235] Johannes Hallier. “Design of an ultrafast laser pulse shaper using a spatial light modulator and evolutionary strategies”. PhD thesis. University of the Federal Armed Forces Hamburg Postfach 70 08 22 22008 Hamburg: Helmut Schmidt University, 2007.
- [236] *CRi Spatial Light Modulator - Data Sheet*. URL: http://www.meadowlark.com/store/data-sheet/CRi_product_announcement.pdf (visited on 03/22/2016).
- [237] *Thorlabs Reflective Gratings - Data Sheet*. URL: http://www.thorlabs.co.uk/newgroupage9.cfm?objectgroup_id=25&pn=GH13-18V (visited on 03/21/2016).
- [238] A. M. Weiner. “Femtosecond pulse shaping using spatial light modulators”. In: *American Institute of physics* (2000).

BIBLIOGRAPHY

- [239] Johannes Hallier. "Design of an ultrafast laser pulse shaper using a spatial light modulator and evolutionary strategies". Helmut Schmidt Universitat, 2007.
- [240] *Introduction to liquid crystals*. 2016. URL: <https://bly.colorado.edu/lcphysics/lcintro/tnlc.html>.
- [241] 2016. URL: <http://www.meadowlark.com/>.
- [242] Thorlabs. *Optical Tables Tutorial*. URL: https://www.thorlabs.com/newgroupage9.cfm?objectgroup__id=8275.
- [243] pbroks13. "FileMass spring damper.svg". In: (). URL: https://commons.wikimedia.org/wiki/File:Mass__spring__damper.svg.
- [244] Prof R S Dow. *Bending Stress*. URL: http://www.dianliwenmi.com/postimg__4545236.html.
- [245] S.E. Urreta de Pereyra, E. Carreño Morelli, and A.A. Ghilarducci. "Grain boundary internal friction peak in aluminium alloys containing particles". In: *Scripta Metallurgica* 23.10 (1989), pp. 1691–1695. ISSN: 00369748. DOI: 10.1016/0036-9748(89)90344-X.
- [246] Jim Burge. "Mounting of Optical Components General approaches to designing optical mounts". In: (2011).
- [247] Katie Schwertz. "Useful Estimations and Rules of Thumb for Optomechanics". In: *University of Arizona* (2010), pp. 1–92.
- [248] Jim Burge. "Mounting of Optical Components: Mounting of lenses". In: *Mounting of Optical Components* (2011).
- [249] [Http://www.olympusmicro.com/](http://www.olympusmicro.com/). *Optical Birefringence*. URL: <http://www.olympusmicro.com/primer/lightandcolor/birefringence.html>.
- [250] [Http://www.anodizing.org](http://www.anodizing.org). *What is Anodizing?* URL: http://www.anodizing.org/?page=what__is__anodizing.
- [251] Alexander Korsunsky. "Finite Element Methods 1". In: 2 (2006), p. 2006.
- [252] Lloyd N Trefethen and David Bau. *Numerical linear algebra*. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1997, p. 361. ISBN: 9780898719574. URL: <http://books.google.com/books?id=JaPtx0ytY7kC>.
- [253] COMSOL. *Meshing Considerations for Linear Static Problems*. 2013. URL: www.comsol.com/blogs/meshing-considerations-linear-static-problems/.

BIBLIOGRAPHY

- [254] [Http://what when how.com](http://whatwhenhow.com). *FEM for 3D Solids (Finite Element Method) Part 4*. URL: <http://what-when-how.com/the-finite-element-method/fem-for-3d-solids-finite-element-method-part-4/>.
- [255] K. Ho-Le. "Finite element mesh generation methods: a review and classification". In: *Computer-Aided Design* (1988). ISSN: 00104485. DOI: 10.1016/0010-4485(88)90138-8.
- [256] [Https://deust.wordpress.com](https://deust.wordpress.com). *h-method & p-method*. URL: <https://deust.wordpress.com/2014/11/30/h-method-p-method/>.
- [257] Andreas Ostendorf and Boris N. Chichkov. "Two-Photon Polymerization: A New Approach to Micromachining". In: *Photonics Spectra* (2006).
- [258] Adina I. Ciuciuaand and Piotr J. Cywinski. "Two-photon polymerization of hydrogels—versatile solutions to fabricate well-defined 3D structures". In: *RSC Advances* (2014).
- [259] ReplicatorG. *ReplicatorG*. 2016. URL: <http://replicat.org/start>.
- [260] Mary L. Boas. *Mathematical Methods in the Physical Sciences*. Kaye Pace, 2006.
- [261] Aerotech. *Linear Stages*. 2016. URL: <https://www.aerotech.co.uk/product-catalog/stages/linear-stage.aspx>.
- [262] tribology abc. *Linear ball slide with limited travel*. 2016. URL: http://www.tribology-abc.com/calculators/c4_8.htm.
- [263] Inc. Del-Tron Precision. *Linear Motion Engineering*. 2016. URL: https://www.deltron.com/search/Linear_Slides_Ball_Slides_Crossed_Roller_Slides_Search.aspx.
- [264] New Way air Bearings. *Air Bearing Application and Design Guide*. 2016. URL: http://www.newwayairbearings.com/sites/default/files/new_way_application_and_design_guide_%20Rev_E_2006-01-18.pdf.
- [265] Matt Reck. *Advantages of Air Bearings*. 2016. URL: <http://machinedesign.com/bearings/advantages-air-bearings>.
- [266] Newport Corporation. *Ultra-Precision Linear Motor Stage, 100 mm Travel, XMS Series*. 2016. URL: http://search.newport.com/?q=*&x2=sku&q2=XMS100.
- [267] Aerotech. *Minimum Incremental Motion and Holding Stability in Beamline Positioning*. 2016. URL: <http://go.aerotech.com/technicalreport>.
- [268] PI. *P-541.Z Piezo Z and Z / Tip / Tilt Stages*. 2016. URL: <http://www.physikinstrumente.com/product-detail-page/p-541z-201540.html>.

BIBLIOGRAPHY

- [269] A. Telba and Wahied G. Ali, eds. *Hysteresis Modeling in a Piezoelectric Nanopositioner Stage*. Vol. II. World Congress on Engineering. 2011.
- [270] Newport Corporation. *XM Series: Ultra-Precision Linear Motor Stages*. Newport Corporation.
- [271] NPTEL. *NPTEL Mechanical Engineering*. 2016. URL: <http://nptel.ac.in/courses/112101005/images/fig2.2.5..JPG>.
- [272] University of Cambridge. *Beam of square cross-section*. 2016. URL: <http://www.doitpoms.ac.uk/tlplib/optimisation-biomaterials/def-derivation-I.htm>.
- [273] Richard G Budynas and Keith J Nisbett. *Shigley's Mechanical Engineering Design*. 1st. McGraw-Hill, 2006.
- [274] University of Cambridge. *Young's Modulus - Density selection map*. 2016. URL: http://www.doitpoms.ac.uk/tlplib/optimisation-biomaterials/modulus_density.php.
- [275] Inc. Silicon Carbide Products. *Silicon Carbide Products*. 2016. URL: <http://www.scprobond.com/about-us/>.
- [276] Y. Zhang, S. Finger, and S. Behrens. *Introduction to Mechanisms*. URL: <https://www.cs.cmu.edu/~rapidproto/mechanisms/chpt4.html>.
- [277] URL: <https://www.unitag.io/qrcode/what-is-a-qrcode>.
- [278] Daniel Smilkov. *Largest Rectangle in a Polygon*. URL: <http://d3plus.org/blog/behind-the-scenes/2014/07/08/largest-rect/>.
- [279] Kelly Kearney. *Matlab File Exchange: bufferm2*. URL: <http://uk.mathworks.com/matlabcentral/fileexchange/11095-bufferm2>.
- [280] Dong Lin et al. “Three dimensional printing of complex structures: man made or toward nature?” In: *American Civil Society* (2014).
- [281] Michael Porter. “How competitive forces shape strategy”. In: *HBR* (1979).