Laser-induced shock wave cleaning of EUV photomasks

Lammers, N.A.

DOI:
10.6100/IR731141

Published: 01/01/2012

Citation for published version (APA):
Laser–Induced Shock Wave Cleaning of EUV Photomasks

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus, prof.dr.ir. C.J. van Duijn, voor een commissie aangewezen door het College voor Promoties in het openbaar te verdedigen op donderdag 5 april 2012 om 16.00 uur

door

Nicolaas Arnoldus Lammers
geboren te Zeven, Duitsland
Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. J.J.A.M. van der Mullen,
en
prof.dr.ir. J.P.H. Benschop

Copromotor:
dr. L. Scaccabarozzi

This research is sponsored by the company ASML Netherlands B.V.

CIP-DATA TECHNISCHE UNIVERSITEIT EINDHOVEN

Lammers, Nicolaas Arnoldus

Laser-Induced Shock Wave Cleaning of EUV Photomasks / by Nicolaas Arnoldus Lammers. –

A catalogue record is available from the Eindhoven University of Technology Library.
NUR 924
Subject headings: laser-induced plasmas / submicron particle removal / microlithography / gas dynamics / shock theory / similarity methods / adhesion theory / turbulence modeling

Copyright © 2012 Nicolaas Arnoldus Lammers

All rights reserved. No part of this book may be reproduced, stored in a database or retrieval system, or published, in any form or in any way, electronically, mechanically, by print, photo-print, microfilm or any other means without prior written permission of the author.
With the arrival of the next generation extreme ultraviolet (EUV) lithography tools, printed features on integrated circuits will continue to shrink in the future. In the absence of a protective pellicle to shield the surface of the photomask from particles, the requirements for contamination control are getting more and more demanding. In some cases the presence of a single particle as small as 30 nm on the surface of a photomask can already lead to costly production errors, which is why the semiconductor industry is searching for a quick, dry, and contact-free method to remove particles from contamination critical surfaces. In light of these developments, this thesis presents an investigation into the feasibility of laser-induced shock wave cleaning as a novel method for removing small submicron particles from the surface of an EUV photomask. This particular cleaning method removes particles by exposing them to a fast moving shock wave, which is generated by the laser-induced breakdown of a gas above the surface.

To study its properties and to validate the proposed shock theory, the shock wave produced by the laser-induced breakdown is visualized with specialized techniques such as shadowgraphy and Schlieren photography. The recorded images clearly reveal the transient behaviour of the shock wave, from which it can be concluded that the laser-induced breakdown process is consistent with the Taylor & Sedov solution of an intense point explosion. By combining this result with particle adhesion theory, a model has been developed that is able to predict the diameter of the smallest removable particle as a function of the main cleaning parameters. The results of this model were found to be in agreement with the experimental results.

To assess the cleaning capabilities of laser-induced shock wave cleaning, a series of cleaning experiments were performed on flat unpatterned 300 mm silicon wafers. The wafers were purposely contaminated with polystyrene latex and silica spheres.
Abstract

with sizes between 60 and 150 nm, after which they were exposed to various laser-induced shock wave cleaning treatments. The results of these cleaning experiments showed that in some cases it is possible to fully remove polystyrene latex spheres with a diameter below 60 nm. Furthermore, it is expected that the removal of particles from patterned surfaces is more problematic, because the particles that are located at the bottom of a trench are able to escape the full force of the shock wave. Unfortunately, damage assessments on the fragile multilayer mirror of the EUV photomask have shown that the laser-induced shock wave cleaning method is too damaging for the surface of the EUV photomask. It is therefore concluded that laser-induced shock wave cleaning can be an effective method of particle removal from flat surfaces, but that it is not suitable for the cleaning of EUV photomasks.
# Contents

Abstract I

Contents III

List of Tables IX

List of Figures XI

1 Introduction 1

1.1 Microlithography ........................................ 2

1.1.1 Optical Lithography .................................. 3

1.1.2 Immersion Lithography ............................... 4

1.1.3 MP Lithography ....................................... 5

1.1.4 EUV Lithography ...................................... 6

1.2 Contamination ........................................... 8

1.2.1 Prevention and Detection ............................. 9

1.2.2 Laser-Induced Shock Wave Cleaning ................. 10

1.3 Outline of This Thesis .................................. 10

References .................................................. 11
2 Laser Induced Breakdown in Gases

2.1 Laser Induced Plasmas

2.1.1 Plasma Basics

2.1.2 Plasma Ignition

2.1.3 Plasma Heating

2.2 Shock Wave Theory

2.2.1 Gas Dynamics

2.2.2 Euler Equations

2.2.3 Normal Shock Relations

2.3 Physics of an Intense Point Explosion

2.3.1 Taylor & Sedov Solution

2.3.2 Post Shock Wave Solution

2.3.3 Shock Strength

2.4 Conclusions

References

3 Time Resolved Imaging

3.1 Flow Visualisation in Transparent Gases

3.1.1 Deflection of Light

3.1.2 Shadowgraphy

3.1.3 Schlieren Photography

3.2 Experimental Setups

3.2.1 Laser Flash Shadowgraphy Setup

3.2.2 Z-Type Schlieren Setup

3.3 Time Resolved Images

3.3.1 Plasma Ignition

3.3.2 Shock Wave Propagation

3.3.3 Hot Gas Core

3.4 Conclusions

References

4 Particle Adhesion Theory

4.1 Microscopic Adhesion Forces
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.1 Electrostatic Interactions</td>
<td>62</td>
</tr>
<tr>
<td>4.1.2 Interactions Involving Polarization</td>
<td>64</td>
</tr>
<tr>
<td>4.1.3 Quantum Mechanical Interactions</td>
<td>65</td>
</tr>
<tr>
<td>4.2 Adhesion Between Macroscopic Bodies</td>
<td>66</td>
</tr>
<tr>
<td>4.2.1 van der Waals Force</td>
<td>66</td>
</tr>
<tr>
<td>4.2.2 Capillary Force</td>
<td>72</td>
</tr>
<tr>
<td>4.2.3 Electric Double Layer Force</td>
<td>75</td>
</tr>
<tr>
<td>4.2.4 Coulombic Image Force</td>
<td>76</td>
</tr>
<tr>
<td>4.3 Simple Particle Removal Model</td>
<td>77</td>
</tr>
<tr>
<td>4.3.1 Model Description</td>
<td>78</td>
</tr>
<tr>
<td>4.3.2 Limitations</td>
<td>79</td>
</tr>
<tr>
<td>4.3.3 Implementation</td>
<td>82</td>
</tr>
<tr>
<td>4.3.4 Results</td>
<td>82</td>
</tr>
<tr>
<td>4.4 Conclusions</td>
<td>82</td>
</tr>
<tr>
<td>References</td>
<td>85</td>
</tr>
<tr>
<td>5 Cleaning Experiments</td>
<td>87</td>
</tr>
<tr>
<td>5.1 Laser–Induced Shock Wave Cleaning Setup</td>
<td>88</td>
</tr>
<tr>
<td>5.1.1 Specifications</td>
<td>88</td>
</tr>
<tr>
<td>5.1.2 Limitations</td>
<td>89</td>
</tr>
<tr>
<td>5.2 Cleaning Experiments</td>
<td>90</td>
</tr>
<tr>
<td>5.2.1 Particle Deposition and Detection</td>
<td>90</td>
</tr>
<tr>
<td>5.2.2 Cleaning Parameters</td>
<td>92</td>
</tr>
<tr>
<td>5.2.3 Removal of Polystyrene Particles</td>
<td>93</td>
</tr>
<tr>
<td>5.2.4 Removal of Silica Particles</td>
<td>98</td>
</tr>
<tr>
<td>5.3 Evaluation of Surface Damage</td>
<td>98</td>
</tr>
<tr>
<td>5.3.1 Damage Experiments</td>
<td>98</td>
</tr>
<tr>
<td>5.4 Conclusions</td>
<td>101</td>
</tr>
<tr>
<td>References</td>
<td>101</td>
</tr>
<tr>
<td>6 Conclusions</td>
<td>103</td>
</tr>
<tr>
<td>6.1 Shock Theory</td>
<td>104</td>
</tr>
<tr>
<td>6.2 Time Resolved Imaging</td>
<td>104</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.3 Particle Removal Model</td>
<td>104</td>
</tr>
<tr>
<td>6.4 Cleaning Capabilities</td>
<td>105</td>
</tr>
<tr>
<td>6.5 Surface Damage</td>
<td>105</td>
</tr>
</tbody>
</table>

**Addendum**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 Turbulence Modeling</td>
<td>109</td>
</tr>
<tr>
<td>7.1 Introduction</td>
<td>110</td>
</tr>
<tr>
<td>7.1.1 Historical Context</td>
<td>111</td>
</tr>
<tr>
<td>7.1.2 The Nature of Turbulence</td>
<td>113</td>
</tr>
<tr>
<td>7.2 The Turbulence Closure Problem</td>
<td>115</td>
</tr>
<tr>
<td>7.2.1 Instantaneous Flow Equations</td>
<td>115</td>
</tr>
<tr>
<td>7.2.2 Reynolds Decomposition</td>
<td>116</td>
</tr>
<tr>
<td>7.2.3 Averaged Flow Equations</td>
<td>118</td>
</tr>
<tr>
<td>7.2.4 Closure Approximations</td>
<td>118</td>
</tr>
<tr>
<td>7.3 Eddy Viscosity Turbulence Modeling</td>
<td>121</td>
</tr>
<tr>
<td>7.3.1 The $k$-Equation</td>
<td>121</td>
</tr>
<tr>
<td>7.3.2 The $k-\varepsilon$ Model</td>
<td>123</td>
</tr>
<tr>
<td>7.4 Near Wall Behaviour</td>
<td>123</td>
</tr>
<tr>
<td>7.4.1 Turbulent Boundary Layers</td>
<td>124</td>
</tr>
<tr>
<td>7.4.2 Wall Damping Functions</td>
<td>129</td>
</tr>
<tr>
<td>7.5 Mixing in Turbulent Flows</td>
<td>130</td>
</tr>
<tr>
<td>7.6 Numerical Implementation</td>
<td>133</td>
</tr>
<tr>
<td>7.6.1 Averaged Flow Equations</td>
<td>133</td>
</tr>
<tr>
<td>7.6.2 The $k-\varepsilon$ Model</td>
<td>134</td>
</tr>
<tr>
<td>7.7 Validation Experiments</td>
<td>135</td>
</tr>
<tr>
<td>7.7.1 Pipe Flow</td>
<td>135</td>
</tr>
<tr>
<td>7.7.2 Channel Flow</td>
<td>138</td>
</tr>
<tr>
<td>7.7.3 Backward Facing Step</td>
<td>139</td>
</tr>
<tr>
<td>7.8 Conclusions</td>
<td>146</td>
</tr>
</tbody>
</table>

References

Acknowledgements

VI
<table>
<thead>
<tr>
<th>Contents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>151</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td>153</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Common contaminations ............................................. 8
3.1 Overview of the energy distribution in laser–induced plasmas . . . . 58
5.1 Comparison predicted and experimental particle removal ............. 97
7.1 Boundary conditions for a fully developed turbulent pipe flow . . . . 136
7.2 Boundary conditions for a fully developed turbulent channel flow . . 139
7.3 Boundary conditions for a turbulent flow over a backward facing step 143
1.1 Transistor count of Intel and AMD microprocessors .......................... 2
1.2 Processing steps in a typical optical lithographic process .................... 3
1.3 Principle of immersion lithography ...................................................... 5
1.4 Simplified representation of the LELE and LFLE processes .................. 6
1.5 Image of the ASML NXE:3100 pre-production EUV tool ..................... 7
1.6 Cross section of an EUV photomask .................................................... 7
1.7 Protective pellicle underneath a photomask ......................................... 9
1.8 Laser-induced shock wave cleaning ..................................................... 10

2.1 Examples of several different types of plasmas .................................... 15
2.2 Classification diagram of a broad range of plasmas ............................ 16
2.3 Laser-induced plasma ignition mechanisms .......................................... 17
2.4 Normal shock wave in a shock–fixed coordinate system ...................... 24
2.5 Overview of the shock relations for an ideal diatomic gas .................... 26
2.6 Error introduced by the approximate shock relations ......................... 26
2.7 Photographs of the first atomic bomb test ........................................ 28
2.8 Self–similar velocity profile behind the shock wave ........................... 31
2.9 Self–similar density profile behind the shock wave ............................ 31
2.10 Self–similar pressure profile behind the shock wave .......................... 32

XI
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.11</td>
<td>Self–similar temperature profile behind the shock wave</td>
<td>32</td>
</tr>
<tr>
<td>3.1</td>
<td>Deflection of a light beam by refractive index gradients</td>
<td>40</td>
</tr>
<tr>
<td>3.2</td>
<td>Simple shadowgraphy experiment</td>
<td>43</td>
</tr>
<tr>
<td>3.3</td>
<td>Deflection of light passing through various refractive index fields</td>
<td>43</td>
</tr>
<tr>
<td>3.4</td>
<td>Simplified Schlieren system</td>
<td>44</td>
</tr>
<tr>
<td>3.5</td>
<td>Realistic Schlieren system</td>
<td>45</td>
</tr>
<tr>
<td>3.6</td>
<td>Sketch of the source images in the knife edge plane</td>
<td>46</td>
</tr>
<tr>
<td>3.7</td>
<td>Sketch of the laser flash shadowgraphy setup</td>
<td>47</td>
</tr>
<tr>
<td>3.8</td>
<td>Sketch of the Z–type Schlieren setup</td>
<td>49</td>
</tr>
<tr>
<td>3.9</td>
<td>Shadowgraphs of the ignition phase of a laser–induced plasma</td>
<td>50</td>
</tr>
<tr>
<td>3.10</td>
<td>Shadowgraphs of laser–induced shock waves in air</td>
<td>52</td>
</tr>
<tr>
<td>3.11</td>
<td>Plot of the the volume $V$ versus the time $t$ for air</td>
<td>53</td>
</tr>
<tr>
<td>3.12</td>
<td>Plot of the the volume $V$ versus the time $t$ for argon</td>
<td>54</td>
</tr>
<tr>
<td>3.13</td>
<td>Relation between $t^{6/5}$ and the volume $V$ for air</td>
<td>56</td>
</tr>
<tr>
<td>3.14</td>
<td>Relation between $t^{6/5}$ and the volume $V$ for argon</td>
<td>57</td>
</tr>
<tr>
<td>3.15</td>
<td>Schlieren photographs of the hot gas core</td>
<td>60</td>
</tr>
<tr>
<td>4.1</td>
<td>Overview of various intermolecular interactions</td>
<td>63</td>
</tr>
<tr>
<td>4.2</td>
<td>Geometry of a molecule and an infinitely large surface</td>
<td>67</td>
</tr>
<tr>
<td>4.3</td>
<td>Geometry of a sphere and an infinitely large surface</td>
<td>69</td>
</tr>
<tr>
<td>4.4</td>
<td>Geometry of a deformed sphere and an infinitely large surface</td>
<td>69</td>
</tr>
<tr>
<td>4.5</td>
<td>Overview of Hertz, JKR, and DMT contact mechanics</td>
<td>71</td>
</tr>
<tr>
<td>4.6</td>
<td>Pressure in the contact area between a sphere and a flat surface</td>
<td>73</td>
</tr>
<tr>
<td>4.7</td>
<td>Capillary condensation around the contact area</td>
<td>73</td>
</tr>
<tr>
<td>4.8</td>
<td>Sketch of an electric double layer</td>
<td>75</td>
</tr>
<tr>
<td>4.9</td>
<td>Comparison of the various adhesion forces</td>
<td>77</td>
</tr>
<tr>
<td>4.10</td>
<td>Drawing of the geometry used for the particle removal model</td>
<td>78</td>
</tr>
<tr>
<td>4.11</td>
<td>Reduced contact area of a rough particle and a rough surface</td>
<td>81</td>
</tr>
<tr>
<td>4.12</td>
<td>Flow diagram of the particle removal model</td>
<td>83</td>
</tr>
<tr>
<td>4.13</td>
<td>Results particle removal model with capillary forces</td>
<td>84</td>
</tr>
<tr>
<td>4.14</td>
<td>Results particle removal model without capillary forces</td>
<td>84</td>
</tr>
<tr>
<td>5.1</td>
<td>Sketch of the laser–induced shock wave cleaning setup</td>
<td>88</td>
</tr>
<tr>
<td>Figure Number</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.2</td>
<td>Clustering of 50 nm polystyrene latex spheres</td>
<td>90</td>
</tr>
<tr>
<td>5.3</td>
<td>KLA Tencor Surfscan SP2</td>
<td>91</td>
</tr>
<tr>
<td>5.4</td>
<td>Meandering cleaning path</td>
<td>92</td>
</tr>
<tr>
<td>5.5</td>
<td>Laser–induced shockwave cleaning of polystyrene latex spheres with varying number of pulses per position</td>
<td>94</td>
</tr>
<tr>
<td>5.6</td>
<td>Laser–induced shockwave cleaning of polystyrene latex spheres with varying laser pulse energy</td>
<td>95</td>
</tr>
<tr>
<td>5.7</td>
<td>Laser–induced shockwave cleaning of polystyrene latex spheres with varying gap distance</td>
<td>96</td>
</tr>
<tr>
<td>5.8</td>
<td>Laser–induced shockwave cleaning of silica particles with varying laser pulse energy</td>
<td>99</td>
</tr>
<tr>
<td>5.9</td>
<td>Types of damage caused by laser–induced shock wave cleaning</td>
<td>100</td>
</tr>
<tr>
<td>5.10</td>
<td>Results of the multilayer damage experiment</td>
<td>100</td>
</tr>
<tr>
<td>7.1</td>
<td>Drawings made by Leonardo da Vinci depicting turbulence</td>
<td>111</td>
</tr>
<tr>
<td>7.2</td>
<td>Paintings made by Vincent van Gogh depicting hints of turbulence</td>
<td>111</td>
</tr>
<tr>
<td>7.3</td>
<td>von Kármán vortex street near the Juan Fernandez Islands</td>
<td>114</td>
</tr>
<tr>
<td>7.4</td>
<td>Turbulent ash plume arising from Mount Cleveland</td>
<td>114</td>
</tr>
<tr>
<td>7.5</td>
<td>Reynolds decomposition</td>
<td>117</td>
</tr>
<tr>
<td>7.6</td>
<td>Formation of a turbulent boundary layer</td>
<td>124</td>
</tr>
<tr>
<td>7.7</td>
<td>Turbulent mixing</td>
<td>131</td>
</tr>
<tr>
<td>7.8</td>
<td>Simulated velocity profiles of fully developed turbulent pipe flows</td>
<td>137</td>
</tr>
<tr>
<td>7.9</td>
<td>Simulated TVR profiles of fully developed turbulent pipe flows</td>
<td>138</td>
</tr>
<tr>
<td>7.10</td>
<td>Simulated velocity profiles of fully developed channel turbulent flows</td>
<td>140</td>
</tr>
<tr>
<td>7.11</td>
<td>Simulated TVR profiles of fully developed turbulent channel flows</td>
<td>141</td>
</tr>
<tr>
<td>7.12</td>
<td>Plot of $y^+$ versus $v^+$ for fully developed turbulent channel flows</td>
<td>142</td>
</tr>
<tr>
<td>7.13</td>
<td>Geometry of a backward facing step</td>
<td>142</td>
</tr>
<tr>
<td>7.14</td>
<td>Simulation results of a turbulent flow over a backward facing step</td>
<td>144</td>
</tr>
<tr>
<td>7.15</td>
<td>Streamlines of a turbulent flow over a backward facing step</td>
<td>145</td>
</tr>
<tr>
<td>7.16</td>
<td>Wall velocity at the base of the backward facing step</td>
<td>146</td>
</tr>
</tbody>
</table>
List of Figures

XIV
CHAPTER 1

Introduction

Abstract. With the arrival of the next generation extreme ultraviolet lithography tools, printed features on integrated circuits will continue to shrink in the future. In the absence of a protective pellicle to shield the surface of the photomask from particles, the requirements for contamination control are getting more and more demanding. This is especially true for lithography, which is a crucial step in the manufacturing process of integrated circuits. In this chapter the lithography process is therefore examined in more detail, after which the influence of contaminations in the form of particles is shortly discussed. In light of these developments, a cleaning method called laser-induced shock wave cleaning is presented, which is a novel cleaning technique that is able to remove particles from a surface by subjecting them to laser-induced plasma shock waves. This chapter then concludes with a short overview of the contents of this thesis.
Figure 1.1: Number of transistors on Intel and AMD microprocessors during the past few decades.

1.1 Microlithography

Going through our daily lives, we often take for granted the advanced electronic devices that have taken a deep-seated place in our modern society. Not only personal computers, but also automobiles, smartphones, and many home appliances benefit from the development of the integrated circuit. By integrating a large collection of miniaturized electronic components onto a single small piece of silicon substrate, the integrated circuit made it possible to mass-produce small and reliable computing devices at relatively low costs.

To keep costs as low as possible, the semiconductor industry has always strived for further miniaturization of electronic components. Over the past few decades, this has led to a dramatic increase in the number of electronic components on an
1.1. Microlithography

1.1.1 Optical Lithography

Optical lithography systems are used extensively by the semiconductor industry to transfer the pattern of an integrated circuit layer from a photomask to a silicon substrate. This is a complex and delicate procedure involving many processing steps, but for the purpose of this thesis the simplified procedure outlined in figure 1.2 will...
suffice. The lithographic process starts by coating the silicon substrate with a very thin photosensitive film called resist. After coating the substrate with resist, light is passed through the transparent parts of a photomask and the resulting image is projected onto the coated substrate. The exposed parts of the resist film undergo a photochemical reaction, after which the exposed or unexposed parts are stripped from the substrate. The result is a patterned film of resist that can be used for further etching or deposition processes. The entire process can be repeated several times to create an integrated circuit that consists of multiple layers.

Two important aspects that limit component miniaturization in lithography, are the resolution and depth of focus of the projected image. Here, the resolution is defined as the minimum line width that can be projected onto the wafer, and the depth of focus gives the extent of the region around the image plane in which the image is properly focused. Both the resolution (RES) and the depth of focus (DOF) are bound by the laws of physics, and may be calculated with the following formulas [4]:

$$RES = k_1 \left( \frac{\lambda}{NA} \right),$$  \hspace{1cm} (1.1)

$$DOF = k_2 \left( \frac{\lambda n}{NA^2} \right),$$  \hspace{1cm} (1.2)

where $\lambda$ is the wavelength of the light, $n$ the refractive index of the medium in which the projection optics operate, and $k_1$ and $k_2$ are process dependent factors. The numerical aperture $NA$ is given by:

$$NA = n \sin \theta,$$  \hspace{1cm} (1.3)

where $\theta$ is the half-angle of the maximum cone of light that can exit the projection optics. Equation (1.1) shows that there are several possibilities to enhance the resolution: decrease the process dependent factor $k_1$, increase the numerical aperture $NA$, or decrease the wavelength $\lambda$.

### 1.1.2 Immersion Lithography

Today's high volume lithography tools use a technique called immersion to increase the numerical aperture of their projection optics. Immersion lithography tools are able to achieve a higher numerical aperture by introducing a liquid between the final element of the projection optics and the wafer, as is shown in the right half of figure 1.3. The introduction of a liquid between the lens and the wafer results in a higher numerical aperture and a better resolution than a projection system without immersion. Unfortunately, the concept of immersion lithography is almost
stretched to its limits. According to the ITRS\textsuperscript{1} \cite{1}, it is expected that the half-pitch\textsuperscript{2} of dynamic random-access memory will drop to 25 nm in the year 2015. To reach this ambitious technology node, and to keep pace with Moore’s Law in the foreseeable future, immersion technology is now being combined with multiple patterning techniques. However, after this period a more rigorous change is required to produce even smaller features. The most likely technology for these more demanding requirements is extreme ultraviolet lithography, which is discussed in more detail in section 1.1.4.

1.1.3 Multiple Patterning Lithography

*Multiple patterning* (MP) is a form of lithography that requires multiple exposures to create higher pattern densities than would be possible with a single exposure technique. This is achieved by splitting the photomask with the final high pattern density into several separate photomasks with a lower pattern density. In its most basic form it requires only two exposures to double the pattern density, and is therefore often termed *double patterning*. As figure 1.4 shows, the most straightforward method to achieve double patterning is the *litho–etch litho–etch* (LELE) technique. Here the first exposure is used to etch the first half of the pattern into the hard mask on top of the substrate. The second exposure is then used to transfer the remaining part of the pattern, after which the final pattern is etched into the substrate. However, this solution is very costly because it involves two exposures as well as two etching steps. To reduce costs, it is also possible to use the second technique

\textsuperscript{1}International Technology Roadmap for Semiconductors.

\textsuperscript{2}Half the distance between identical features.
in figure 1.4, which is called *litho-freeze litho-etch*. In this case the resist is chemically frozen after the first exposure, which reduces the number of etching steps from two to one. Besides these two techniques, there are several other double patterning techniques (e.g. *self-aligned double patterning* or *dual tone development*) available, which are discussed in more detail in [12, 8].

### 1.1.4 Extreme Ultraviolet Lithography

At this moment, *extreme ultraviolet* (EUV) lithography represents the pinnacle of lithographic technologies, and is considered to be the prime candidate for the next generation of lithography tools. Figure 1.5 shows a drawing of the ASML NXE:3100 pre-production EUV tool, which uses a plasma light source to produce extreme ultraviolet radiation with a much shorter wavelength of 13.5 nm to decrease the feature size even further. Unfortunately, the mean free path length of the extreme ultraviolet radiation is extremely small for most gases, and the entire system is therefore operated under vacuum conditions\(^3\). Furthermore, it is also necessary to use reflective multilayer optics instead of transmissive optics, because transmissive optics would absorb too much EUV radiation. As figure 1.6 shows, even the photomask itself is no longer transmissive, but has been replaced by a reflective multilayer mirror coated with a patterned absorber layer that represents the integrated circuit layer.

The multilayer mirrors in the EUV lithography tool are typically composed of 40 silicon and molybdenum bilayers to form a *Bragg reflector*, and have a reflectivity

\(^3\)The typical pressure at the exposure position is approximately \(10^{-2}\) mBar \(\text{H}_2\).
1.1. Microlithography

Figure 1.5: Image of the ASML NXE:3100 pre-production EUV tool.

Figure 1.6: Cross section of an EUV photomask consisting of a patterned absorber layer on top of a multilayer reflector. Legend:  
- Molybdenum/silicon multilayer mirror
- TaN layer
- TaNO layer
- Low thermal expansion substrate.
Introduction

<table>
<thead>
<tr>
<th>Particle contamination</th>
<th>Chemical contamination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin flakes</td>
<td>Lubricants</td>
</tr>
<tr>
<td>Fibres</td>
<td>Adhesives</td>
</tr>
<tr>
<td>Saliva</td>
<td>Cleaning agents</td>
</tr>
<tr>
<td>Cosmetics</td>
<td>Outgassing of plastics</td>
</tr>
<tr>
<td>Metal particles</td>
<td>Skin grease</td>
</tr>
<tr>
<td>Smoke</td>
<td>Transpiration</td>
</tr>
<tr>
<td>Pollen</td>
<td>Finger prints</td>
</tr>
<tr>
<td>Sand</td>
<td>Saliva</td>
</tr>
<tr>
<td></td>
<td>Cosmetics</td>
</tr>
</tbody>
</table>

Table 1.1: List of common particle and chemical contaminations. Some types of contamination fall into both categories.

of approximately 70% at the operating wavelength of 13.5 nm. When it is kept in mind that the light in the lithography tool is reflected more than ten times by multilayer mirrors, it becomes apparent that the total transmission of the projection optics is extremely low. The plasma light source is therefore required to produce a very high output power to compensate for the losses experienced by the multilayer optics. However, at this moment the available plasma light sources are not yet able to produce the required output power necessary for the high volume production of integrated circuits. This has delayed the readiness of EUV lithography, and the gap to the next technology node is currently being bridged with double patterning techniques as discussed in the previous section. While the output power of the light source remains a cause for concern, there are also other matters that need to be resolved, such as the production of defect free EUV photomasks and the development of better resists. Also the contamination of critical areas such as the photomask can pose a serious problem, and needs to be controlled to avoid costly production errors. This particular issue is addressed in more detail in the next section, where the focus is placed on the prevention, detection, and removal of unwanted contaminants from the surface of critical areas.

1.2 Contamination

In the semiconductor industry contamination control plays an important role to ensure a clean production process with as little defects as possible. In general, the three key aspects of any successful contamination control strategy are prevention, detection and cleaning. With the arrival of the next generation EUV lithography tools, the requirements for each of these three key aspects are becoming increasingly
demanding. In principle, there are two distinct types of contaminations: particles and chemical contaminations. Examples of particle contaminations are skin flakes or fibres, such as hair or clothing. Chemical contaminations are caused by exposure to contaminated gases or fluids, which can lead to film like deposits on exposed surfaces. Table 1.1 gives a list of examples of both particle and chemical contaminations. Note that some contaminations, such as cosmetics, fall into both categories of contamination.

1.2.1 Prevention and Detection

The presence of particles on the surface of the photomask can lead to serious imaging errors. To counteract this problem, very thin transparent pellicles are used to protect the surface of the photomask, as can be seen in figure 1.7. The pellicle is located a few millimeters underneath the photomask, which ensures that the particles do not come into contact with the photomask, while the particles on the pellicle are kept out of focus to avoid imaging errors. By using this strategy, only relatively large particles with a diameter of 5 μm and above pose a problem to the integrity of the projected image. Furthermore, it is also relatively straightforward to quickly and reliably detect these relatively large particles with e.g. direct optical inspection or some form of scatterometry. Unfortunately, it is almost impossible to apply this proven technology to EUV lithography, because a pellicle would absorb too much EUV radiation. This means that EUV photomasks probably need to be operated without a pellicle, thereby allowing particles to settle directly on the surface of the photomask itself. In this case, particles as small as 30 nm could already cause killer defects. In light of these developments, it is therefore paramount to have a quick and reliable cleaning method, that is capable of removing such small particles.
1.2.2 Laser–Induced Shock Wave Cleaning

The main subject of this thesis is laser–induced shock wave cleaning, which was invented more than two decades ago [11], and has been investigated by the groups of Lee [6], Watkins [5] and Cetinkaya [3, 2, 10]. Recently it has attracted a renewed interest as a cleaning method for EUV photomasks, because the conventional cleaning methods (wet cleaning, chemical cleaning, megasonics, and CO₂ snow cleaning) are having difficulties to meet the new cleaning requirements. Laser–induced shock wave cleaning removes particles by exposing them to a fast moving shock wave generated by a laser–induced plasma in a gaseous medium. Laser–induced breakdown is achieved by exposing a pocket of gas to the field of a very intense electromagnetic wave by focusing the light of a high energy laser pulse into a very small volume. The plasma that is formed has an extremely high pressure, which causes it to expand at a supersonic velocity. The shock wave that is generated in the process impinges on the surface and exerts a cleaning force on any particles that may have settled on it. It is important to note, that the plasma pocket is created above the surface at a certain gap distance \( g \), as can be seen in figure 1.8.

1.3 Outline of This Thesis

The main subject of this thesis is to determine the feasibility of laser–induced shock wave cleaning as a method to remove unwanted particles from EUV photomasks. In the next chapter this investigation starts by studying the properties and formation mechanisms of laser–induced plasmas, followed by a discussion of shock theory and the Taylor & Sedov solution for an intense point explosion to describe the shock waves produced by the expanding plasma pocket. In chapter 3 this theory is val-
idated after investigating the formation and propagation of the shock wave with the help of high speed imaging techniques. This is followed by a discussion of the adhesive properties and contact mechanics of small submicron particles in chapter 4. Combining particle adhesion theory with shock theory then leads to the development of a numerical model that is capable of predicting the cleaning performance of laser–induced shock wave cleaning as a function of its main cleaning parameters. In chapter 5 the particle removal capabilities of laser–induced shock wave cleaning are experimentally tested and compared to the particle removal model developed earlier. Chapter 5 also addresses the issue of surface damage, which is evaluated by optical inspection of multilayer mirrors. The investigation is then concluded with chapter 6, which contains both the conclusions as well as some remarks and further recommendations.

While the results presented in this thesis support the use of the Taylor & Sedov solution, it would also be interesting to perform a more complete numerical ab initio simulation of the laser–induced plasma. Such a numerical simulation can provide much more insight into the properties of the laser–induced plasma, and would need to encompass several important components, such as the propagation, deflection and absorption of the laser beam, as well as the creation and expansion of the plasma. Furthermore, it is also important to include turbulence into such a simulation, since it may have a profound impact on the transport of particles after they have been removed from the surface by the blast wave. As a first step towards such a numerical model, chapter 7 discusses these matters in more detail, and places particular emphasis on the implementation of the $k$–$\varepsilon$ turbulence model into the plasma simulation toolkit Plasimo.

References


Introduction


CHAPTER 2

Laser Induced Breakdown in Gases

Abstract. In this chapter a theoretical investigation is made to identify the fundamental parameters that play an important role in the removal of submicron particles by laser–induced shock waves. Several aspects of laser–induced plasmas are examined, such as the basic mechanisms involved in the ignition and heating of the plasma. Particular emphasis has been placed on the dynamics of the gas after breakdown, and it is theorized that the gas behaves as predicted by the self–similar Taylor & Sedov solution of an intense point explosion. By combining shock theory and the Taylor & Sedov solution, it is shown that the strength of the shock wave depends on three key parameters: the gap distance, the laser pulse energy absorbed during laser–induced breakdown, and the specific heat capacity ratio of the gas. These parameters are also expected to have a profound influence on the cleaning capabilities of laser–induced shock wave cleaning.
2.1 Laser Induced Plasmas

The formation of laser-induced plasmas in gases is of crucial importance to the cleaning method described in this thesis. Due to its high pressure, the small laser-induced plasma expands at a phenomenal rate, thereby producing a shock wave that can be used to remove small particles from a surface. But before concerning ourselves with the properties of the shock wave, it is necessary to take a closer look at the laser-induced plasma itself. After a short informal introduction into the different types of plasmas, the focus is placed on the various ignition and breakdown mechanisms that are responsible for the creation of laser-induced plasmas.

2.1.1 Plasma Basics

Plasmas are ionized gases, and are therefore characterized by the presence of charged particles, such as ions and electrons. Perhaps without knowing it, many of us encounter plasmas on a regular basis in our daily lives. The sun is probably the most well-known example of a plasma, but there are plenty of other examples as can be seen in figure 2.1. Because of their unique physical properties, plasmas are often viewed as the fourth state of matter. The presence of charged particles is what gives the plasma its unique properties, and sets it apart from ordinary non-ionized gases. The separation of charged particles in the plasma produces electric fields, and the flow of charged particles gives rise to currents and magnetic fields. These electric and magnetic fields allow the charged particles in the plasma to interact over much longer distances than the particles in ordinary non-ionized gasses.

Plasmas may be produced in a number of ways, i.e. through electrical discharge, microwave radiation, or nuclear fusion like in the center of the sun. The laser-induced plasmas that are discussed in this chapter are somewhat more exotic than the examples shown in figure 2.1. These man-made plasmas are created by exposure to intense laser radiation, and are usually the exclusive domain of scientific research laboratories.

Generally speaking, the most important parameter for any plasma is the electron density \( n_e \). The second most important parameter is the electron temperature \( T_e \), which is often expressed in electron volts \((1 \text{ eV} = 1.16 \cdot 10^4 \text{ K})\). With these two parameters it is possible to classify the various types of plasma, as can be seen in figure 2.2. The classification diagram shows that laser-induced plasmas have a relatively high electron density and a moderate electron temperature.
2.1. Laser Induced Plasmas

Figure 2.1: Examples of several different types of plasmas. (a) Image of the chromosphere of the sun taken by the Japanese solar observation satellite Hinode. (b) Example of an energy saving fluorescent lamp. (c) Picture of a lightning storm showing multiple discharges. (d) Image of a neon sign emitting its characteristic red light. (e) Flame of a burning candle. (f) Picture of the aurora borealis over the skies of Alaska.
Figure 2.2: Classification diagram of a broad range of plasmas according to their electron density and electron temperature [9].
2.1. Laser Induced Plasmas

2.1.2 Plasma Ignition

In general there are two mechanisms responsible for the ignition of laser-induced plasmas: multi photon ionization and tunneling. Both processes are self sufficient, which means that they do not require the presence of free electrons. This is of particular importance, since the probability that a free electron is produced by natural background radiation in the focal volume during the arrival of the laser pulse is negligible [7]. Multi photon ionization and tunneling therefore play an important role in the liberation of the first free electrons, after which more efficient mechanisms such as inverse Bremsstrahlung become more predominant.

Multi Photon Ionization

In most circumstances it is not possible to produce free electrons by direct photoionization, because the photon energy of the laser radiation is typically less than the ionization energy. The formation of the first free electrons in the focal volume must therefore involve the absorption of multiple photons, hence the name multi photon ionization. Figure 2.3 shows a schematic depiction of the multi photon ionization process. After absorbing one photon, the electron makes the transition from the ground state into a virtual state. The lifetime of this first virtual state is extremely short (\( \sim 10^{-15} \) s), and can be calculated with Heisenberg’s uncertainty principle, which states that it is impossible to determine both the energy and the lifetime of the virtual state at exactly the same moment [4]:

\[
\Delta t_1 \Delta E_1 \leq \frac{\hbar}{4\pi}, \tag{2.1}
\]
where $h$ is Planck’s constant, and $\Delta t_1$ and $\Delta E_1$ represent the uncertainties in the lifetime and energy of the first virtual state. When the uncertainty $\Delta E_1$ is taken equal to the energy of the absorbed photon, it is possible to rewrite equation (2.1) into the following form:

$$\Delta t_1 \leq \frac{h}{4\pi h\nu} = \frac{1}{4\pi \nu},$$

(2.2)

where $h\nu$ is the photon energy, and $\nu$ the frequency of the photon. In other words, this result shows that the virtual state with energy $\Delta E_1$ is allowed to exist on the condition that it disappears again after a time $\Delta t_1$. The electron in the first virtual state can be excited to the second virtual state by absorbing another photon. The lifetime of the second virtual state is however different from the first virtual state:

$$\Delta t_2 \leq \frac{h}{8\pi h\nu} = \frac{1}{8\pi \nu}.$$  

(2.3)

For a high enough photon flux this process can continue until the electron has gained enough energy to reach the ionized state. In short, the multi photon ionization process may be summarized by the following reaction scheme:

$$A + k h\nu \rightarrow A^+ + e^-,$$

(2.4)

where $A$ represents a neutral gas atom, $h\nu$ the energy of a single photon, $A^+$ a positively charged ion, $e^-$ an electron, and $k$ the minimum number of photons necessary to achieve multi photon ionization. The energy of the electron after multi photon ionization is equal to $kh\nu - E_i$, where $E_i$ is the ionization potential.

For multi photon ionization to occur, an extremely high photon flux is needed to supply each virtual state with a photon during its limited lifetime. When the photon flux drops below a certain threshold value, it becomes impossible to reach the ionized state. The threshold photon flux for multi photon ionization can be estimated with the following expression [7]:

$$F_{\text{th}} = \frac{\nu}{\sigma \left[ \frac{\delta (k - 1)!}{\nu \tau} \right]^{1/k}},$$

(2.5)

where $\sigma$ is the photon absorption cross section, $\delta$ the ionization fraction characteristic for breakdown, and $\tau$ the duration of the laser pulse. Filling in typical values for each variable ($\nu \sim 10^{15}$ s$^{-1}$, $\sigma \sim 10^{-16}$ cm$^2$, $\delta \sim 10^{-3}$, $k \sim 10$, $\tau \sim 10^{-8}$ s) gives a threshold photon flux of $10^{30}$ cm$^{-2}$ s$^{-1}$, which may be produced by focusing the beam of a high energy laser pulse.
2.1. Laser Induced Plasmas

Tunneling

The second ionization mechanism that is of importance to the ignition of laser-induced plasmas is tunneling. Tunneling may occur when the electric field of the laser radiation is strong enough to shift the energy levels of the bound states. This energy shift is known as the AC Stark shift, and its influence on the most weakly bound states can be significant. The result is a tilted atomic potential as shown in figure 2.3, which makes it possible for an electron to tunnel through the potential barrier from the ground state to the ionized state. In this case the reaction scheme is rather simple:

\[ A \rightarrow A^+ + e^-, \]  

(2.6)

where the laser radiation is not an integral part of the reaction scheme, but creates the conditions suitable for this reaction to occur.

Tunneling typically occurs at much higher laser intensities than multi photon ionization. To determine which of the two mechanisms is the most important, it is often useful to examine the Keldysh parameter, which is defined as the square root of the ratio of the ionization potential and the ponderomotive energy of the electron:

\[ \Gamma = \frac{\omega}{e\mathcal{E}} \sqrt{2mE_i}, \]  

(2.7)

where \( \omega \) is the angular frequency of the laser radiation, \( e \) the electron charge, \( \mathcal{E} \) the peak electric field strength, \( m \) the electron mass, and \( E_i \) the ionization potential. When the Keldysh parameter is much larger than unity, multi photon ionization is the dominant ionization mechanism. For values much smaller than unity, the electron has enough energy to cross the potential barrier before the electric field of the laser radiation changes its sign and magnitude. In this case tunneling is the more predominant ionization mechanism.

2.1.3 Plasma Heating

Even though multi photon ionization and tunneling are capable of creating free electrons, their efficiency is usually not high enough to cause complete breakdown of the gas. To achieve this, an additional mechanism called inverse Bremsstrahlung is needed to efficiently heat the free electrons in the plasma. Once the process of inverse Bremsstrahlung becomes efficient, the electrons may become energetic enough to ionize other neutral gas atoms by impact. Both mechanisms are discussed below.
After the first electrons have been liberated by multi photon ionization and the tunneling effect, the process of inverse Bremsstrahlung becomes more dominant. Inverse Bremsstrahlung increases the electron temperature of the free electrons in the plasma, and is characterized by the following reaction scheme:

\[ A^+ + e^- + h\nu \rightarrow A^+ + e^- \]  

(2.8)

Please note that the above reaction takes place when the electron absorbs a photon during a collision with an ion. The presence of the ion during the absorption of the photon is crucial in the process for reasons of momentum conservation. The strength of the inverse Bremsstrahlung absorption depends on the properties of the plasma and the laser radiation. The amount of intensity lost due to inverse Bremsstrahlung at each point along the beam path can be calculated with the following expression:

\[ dI(z) = -\alpha(\lambda, z) I(z) \, dz, \]  

(2.9)

where \( I(z) \) is the intensity of the laser radiation at position \( z \) along the beam path. The absorption coefficient \( \alpha \) for inverse Bremsstrahlung is given by [2]:

\[ \alpha = \frac{13.49}{\lambda^2} \bar{Z} \left( \frac{n_e}{n_c} \right)^2 \frac{\ln \Lambda}{\sqrt{1 - \frac{n_e}{n_c}}} \frac{1}{T_e^{3/2}}, \]  

(2.10)

where \( \lambda \) is the wavelength, \( \bar{Z} \) the average charge state, \( n_e \) the electron density, \( \Lambda \) the Debye number\(^1\), and \( T_e \) the electron temperature in Kelvin. The critical electron density \( n_c \) is given by:

\[ n_c = \frac{\epsilon_0 m \omega^2}{e^2}, \]  

(2.11)

where \( \epsilon_0 \) is the permittivity of vacuum, \( m \) the mass of the electron, \( \omega \) the angular frequency of the laser radiation, and \( e \) the charge of the electron. Equation (2.10) clearly shows that the expression for the absorption coefficient becomes singular when \( n_e = n_c \). When this happens the plasma becomes opaque, and much of the laser radiation is no longer absorbed but reflected, which greatly diminishes the possibility to couple laser power into the plasma.

\(^1\)The Debye number is the number of particles within a sphere of radius the Debye length, where the Debye length is defined as the distance at which the free electrons in the plasma are able to shield the positive charge of an ion.
Impact Ionization

By absorbing photons through inverse Bremsstrahlung the electrons are heated until they have gained enough energy to ionize other atoms by impact. This process is called *impact ionization* and can be summarized as follows:

\[ A + e^- \rightarrow A + e^- + e^- \quad (2.12) \]

Strictly spoken this is not a heating mechanism, but this process does liberate more electrons, whose temperature may be increased by inverse Bremsstrahlung until they have also gained enough energy to cause ionization by impact. Under the right conditions this will lead to an exponential growth of free electrons, and this effect is often referred to as *cascade ionization*.

2.2 Shock Wave Theory

The interaction with the shock wave produced by the laser–induced plasma provides the actual cleaning force by which the particles are removed from the surface. It is therefore necessary to investigate the properties and behaviour of shock waves, starting with a short introduction to *gas dynamics*, followed by a more in–depth discussion of the *Euler equations* and the *shock relations*.

2.2.1 Gas Dynamics

In general, the transport of heat and mass by fluids is governed by conservation laws, which describe the conservation of mass, momentum and energy. Under most circumstances it is possible to neglect the compressibility of the fluid, which simplifies the system of equations considerably by decoupling the energy equation from the mass and momentum equations. This assumption is mostly valid for liquids, but also for gases flowing at a relatively low velocity. Gases flowing at higher velocities on the other hand, experience considerable pressure gradients, resulting in significant density variations that may not be neglected. Under these circumstances the transport and thermodynamics of the gas are intimately linked, and it is not possible to decouple the energy equation from the mass and momentum equation. Since these conditions usually apply to gasses, the related field of research is called *gas dynamics*.

The compressibility of gasses gives rise to interesting phenomena, which are not present in incompressible fluids. The most interesting phenomenon pertaining to this thesis is the existence of *shock waves*, which are propagating disturbances characterized by a nearly stepwise change in the properties of the gas. Shock waves
are formed when the gas undergoes a velocity change that exceeds the *speed of sound*, which is defined as the velocity at which small infinitesimal disturbances propagate through the gas. The speed of sound is closely related to the mean velocity of the gas molecules, and is therefore relatively insensitive to the pressure or density, but depends strongly on the temperature and gas species. For a perfect gas the speed of sound $a$ is given by:

$$a = \sqrt{\frac{\gamma R T}{\rho}},$$

where $\gamma$ is the specific heat ratio, $R$ the specific gas constant, and $T$ the temperature in Kelvin. For many gas dynamical problems it is often convenient to express the velocity in terms of the speed of sound. The resulting dimensionless number is called the *Mach number*, and is given by:

$$M = \frac{v}{a},$$

where $M$ is the Mach number, and $v$ the velocity. For velocities above the speed of sound ($M > 1$) the flow is said to be *supersonic*, whereas flows with a velocity below the speed of sound ($M < 1$) are termed *subsonic*. As mentioned earlier, shock waves only occur at velocities above the speed of sound, and are thus exclusive to supersonic flows only.

### 2.2.2 Euler Equations

As mentioned in the previous section, the transport of mass, momentum and energy is governed by conservation laws. For Newtonian fluids these conservation laws take the following form:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0,$$

$$\frac{\partial}{\partial t} (\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \mathbf{v}) = -\nabla p + \nabla \cdot \mathbb{S},$$

$$\frac{\partial}{\partial t} (\rho e) + \nabla \cdot (\rho \left(h + \frac{1}{2} \mathbf{v} \cdot \mathbf{v}\right) \mathbf{v}) = -\nabla \cdot \mathbf{q} + \nabla \cdot (\mathbf{v} \cdot \mathbb{S}),$$

where $\rho$ is the density, $\mathbf{v}$ the velocity, $p$ the pressure, $\mathbb{S}$ the deviatoric stress tensor, $e$ the specific total energy, $h$ the specific enthalpy, $\mathbf{q}$ the heat flux vector, $\nabla$ the vector differential operator, and $t$ the time. For typical situations encountered in gas dynamics, it is customary to use a simplified version of these conservation laws, in which the viscosity and heat conduction have been neglected. This can be understood by examining the *Reynolds number* and the *Péclet number*. The Reynolds
number is defined as the ratio of inertial forces and viscous forces:

\[ \text{Re} = \frac{\rho v L}{\mu}, \]  

where \( v \) is the velocity, \( L \) the characteristic length scale of the flow under study, and \( \mu \) the viscosity. The Péclet number is defined as the ratio of advection of heat and conduction of heat:

\[ \text{Pe} = \frac{v L}{k}, \]

where \( k \) is the heat conduction coefficient. For the fast majority of gas dynamical problems both the Reynolds and the Péclet number are much larger than unity, which means that the viscosity as well as the heat conduction play a role of negligible importance. When the viscosity and heat conduction are neglected, the conservation laws may be reduced to the much simpler set of Euler equations:

\[ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \]  
\[ \frac{\partial}{\partial t} (\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \mathbf{v}) + \nabla p = 0, \]  
\[ \frac{\partial \rho e}{\partial t} + \nabla \cdot (\rho (h + \frac{1}{2} \mathbf{v} \cdot \mathbf{v}) \mathbf{v}) = 0. \]

It should be remembered however, that the Euler equations are not valid in close proximity to walls and shock waves, where the Reynolds and Péclet numbers are not much larger than unity. In such cases a more complete description including viscosity and heat conduction is required, as is shown in the following subsection.

### 2.2.3 Normal Shock Relations

During laser induced breakdown a small but highly pressurized plasma is formed in the focal volume of the lens. Due to its high pressure the plasma will expand into the surrounding gas with a velocity well above the speed of sound, which means that the gas is unable to react to the expanding plasma before it arrives. The surrounding gas is therefore swept up by the expanding plasma and compressed into a thin shell at the plasma boundary. The thin compressed shell of gas is called a shock wave, and is characterized by very abrupt changes in the properties of the gas.

In this work the discussion of shock waves is mostly limited to normal shock waves, whose orientation is normal to the direction of propagation. Figure 2.4 shows an example of a normal shock wave in a shock–fixed coordinate system. By applying the conservation laws of mass, momentum, and energy to the control volume drawn in figure 2.4, it is possible to derive a number of very useful relationships. For a
steady flow without gravitational forces, the conservation laws take the following integral form:

\[ \int_A \rho (\mathbf{v} \cdot \mathbf{n}) \, dA = 0, \quad (2.23) \]
\[ \int_A \rho \mathbf{v} (\mathbf{v} \cdot \mathbf{n}) \, dA = \int_A \mathbf{t} \, dA, \quad (2.24) \]
\[ \int_A \rho (\mathbf{v} \cdot \mathbf{n}) \left( \frac{1}{2} \mathbf{v} \cdot \mathbf{v} + e \right) \, dA = \int_A - (\mathbf{q} \cdot \mathbf{n}) \, dA + \int_A (\mathbf{t} \cdot \mathbf{v}) \, dA, \quad (2.25) \]

where \( \mathbf{n} \) is the unit vector normal to the control surface \( A \). The surface force \( \mathbf{t} \) acting on the control surface \( A \) may be expressed as follows:

\[ \mathbf{t} = \mathbf{n} \cdot \sigma, \quad \text{with} \quad \sigma = -p^\parallel + S, \quad (2.26) \]

where \( \sigma \) is the stress tensor, \( p \) the pressure, and \( S \) the deviatoric stress tensor. Here it is assumed that the properties of the gas jump discontinuously at the shock wave, and that they have steady uniform values in the regions immediately before and after the shock wave. This ensures that the heat flux \( \mathbf{q} \) and the stress tensor \( S \) are equal to zero, which reduces equations (2.23) to (2.25) to the following form:

\[ \rho_0 v_0 = \rho_1 v_1, \quad (2.27) \]
\[ p_0 + \rho_0 v_0^2 = p_1 + \rho_1 v_1^2, \quad (2.28) \]
\[ h_0 + \frac{1}{2} v_0^2 = h_1 + \frac{1}{2} v_1^2, \quad (2.29) \]

where the specific enthalpy \( h \) is given by \( h = e + p/\rho \). Equations (2.27) to (2.29)
can be condensed into one single equation called the Rankine–Hugoniot equation:

\[ 2 \left( h_1 - h_0 \right) = \left( \frac{1}{\rho_1} + \frac{1}{\rho_0} \right), \quad (2.30) \]

which is valid for arbitrary gasses. In addition to these conservation laws, it may also be assumed that the gas behaves as a perfect gas for shock waves typically encountered in strong explosions [8]. Perfect gasses have constant specific heat capacities \( c_p \) and \( c_v \), which leads to the following equation of state for a perfect gas:

\[ h = \frac{\gamma}{\gamma - 1} \left( \frac{p}{\rho} \right), \quad (2.31) \]

where \( \gamma = c_p/c_v \) is the specific heat capacity ratio. Combining equations (2.30) and (2.31) then yields the Rankine–Hugoniot relations for perfect gasses:

\[ \frac{p_1}{p_0} = \frac{\gamma + 1}{\gamma - 1} \frac{\rho_1}{\rho_0}, \quad (2.32) \]

\[ \frac{\rho_1}{\rho_0} = \frac{1 + \frac{\gamma + 1}{\gamma - 1} \frac{p_1}{p_0}}{\frac{\gamma + 1}{\gamma - 1} + \frac{p_1}{p_0}}, \quad (2.33) \]

Interestingly, the shock adiabat described by the Rankine–Hugoniot relations has an asymptote at \( \rho_1/\rho_0 = (\gamma + 1) / (\gamma - 1) \). The density ratio \( \rho_1/\rho_0 \) therefore experiences an upper limit regardless of the strength of the shock wave. In practice this means that \( \rho_1/\rho_0 < 4 \) for an ideal monoatomic gas \( (\gamma = 5/3) \) and \( \rho_1/\rho_0 < 6 \) for an ideal diatomic gas \( (\gamma = 7/5) \).

Several more interesting relationships known as the normal shock relations may be derived from the Rankine–Hugoniot relationships and conservation laws. When combined with equation (2.14) for the Mach number, the following relationships are found:

\[ \frac{p_1}{p_0} = \frac{2\gamma M_0^2 - (\gamma - 1)}{\gamma + 1}, \quad (2.34) \]

\[ \frac{\rho_1}{\rho_0} = \frac{(\gamma + 1) M_0^2}{(\gamma - 1) M_0^2 + 2}, \quad (2.35) \]

\[ \frac{T_1}{T_0} = \frac{(2\gamma M_0^2 - (\gamma - 1)) \left( (\gamma - 1) M_0^2 + 2 \right)}{(\gamma + 1)^2 M_0^2}, \quad (2.36) \]

where \( M_0 \) is the upstream Mach number. As predicted earlier, figure 2.5 shows that the density ratio is indeed limited to an asymptotical value of \( (\gamma + 1) / (\gamma - 1), \)
Figure 2.5: Overview of the shock relations (2.34) to (2.36) for an ideal diatomic gas with a specific heat capacity ratio $\gamma = 7/5$.

Figure 2.6: Overview of the relative error introduced by the approximate shock relations (2.38) to (2.40) for an ideal diatomic gas with a specific heat capacity ratio $\gamma = 7/5$. 
whereas the pressure and temperature ratios are allowed to reach much higher values. Furthermore, it is possible to find the following expression for the downstream Mach number:

\[ M_1^2 = \frac{(\gamma - 1) M_0^2 + 2}{2\gamma M_0^2 - (\gamma - 1)}, \] (2.37)

from which it may be concluded that \( M_1 < 1 \). The flow behind a shock wave is therefore always subsonic. For very strong shock waves the shock relations may be approximated by:

\[ \frac{p_1}{p_0} = \frac{2\gamma M_0^2}{\gamma + 1}, \] (2.38)

\[ \frac{\rho_1}{\rho_0} = \frac{\gamma + 1}{\gamma - 1}, \] (2.39)

\[ \frac{T_1}{T_0} = \frac{2\gamma (\gamma - 1) M_0^2}{(\gamma + 1)^2}, \] (2.40)

which are reasonably accurate for high Mach numbers as can be seen in figure 2.6. Especially equation (2.38) for the pressure ratio \( p_1/p_0 \) is already accurate to within 1% for \( M_0 > 4 \). Finally, it should be noted that the shock relations are only applicable to supersonic flows for which \( M_0 \geq 1 \). This can be illustrated by examining the following expression for the entropy change across the shock wave [5]:

\[ s_1 - s_0 = c_p \ln \left( \frac{T_1}{T_0} \right) - \tilde{R} \ln \left( \frac{p_1}{p_0} \right), \] (2.41)

where \( s \) is the entropy. For subsonic flows (\( M_0 < 1 \)) this equation yields a negative change in entropy, which is in violation of the second law of thermodynamics.

### 2.3 Physics of an Intense Point Explosion

In the early autumn of 1940, at the beginning of the Second World War, the English physicist and mathematician Sir Geoffrey Ingram Taylor was approached by the British MAUD\(^2\) Committee to discuss the mechanical effects of a bomb, in which a very large amount of energy would be released by nuclear fission [1]. Less than a year later, Taylor had solved the problem and submitted his paper on June 27th 1941. In his paper Taylor had reduced the highly complex problem of a nuclear explosion to the much simpler idealized problem of an intense point explosion. Through the clever application of dimensional and similarity methods, Taylor was able to derive

\(^2\)MAUD = Military Application of Uranium Detonation
an expression for the radius of the shock wave, as well as a solution for the gas flow behind the shock wave. After the Second World War the same solution was also discovered by the Soviet scientist L.I. Sedov [10], and the solution is therefore often called the *Taylor & Sedov solution*.

Several years later in 1950, Taylor was proven to be right when the United States department of Defense declassified photographs of the first atomic bomb test. By applying his solution to the photographs shown in figure 2.7, Taylor was able to determine that the explosive yield of the first atomic bomb was equivalent to 17 ktons of TNT [11, 12]. As it turned out, his results were very close to the official explosive yield of 20 ktons of TNT, which also included the energy lost due to radiation. At the time this was still classified information, and Taylor’s publication caused a fair amount of turmoil as the United States and the Soviet Union embarked on a path to nuclear armament.

---

3 Only three days after Taylor published his results, the American mathematician *John von Neumann* also published his paper. Von Neumann’s results complemented Taylor’s solution, and it would be more appropriate to speak of the Taylor–von Neumann–Sedov solution.
of escalating Cold War tensions.

At a first glance it might seem that the detonation of a nuclear device and the laser induced–breakdown of gases do not have much in common. But the laser–induced breakdown of gases is accomplished by the sudden release of a relatively large amount of energy within a very small volume. These characteristics match almost perfectly with the description of an intense point explosion, and it is therefore not surprising that the same basic principles may also be applied to the laser–induced breakdown of gases, albeit on a much smaller scale.

2.3.1 Taylor & Sedov Solution

Explosions are very complex phenomena that can be created in a number of ways, each with its own properties and peculiarities. In practice it is nearly impossible to find a detailed analytical solution for each type of explosion, and in order to reduce the problem to manageable proportions it makes sense to make the following assumptions:

- A finite amount of energy is suddenly released in an infinitely concentrated form.
- The maximum pressure of the gas at the shock wave is much larger than the pressure of the ambient gas.

The first assumption defines the explosion as the instantaneous release of a finite amount of energy within a point. This assumption is particularly useful, since it reduces explosions of different types to the same basic problem. The second assumption makes it possible to neglect the pressure of the surrounding gas, which proves to be very useful in the derivation of the Taylor & Sedov solution.

Under the above assumptions, the shock wave radius $R$ of an idealized point explosion depends only on the energy $E$ liberated during the explosion, the density $\rho_0$ of the surrounding gas, the time $t$, and the specific heat capacity ratio $\gamma$:

$$R = R(E, \rho_0, t, \gamma).$$

(2.42)

The precise form of the function $R(E, \rho_0, t, \gamma)$ may be determined through the application of dimensional methods. The units of each of the dimensional parameters on which the function $R$ depends may be decomposed into fundamental units:

$$[E] = ML^2T^{-2},$$

(2.43)

$$[\rho_0] = ML^{-3},$$

(2.44)

$$[t] = T,$$

(2.45)
Laser Induced Breakdown in Gases

where M, L, and T are the units of mass, length and time. Dimensional analysis dictates that dimensional formulas of physical quantities must be monomial powers [10], which leads to the following expression:

\[ R \sim E^a \rho_0^b t^c, \tag{2.46} \]

where \( a, b, \) and \( c \) are exponents yet to be determined. By combining equations (2.43) to (2.46), it is possible to find the following expression:

\[ [R] = M^{a+b}L^{2a-3b}T^{-2a+c} = L, \tag{2.47} \]

from which it may be concluded that \( a = 1/5, \quad b = -1/5 \) and \( c = 2/5 \). Substituting these exponents back into equation (2.46) yields:

\[ R \sim t^{2/5} \left( \frac{E}{\rho_0} \right)^{1/5}, \tag{2.48} \]

which describes the shock wave radius to within a nondimensional proportionality constant. Until now the specific heat capacity ratio \( \gamma \) has not entered the problem yet, mainly because it is a nondimensional quantity. The proportionality constant is therefore dependent on \( \gamma \), which gives:

\[ R = S(\gamma) t^{2/5} \left( \frac{E}{\rho_0} \right)^{1/5}, \tag{2.49} \]

where the exact value of the proportionality constant \( S(\gamma) \) can be determined from energy considerations. This is done in the next section, but for now it suffices to know that the value of \( S(\gamma) \) is close to unity. Equation (2.49) now fully describes the shock wave radius as a function of time, and is called the Taylor & Sedov solution.

\[ \text{2.3.2 Post Shock Wave Solution} \]

In the previous section the Taylor & Sedov solution was derived purely on the basis of dimensional arguments. To solve the flow within the shock wave a more rigorous
2.3. Physics of an Intense Point Explosion

Figure 2.8: Self-similar velocity profile behind the shock wave.

Figure 2.9: Self-similar density profile behind the shock wave.
Figure 2.10: Self-similar pressure profile behind the shock wave.

Figure 2.11: Self-similar temperature profile behind the shock wave.
approach is needed, which involves solving the radially symmetric Euler equations:

\[
\begin{align*}
\frac{\partial \rho}{\partial t} + \frac{2\rho v}{r} + \frac{\partial}{\partial r} (\rho v) &= 0, \\
\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial r} + \frac{1}{\rho} \frac{\partial p}{\partial r} &= 0, \\
\frac{\partial}{\partial t} \left( \frac{p}{\rho^\gamma} \right) + v \frac{\partial}{\partial r} \left( \frac{p}{\rho^\gamma} \right) &= 0,
\end{align*}
\]

(2.50) (2.51) (2.52)

where \( v \) represents the radial velocity. In this case the boundary conditions at the shock wave are given by equations (2.27) to (2.29), and the initial starting conditions take the following form:

\[
\rho (r, 0) = \rho_0, \quad p (r, 0) = p_0, \quad v_r (r, 0) = 0.
\]

(2.53)

This system of equations may be solved by assuming that the motion of the gas is self-similar, which means that only the scaling of the spatial distributions vary with time, and not the shape itself. Introducing the dimensionless radial coordinate \( \xi = r/R \), the foregoing may be summarized as follows:

\[
p = p_1 (t) \tilde{p} (\xi, \gamma), \quad \rho = \rho_1 (t) \tilde{\rho} (\xi, \gamma), \quad v = v_1 (t) \tilde{v} (\xi, \gamma),
\]

(2.54)

where \( p_1, \rho_1 \) and \( v_1 \) are the pressure, density and velocity directly behind the shock wave. Furthermore, \( \tilde{p}, \tilde{\rho}, \) and \( \tilde{v} \) are dimensionless functions that define the shape of their respective spatial distributions. Substituting equation (2.54) into the radially symmetric Euler equations results in a system of three ordinary first-order differential equations for the functions \( \tilde{p}, \tilde{\rho}, \) and \( \tilde{v} \). The process of solving this system of ordinary differential equations is a lengthy and nontrivial task, and the solution procedure is therefore not included in this thesis. The full solution and the resulting expressions may be found in [6, 10], but for the purpose of this thesis the end results shown in figures 2.8 to 2.11 should be sufficient.

Now that the pressure, density and velocity distributions within the shock wave are known, it is possible to calculate the value of the proportionality constant \( S (\gamma) \) in equation (2.49) by solving the following energy equation:

\[
E = \int_0^R \frac{\rho v^2}{2} 4\pi r^2 \, dr + \int_0^R \frac{p}{\gamma - 1} 4\pi r^2 \, dr,
\]

(2.55)

where the first term represents the kinetic energy, and the second term the internal or thermal energy. This equation is often solved numerically, but it is also possible to find an analytical expression for \( S (\gamma) \) by using a few simple approximations.
In this case it is assumed that the fast majority of the gas is compressed into an infinitesimal thin shell directly behind the shock wave. The kinetic energy of the gas in the shell is then equal to \( \frac{1}{2} \mathcal{M} v_1^2 \), where \( \mathcal{M} = \frac{4\pi}{3} R_0^3 \rho_0 \) is the mass of the shell, and \( v_1 \) the velocity of the gas directly behind the shock wave (not to be confused with the velocity \( v_0 \) of the shock wave itself). From the conservation of mass it is known that \( \rho_0 v_0 = \rho_1 v_1 \), and when combined with the strong shock relation (2.39) this yields:

\[
v_1 = \left( \frac{\gamma - 1}{\gamma + 1} \right) v_0. \tag{2.56}
\]

The velocity \( v_0 \) of the shock wave can be found in a relatively straightforward manner by taking the time derivative of the shock wave radius found in equation (2.49):

\[
v_0 = \frac{dR}{dt} = \frac{2}{5} \frac{R}{t}. \tag{2.57}
\]

This expression can be rewritten into a more convenient form by combining equations (2.49) and (2.57):

\[
v_0^2 = \frac{4}{25} \left( \frac{E}{\rho_0} \right) S^5. \tag{2.58}
\]

Putting together the above relationships then gives the following expression for the kinetic energy term:

\[
E_{\text{kin}} = \frac{8\pi}{75} \left( \frac{\gamma - 1}{\gamma + 1} \right)^2 E S^5. \tag{2.59}
\]

The thermal energy term may be approximated by considering the remaining small portion of gas that is located inside the shock wave cavity:

\[
E_{\text{th}} = \frac{1}{\gamma - 1} \left( \frac{4\pi R_0^3}{3} \right) p_c, \tag{2.60}
\]

where \( p_c \) is the average pressure inside the cavity. A quick inspection of figure 2.10 reveals that \( p_c = \frac{1}{2} p_1 \) is probably a good approximation. A suitable expression for the pressure \( p_1 \) may be found by rewriting equation (2.38):

\[
p_1 = \left( \frac{2}{\gamma + 1} \right) \rho_0 v_0^2. \tag{2.61}
\]

Using \( p_c = \frac{1}{2} p_1 \) and putting together the above equations then yields the following approximate expression for the thermal energy of the gas inside the cavity:

\[
E_{\text{th}} = \frac{16\pi}{75} \frac{E S^5}{(\gamma - 1)(\gamma + 1)}. \tag{2.62}
\]
The proportionality constant \( S(\gamma) \) may now be found by ensuring that the sum of the kinetic and thermal energy is equal to the total energy \( E \). This results in the following expression:

\[
S(\gamma) = \left[ \frac{75 (\gamma + 1) (\gamma - 1)}{8\pi (\gamma - 1)^3 + 2} \right]^{1/5},
\]

which is quite accurate when compared to the more precise numerical solution. The above expression predicts a proportionality constant of 0.92 and 1.07 for \( \gamma = 6/5 \) and \( \gamma = 7/5 \), while the more exact numerical method yields 0.89 and 1.03 for the same values of \( \gamma \).

### 2.3.3 Shock Strength

A good measure of the strength of the shock wave is the pressure jump across the shock wave. For strong shock waves \( (M_0 \gg 1) \) it is possible to derive an expression for this pressure jump with the help of the strong shock relation in equation (2.38):

\[
\Delta p = \left( \frac{2\gamma M_0^2}{\gamma + 1} \right) p_0,
\]

where \( M_0^2 \) may be calculated by scaling equation (2.58) with the speed of sound:

\[
M_0^2 = \frac{4}{25a_0^2} \left( \frac{E}{\rho_0} \right) \frac{S^5}{R^3}.
\]

The resulting expression for the pressure jump can be simplified considerably by introducing the expressions \( a_0^2 = \gamma RT_0 \) and \( p_0 = \rho_0 RT_0 \) for a perfect gas. This gives:

\[
\Delta p = \frac{8}{25} \left( \frac{S^5}{\gamma + 1} \right) \frac{E}{R^3}.
\]

Introducing equation (2.63) for the constant \( S(\gamma) \) then yields the following result:

\[
\Delta p = \frac{3}{\pi} \frac{\gamma - 1}{(\gamma - 1)^3 + 2} \frac{E}{R^3}.
\]

This expression for the pressure jump shows that the three most influential parameters on the strength of a shock wave produced by an intense point explosion are the distance to the explosion, the energy liberated during the explosion, and the specific heat capacity ratio of the gas.

It should be remembered that equation (2.67) is only valid for strong shock waves. For weaker shock waves approaching the speed of sound many of the used
expressions and assumption are no longer valid. As the shock wave travels further away, its velocity will decrease until it starts to behave as a sound wave. The propagation of a sound wave is described by the well-known wave equation [3]:

$$\nabla^2 p - \frac{1}{a^2} \frac{\partial^2 p}{\partial t^2} = 0,$$

(2.68)

where $p$ is the pressure, and $a$ the speed of sound. For a spherically symmetric case the wave equation can be rewritten into:

$$\frac{\partial^2 p}{\partial t^2} - a^2 \left( \frac{\partial^2 p}{\partial r^2} + \frac{2}{r} \frac{\partial p}{\partial r} \right) = 0.$$  

(2.69)

The general solutions that satisfy the spherical wave equation are of the form:

$$p(r,t) = \frac{1}{r} F(r - at) + \frac{1}{r} G(r + at),$$

(2.70)

where $F$ and $G$ are arbitrary functions. From equation (2.70) it can be concluded that sound waves travel at a constant speed equal to the speed of sound, and that the pressure attenuation of the sound wave is proportional to the inverse distance $r^{-1}$.

### 2.4 Conclusions

Based on the theoretical investigation in this chapter it is possible to draw a number of interesting conclusions. The most important conclusions are related to equation (2.67), which shows that the strength of a shock wave produced by the laser-induced breakdown in a gas is dependent on three key parameters. The most influential of these three parameters is the distance traveled by the shock wave, while the laser energy absorbed by the plasma and the specific heat ratio of the gas play a somewhat lesser role. When the shock wave is used to remove particles from a surface it is therefore important to create a high energy laser-induced plasma as close to the surface as possible without causing damage to the surface itself.

### References


2.4. References


Laser Induced Breakdown in Gases
Abstract. The creation and propagation of shock waves produced by the laser-induced breakdown of gases is a highly transient phenomenon that lies far beyond the capabilities of human perception. In most cases the gas is also transparent, which means that specialized techniques such as shadowgraphy and Schlieren photography must be combined with high speed imaging to visualize the shock wave. The time resolved images obtained by these techniques reveal the shape, size, and position of the shock wave as it travels through the gas. Further analysis of the images clearly shows that the propagation of the shock wave is consistent with the self–similar Taylor & Sedov solution for an intense point explosion. The results in this chapter may therefore be interpreted as a validation of the theoretical investigation made in the previous chapter.
3.1 Flow Visualisation in Transparent Gases

When working with transparent gases it is often necessary to use specialized visualization techniques. In this chapter two closely related visualization techniques called shadowgraphy and Schlieren photography are used to investigate the properties of the shock wave produced by the laser-induced breakdown of a gas. Both techniques have in common that they are based on the deflection of light due to refractive index gradients. The discussion on flow visualization techniques therefore starts with a more in-depth analysis of the propagation of light through an inhomogeneous refractive index field, after which the shadowgraphy and Schlieren methods are discussed in greater detail.

3.1.1 Deflection of Light

Light traveling through a gas with refractive index gradients does not follow a straight line. According to Fermat’s principle, light always takes the route with the shortest optical path length. The first order variation of the optical path length along the beam path must therefore vanish:

\[ \delta \int n(x(s), y(s), z(s)) \, ds = 0, \]

where \( n(x, y, z) \) is the refractive index field in a cartesian coordinate system, and \( ds \) represents an infinitesimal displacement along the beam path. The solution to this variational problem must satisfy the corresponding system of Euler–Lagrange
equations:

\[
\frac{d}{ds} \left( n \frac{dx}{ds} \right) = \frac{\partial n}{\partial x}, \quad (3.2)
\]

\[
\frac{d}{ds} \left( n \frac{dy}{ds} \right) = \frac{\partial n}{\partial y}, \quad (3.3)
\]

\[
\frac{d}{ds} \left( n \frac{dz}{ds} \right) = \frac{\partial n}{\partial z}. \quad (3.4)
\]

When the problem is restricted to beam paths which always have a component in the \( z \)-direction, it is possible to describe the beam path by \( x = x(z) \) and \( y = y(z) \). This is not a very harsh restriction, but it merely ensures that the light beam is able to cross the refractive index field. After careful algebraic calculation equations (3.2) to (3.4) can be rewritten into:

\[
\frac{d^2 x}{dz^2} = \left[ 1 + \left( \frac{dx}{dz} \right)^2 + \left( \frac{dy}{dz} \right)^2 \right] \left[ \frac{1}{n} \frac{\partial n}{\partial x} - \frac{dx}{dz} \frac{1}{n} \frac{\partial n}{\partial z} \right], \quad (3.5)
\]

\[
\frac{d^2 y}{dz^2} = \left[ 1 + \left( \frac{dx}{dz} \right)^2 + \left( \frac{dy}{dz} \right)^2 \right] \left[ \frac{1}{n} \frac{\partial n}{\partial y} - \frac{dy}{dz} \frac{1}{n} \frac{\partial n}{\partial z} \right]. \quad (3.6)
\]

This system of differential equations can be simplified considerably by assuming that the slope of the light beam and the refractive index gradient are much smaller than unity. Equations (3.5) and (3.6) are then reduced to:

\[
\frac{d^2 x}{dz^2} = \frac{1}{n} \frac{\partial n}{\partial x}, \quad (3.7)
\]

\[
\frac{d^2 y}{dz^2} = \frac{1}{n} \frac{\partial n}{\partial y}. \quad (3.8)
\]

This effectively decouples both differential equations, and the deflection angle of a light beam passing through the refractive index field may be calculated by integrating equations (3.7) and (3.8):

\[
\epsilon_x = \int \frac{1}{n} \frac{\partial n}{\partial x} \, dz, \quad (3.9)
\]

\[
\epsilon_y = \int \frac{1}{n} \frac{\partial n}{\partial y} \, dz, \quad (3.10)
\]

where \( \epsilon_x \) and \( \epsilon_y \) are the deflection angles as shown in figure 3.1. For a twodimensional refractive index field \( n(x, y) \) of extent \( L \) these two equations can be further reduced
to:

\[ \varepsilon_x = \frac{L}{n} \frac{\partial n}{\partial x}, \]  
\[ \varepsilon_y = \frac{L}{n} \frac{\partial n}{\partial y}. \]  

These two equations show that the deflection angle is directly proportional to the refractive index gradient. With the help of the Gladstone–Dale relationship it is possible to rewrite the refractive index gradients in equations (3.11) and (3.12) into density gradients:

\[ n - 1 = K \rho, \]  

where \( \rho \) is the mass density and \( K \) the Gladstone–Dale constant. Inserting the Gladstone–Dale relationship into equations (3.11) and (3.12) yields:

\[ \varepsilon_x = \left( \frac{KL}{1+K\rho} \right) \frac{\partial \rho}{\partial x}, \]  
\[ \varepsilon_y = \left( \frac{KL}{1+K\rho} \right) \frac{\partial \rho}{\partial y}. \]  

Density gradients within a transparent medium are thus capable of deflecting an incident beam of light over an angle that is proportional to the density gradient. This behaviour is exploited in flow visualization techniques such as shadowgraphy and Schlieren photography, which are the subject of the next two subsections.

### 3.1.2 Shadowgraphy

One of the simplest methods to visualise density gradients in transparent gases is shadowgraphy. The basic principle of this technique is illustrated in figure 3.2, which shows how light from a single point light source is deflected by a density gradient in its beam path. Some parts of the screen in figure 3.2 are therefore less illuminated than other parts, and the resulting image with lighter and darker regions is called a shadowgraph. Unfortunately the simplicity of shadowgraphy comes at the cost of reduced sensitivity. Shadowgraphy is only sensitive to the second derivative of the density, which means that this method is not very well suited for the visualization of detailed flow structures. On the other hand, it is ideally suited for the visualization of strong density gradients such as shock waves.

To demonstrate the sensitivity of shadowgraphy, consider a transparent plate of a homogeneous material and constant thickness, as illustrated in figure 3.3a. The light that enters from the left is not affected by the constant density field, and the screen located at the right of the plate is uniformly illuminated. In figure 3.3b the
3.1. Flow Visualisation in Transparent Gases

Figure 3.2: Example of a simple shadowgraphy experiment [2], consisting of a point light source (S) and a screen. (a) When there are no inhomogeneities present, the light travels in straight lines from the point light source to the screen. Each section of the screen is illuminated by one beam of light. (b) The introduction of an inhomogeneity in one of the beam paths causes a deflection of the light beam. As a consequence, the light is inhomogeneously redistributed over the screen.

Figure 3.3: Analogy to the deflection of light passing through a density field [2] with a (a) constant value, (b) constant derivative, (c) constant second derivative, and (d) variable second derivative.
Figure 3.4: Simple Schlieren photography system consisting of a point light source, two lenses, and a screen [4].

plate has been replaced by a wedge with flat surfaces. In this case the first derivative of the density is constant. Light entering from the left is now deflected by a constant angle, and the screen on the right is still uniformly illuminated. Replacing one flat surface of the wedge with a curved surface results in a density field with a constant second derivative. The wedge converges the light like a positive lens, as can be seen in figure 3.3c. This results in a more intense, but still uniform illumination of the screen. To achieve non-uniform illumination of the screen, the second derivative of the density has to be non-zero. Only then will it be possible to derive information about the density distribution from the image on the screen. Figure 3.3d shows an arrangement in which this is the case. From this it can be concluded that a shadowgraph is only sensitive to changes in the second derivative of the density.

3.1.3 Schlieren Photography

Named after the German word Schliere, which means striation or streak, Schlieren photography is another flow visualization technique that can be used to visualize density gradients in transparent gases. Unlike shadowgraphy, Schlieren photography systems are usually much more complicated to design and build. But perhaps more importantly, Schlieren methods are also more sensitive than shadowgraphy, which makes them more suitable for detailed flow visualization.

Figure 3.4 shows an example of a very simple Schlieren system with a point light source. The beam from the point light source is collimated by the first lens, after which a second lens refocuses the light again to form an image of the point light source. Next, the light is projected onto the screen, where a real inverted image of the Schlieren object in the test area is formed. At this moment the Schlieren system behaves as a simple projector, and is not yet capable of visualizing transparent Schlieren objects. This may be solved by adding a knife edge or any other type of
cut-off in the focal plane of the second lens. The knife edge can then be used to block the image of the point light source to produce a dark screen. The knife edge is also used to selectively block the light that is deflected by the Schlieren object in the test area. As figure 3.4 shows, the upward deflected ray is allowed to pass unhindered, whereas the downward deflected ray is blocked by the knife edge. By introducing a knife edge it is thus possible to visualize the transparent Schlieren object in the test area.

In a more realistic Schlieren system the illumination is provided by an extended light source, as is shown in figure 3.5. The extended light source may be interpreted as an array of individual point light sources, and in this case the first lens is not able to produce a collimated beam of perfectly parallel rays in the test area. As a consequence, the Schlieren object is illuminated by a bundle of rays coming from all point light sources, and all of the rays within this bundle are deflected over the same refraction angle $\epsilon$, as indicated by the dashed lines in figure 3.5. The deflected ray bundle is then refocused by the second lens to produce a weak source image in the knife edge plane that is displaced from the original undisturbed source image. Finally, the light passes through a third lens to produce a properly focused Schlieren image on the screen.

Figure 3.6 shows a sketch of the light source images in the knife edge plane. The original undisturbed source image has an unobscured height of $a$, whereas the weak displaced source image has an unobstructed height of $a + \Delta a$. This means that extra light is allowed to pass the knife edge to its corresponding point in the Schlieren image on the screen. For small deflection angles the amount of displacement $\Delta a$ may be calculated with geometrical optics:

$$\Delta a = f \epsilon,$$

where $f$ is the focal distance of the second lens. When it is also assumed that the
Figure 3.6: Sketch of the source images in the knife edge plane [4]. In this example the rectangular undisturbed source image is cut off by a half, but the weak displaced source image has been deflected upward and has only been cut off by a quarter.

The intensity distribution of the light source is uniform, the image illuminance is linearly proportional to the amount of cut–off, which makes it relatively straightforward to find the following expression for the image contrast:

\[ C = \frac{\Delta I}{I} = \frac{\Delta a}{a} = \frac{f \epsilon}{a}, \]  

(3.17)

where \( I \) is the image illuminance. Also note that according to equation (3.17) the image contrast can become infinitely large for a completely obstructed source image. In practice however, this is never achieved because diffraction effects will limit the amount of image contrast. Furthermore, the image contrast \( C \) is the output of the Schlieren system, and when combined with equations (3.14) and (3.15) it becomes clear that Schlieren photography is sensitive to the first derivative of the density. Having already established that shadowgraphy is sensitive to the second derivative of the density, Schlieren photography is indeed a more sensitive method than shadowgraphy.

3.2 Experimental Setups

Besides the visualization of density gradients within the gas, the imaging methods must also be capable of capturing the fast moving shock wave. This can be achieved by using a high–speed recording device, or by using a pulsed light source with an extremely short pulse length. Other practical issues, such as affordability and the required laboratory space, also play an important role in the design of an imag-
3.2. Experimental Setups


...ing setup. For instance, the shadowgraphy setup discussed in section 3.2.1 uses a Q-switched laser as a light source, since there are no cameras available that can meet the stringent demands of some of the experiments. Another example is the Schlieren setup discussed in section 3.2.2, which uses parabolic mirrors to save laboratory space and cut costs. In general, the laser flash shadowgraphy setup is the most suitable setup to capture very fast processes due to the use of a Q-switched laser as a light source, and is used in this thesis to visualize the formation and initial expansion of the laser-induced plasma. The Schlieren setup on the other hand, is also capable of visualizing much finer details, whereas the laser flash shadowgraphy setup is only able to visualize very strong density gradients such as shock waves. The use of large diameter parabolic mirrors in the Schlieren setup also makes it possible to study a larger field of view, but this advantage is not specific to the Schlieren method.
3.2.1 Laser Flash Shadowgraphy Setup

The shadowgraphy setup that is discussed here uses a Q-switched Nd:YAG laser manufactured by Ekspla for illumination. It operates at a visible wavelength of 532 nm, and has a sufficiently short FWHM pulse width of 150 ps to capture the highly transient shock wave. The laser–induced breakdown of the gas is produced with another Nd:YAG Q-switched laser, which has an infrared wavelength of 1064 nm and a FWHM pulse width of 10 ns. Both lasers are arranged as shown in figure 3.7, which shows a schematic sketch of the laser flash shadowgraphy setup that was used to study the early stages of laser–induced breakdown in gases. The light from the illumination laser is passed through a beam expander to produce a collimated beam that is sufficiently large to contain the shock wave. The laser–induced breakdown is produced at the center of the collimated beam by focusing the light from the other laser with the help of a single plano–convex lens that has a focal length of 75 mm. The light of the collimated beam is then deflected by the density gradients introduced by the laser–induced breakdown, thereby producing a shadowgraph which is recorded by the CCD camera. A high precision digital delay generator (Stanford Research Systems, DG535) was used for the correct timing of both lasers and the CCD camera. Furthermore, beam splitters and photo diodes with a rise time of 1 ns were also placed in the beam path of each laser to ensure the correct timing of the complete setup. Finally, a power meter was used to determine the amount of laser energy that is absorbed during the laser-induced breakdown process.

3.2.2 Z–Type Schlieren Setup

The Schlieren setup that was used to image the laser–induced breakdown of gases is shown in figure 3.8. This type of Schlieren system saves laboratory space by folding the beam path, and is called a \textit{Z-type Schlieren system} since the shape of the beam path resembles the letter Z. Here the illumination is provided by a \textit{Schlieren illuminator}, which consists of a halogen light source, a condenser lens, and an adjustable slit. The condenser lens produces an intermediate image of the halogen light source, which is then cropped by the adjustable slit to produce an approximately uniform light source with well defined edges. After the light leaves the Schlieren illuminator, it is collimated and refocused by a set of 108 mm f/8 parabolic mirrors. The knife edge is mounted on a translation stage, and its orientation is identical to the orientation of the adjustable slit in the Schlieren illuminator. The camera that is used in this setup is an intensified CCD camera, which is capable of very short exposure times below 10 ns. As with the laser flash shadowgraphy setup, the correct timing of both the laser and CCD camera is controlled by a high precision digital delay generator.
3.3 Time Resolved Images

After discussing the shadowgraphy and Schlieren techniques in the previous chapter, it is now time to put these techniques into practice and study the laser-induced breakdown of gases in more detail. Both flow visualization techniques are used, but because of its extremely short exposure time of 150 ps, the laser flash shadowgraphy setup is primarily used for those experiments that involve the highly transient early stages of plasma formation and shock wave propagation. The more sensitive Z-type Schlieren setup is used for the imaging experiments that require a larger field of view or higher sensitivity. This includes the propagation of the shock wave far away from the focal point, as well as the behaviour of the hot gas core left after breakdown.

3.3.1 Plasma Ignition

The first step in the laser-induced breakdown of gases is the ignition phase. As was already discussed in the previous chapter, ignition of the laser-induced plasma
may occur when the laser fluence is sufficiently high to produce free electrons by multi photon ionization. Once free electrons are present, inverse Bremsstrahlung will further heat the free electrons, until they have gained enough energy to ionize other neutrals by impact. It was then also theorized that the expansion of the laser–induced plasma follows the Taylor & Sedov solution for an intense point explosion. This assumption may now be checked by examining the time resolved shadowgraphs in figure 3.9, which shows the ignition and subsequent evolution of the laser–induced plasma during the laser pulse. These images show very clearly that the plasma is not ignited in one particular place, but that there are many small ignition sites along the optical axis. This behaviour may be caused by a number of things, such as the presence of dust particles or the formation of hot spots along the optical axis due to non–perfect focusing. In any case, the assumption that the laser–induced plasma behaves as a point explosion is thus not valid during the laser pulse.

At these very early stages of plasma formation and expansion, it is perhaps not surprising that the Taylor & Sedov formula is not valid, since the mathematical concept of a point cannot be reproduced in the real world. Another reason why the Taylor & Sedov solution should not apply at this stage during the expansion, is the fact that laser energy is still being deposited into the plasma. This also violates one of the other assumption of the Taylor & Sedov solution, which states that the energy of the explosion is constant. In this case the shock relations would have to be modified to take the energy deposition behind the shock front into account. The resulting theory is called the Chapman–Jouguet theory, which is discussed in more detail in [3]. In practice however, the Chapman–Jouget is of limited importance to laser–induced shock wave cleaning, because the energy deposition into the plasma will have seized well before the shock wave reaches the surface.

### 3.3.2 Shock Wave Propagation

After ignition, the laser–induced plasma continues to expand into the surrounding gas. As is shown by the shadowgraphs in the previous section, the deposition of the laser energy does not occur at one specific location, but in numerous sites along the optical axis. This results in an elongated plasma that produces a shock wave with an oblate shape, as can be seen in the shadowgraphs in figure 3.10. Since the shape of the shock wave is not spherical, it is not possible to apply the spherically symmetric Taylor & Sedov solution directly to the oblate shock wave. Another option is to consider the volume that has been engulfed by the shock wave, which may be calculated from equation (2.49) in a relatively straightforward manner:

\[
V = \frac{4\pi R^3}{3} = t^{6/5} \left[ \frac{4\pi}{3} \left( \frac{E_{bw}}{\rho_0} \right)^{3/5} S^3 \right],
\]  

(3.18)
Figure 3.10: Shadowgraphs of the shock wave produced by the laser-induced breakdown in air. The laser pulse is incident from the right side and carries an average energy of 169 mJ.
Figure 3.11: Double logarithmic plot of the volume $V$ engulfed by the shock wave in air at atmospheric pressure and room temperature versus the time $t$. Each data point represents the average value calculated over 10 separate measurements.
$E_{\text{abs}} = (33 \pm 1) \text{ mJ}$  \hspace{1cm} \text{Slope} = 1.09 \pm 0.04

$E_{\text{abs}} = (62 \pm 1) \text{ mJ}$  \hspace{1cm} \text{Slope} = 1.14 \pm 0.04

**Figure 3.12:** Double logarithmic plot of the volume $V$ engulfed by the shock wave in argon at atmospheric pressure and room temperature versus the time $t$. Each data point represents the average value calculated over 10 separate measurements.
where $V$ is the volume engulfed by the shock wave, $R$ the radius of the shock wave, $E_{bw}$ the blast wave energy, $\rho_0$ the density of the undisturbed gas, and $S$ a dimensionless constant given by equation (2.63). Strictly spoken, equation (3.18) is derived on the basis of a spherically symmetric shock wave, but it will be shown here that this relationship is also valid for other shapes that do not deviate too much from a sphere. This may be checked by rewriting equation (3.18) into the following form:

$$\ln V = \frac{6}{5} \ln t + \ln \left[ \frac{4\pi}{3} \left( \frac{E_{bw}}{\rho_0} \right)^{3/5} \right], \quad (3.19)$$

from which it may be concluded that a plot of $\ln V$ against $\ln t$ should yield a straight line with a slope equal to $6/5$. This is confirmed by figures 3.11 and 3.12, which show that the slope is indeed equal to $6/5$ for both argon and air at several different amounts of absorbed laser energy. A more careful examination of both figures also reveals that the experimental data points at very small times start to deviate from the linear relationship predicted by equation (3.19). For these small times the volume is larger than expected, which may be explained by the fact that the laser-induced plasma is not ignited in one specific point, but at multiple locations along the optical axis.

Equation (3.18) also predicts that there should be a linear relationship between the shock wave volume $V$ and $t^{6/5}$. The experimental data supports this prediction, as can be seen in figures 3.13 and 3.14. This leads to the conclusion that the dimensionless constant $S$ is indeed constant, and that the specific heat capacity ratio $\gamma$ on which it depends is also constant. The specific heat capacity ratio therefore has some constant effective value, which is surprising, since the gas undergoes large pressure and temperature variations, as well as dissociation in the case of a diatomic gas. For a monoatomic gas such as argon, the specific heat capacity ratio is relatively constant under most circumstances, and may be taken equal to $5/3$. Air on the other hand, is primarily composed of diatomic molecules, which makes it much more difficult to obtain a value for the effective specific heat capacity ratio. As an estimation, it is possible to take the value of $\gamma$ that corresponds to the gas state when the shock wave has traveled to half of the maximum measured distance. From the shadowgraphs in figures 3.13 and 3.14 it is then possible to deduce the shock wave velocity at this specific position, which may be used in equation (2.36) to calculate the gas temperature behind the shock wave. Once the temperature is known, the specific heat capacity ratio $\gamma$ can be calculated with the help of the Shomate equation:

$$c_p = \sum_{i=0}^{4} c_i \left( T/1000 \right)^i, \quad (3.20)$$
Figure 3.13: Relation between $t^{6/5}$ and the volume $V$ engulfed by the shock wave in air at atmospheric pressure and room temperature. Each data point represents the average value calculated over 10 separate measurements.
Figure 3.14: Relation between $t^{6/5}$ and the volume $V$ engulfed by the shock wave in argon at atmospheric pressure and room temperature. Each data point represents the average value calculated over 10 separate measurements.
Table 3.1: Overview of the energy distribution in laser-induced plasmas in air and argon for several different amounts of absorbed laser energy.

<table>
<thead>
<tr>
<th>Gas</th>
<th>$E_{\text{abs}}$ (mJ)</th>
<th>$E_{\text{bw}}$ (mJ)</th>
<th>$E_{\text{rad}}$ (mJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air</td>
<td>122 ± 1</td>
<td>62 ± 7</td>
<td>60 ± 8</td>
</tr>
<tr>
<td>Air</td>
<td>67 ± 2</td>
<td>27 ± 3</td>
<td>40 ± 5</td>
</tr>
<tr>
<td>Argon</td>
<td>62 ± 1</td>
<td>42 ± 5</td>
<td>20 ± 6</td>
</tr>
<tr>
<td>Argon</td>
<td>33 ± 1</td>
<td>26 ± 4</td>
<td>7 ± 5</td>
</tr>
</tbody>
</table>

where $c_p$ is the specific heat capacity at constant pressure, and $c_i$ a collection of coefficients which may be obtained from NIST\(^1\) [1]. If the temperature behind the shock wave is below the dissociation temperature, the specific heat capacity ratio may be calculated as follows:

$$\gamma = \frac{c_p}{c_p - R},$$  \hspace{1cm} (3.21)

where $R$ represents the gas constant. By bringing equations (2.36), (3.20) and (3.21) into agreement, the following values for the effective specific heat capacity ratio of air are found: $\gamma = 1.34$ at $T = 990$ K and $E_{\text{abs}} = 67$ mJ, and $\gamma = 1.32$ at $T = 1200$ K and $E_{\text{abs}} = 122$ mJ. Fortunately, the calculated temperatures are not high enough to cause dissociation, and it can be concluded that the use of equation (3.21) is justified.

Now that the effective value of $\gamma$ is known, figures 3.13 and 3.14 may also be used to calculate the amount of blast wave energy. This may be done in the following way:

$$E_{\text{bw}} = \left( \frac{V}{t^{6/5}} \right)^{5/3} \left[ \rho_0 \left( \frac{3}{4\pi} \right)^{5/3} S^{5} \right],$$  \hspace{1cm} (3.22)

where $V/t^{6/5}$ represents the slope of the linear fit. The resulting blast wave energies have been collected in table 3.1, which shows that the blast wave energies are significantly smaller than the absorbed laser pulse energies. This may be explained by examining the following energy balance:

$$E_{\text{abs}} = E_{\text{bw}} + E_{\text{rad}},$$  \hspace{1cm} (3.23)

where $E_{\text{abs}}$ is the amount of absorbed laser energy, $E_{\text{bw}}$ the blast wave energy, and $E_{\text{rad}}$ the energy lost due to radiation. According to equation (3.22) and the results in figures 3.13 and 3.14, the blast wave energy has a constant value during the measured time range. From this it may be concluded that the largest loss of energy

\(^1\)NIST = National Institute of Standards and Technology.
due to radiation occurs before the first measurement, within a time period smaller than 100 ns.

3.3.3 Hot Gas Core

The previous sections were concerned with the ignition of the laser-induced plasma and the propagation of the shock wave into the surrounding gas. To complete the investigation, this section very briefly focuses on the behaviour of the hot gas core, which is left after laser-induced breakdown. Figure 3.15 shows a collection of time resolved Schlieren images that depict the evolution of the hot gas core. As can be seen in the images, the laser-induced breakdown produces a highly turbulent hot gas core, which expands with a velocity that lies well below the speed of sound. For instance, the average velocity of the plume of hot gas that is ejected from the main body is approximately equal to 10 m/s.

3.4 Conclusions

The most important conclusion is that the time resolved images presented in this chapter confirm that the propagation of the shock wave produced by the laser-induced breakdown in air and argon is consistent with the Taylor & Sedov solution discussed in the previous chapter. In addition to the validation of the Taylor & Sedov solution, the images also provide more insight into the energy distribution of the laser-induced breakdown process. As is shown in the previous section, the radiative losses can be significant, and must occur somewhere during the first 100 ns after the ignition of the laser-induced plasma. When the laser-induced breakdown of gases is applied to the removal of small particles from a surface, it may therefore also be important to consider the possibility of surface damage caused by the exposure to the intense radiation emitted directly after plasma ignition.

References


Figure 3.15: Schlieren photographs of the hot gas core in air after the passage of the shock wave.
CHAPTER 4

Particle Adhesion Theory

Abstract. Small submicron particles on a flat surface experience relatively large adhesion forces, which may be orders of magnitude greater than the weight of the particle itself. To identify and characterize the most important contributions to the adhesion force, this chapter investigates the interaction mechanisms between the submicron particles and the surface in more detail. This leads to the conclusion that the strongest adhesion force results from the capillary condensation of moisture into the cracks and crevices surrounding the contact area. Furthermore, it is also shown that for dry environments in which the capillary force is absent, the van der Waals force becomes the most dominant adhesion force. This chapter then concludes by combining adhesion theory with shock theory to create an order accurate model, which is able to predict the cleaning efficiency of laser–induced shock wave cleaning.
4.1 Microscopic Adhesion Forces

Nowadays it is commonly accepted that all phenomena in our universe originate from one of the four fundamental forces of nature. These forces are the electromagnetic force, gravitational force, and the strong and weak nuclear forces. As their name suggests, the strong and weak nuclear forces are responsible for the interactions between neutrons, protons, and other elementary particles within the nucleus of an atom. But these two forces are not responsible for the adhesion of small particles to a surface, since their range of interaction is much smaller than the typical intermolecular distance. The remaining gravitational and electromagnetic forces both have an infinite range of interaction, but the gravitational force lacks the necessary repulsive component to accurately describe the contact mechanics between the particle and the surface. In contrast to the gravitational force, the electromagnetic force has an attractive as well as a repulsive component, which leads to the general conclusion that adhesion forces are electromagnetic in nature.

The intermolecular forces can be organized into three distinct categories: electrostatic forces, polarization forces, and quantum mechanical forces. The electrostatic forces involve all interactions between charges, dipoles, quadrupoles, etc. The electric field of such charges and permanent dipoles are also capable of inducing a dipole moment in nearby atoms or molecules. The forces between the induced dipole and the charge or dipole that induced it, are called polarization forces. These two categories of intermolecular forces lead to physical bonding. The last category, which consists of quantum mechanical forces, are responsible for chemical or covalent bonding. These three categories together comprise a wide variety of intermolecular interactions, as can be seen in figure 4.1.

4.1.1 Electrostatic Interactions

The most simple electrostatic interaction is the well-known Coulomb interaction between two charged atoms or molecules. This type of interaction has an infinite range, as is illustrated by the following expression for the Coulomb energy:

\[
U = -\frac{q_1 q_2}{4\pi \varepsilon_0 r},
\]

where \(\varepsilon_0\) is the permittivity of free space, and \(q_1\) and \(q_2\) are two charges seperated by a distance \(r\). An important example of this intermolecular force is the ionic bond between two oppositely charged ions. However, in most cases the Coulomb energy reduces to zero, because the majority of molecules is electrically neutral. But in this case, it is still possible that some atoms or groups within the molecule have a higher affinity for electrons than the other groups or atoms. This property is called
4.1. Microscopic Adhesion Forces

<table>
<thead>
<tr>
<th>Interaction Type</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charge, Charge</td>
<td>[ U = -\frac{q_1q_2}{4\pi\varepsilon_0r} ]</td>
</tr>
<tr>
<td>Charge, Dipole</td>
<td>[ U = -\frac{q^2\mu_2^2}{6(4\pi\varepsilon_0)^2kT r^4} ]</td>
</tr>
<tr>
<td>Dipole, Dipole</td>
<td>[ U = -\frac{\mu_1^2\mu_2^2}{3(4\pi\varepsilon_0)^2kT r^6} ]</td>
</tr>
<tr>
<td>Dipole, Non–Polar</td>
<td>[ U = -\frac{\alpha\mu^2}{(4\pi\varepsilon_0)^2r^6} ]</td>
</tr>
<tr>
<td>Charge, Non–Polar</td>
<td>[ U = -\frac{\alpha q^2}{2(4\pi\varepsilon_0)^2r^4} ]</td>
</tr>
<tr>
<td>Non–Polar, Non–Polar</td>
<td>[ U = -\frac{3}{4} \frac{\hbar\nu\alpha^2}{(4\pi\varepsilon_0)^2r^6} ]</td>
</tr>
<tr>
<td>Hydrogen Bonding</td>
<td>[ U \sim -\frac{1}{r^2} ]</td>
</tr>
<tr>
<td>Covalent Bonding</td>
<td>Short Range</td>
</tr>
</tbody>
</table>

**Figure 4.1:** Overview of various intermolecular interactions and their interaction energies [6].
the *electronegativity*, and it causes a shift within the electron distribution of the molecule. The result is a so called *polar molecule*, which has a dipole moment due to its non–uniform charge distribution. In general, the dipole moment of a polar molecule is given by the following integral:

$$\mu = \int \rho (r) \cdot r \, dV,$$

where $\rho (r)$ is the charge density at position $r$, and $dV$ an infinitesimal volume element. The interaction energy between two dipoles depends strongly on their mutual orientation, and to minimize their interaction energy, the dipoles have a tendency to align themselves parallel to each other. On the other hand, the alignment of the dipoles is randomized by the thermal energy. This results in the following expression for the interaction energy between two freely rotating dipoles:

$$U = -\frac{C_{\text{ori}}}{r^6} = -\frac{\mu_1^2 \mu_2^2}{3 (4\pi \varepsilon_0)^2 kT r^6},$$

where $\mu_1$ and $\mu_2$ are the dipole moments of both molecules, $k$ the Boltzmann constant, $T$ the temperature, and $r$ the distance between the two dipoles. This type of interaction is often called the *Keesom or orientation* interaction. An example of a very strong Keesom interaction is *hydrogen bonding* between polar water molecules. As can be seen in figure 4.1, the hydrogen bond in water is located between the negatively charged oxygen atom and the positively charged hydrogen atom. Because of the small size of the hydrogen atom, the distance between both atoms is much smaller than usual, which results in a very strong bond.

Another type of electrostatic interaction, is the interaction between a charge $q$ and a dipole $\mu$. In this case the dipole has the tendency to minimize its interaction energy by aligning itself towards or away from the charge, depending on the sign of the charge. This behavior is also randomized by thermal energy, which leads to the following expression for the interaction energy:

$$U = -\frac{q^2 \mu^2}{6 (4\pi \varepsilon_0)^2 kT r^4},$$

where $r$ is the distance between the charge and the dipole.

### 4.1.2 Interactions Involving Polarization

At first, it might seem that there is no interaction between a dipole and a non–polar molecule. But the electric field produced by the dipole causes the electron
distribution of the non–polar molecule to shift. This produces an induced dipole moment in the non–polar molecule, which can be calculated with:

$$\mu_{\text{ind}} = \alpha E,$$

where $\alpha$ is the polarizability, and $E$ the electric field experienced by the non–polar molecule. The interaction between the permanent and the induced dipole is often called the Debye or induction interaction, and the corresponding interaction energy is given by:

$$U = -\frac{C_{\text{ind}}}{r^6} = -\frac{\alpha \mu^2}{(4\pi\varepsilon_0)^2 r^6},$$

(4.6)

where $\mu$ is the dipole moment, and $r$ the distance between the dipole and the non-polar molecule. Another type of interaction that involves polarization, is the interaction between a charge and a non-polar molecule. In this case the charge is also able to induce a dipole moment in the non-polar molecule, and the interaction energy between the charge and the induced dipole is given by:

$$U = -\frac{\alpha q^2}{2 (4\pi\varepsilon_0)^2 r^4},$$

(4.7)

where $q$ is the charge, and $r$ the distance between the charge and the non–polar molecule.

### 4.1.3 Quantum Mechanical Interactions

Until now it was possible to describe the various intermolecular interactions with the help of classical physics, but they fail to explain the intermolecular interaction between two non–polar molecules or atoms. The origin of this interaction can be understood by considering a non–polar atom, which consists of a positively charged nucleus and a negatively charged electron cloud. The electrons circle around the nucleus with a high frequency of approximately $10^{15}$ Hz, and at every instant the atom is therefore polar. When two of these oscillators are brought together, they will start to influence each other. In this case attractive orientations have a higher probability than repulsive ones, which leads to an attractive force. This type interaction is called the London or dispersion interaction, and the corresponding interaction energy can be calculated with quantum mechanical perturbation theory [9]:

$$U = -\frac{C_{\text{disp}}}{r^6} = -\frac{3h}{2} \left( \frac{\nu_1 \nu_2}{\nu_1 + \nu_2} \right) \frac{\alpha_1 \alpha_2}{(4\pi\varepsilon_0)^2 r^6},$$

(4.8)
where $\alpha_1$ and $\alpha_2$ are the electric polarizabilities of the non-polar molecules, $r$ the distance between the non-polar molecules, $\hbar$ Planck’s constant, and $h\nu_1$ and $h\nu_2$ are the ionization potentials of the molecules or atoms.

Covalent bonding is also quantum mechanical in nature, and it differs greatly from the previously discussed intermolecular interactions. These types of interactions were mostly based on electrostatics and polarization, whereas the covalent bond involves changes in the electron configuration of the two binding atoms. Covalent bonding is characterized by the sharing of pairs of electrons between atoms, as is illustrated in figure 4.1. The resulting bond is usually a very strong bond, with a typical energy that is well above the thermal energy $kT$.

4.2 Adhesion Between Macroscopic Bodies

In this section it is shown that the intermolecular interactions that have been summarized in the previous section can lead to physical bonding between macroscopic bodies. Furthermore, it is also shown that there are several types of macroscopic bonding or adhesion forces that may be of interest to the adhesion of small sub-micron particles to a flat surface. Each of these adhesion forces is therefore investigated in more detail to determine its characteristics and adhesion strength.

4.2.1 van der Waals Force

One of the adhesion forces that is responsible for the bonding of two macroscopic bodies is the van der Waals force. This force is a combination of the orientation, induction, and dispersion forces, which are all particularly strong at small distances due to their $r^{-6}$ dependence. When all three interactions are summed, the following expression is found for the van der Waals energy between two single molecules:

$$U_{vdW} = -\frac{C_{ori}}{r^6} - \frac{C_{ind}}{r^6} - \frac{C_{disp}}{r^6} = -\frac{C}{r^6}, \quad (4.9)$$

where $r$ is the distance between both molecules. To calculate the van der Waals energy between two macroscopic bodies, it is convenient to assume that the van der Waals interaction energies are additive. This can be illustrated by calculating the van der Waals energy of a single molecule at a distance $D$ away from a flat surface. The molecule interacts with the other molecules in the surface, and the sum of all the individual interactions gives the total van der Waals energy. When the molecular density of the surface is taken equal to $\rho$, it is possible to express the total van der Waals energy in a single integral. Figure 4.2 shows that the surface can be divided into a collection of infinitesimal thick rings. Each ring contains $2\pi x \rho \, dx \, dz$.
4.2. Adhesion Between Macroscopic Bodies

Figure 4.2: Geometry of a molecule and an infinitely large surface separated by a distance $D$.

molecules, and the interaction energy can be calculated with the following integral:

$$
U = \int_{D}^{\infty} dz \int_{0}^{\infty} 2\pi x \rho \left( -\frac{C}{r^6} \right) dx
$$

$$
= -2\pi \rho C \int_{D}^{\infty} dz \int_{0}^{\infty} \frac{x}{(x^2 + z^2)^3} dx
$$

$$
= -2\pi \rho C \int_{D}^{\infty} \frac{1}{4z^4} dz
$$

$$
= -\frac{\pi \rho C}{6D^3}. \quad (4.10)
$$

Equation (4.10) can then be used to calculate the van der Waals energy between a plane surface and a spherical particle. In this case the spherical particle is divided into a collection of slices, as can be seen in figure 4.3. The number of molecules contained in such an infinitesimal thin slice is equal to $z (2R - z) \pi \rho_1 dz$. This gives
the following integral:

$$U = \int_0^{2R} z (2R - z) \pi \rho_1 \left( -\frac{\pi \rho_2 C}{6 (z + D)^3} \right) dz$$

$$= -\frac{\pi^2 \rho_1 \rho_2 C}{6} \int_0^{2R} \frac{z (2R - z)}{(z + D)^3} dz,$$

(4.11)

where $\rho_1$ and $\rho_2$ are the molecular densities of the surface and the sphere. The distance $D$ is typically in the order of 4 Å, which is much smaller than the radius $R$ of the particle itself. Therefore, only small values of $z$ will contribute significantly to the integral. In this case the integral can be simplified as follows:

$$U = -\frac{\pi^2 \rho_1 \rho_2 C}{6} \int_0^{2R} \frac{2Rz}{(z + D)^3} dz,$$

$$= -\frac{\pi^2 \rho_1 \rho_2 CR}{6D}.$$

(4.12)

The van der Waals force is then calculated from equation (4.12) in a relatively straightforward manner:

$$F_{vdW} = -\frac{\partial U}{\partial D} = \frac{AR}{6D^2},$$

(4.13)

where $A = \pi^2 \rho_1 \rho_2 C$ is the Hamaker constant, which does not vary much for most materials and has a typical value of $\sim 10^{-19}$ J. Unfortunately, equation (4.13) underestimates the van der Waals force, because it fails to recognize that the van der Waals force can be strong enough to deform the particle. This creates a circular contact area with radius $a$ at the interface between the particle and the surface, which greatly enhances the short range van der Waals interaction. The interaction energy can be calculated by adjusting equation (4.12) to the geometry shown in figure 4.4:

$$U = -\frac{\pi^2 \rho_1 \rho_2 C}{6} \int_0^{2R} \frac{2R (z + \delta)}{(z + D)^3} dz,$$

$$= -\frac{\pi^2 \rho_1 \rho_2 CR}{6} \left( \frac{1}{D} + \frac{\delta}{D^2} \right).$$

(4.14)

For small values of the deformation length $\delta$ it is possible to write down the following simple geometric relationship between the contact radius $a$ and the particle radius $R$:

$$a^2 = R^2 - (R - \delta)^2 = 2\delta R - \delta^2 \approx 2\delta R \quad \Rightarrow \quad \delta = \frac{a^2}{2R}.$$
4.2. Adhesion Between Macroscopic Bodies

Figure 4.3: Geometry of a sphere and an infinitely large surface separated by a distance \( D \).

Figure 4.4: Geometry of a deformed sphere and an infinitely large surface separated by a distance \( D \).
Particle Adhesion Theory

Substituting this into equation (4.14) gives the following expression for the van der Waals force between a deformed sphere and a flat surface:

\[ F_{\text{vdW}} = -\frac{\partial U}{\partial D} = \frac{AR}{6D^2} \left(1 + \frac{a^2}{RD}\right), \tag{4.16} \]

where the second term between the parentheses accounts for the deformation of the particle. As is shown in the following discussion on contact mechanics, the additional deformation term in equation (4.16) can be become quite significant for sub-micron particles because of their relatively large contact area, and should not be neglected when calculating the van der Waals force.

Contact Mechanics

The amount of deformation may be calculated by examining the strains and stresses that occur when the particle is brought into contact with the surface. This field of research is now known as contact mechanics, and was first explored more than a century ago by the well-known physicist Heinrich Hertz. Especially his 1881 publication [5], in which he investigates the deformation of two contacting elastic bodies under the influence of an external load, is considered by many as the foundation of modern contact mechanics. His theory is nowadays known as Hertzian contact mechanics, and can be used to calculate the radius of the contact area between a particle and a flat surface:

\[ a^3 = \frac{3RF}{4E^*}, \tag{4.17} \]

where \( F \) is the external load applied to the particle. The parameter \( E^* \) depends on the material properties of the particle and the surface, and is given by:

\[ \frac{1}{E^*} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2}, \tag{4.18} \]

where \( \nu \) represents the Poisson coefficient, \( E \) the Young modulus, and the subscripts refer to the particle (1) and the surface (2). Hertz theory also predicts that the pressure in the circular contact area has a parabolic profile:

\[ p(r) = p_0 \left(1 - \frac{r^2}{a^2}\right)^{1/2}, \quad p_0 = \frac{3F}{2\pi a^2}, \tag{4.19} \]

where \( r \) is the radial position. Hertz theory proved very successful and remained unchallenged for almost a century, until it was discovered that the contact area is larger than predicted by Hertz. It was also observed that particles with no applied external
4.2. Adhesion Between Macroscopic Bodies

Hertz JKR DMT

Figure 4.5: Schematic overview of the Hertz, JKR, and DMT contact mechanics theories. The arrows indicate the van der Waals interactions that are considered by each model.

load still experienced a non-zero contact area, which is not in agreement with equation (4.17). These observations suggested the presence of adhesive forces, which are not considered in Hertzian contact mechanics. In response to these shortcomings, more advanced theories were developed of which the Johnson–Kendall–Roberts [11] (JKR) and the Derjaguin–Toporov–Muller [2, 3] (DMT) theories are probably the most well known.

Both JKR and DMT theory take into account the adhesive van der Waals forces between two contacting bodies, but do so in an entirely different manner. As is schematically depicted in figure 4.5, JKR theory only considers the van der Waals force in the contact area, which is larger than the Hertzian contact area due to necking of the particle. DMT theory on the other hand, adopts a Hertzian contact area, but also considers additional van der Waals interactions outside the contact area. Both approaches lead to an increased contact area, as can be seen in the following expressions for the contact radius:

\[
a^3 = \frac{3R}{4E^*} \left( F + 3\pi WR + \sqrt{6\pi WRF + (3\pi WR)^2} \right), \quad \text{(JKR)} \quad (4.20)
\]

\[
a^3 = \frac{3R}{4E^*} (F + 2W\pi R), \quad \text{(DMT)} \quad (4.21)
\]

where \( W = A / (12\pi D^2) \) represents the work of adhesion. In general, JKR theory is only applicable to relatively large and elastic bodies, while DMT theory is better suited for smaller and more rigid bodies. To decide which of these two theories
Particle Adhesion Theory

should be used, it is often convenient to examine the Tabor coefficient [12]:

\[ \mu = \left( \frac{RW^2}{E^*2D3} \right)^{1/3}, \]

where large values (\( \mu \sim 5 \)) favor JKR theory, and small values (\( \mu \sim 0.005 \)) correspond closely to DMT theory [4]. Due to the inclusion of adhesive forces, both theories also predict higher pressures in the contact area than Hertz theory. For example, the JKR theory predicts the following pressure profile:

\[ p(r) = p_0 \left( 1 - \frac{r^2}{a^2} \right)^{1/2} + p'_0 \left( 1 - \frac{r^2}{a^2} \right)^{-1/2}, \]

where \( p_0 \) and \( p'_0 \) are given by:

\[ p_0 = \frac{2aE^*}{\pi R}, \]

\[ p'_0 = -\left( \frac{2WE^*}{\pi a} \right)^{1/2}. \]

Equation (4.23) may be used to show that the peak pressure in the contact area can reach very high values for small unloaded \((F = 0)\) sub-micron particles. For instance, figure 4.6 shows that the peak pressure of polystyrene latex spheres on a silicon surface can reach values of 100 MPa and above for very small particles. These high pressures in the contact area explain why it is usually very difficult to remove very small particles from a surface.

4.2.2 Capillary Force

Another important macroscopic adhesion force is the capillary force. This force is the result of capillary condensation of moisture in the small cracks and crevices that surround the contact area, as can be seen in figure 4.7. Capillary condensation is a phenomenon that allows vapor to condensate, even though the vapor pressure is well below the equilibrium vapor pressure of the liquid. This is possible because the surface of the liquid in figure 4.7 is curved, which causes the value of the equilibrium vapor pressure to change. The change in equilibrium vapor pressure of a curved liquid surface can be calculated with the Kelvin equation, which gives the ratio \( p_c/p_p \) between the equilibrium vapor pressure of a curved and a planar surface:

\[ \ln \left( \frac{p_c}{p_p} \right) = \frac{\gamma V_m}{RT} \left( \frac{1}{R_1} + \frac{1}{R_2} \right), \]
4.2. Adhesion Between Macroscopic Bodies

**Figure 4.6:** Calculated peak pressure in the contact area between polystyrene spheres and a flat silicon surface.

**Figure 4.7:** Capillary condensation of moisture in the cracks and crevices around the contact area.
where $R_1$ and $R_2$ are the two principal radii of curvature, $\gamma$ the surface tension of the liquid, $V_m$ the molar volume of the liquid, $R$ the gas constant, and $T$ the temperature in Kelvin. The radius of curvature is defined as positive when the interface is curved towards the liquid, and negative when the interface is curved in the opposite direction. Applying the Kelvin equation to the situation shown in figure 4.7 yields the following expression:

$$\ln \left( \frac{p_c}{p_p} \right) = \frac{\gamma V_m}{RT} \left( \frac{1}{x} - \frac{1}{r} \right). \tag{4.27}$$

In general it is safe to assume that $x \gg r$, and the previous equation can be rewritten into the following form:

$$\ln \left( \frac{p_c}{p_p} \right) = -\frac{\gamma V_m}{RT} \frac{1}{r}, \tag{4.28}$$

which shows that the equilibrium vapor pressure of the negatively curved liquid is smaller than that of a planar surface. This explains why vapor is able to condensate in small cracks and crevices at vapor pressures well below the equilibrium vapor pressure $p_p$.

The curvature of the fluid is the result of a pressure difference inside and outside the liquid. This pressure difference is called the Laplace pressure, and it is calculated in the following way:

$$\Delta p = \gamma \left( \frac{1}{R_1} + \frac{1}{R_2} \right), \tag{4.29}$$

where $\gamma$ is the surface tension of the liquid. Applying this to the curved liquid in figure 4.7 yields the following expression for the Laplace pressure:

$$\Delta p = \gamma \left( \frac{1}{x} - \frac{1}{r} \right) \approx -\left( \frac{\gamma}{r} \right). \tag{4.30}$$

This shows that the pressure inside the liquid is smaller than the pressure of the surrounding air. The pressure difference acts on an area $\pi x^2 \approx 2\pi Rd$ between the two surfaces, and the resulting adhesion force $2\pi \gamma Rd/r$ is called the capillary force. For small values of $\phi$ the approximation $d \approx 2r \cos \theta$ can be used to find the following expression for the capillary force between a sphere and a flat surface:

$$F_{\text{cap}} = 4\pi \gamma R \cos \theta \tag{4.31}$$

where $\theta$ is the contact angle between the liquid and the surface of the sphere, as indicated in figure 4.7.
4.2. Adhesion Between Macroscopic Bodies

### 4.2.3 Electric Double Layer Force

Besides the van der Waals and capillary forces, the electric double layer force can also lead to the adhesion of small particles to a surface. This force arises when two solids charge each other electrostatically when they come into contact with one another. This phenomenon is often called the *triboelectric effect*, because in some cases it can be enhanced by rubbing the two materials together (*tribos* is Greek for rubbing). The result of the electrostatic charging is a so called electronic double layer at the interface between the two solids. The mechanism responsible for the charge transfer between the two solids is diverse, and it depends on a wide variety of properties such as the material, surface roughness and temperature.

A good example to demonstrate the formation of an electric double layer is a p–n junction between two semiconductor materials. Upon contact, the electrons and holes will migrate from one material to the other, due to the difference in concentration. This process will continue until the potential difference, that develops between the two materials due to the charge separation, has grown large enough to stop any further migration of charge carriers. The result is an electronic double layer at the interface between the two solids, as can be seen in figure 4.8. The potential difference between the two materials is called the contact potential, and its maximum value is about 0.5 V [7]. The resulting electric double layer force at the contact area between a spherical particle and a flat surface can be calculated with [7]:

\[
F_{\text{edl}} = \frac{\pi a^2 \varepsilon_0 U^2}{2D^2},
\]

where \(a\) is the contact radius, \(\varepsilon_0\) the permittivity of free space, and \(U\) the contact potential.

**Figure 4.8:** Schematic sketch of the electric double layer that forms at the interface of a p–n junction between two semiconductors.
4.2.4 Coulombic Image Force

The last adhesion force that may be of influence, is the Coulomb force that exists between a charged sphere and the surface charge that it induces in an electrically neutral flat surface. From electrostatic theory it is known, that the electric field produced by the electrostatic induction and the charged sphere is identical to the electric field produced by two oppositely charged point charges. The force between the surface and the sphere is therefore equal to the force between two oppositely charged point charges separated by a distance $2R$. This force is often called the *Coulombic image force*, and it can be calculated with:

$$F_{\text{cif}} = \frac{q^2}{16\pi\epsilon_0 R^2}, \quad (4.33)$$

where $q$ is the charge of the sphere, $\epsilon_0$ the permittivity of free space, and $R$ the radius of the sphere. To get a rough estimate of the maximum charge a spherical particle may carry, it is assumed that the particle is uniformly charged and completely surrounded by a uniform isolating medium. In that case the electric field strength at the surface of the particle is given by:

$$E = \frac{q}{4\pi\epsilon_0 R^2}, \quad (4.34)$$

which should not exceed the *dielectric strength* $E_{\text{br}}$ of the surrounding medium to prevent breakdown. If breakdown does occur, the surrounding isolating medium suddenly becomes conductive, and the particle loses all or most of its charge. The maximum amount of charge that a uniformly charged sphere can carry is therefore given by:

$$Q = 4\pi\epsilon_0 R^2 E_{\text{br}}, \quad (4.35)$$

which may be substituted into equation (4.33) to obtain the following expression for the Coulombic image force:

$$F_{\text{cif}} = \pi\epsilon_0 R^2 E_{\text{br}}^2. \quad (4.36)$$

In practice, the particle will not be able to carry the maximum amount of charge predicted by equation (4.35), since it is in contact with a surface through which it may lose charge due to leak currents. Equation (4.36) must therefore be interpreted as an upper limit of the Coulombic image force.
4.3 Simple Particle Removal Model

After discussing the various adhesion forces in the previous section, it is now possible to combine adhesion theory with shock theory to create a model that predicts the cleaning capabilities of laser-induced shock wave cleaning. But rather than including each and every adhesion force into the model, it is worthwhile to investigate which adhesion forces are the strongest. To this purpose, figure 4.9 shows a comparison between the various adhesion forces. In this case the van der Waals force is calculated with equation (4.16) for polystyrene latex spheres on a silicon surface. The capillary force due to water condensation is calculated with equation (4.31), where it is assumed that the capillary force reaches its maximum value. Furthermore, the electric double layer force was calculated with equation (4.32) for a contact potential of 0.5 V. The maximum Coulombic image force is calculated with equation (4.36), where the maximum charge on the particle is determined by the dielectric strength of air, which is equal to \(3.00 \cdot 10^6 \text{ V/m} \) [10]. As figure 4.9 shows, it is possible to conclude that for sub-micron particles there are two major adhesion forces of interest: the capillary force and the van der Waals force. Therefore, only these two adhesion forces are considered in the particle removal model, while the remaining two adhesion forces are ignored.

Figure 4.9: Comparison of the various adhesion forces that can act on a spherical polystyrene particle on a silicon surface.
4.3.1 Model Description

The particle removal model that is presented here is based on the geometry that is shown in figure 4.10. As the figure shows, the laser–induced shock wave is created in atmospheric air at a distance $g$ above the surface, after which it expands spherically until it reaches the particle. When the shock wave passes the particle, it experiences a pressure jump $\Delta p$, which depends on the blast wave energy $E$ and the distance the shock wave has to travel before it reaches the particle. The resulting shock wave force $F_{sw}$ on a spherical particle with radius $R$ is then given by:

$$F_{sw} = \pi R^2 \Delta p,$$  \hspace{1cm} (4.37)

where the pressure jump $\Delta p$ is given by equation (2.67). The adhesion force $F_{adh}$ that acts on the particle is composed of the van der Waals force $F_{vdW}$ and in some cases also the capillary force $F_{cap}$ due to condensation of water around the contact area. As equation (4.31) shows, the capillary force depends strongly on the contact angle between the water and the surface of the particle. Unfortunately, the contact angle is not known, and in this model a worst case scenario is adopted in which the capillary force reaches its maximum value at a contact angle of 0 degrees.

All these forces together result in a cleaning moment $M_c$ and an adhesion moment $M_a$ which both try to rotate the particle in opposite directions about the pivot.
4.3. Simple Particle Removal Model

Point P:

\[ M_c = F_p \cdot l_y, \]  
\[ M_a = F_n \cdot (l_x + a) + (F_{vdW} + F_{cap}) \cdot a, \]

where \( a \) is the contact radius, \( F_n \) and \( F_p \) are the normal and parallel component of the shock wave force, and \( l_x \) and \( l_y \) are the lever arms as shown in figure 4.10. To initiate a rolling motion of the particle, the cleaning moment must be larger than the adhesion moment, which leads to the following cleaning condition:

\[ \frac{M_c}{M_a} > 1. \]

As the results in section 4.3.4 show, this cleaning condition becomes exceedingly difficult to satisfy for smaller particles. The main goal of this particle removal model is therefore to find the smallest removable particle for a given gap distance \( g \) and blast wave energy \( E \).

Finally, this model assumes that the particles are made from polystyrene latex, and that the surface from which the particles must be removed is made from silicon. In this case equation (4.22) shows that the Tabor coefficient is equal to 0.3 for a 10 nm particle, and increases for particles with a larger diameter. This means that the contact mechanics between the particle and the surface can be described by JKR theory. It should also be remembered that under certain circumstances the particle is also subjected to a capillary force, which is interpreted as an external load \( F \) on the particle. The capillary force therefore does not only contribute to a stronger bonding by providing an additional adhesion force, but it also enhances the bond by increasing the contact area due to the additional deformation of the particle.

4.3.2 Limitations

By using equation (4.37) to calculate the shock wave force, it is implicitly assumed that the air behaves as a continuum. However, this principle is clearly violated when the model is applied to particles which have a diameter comparable to or smaller than the mean free path length of the air molecules. To determine the limiting effect of the continuum approach, it is therefore of interest to examine the mean free path length \( l_1 \) of the air molecules directly behind the shock wave:

\[ l_1 = \frac{1}{n_1 \sigma}, \]

where \( \sigma \) is the collisional cross section, and \( n_1 \) the number of air molecules per unit volume directly behind the shock wave. The molecular density \( n_1 \) can be rewritten
Particle Adhesion Theory

into a more convenient form:

\[ n_1 = \frac{\rho_1}{m} = \frac{\rho_0}{m} \left( \frac{\rho_1}{\rho_0} \right) = n_0 \left( \frac{\rho_1}{\rho_0} \right), \quad (4.42) \]

where \( m \) is the averaged molecular mass, and the density ratio \( \rho_1/\rho_0 \) is given by shock relation (2.35). This yields the following expression:

\[ l_1 = l_0 \left( \frac{\gamma - 1}{\gamma + 1} \right) M_0^2 + 2, \quad (4.43) \]

where \( l_0 \) is the mean free path length of the molecules in the undisturbed air, and \( M_0 \) the Mach number of the shock wave. To determine the validity of the continuum approach, the particle removal model uses equation (4.43) to compare the mean free path length with the diameter of the smallest removable particle.

Furthermore, the model is also affected by the restrictions placed on the use of equation (2.38), which is part of the derivation that results in equation (2.67) for the pressure jump \( \Delta p \) across the shock wave. As is indicated in chapter 2, equation (2.38) is an approximate shock relation which is only accurate for strong shock waves with \( M_0 > 4 \). It should also be kept in mind that the expression for the pressure jump \( \Delta p \) is also based on the Taylor & Sedov solution. One of the key assumptions of the Taylor & Sedov solution is that the pressure \( p_0 \) of the undisturbed air may be neglected when compared to the peak pressure \( p_1 \) directly behind the shock wave. Therefore, there should be at least one order of magnitude difference between both pressures, which results in the requirement \( p_1/p_0 > 10 \). This requirement may be rewritten in a more useful form by introducing shock relation (2.34):

\[ M_0^2 > \frac{11 \gamma + 9}{2 \gamma}, \quad (4.44) \]

where \( \gamma \) is the specific heat capacity ratio. Using \( \gamma = 7/5 \) for air then yields \( M_0 > 3 \), which is less stringent than the requirement \( M_0 > 4 \) introduced by the use of equation (2.38).

Finally, equation (4.37) is also based on the assumption that the particle experiences the full pressure jump generated by the shock wave. While this is always true for an ideal shock wave with a discontinuous pressure profile, it is certainly not true for real gases which exhibit shock wave thicknesses in the order of the mean free path length. For the particle to experience the full shock wave force, the shock wave thickness must therefore not exceed the diameter of the smallest removable particle. In summary, the particle removal model is thus limited by the following
4.3. Simple Particle Removal Model

Figure 4.11: Reduced contact area of a rough particle and a rough surface [1].

three requirements:

- Continuum limit: \( l_1 < 2R_{\text{min}} \), (4.45)
- Taylor & Sedov: \( M_0 > 4 \), (4.46)
- Shock wave thickness: \( \lambda < 2R_{\text{min}} \), (4.47)

where \( \lambda \) is the shock wave thickness. For strong shock waves the shock wave thickness approaches the mean free path length [8], which leads to the conclusion that equations (4.45) and (4.47) are both equivalent, and that there are only two unique requirements. These two requirements limit the ability of the particle removal model to accurately predict the removal capabilities of laser induced shock wave cleaning. But as section 4.3.4 shows, the particle removal model is still able to produce very interesting results despite these restrictions.

In this particular particle removal model the particle is also assumed to be spherical, which simplifies the geometric relations and allows the direct application of the adhesion theory discussed in this chapter. In reality however, particles are not perfectly spherical, and their shape may have a profound influence on the particle removal capabilities of laser-induced shock wave cleaning. Especially particles with a high aspect ratio, such as flakes, are expected to be more difficult to remove due to their large contact area and relatively small frontal area on which the shock wave can exert a cleaning force. The material choice is expected to be less influential, since the Hamaker constant is relatively constant for most materials [13]. Furthermore, the roughness of the surface is neglected in the particle removal model, which has a tendency to reduce the strength of both the van der Waals and the capillary
force [1]. The main reason for this is the reduced contact area between the particle and the surface, as is illustrated in figure 4.11.

### 4.3.3 Implementation

As mentioned before, the main goal of this particle removal model is to predict the diameter of the smallest particle that can be removed for a given gap distance $g$ and blast wave energy $E$. The algorithm is implemented in MATLAB, and its structure is summarized in a flow diagram, which can be found in figure 4.12. This flow chart shows that the particle removal model iterates through a range of gap distances and blast wave energies, and determines the smallest particle diameter at which cleaning condition (4.40) is still satisfied. Furthermore, the particle removal model also checks whether or not the requirements that were discussed in the previous section are satisfied.

### 4.3.4 Results

The particle removal model presented in this chapter was applied to polystyrene latex spheres on a silicon surface, because these materials are also used for the particle removal experiments presented in the next chapter. Figure 4.13 shows the results of the simulation when the particles are subjected to both the van der Waals force and the capillary force. From these results it may be concluded that for a typical set of cleaning parameters ($g = 1.0$ mm and $E = 100$ mJ) the model predicts that the diameter of the smallest removable particle is equal 218 nm. It should be remembered however, that the results in figure 4.13 are based on the worst case scenario, in which the capillary force reaches its maximum value. Therefore, the particle removal model was also applied to polystyrene latex spheres on a silicon surface in the absence of capillary forces. The results of this simulation can be found in figure 4.14, and it shows that for this particular set of cleaning parameters the diameter of the smallest removable particle reduces to 54 nm. Therefore, it is expected that in this particular case the diameter of the smallest removable particle lies somewhere in the range (54–218) nm. It may therefore be concluded, that the particle removal model predicts the removal of sub–micron particles and warrants further experimental investigation.

### 4.4 Conclusions

As this chapter shows, it may be concluded that a broad range of electrostatic interactions at the molecular level is responsible for the macroscopic adhesion forces that bind small sub–micron particles to a surface. While there are various types of
Figure 4.12: Flow diagram of the particle removal model.
Figure 4.13: Diameter of the smallest removable particle as predicted by the particle removal model in the presence of van der Waals and capillary forces.

Figure 4.14: Diameter of the smallest removable particle as predicted by the particle removal model in the presence of van der Waals forces only.
adhesion forces responsible for this binding, it was also shown that the capillary and van der Waals force are the strongest for sub–micron particles. The particle removal model that is presented at the end of this chapter, therefore only considers these two adhesion forces and combines them with shock theory to predict the diameter of the smallest removable particle for a given amount of blast wave energy $E$ and gap distance $g$. The results obtained with this model clearly predict that laser–induced shock wave cleaning should be capable of removing sub–micron particles from a surface, which will be confirmed later by the particle removal experiments in the next chapter.

References


Particle Adhesion Theory


Abstract. In this chapter the laser–induced shock wave cleaning method is evaluated by removing polystyrene latex and silica particles from silicon wafers. The results of these experiments show that it is very well possible to remove particles with a diameter of 60 nm and below from the surface of the wafer. Unfortunately, it also has to be concluded that the laser–induced shock wave cleaning method is too harsh for the surface of an EUV photomask. Damage evaluation on multilayer mirrors shows that for effective cleaning parameters the cleaning process leads to irreversible damage to the multilayer structure, or an unacceptable decrease in reflectivity. In light of these results, it has to be concluded that laser–induced shock wave cleaning is an effective method to remove small sub-micron particles from a surface, but that it is not suitable for the removal of particles from the surface of an EUV photomask.
5.1 Laser–Induced Shock Wave Cleaning Setup

Since this thesis is about particle removal, it would not be complete without a series of cleaning experiments. In this chapter the laser–induced shock wave cleaning method is tested by removing polystyrene latex and silica particles from the surface of flat unpatterned wafers. For this purpose a laser–induced shock wave cleaning setup has been build, and its capabilities and limitations are discussed in this section. Furthermore, the deposition and detection of submicron particles are also an important part of the cleaning experiments, and are therefore also discussed in more detail.

5.1.1 Specifications

The setup that was used for the laser–induced shock wave cleaning experiments is shown in figure 5.1. This particular cleaning setup is intended for the cleaning of 300 mm silicon wafers, since EUV photomasks are deemed too costly to use for these potentially damaging experiments. Furthermore, it is also very difficult
and time consuming to inspect the surface of a patterned photomask, whereas the inspection of a flat wafer can be performed within a few minutes with the help of industrial inspection tools. As the figure shows, the setup consists of a Nd:YAG laser, an optical delivery system, and a high precision programmable wafer table. The Nd:YAG Q-switched laser in this setup is used to produce a laser-induced plasma above the surface of the wafer, and operates at a wavelength of 1064 nm with a pulse length of 10 ns. The optical delivery system consists of three mirrors, a focusing lens, and a prism to deliver the focused beam at the proper position and height. The focusing lens has a focal length of 75 mm and a plano-convex shape for optimal focusing. The motorized wafer table is controlled by software that was developed in-house for this particular setup, and makes it possible to quickly and easily move the wafer table in an automated fashion. The entire setup is mounted on a standard optical table, which is located inside an ISO class 6 cleanroom.

5.1.2 Limitations

The cleaning setup depicted in figure 5.1 is limited in several ways. One of these limitations is the inability of the setup to produce laser-induced breakdown at low laser pulse energies. As was shown in chapter 2, there is a minimum threshold intensity that has to be achieved before laser-induced breakdown can be initiated. This limitation may be alleviated by using a laser pulse with a shorter pulse length or a focusing lens with a shorter focal length.

Besides a minimum laser pulse energy, there also exists a minimum gap distance at which the laser-induced plasma may be created above the surface of the wafer. If the gap distance is too small, the laser-induced plasma will make contact with the surface and cause irreversible damage. Furthermore, it is also possible that the surface of the wafer is exposed to direct laser irradiation because the laser-induced plasma does not completely absorb all laser radiation. The transmitted portion of the laser light may also damage the surface of the wafer, which is only aggravated further by the necessity to tilt the beam towards the surface to produce very small gap distances.

Another important limitation of the cleaning setup is the low repetition rate at which a reliable laser-induced plasma can be created. Experiments have shown that in normal air at atmospheric pressure and room temperature, the maximum repetition frequency for this particular cleaning setup is approximately 4 Hz. When the repetition frequency approaches this maximum value, the breakdown probability decreases very rapidly. This limitation in repetition frequency is probably caused by the turmoil left by the previous laser-induced breakdown, which prevents proper

---

1 The air in an ISO class 6 cleanroom contains less than $10^6$ particles/m$^3$ with a diameter larger than 100 nm.
focusing of the next laser pulse. The low maximum repetition frequency of the setup makes it impossible to clean a complete photomask of \((15 \times 15) \text{ cm}^2\) within a reasonable amount of time (less than 12 minutes). This problem may be solved by using fast opto-mechanics to move the laser-induced plasma across the surface, rather than moving the surface underneath the laser-induced plasma, as is the case with the setup in figure 5.1. This makes it possible to create each successive breakdown in undisturbed air, and the maximum repetition frequency is now limited by the maximum repetition frequency of the laser and the speed of the opto-mechanics.

5.2 Cleaning Experiments

In this section the cleaning setup in figure 5.1 is used to experimentally determine the feasability of laser-induced shock wave cleaning as a method to remove polystyrene latex and silica particles from the surface of a silicon wafer. But before presenting the results of these experiments, this section first discusses the method by which the particles are deposited and detected, followed by a short discussion on the most important cleaning parameters.

5.2.1 Particle Deposition and Detection

Before any cleaning experiment can take place, it is important to have a method for depositing and detecting particles on the surface of a 300 mm silicon wafer. When depositing particles, great care has to be taken to ensure that the particles on the...
5.2. Cleaning Experiments

Figure 5.3: Principle of the KLA Tencor Surfscan SP2 inspection tool.

surface are nicely monodispersed, and not clustered as can be seen in figure 5.2. A very simple method to achieve this is spin coating, where the wafer is coated with particles by introducing a particle suspension to the surface of a spinning wafer. Under the influence of the centrifugal force, the particle suspension spreads out very rapidly over the surface of the wafer and usually produces very good samples with nicely monodispersed particles. Unfortunately, this method could not be used since it is rather time consuming and requires a large and impractical setup to accommodate the large 300 mm wafers.

Another much more practical and convenient particle deposition method is aerosol spraying, which was used to prepare the samples for the cleaning experiments presented in this thesis. This method of particle deposition is much simpler and faster than spin coating, once the equipment has been set up. The main component of the aerosol spraying equipment is the atomizer, which transforms the particle suspension into an aerosol. This aerosol is then brought into contact with the wafer, where a mask is used to deposit the particles in a predetermined circular area. This method is very suitable for the deposition of very small particles below 100 nm, and it also produces very good samples with monodispersed particles.

Once the particles have been deposited, it is necessary to determine how many particles have been deposited, as well as their size and position. For these type of experiments most researchers typically use a scanning electron microscope as their main inspection tool. The advantage of this method is a very low detection threshold, and a direct visual conformation of the cleaning results. On the other hand, the limited field of view renders this method impractical for the inspection
of large surface areas. Therefore, the results presented in this thesis where not obtained with a scanning electron microscope, but with an industrial KLA–Tencor Surfscan SP2 inspection tool [1]. The main advantage of this particular inspection tool is its ability to fully inspect the surface of a flat 300 mm wafer within a couple of minutes. As figure 5.3 shows, the Surfscan SP2 is able to achieve this high throughput by rotating the wafer while it is illuminated by a beam of light from a laser source. Particles present on the surface will scatter a portion of the incident laser light, thereby revealing their position and size. With this method it is possible to detect particles that are much smaller than the wavelength of the laser that is used for illumination.

5.2.2 Cleaning Parameters

In chapter 2 it is theorized that the most important cleaning parameters of laser–induced shock wave cleaning are the amount of absorbed laser energy $E$ and the distance between the focal point and the surface, also called the gap distance $g$. While the results in this chapter support this conclusion, it has also become apparent that the are several more parameters that can have a significant influence on the cleaning result. When laser–induced shock wave cleaning is used to clean a large surface, it is necessary to create multiple laser–induced plasmas across the surface. Therefore, a meandering cleaning path is used as shown in figure 5.4, where the spacings $\Delta x$ and $\Delta y$ between each point along the cleaning path must be small.
5.2. Cleaning Experiments

enough to ensure good cleaning between the points. Furthermore, it is also possible to create multiple successive laser–induced plasmas at the same point along the cleaning path to enhance the cleaning capabilities.

5.2.3 Removal of Polystyrene Particles

In practice, the surface of the EUV photomask is contaminated with particles of different sizes, materials and shapes. While these properties can have an influence on the cleaning capabilities of laser–induced shock wave cleaning, it is important to ensure that each cleaning experiment is performed under the same conditions. Therefore, the performance of laser–induced shock wave cleaning was examined by cleaning flat silicon wafers purposely contaminated with polystyrene latex spheres (manufacturer: Duke Scientific, Inc.), which have a very well–defined shape and size.

Before each cleaning experiment approximately 2000–3000 polystyrene latex spheres with sizes of 60–150 nm were deposited in a 40 mm circular area on a 300 mm wafer by means of the aerosol method. Directly after deposition\(^2\) the wafer was inspected with the KLA Tencor Surfscan SP2, and a square area of 10×10 mm\(^2\) within the contaminated area was treated with laser–induced shock wave cleaning. During cleaning, a prescribed number of laser–induced plasma shockwaves were created at equidistant points along a meandering path, as indicated in figure 5.4. Afterwards, the wafer was inspected again to determine the cleaning performance in terms of the Particle Removal Efficiency (PRE), which is defined as follows:

\[
\text{PRE} = \left( \frac{N_0 - N_1}{N_0} \right) \times 100\%, \quad (5.1)
\]

where \(N_0\) and \(N_1\) are the number of particles present before and after the cleaning procedure within a specified area.

Three different measurement series were performed to investigate the effect of (1) the number of pulses per position, (2) the laser pulse energy \(E\), and (3) the gap distance \(g\). During each cleaning experiment the horizontal and vertical spacings between each cleaning position along the meandering cleaning path were kept at a constant value of 1.0 mm. Furthermore, the Q–switched Nd:YAG laser operated at a repetition frequency of 2.0 Hz to create the laser–induced plasma. The results of the cleaning experiments are summarized in figures 5.5–5.7, which show the wafer map after cleaning and a size dependent PRE histogram for each separate laser–induced shock wave cleaning treatment.

As expected, the best cleaning occurs after more intense laser–induced shock

\(^2\)Please note that aging can have a large influence on the adhesion force between particle and surface, see for example [2].
Cleaning Experiments

\[ g = 1.0 \, \text{mm}, \, E = 206 \, \text{mJ}, \, \Delta x = 1.0 \, \text{mm}, \, \Delta y = 1.0 \, \text{mm}, \, 1 \, \text{pulse/position} \]

\[ g = 1.0 \, \text{mm}, \, E = 206 \, \text{mJ}, \, \Delta x = 1.0 \, \text{mm}, \, \Delta y = 1.0 \, \text{mm}, \, 5 \, \text{pulses/position} \]

\[ g = 1.0 \, \text{mm}, \, E = 206 \, \text{mJ}, \, \Delta x = 1.0 \, \text{mm}, \, \Delta y = 1.0 \, \text{mm}, \, 10 \, \text{pulses/position} \]

**Figure 5.5:** Wafer maps and PRE histograms resulting from laser-induced shock wave cleaning of polystyrene latex spheres with varying number of pulses per position. The PRE histograms correspond to the 10×10 mm² cleaned area indicated in the wafer map.
5.2. Cleaning Experiments

\[ g = 1.0 \text{ mm}, \ E = 101 \text{ mJ}, \ \Delta x = 1.0 \text{ mm}, \ \Delta y = 1.0 \text{ mm}, \ 10 \text{ pulses/position} \]

\[ g = 1.0 \text{ mm}, \ E = 151 \text{ mJ}, \ \Delta x = 1.0 \text{ mm}, \ \Delta y = 1.0 \text{ mm}, \ 10 \text{ pulses/position} \]

\[ g = 1.0 \text{ mm}, \ E = 206 \text{ mJ}, \ \Delta x = 1.0 \text{ mm}, \ \Delta y = 1.0 \text{ mm}, \ 10 \text{ pulses/position} \]

Figure 5.6: Wafer maps and PRE histograms resulting from laser-induced shock wave cleaning of polystyrene latex spheres with varying laser pulse energy. The PRE histograms correspond to the 10×10 mm² cleaned area indicated in the wafer map.
Cleaning Experiments

\( g = 2.0 \text{ mm}, E = 206 \text{ mJ}, \Delta x = 1.0 \text{ mm}, \Delta y = 1.0 \text{ mm}, 10 \text{ pulses/position} \)

\[ \text{Particle Size [nm]} \]

\[ \text{PRE [%]} \]

Figure 5.7: Wafer maps and PRE histograms resulting from laser–induced shock wave cleaning of polystyrene latex spheres with varying gap distance. The PRE histograms correspond to the 10×10 mm² cleaned area indicated in the wafer map.
wave cleaning treatments with a large number of pulses per position, high pulse energy and small gap distance. In fact, using 10 pulses per position, 206 mJ of laser pulse energy and a gap distance of 1.0 mm resulted in the complete removal of all 60–150 nm PSL spheres from the treated area. The results also clearly show that the cleaning result is indeed very sensitive to the gap distance, as was predicted by equation 2.66. For example, figure 5.7 shows that the PRE improved significantly when the gap distance was lowered by only 25% from 2.0 mm to 1.5 mm. Furthermore, figure 5.6 confirms the importance of the laser pulse energy by showing a significant increase in PRE after the pulse energy was increased by 50% from 101 mJ to 151 mJ. But perhaps more surprisingly, the number of pulses per position has a strong influence on the cleaning result, as can be seen in figure 5.5. This result seems to suggest that when the shock wave is too weak to fully detach the particle, it is still able to partially detach or loosen the particle from its surface. Another possibility is that the laser–induced plasma interacts with the polystyrene latex spheres, and (partially) removes them by means of plasma chemistry or through interaction with the ultraviolet radiation produced by the plasma. In any case, cumulative effects also play an important role in the laser–induced shock wave cleaning process.

The cleaning results presented in figures 5.5–5.7 also present an opportunity to directly compare the experimentally found cleaning capabilities with the the predictions of the particle removal model presented in the previous chapter. The results of this comparison are summarized in table 5.1, and show that the particle removal model is fairly accurate in predicting the correct diameter of the smallest removable particle. Please note that the experimental removal criterion is rather arbitrary, and in this particular case a PRE of 50% was chosen to ensure that more particles were removed from the cleaned area than that there were left. Also note that some of the experimental results are not included in the comparison, because the particle removal model is not able to give an accurate prediction for these cleaning parameters due to the limitations described in section 4.3.2.

<table>
<thead>
<tr>
<th>$g$ [mm]</th>
<th>$E$ [mJ]</th>
<th>Predicted $d_{\text{min}}$</th>
<th>Experimental $d_{\text{min}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>206</td>
<td>32–135</td>
<td>&lt; 60</td>
</tr>
<tr>
<td>1.0</td>
<td>151</td>
<td>40–164</td>
<td>&lt; 60</td>
</tr>
<tr>
<td>1.0</td>
<td>101</td>
<td>53–216</td>
<td>95 ± 5</td>
</tr>
<tr>
<td>1.5</td>
<td>206</td>
<td>80–319</td>
<td>75 ± 5</td>
</tr>
</tbody>
</table>

Table 5.1: Comparison between the diameter $d_{\text{min}}$ of the smallest removable particle as predicted by the particle removal model presented in the previous chapter and the cleaning results presented in figures 5.5–5.7
5.2.4 Removal of Silica Particles

The results in the previous subsection show that laser–induced shock wave cleaning is able to remove small polystyrene latex spheres with a diameter of 100 nm or less from the wafer surface. It was also observed that the number of cleanings per position has a large influence on the PRE, which may be attributed to the interaction between the plasma and the polystyrene latex spheres. To test this hypothesis, laser–induced shock wave cleaning was also tested with inorganic silica particles, which are much less likely to be removed due to plasma chemistry than the organic polystyrene spheres. As figure 5.8 shows, the results of these experiments show that the cleaning of silica particles is more difficult than the cleaning of polystyrene latex spheres. There is however a substantial amount of cleaning at a gap distance of 1.0 mm and a laser pulse energy of 175 mJ, which can be attributed to the interaction between the shock wave and the particles, and not plasma chemistry.

5.3 Evaluation of Surface Damage

A major concern for the practical applicability of laser–induced shock wave cleaning, is the possibility of damage to the surface of the photomask. As was shown in figure 1.6, the EUV photomask is composed of a substrate covered with a multilayer mirror. The patterned absorber on top of the multilayer mirror contains the pattern of the integrated circuit layer. Especially the multilayer mirror, as well as the patterned absorber, are particularly fragile and thus vulnerable to damage. To determine the suitability of laser–induced shock wave cleaning as a method of removing particles from an EUV photomask, it is therefore also necessary to verify that the cleaning process does not damage its surface.

5.3.1 Damage Experiments

Unfortunately, the amount of damage that an EUV photomask is allowed to endure is extremely small. Its main task is to reflect back the image of the patterned circuit layer as intended by its design, and any changes in reflectivity can cause serious imaging errors. The amount of reflectivity the multilayer on an EUV photomask may lose during 20–100 cleanings during its lifetime must therefore be smaller than 1% globally and 0.1% locally. To test whether or not these stringent tolerances can be met, a series of damage experiments were performed on small multilayer mirrors that are identical to the multilayer mirror on an EUV photomask. The results that have been compiled in figures 5.9 and 5.10 show very clearly that effective cleaning parameters lead to irreversible damage to the multilayer mirror. In some cases were the gap distance is small and the laser pulse energy is high, the laser–induced
5.3. Evaluation of Surface Damage

\( g = 1.0 \text{ mm}, E = 107 \text{ mJ}, \Delta x = 1.0 \text{ mm}, \Delta y = 1.0 \text{ mm}, 10 \text{ pulses/position} \)

![Wafer map and PRE histogram for laser-induced shock wave cleaning with 107 mJ pulse energy.](image)

\( g = 1.0 \text{ mm}, E = 143 \text{ mJ}, \Delta x = 1.0 \text{ mm}, \Delta y = 1.0 \text{ mm}, 10 \text{ pulses/position} \)

![Wafer map and PRE histogram for laser-induced shock wave cleaning with 143 mJ pulse energy.](image)

\( g = 1.0 \text{ mm}, E = 175 \text{ mJ}, \Delta x = 1.0 \text{ mm}, \Delta y = 1.0 \text{ mm}, 10 \text{ pulses/position} \)

![Wafer map and PRE histogram for laser-induced shock wave cleaning with 175 mJ pulse energy.](image)

Figure 5.8: Wafer maps and PRE histograms resulting from laser-induced shock wave cleaning of silica particles with varying laser pulse energy. The PRE histograms correspond to the 10×10 mm² cleaned area indicated in the wafer map.
Figure 5.9: Types of damage caused by laser–induced shock wave cleaning. The microscope image on the left side shows mild damage in the form of discoloration and a rippling pattern across the surface. The microscope image on the right side shows severe structural damage to the multilayer mirror.

Figure 5.10: Results of the multilayer damage experiment. The red and yellow dots refer to the severe and mild damage shown in figure 5.9, the green dots correspond to no observable damage. A selected number of dots also show the experimentally determined diameter $d_{\text{min}}$ of the smallest removable particle that can be removed with a PRE above 50%.
cleaning process can lead to structural damage of the multilayer. On other cases it leads to discoloration of the surface, which is clearly observable by the naked eye. In all these cases it can therefore be concluded that laser–induced shock wave cleaning is unable to meet the reflectivity tolerances, and that the cleaning method is too harsh for the fragile surface of an EUV photomask.

5.4 Conclusions

The results in this chapter show that laser–induced shock wave cleaning is capable of removing both polystyrene latex and silica particles smaller than 60 nm from the surface of a flat unpatterned silicon wafer. The damage experiments however, have shown that the exposure of an EUV photomask to an effective laser–induced shock wave treatment would result in unacceptable damage to its surface. This inevitably leads to the conclusion that even though laser–induced shock wave cleaning can be an effective method to remove particles, it is not suitable for the cleaning of EUV photomasks.

References


Cleaning Experiments
CHAPTER 6

Conclusions
6.1 Shock Theory

In chapter 2 the properties of the shock wave were theoretically investigated with the help of shock theory and the Taylor & Sedov solution. From the results of this investigation it was possible to conclude that the strength of a shock wave produced by the laser-induced breakdown in a gas depends on three key parameters. The most influential of these three parameters is the distance traveled by the shock wave, followed by the amount of absorbed laser pulse energy and the specific heat capacity ratio of the gas. It was therefore theorized that the effectiveness of laser-induced shock wave cleaning can be increased first and foremost by decreasing the gap distance. This was later confirmed by both the particle removal model in chapter 4 and the particle removal experiments in chapter 5.

6.2 Time Resolved Imaging

The time resolved images obtained by high-speed shadowgraphy and Schlieren photography reveal the shape, size, and position of the shock wave as it travels through the gas. From these images it was concluded that the propagation of the shock wave produced by the laser-induced breakdown in air and argon is consistent with the Taylor & Sedov solution discussed in chapter 2. In addition to the validation of the Taylor & Sedov solution, the images also provide more insight into the energy distribution of the laser-induced breakdown process. By applying the Taylor & Sedov solution it was also possible to show that the radiative losses can be substantial, and that the majority of the emission must occur somewhere during the first 100 ns after the ignition of the laser-induced plasma. When the laser-induced breakdown of gases is applied to the removal of small particles from a surface, it may therefore also be important to consider the possibility of surface damage caused by the exposure to the intense radiation emitted directly after plasma ignition. Unfortunately, the damage experiments in chapter 5 have shown that this is indeed the case.

6.3 Particle Removal Model

In chapter 4 it was shown that small submicron particles on a flat surface experience relatively large adhesion forces, which may be orders of magnitude greater than the weight of the particle itself. Therefore, the interaction mechanisms between the submicron particles and the surface have been investigated in more detail in chapter 4, from which it was concluded that the strongest adhesion force results from the capillary condensation of moisture into the cracks and crevices surrounding the contact area. Furthermore, it was also shown that for dry environments in which
the *capillary force* is absent, the *van der Waals force* becomes the most dominant adhesion force. The particle removal model that is presented at the end of this chapter, therefore only considers these two adhesion forces and combines them with shock theory to predict the diameter of the smallest removable particle for a given amount of absorbed laser pulse energy and gap distance. The results obtained with this model clearly predict that laser–induced shock wave cleaning should be capable of removing sub–micron particles from a surface, which is confirmed by the particle removal experiments in chapter 5.

### 6.4 Cleaning Capabilities

The cleaning experiments in chapter 5 have shown that laser–induced shock wave cleaning can be an effective method for the removal of small submicron particles. The results show that it is possible to fully remove 60 nm polystyrene latex spheres from the surface of a silicon wafer at a gap distance of 1.0 mm and a laser pulse energy of 200 mJ. In this particular case the inspection of the wafer surface was limited by the detection threshold of the KLA Tencor SP2 Surfscan inspection tool, and it is very well possible that particles with a diameter well below 60 nm are also removed. Furthermore, it can be concluded that the gap distance and the laser pulse energy have a large influence on the particle removal efficiency of laser–induced shock wave cleaning, as was predicted in chapters 2 and 4.

Besides the gap distance and laser pulse energy, it was also found that the number of cleanings per position has a large influence on the particle removal efficiency. From this it can be concluded that cleaning by laser–induced shock wave cleaning is a cumulative process, where the particles are partially detached by the shock wave before they are removed. To ensure that the cumulative effect is not solely caused by chemistry between the plasma and the organic polystyrene latex spheres, the particle removal experiments were repeated with inorganic silica particles. The results of these experiments show that the particle removal efficiency of silica particles is substantial, but less than the particle removal efficiency of polystyrene latex spheres. Therefore, it is possible that the polystyrene latex particles are partially removed by plasma chemistry or UV radiation, but the removal of silica particles also clearly shows that the interaction with the shock wave is the main removal mechanism.

### 6.5 Surface Damage

The results of the damage experiments in chapter 5 clearly show that a laser–induced shock wave treatment that is able to remove sub 100 nm particles, would result in unacceptable damage to the surface of an EUV photomask. In this case the damage
Conclusions

to the surface was clearly visible by the naked eye, ranging from discoloration to structural damage of the multilayer mirror. In this case it also very difficult, if not impossible, to apply an effective damage mitigation strategy, since this would most probably involve blocking some parts of the incident shock wave and radiation. While such a strategy may decrease damage, it also prevents the shock wave from removing the particles from the surface. In a more elaborate strategy, the direct line of sight between the plasma and the surface would be blocked, and the shock wave is redirected along an indirect route to the surface by means of reflection. Unfortunately, this strategy is also flawed, because the shock wave has to travel over a much longer distance, and has lost much of its potency when it reaches the surface. This inevitably leads to the conclusion that even though laser-induced shock wave cleaning can be an effective method to remove particles, it is not very suitable for the cleaning of EUV photomasks.
Addendum
7.1 Introduction

Even though the simple particle removal model presented in chapter 4 performs relatively well, it is also interesting to further explore the particle removal capabilities of laser–induced plasmas through the use of an ab initio numerical model. Such a numerical simulation would require several key components, which can be summarized as follows:

- Beam propagation code.
- Plasma model.
- Compressible flow solver.
- Turbulence model.

A very important aspect of the numerical simulation is the propagation, deflection and absorption of the laser beam as it propagates through the simulation domain, since this determines the local intensity of the laser light, and therefore also the rate of ionization. To describe the evolution of the plasma, it is often convenient to assume Local Thermal Equilibrium (LTE), which means that both the electrons and heavy particles may be described by the same temperature. While this results in a relatively simple plasma model, it should always be remembered that the LTE assumption must be validated before drawing any conclusions from the simulation results. The next component in the list is the compressible flow solver, which describes the transport of species within the plasma. In this particular case it is very important that the flow solver is able to simulate compressible flows, since shock waves only appear in compressible media. Finally, the numerical model also needs to include a turbulence model to predict the swirling motion of the gas produced by the laser–induced breakdown. This is of particular importance to predict the transport of particles that may have been dislodged from the surface during the laser–induced shock wave cleaning process.

The simulation platform that was chosen for this task was Plasimo [2, 16], which is a framework for modeling plasmas developed at the Eindhoven University of Technology. One of the shortcomings of the Plasimo framework is the absence of turbulence, which may play an important role in the flows generated by laser–induced plasmas. Since almost all flows of engineering importance involve some degree of turbulence modeling, the implementation of a turbulence model within the Plasimo framework was therefore long overdue. In this chapter the turbulence problem is addressed by the introduction of the averaged Navier–Stokes equations, which together form an unclosed system with many unknown quantities. Turbulence modeling is therefore key in solving the averaged Navier–Stokes system by reducing the
7.1. Introduction

Figure 7.1: Drawings made by Leonardo da Vinci depicting turbulence generated by a flow into a pool and turbulence generated by objects placed in a flow.

Figure 7.2: Paintings depicting hints of turbulence made by the Dutch painter Vincent van Gogh entitled Starry Night and Wheat Fields with Stacks.

number of unknowns. The modeling approach taken in this thesis is based on the Boussinesq approximation in combination with an eddy–viscosity model called the \( k-\varepsilon \) model. This well–established turbulence model is a popular choice for many research scientists, because it is self–consistent and widely validated. But before concerning ourselves with the details of turbulence modeling, the remainder of this introduction is dedicated to a more general discussion of turbulence and a brief historical overview of turbulence modeling.

7.1.1 Historical Context

One of the first persons that systematically investigated turbulence was the Italian Renaissance icon Leonardo da Vinci. Being an artist as well as a scientist, he was able to make very detailed sketches of turbulent flows. Figure 7.1 shows two of
his now famous sketches, depicting the turbulence generated by a flow into a pool and the turbulence generated by objects placed in a flow. But Leonardo da Vinci may not have been the only artist that recognized turbulent phenomena in nature. As figure 7.2 shows, the Dutch painter van Gogh made several paintings that show evidence of turbulence. In fact, some researchers [1] claim that some of the paintings that he made during his more turbulent years, are so accurate that they agree remarkably well with modern turbulence theories.

The first real scientific advances towards an understanding of turbulence were made by Osbourne Reynolds in the late nineteenth century. He realized that the seemingly irregular motion of turbulent flows could be decomposed into a mean part and a fluctuating part. This idea was of such profound importance for the further development of more advanced turbulence theories, that it is now simply referred to as the Reynolds decomposition. Not surprisingly the Reynolds decomposition also forms the basis of the turbulence theory in this thesis, and is therefore discussed in more detail in section 7.2.2. Joseph Boussinesq, a contemporary of Osbourne Reynolds, also immortalized himself in turbulence literature, by introducing the concept of an eddy viscosity in an attempt to model the turbulent stresses as a molecular gradient–diffusion process. His modeled expression for the turbulent stresses is nowadays known as the Boussinesq approximation, of which a more detailed discription can be found in section 7.2.4.

Further progress in the field of turbulence modeling was made in the beginning of the twentieth century, when Ludwig Prandtl discovered the boundary layer. Later in his career Prandtl also introduced the important concept of mixing length, which describes the distance a fluid parcel can travel before it loses its properties due to mixing with the surrounding fluid. Combined with a straightforward recipe to calculate the turbulence viscosity, the mixing length model became the dominant turbulence model for the following two decades. Unfortunately, the mixing length varies greatly for different flow types, and must be specified manually for each individual calculation. The mixing length model is therefore not complete, which means that is not possible to obtain a solution by merely specifying the initial and boundary conditions.

The first complete turbulence model was formulated by Andrei Kolmogorov in 1942. His model is now known as the $k-\omega$ model, and features a set of modeled differential equations for the turbulence kinetic energy $k$ and the specific dissipation $\omega$. This type of model is often labeled as a two–equation model, because of the number of differential equations it contains. At the time however, the $k-\omega$ model was of little practical use because of the lack of adequate computing power. Over the years different types of two–equation models have been formulated, of which Laun- der and Sharma’s $k-\varepsilon$ model from 1974 is probably the most well–known. This
particular turbulence model is often called the standard $k-\varepsilon$ model and is discussed in section 7.3.2.

With the availability of high-performance computers, it is nowadays possible to use more complex models and methods such as Large Eddy Simulation (LES) and Direct Numerical Simulation (DNS). In principle DNS can be regarded as the holy grail of turbulence modeling, since it tries to solve the instantaneous Navier–Stokes equations for turbulent flows. This approach requires enormous computational resources, as the grid has to be extremely fine and the time step extremely short. As a compromise, one may choose to use LES to solve the Navier–Stokes equations for the larger turbulence scales, and use turbulence modeling such as a two-equation model to solve the smallest turbulence scales. Both methods are considered more advanced, and will not be covered in this thesis.

### 7.1.2 The Nature of Turbulence

Looking at our surroundings, we can conclude that turbulence is all around us. The mixing of milk in a cup of coffee, the flow of water around the pillar of a bridge, or the wake behind a boat, are all examples of turbulence. On a more grand scale there are also beautiful examples of less mundane manifestations of turbulence, such as the atmospheric von Kármán vortex street or the vulcanic ash plume shown in figures 7.3 and 7.4.

But what exactly makes a turbulent flow turbulent? What is needed is a formal definition of turbulence. Unfortunately, until this day turbulence is still a somewhat enigmatic phenomenon, and many definitions are in use. One of the most generally accepted definitions of turbulence was given by G.I. Taylor in 1937:

> “Turbulence is an irregular motion which in general makes its appearance in fluids, gaseous or liquid, when they flow past solid surfaces or even when neighboring streams of the same fluid flow past or over one another.”

This definition implicitly states that high shear stresses, as commonly found near walls and neighbouring streams, are a prerequisite for the development of turbulence. This is indeed confirmed in section 7.3.1, where it is shown that the production of turbulence kinetic energy is closely linked to the shear stresses.

One of the most striking features of turbulence is the presence of local swirling motions called eddies. As Leonardo da Vinci already observed hundreds of years ago, turbulent flows usually contain a few large macroscopic eddies while the smaller ones are too numerous to count. This observation gave the English physicist and meteorologist Lewis Fry Richardson the idea of an energy cascade. He proposed that the energy is transported from larger eddies to smaller ones without the loss
Figure 7.3: Landsat 7 image of a von Kármán vortex street off the Chilean coast near the Juan Fernández Islands (also known as the Robinson Crusoe Islands). Source: NASA Earth Observatory.

Figure 7.4: Photograph taken from the International Space Station of an ash plume arising from Mount Cleveland in Alaska. Source: NASA Earth Observatory.
of energy, until the eddy becomes so small that the turbulence kinetic energy is converted into heat by viscous dissipation.

Besides the production and dissipation of turbulence kinetic energy, another important feature of turbulence is the enhanced diffusivity. This is what gives turbulent flows their enhanced mass, momentum and energy transport, otherwise known as mixing. The improved mixing behaviour of turbulent flows can have a significant impact on inhomogeneous and chemically reactive fluids, and will therefore be discussed in greater detail in section 7.5.

7.2 The Turbulence Closure Problem

After some introductory remarks, it is now time to put the description of turbulence on a firmer foundation. Since turbulent flows are highly irregular and chaotic, statistical methods are often used to analyze their behaviour. A common approach is to express the flow quantities as the sum of mean and fluctuating parts. These are then substituted into the instantaneous Navier-Stokes equations, after which the equations are averaged. Unfortunately, the decomposition into mean and fluctuating parts has introduced a large number of additional variables, while the number of equations in the averaged Navier-Stokes system has remained the same. Since the number of unknowns exceeds the number of equations, one cannot hope to find a solution. It is thus necessary to use closure approximations to reduce the number of unknowns. These matters and more will be addressed in this section, starting with a short review of the instantaneous Navier-Stokes equations.

7.2.1 Instantaneous Flow Equations

In the absence of external forces, the flow of unsteady compressible fluids is described by the following Navier-Stokes (NS) system:

\[ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \]
\[ \frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla \cdot (\rho \mathbf{v} \mathbf{v}) = -\nabla p + \nabla \cdot \mathbf{S}, \]
\[ \frac{\partial (\rho e)}{\partial t} + \nabla \cdot (\rho (h + \frac{1}{2} \mathbf{v} \cdot \mathbf{v}) \mathbf{v}) = -\nabla \cdot \mathbf{q} + \nabla \cdot (\mathbf{v} \cdot \mathbf{S}), \]

where \( \rho \) is the density, \( \mathbf{v} \) the velocity, \( p \) the pressure, \( e \) the specific total energy, \( h \) the specific enthalpy, \( \mathbf{q} \) the heat flux, and \( \mathbf{S} \) the deviatoric stress tensor. The specific
total energy $e$ and the specific enthalpy $h$ are defined as follows:

$$
e = u + \frac{1}{2} (\mathbf{v} \cdot \mathbf{v}), \quad (7.2)$$

$$h = e + p/\rho, \quad (7.3)$$

where $u$ represents the specific internal energy. For compressible Newtonian fluids the deviatoric stress tensor $\mathbb{S}$ can be written in the following form:

$$
\mathbb{S} = \mu \left( \nabla \mathbf{v} + \nabla \mathbf{v}^T \right) + \zeta (\nabla \cdot \mathbf{v}) \mathbb{I} \\
= 2\mu \left( \mathbb{E} - \frac{1}{3} (\nabla \cdot \mathbf{v}) \mathbb{I} \right), \quad (7.4)
$$

where $\mu$ is the shear viscosity, $\zeta = -\frac{2}{3} \mu$ the bulk viscosity\(^1\), and $\mathbb{E} = \frac{1}{2} (\nabla \mathbf{v} + \nabla \mathbf{v}^T)$ the rate-of-strain tensor. Furthermore, it is possible to rewrite the heat flux $\mathbf{q}$ with Fourier’s law of heat conduction:

$$
\mathbf{q} = -k \nabla T, \quad (7.5)
$$

where $k$ is the heat conduction coefficient, and $T$ the temperature. Together with an equation of state (e.g. the ideal gas law) the above set of equations forms a fully closed system. For instance, for a perfect gas the relevant equations of state are:

$$
p = \rho RT, \quad (7.6)$$

$$u = c_V T, \quad (7.7)$$

where $R$ is the specific gas constant, and $c_V$ the heat capacity at constant volume.

The closed Navier–Stokes system fully describes the flow of compressible Newtonian fluids. This also includes turbulence, and in principle it is possible to use these equations to simulate turbulent flows. Unfortunately, such simulations would require very fine grids and very short time steps to capture every facet of the turbulent flow. In general one can say that the computational effort required for the Direct Numerical Simulation (DNS) of turbulent flows is too high to be of any practical importance for all but the simplest of cases.

### 7.2.2 Reynolds Decomposition

Because of the random and chaotic nature of turbulent flows, it is often assumed that a fluctuating flow variable $\phi(\mathbf{x}, t)$ at position $\mathbf{x}$ and time $t$, can be decomposed into a mean and fluctuating component. This technique is illustrated in figure 7.5, and is often referred to as Reynolds decomposition. The mean component of the

\(^1\)Please note that this assumption is only valid for mono–atomic gasses.
7.2. The Turbulence Closure Problem

Figure 7.5: Reynolds decomposition.

flow variable can be obtained in a number of ways, depending on the type of flow involved. For example, for steady flows the mean value is usually calculated by time averaging (also called Reynolds averaging). However, by its nature, time averaging is less suitable to describe unsteady flows. In these cases ensemble averaging is a much more appropriate method:

\[
\phi(x, t) = \lim_{N \to \infty} \sum_{n=1}^{N} \phi_n(x, t) + \phi'(x, t) = \langle \phi \rangle + \phi',
\]

(7.8)

where \( \phi_n(x, t) \) is the \( n \)th sample measurement in a series of \( N \) identical experiments. Furthermore, the ensemble average of the fluctuating component is always equal to zero, i.e. \( \langle \phi' \rangle = 0 \). Another particularly useful averaging method is density weighted averaging:

\[
\phi(x, t) = \frac{1}{\langle \rho \rangle} \lim_{N \to \infty} \sum_{n=1}^{N} \rho_n(x, t) \phi_n(x, t) + \phi''(x, t) = \{ \phi \} + \phi'',
\]

(7.9)

where the average value of the fluctuating component is also equal to zero, i.e. \( \{ \phi'' \} = 0 \). This averaging method is called Favre averaging, and it is often used for compressible flows, where it greatly simplifies the derivation of the averaged flow
Turbulence Modeling

equations.

7.2.3 Averaged Flow Equations

Favre averaging of the instantaneous Navier–Stokes system makes it possible to derive a system of flow equations that describes the behaviour of the mean flow variables. The sometimes long and cumbersome derivations of these averaged flow equations have been omitted here. The resulting averaged flow equations closely resemble their instantaneous counterparts:

\[
\frac{\partial}{\partial t} \langle \rho \rangle + \nabla \cdot ( \langle \rho \rangle \{ \mathbf{v} \}) = 0, \quad (7.10a)
\]

\[
\frac{\partial}{\partial t} (\langle \rho \rangle \{ \mathbf{v} \}) + \nabla \cdot (\langle \rho \rangle \{ \mathbf{v} \} \{ \mathbf{v} \}) = -\nabla \langle p \rangle + \nabla \cdot (\langle \mathbf{S} \rangle - \langle \rho \mathbf{v}'' \mathbf{v}'' \rangle), \quad (7.10b)
\]

\[
\frac{\partial}{\partial t} (\langle \rho \rangle \mathbf{E}) + \nabla \cdot (\langle \rho \rangle \{ \mathbf{v} \} \mathbf{H}) = \nabla \cdot \left( -\langle \mathbf{q} \rangle - \langle \rho \mathbf{v}'' \mathbf{v}'' \rangle \right) + \\
\nabla \cdot (\langle \mathbf{v}'' \cdot \mathbf{S} \rangle - \frac{1}{2} \langle \rho (\mathbf{v}'' \cdot \mathbf{v}'') \mathbf{v}'' \rangle) + \\
\nabla \cdot (\{ \mathbf{v} \} \cdot (\langle \mathbf{S} \rangle - \langle \rho \mathbf{v}'' \mathbf{v}'' \rangle)), \quad (7.10c)
\]

where \( E \) and \( H \) represent the total energy and total enthalpy:

\[
E = \{ u \} + \frac{1}{2} \{ \mathbf{v} \} \cdot \{ \mathbf{v} \} + \frac{1}{2} \{ \mathbf{v}'' \cdot \mathbf{v}'' \}, \quad (7.11)
\]

\[
H = \{ h \} + \frac{1}{2} \{ \mathbf{v} \} \cdot \{ \mathbf{v} \} + \frac{1}{2} \{ \mathbf{v}'' \cdot \mathbf{v}'' \}. \quad (7.12)
\]

As one can see, the averaged continuity equation has the same basic structure as the instantaneous continuity equation. The same is true for the momentum equation, with the exception of an additional term at the right hand side of the equation. The averaged energy equation has also retained its basic structure, but in this case there are five additional terms present. All of these additional terms are in principle unknown, and will be discussed in the next section.

7.2.4 Closure Approximations

After a brief look at the averaged flow equations derived in the previous section, it becomes clear that this system of equations is not fully closed. There are in total
five additional terms present that are unknown and need to be resolved:

\[
\begin{align*}
\langle \rho \boldsymbol{v}'' \cdot \boldsymbol{v}'' \rangle & : \text{Turbulence viscous stresses}, \\
\frac{1}{2} \langle \rho \boldsymbol{v}'' \cdot \boldsymbol{v}'' \rangle & : \text{Turbulence kinetic energy}, \\
\langle \rho h'' \boldsymbol{v}'' \rangle & : \text{Turbulent heat transport}, \\
\langle \boldsymbol{v}'' \cdot \overline{\mathbb{S}} \rangle & : \text{Molecular diffusion}, \\
\frac{1}{2} \langle \rho (\boldsymbol{v}'' \cdot \boldsymbol{v}'') \boldsymbol{v}'' \rangle & : \text{Turbulent transport of turbulence kinetic energy}.
\end{align*}
\]

The term related to the turbulence viscous stresses appears in both the momentum and the energy equation. It is often modeled with the Boussinesq approximation, which interprets the term as a turbulence viscous stress tensor:

\[
\mathbb{T} = -\langle \rho \boldsymbol{v}'' \cdot \boldsymbol{v}'' \rangle = \mu_t \left( \nabla \langle \boldsymbol{v} \rangle + \nabla \langle \boldsymbol{v} \rangle^T \right) - \frac{2}{3} \mu_t (\nabla \cdot \langle \boldsymbol{v} \rangle) \mathbb{I} - \frac{2}{3} \langle \rho \rangle \{k\} \mathbb{I}
\]

\[
= 2\mu_t \left( \langle \overline{E} \rangle - \frac{1}{3} (\nabla \cdot \langle \boldsymbol{v} \rangle) \mathbb{I} \right) - \frac{2}{3} \langle \rho \rangle \{k\} \mathbb{I}, \tag{7.13}
\]

where \(\mathbb{T}\) is called the Reynolds stress tensor, \(\mu_t\) the turbulence viscosity (also named the eddy viscosity), and \(\{k\} = \frac{1}{2} \{\boldsymbol{v}'' \cdot \boldsymbol{v}''\}\) is the averaged turbulence kinetic energy. This expression looks very similar to the expression for the stress tensor \(\mathbb{S}\) in equation (7.4), but the viscosity \(\mu\) has been replaced by a turbulence viscosity \(\mu_t\), and an additional term \(-\frac{2}{3} \langle \rho \rangle \{k\} \mathbb{I}\) has been added to include normal Reynolds stresses. The sum of these normal Reynolds stresses is equal to \(-2 \langle \rho \rangle \{k\}\), which is easily confirmed by taking the trace of \(-\langle \rho \boldsymbol{v}'' \cdot \boldsymbol{v}'' \rangle\):

\[
\text{Tr} \left( -\langle \rho \boldsymbol{v}'' \cdot \boldsymbol{v}'' \rangle \right) = -\langle \rho \rangle \text{Tr} \{\boldsymbol{v}'' \cdot \boldsymbol{v}''\} = -\langle \rho \rangle \{\boldsymbol{v}'' \cdot \boldsymbol{v}''\} = -2 \langle \rho \rangle \{k\}. \tag{7.14}
\]

Clearly, the additional term \(-\frac{2}{3} \langle \rho \rangle \{k\} \mathbb{I}\) in equation (7.13) guarantees that the trace of the modeled expression evaluates to the correct value. Furthermore, it also shows that the normal Reynolds stresses are distributed equally over all three spatial components. This ensures that their sum is always equal to the physically correct value, but it should also be noted that it implicitly assumes an isotropic distribution, which is generally not true.

Having already defined the averaged turbulence kinetic energy as \(\{k\} = \frac{1}{2} \{\boldsymbol{v}'' \cdot \boldsymbol{v}''\}\), it is now relatively easy to rewrite the term that involves the turbulence kinetic energy:

\[
\frac{1}{2} \langle \rho \boldsymbol{v}'' \cdot \boldsymbol{v}'' \rangle = \langle \rho \rangle \frac{1}{2} \{\boldsymbol{v}'' \cdot \boldsymbol{v}''\} = \langle \rho \rangle \{k\}, \tag{7.15}
\]

The term responsible for turbulent heat transport is usually interpreted as a turbulence heat flux vector, which is assumed to be proportional to the mean temperature
Turbulence Modeling

gradient:

\[ \mathbf{q}_t = \langle \rho h'' \mathbf{v}'' \rangle = - \left( \frac{\mu_t}{\Pr_t} \right) \nabla \{ h \} = - \left( \frac{\mu_t c_p}{\Pr_t} \right) \nabla \{ T \}. \quad (7.16) \]

The turbulent Prandtl number \( \Pr_t \) that appears in this equation is defined as the ratio between the momentum eddy diffusivity and the heat transfer eddy diffusivity, and typically has a value between 0.5 and 1.0. These eddy diffusivities are strongly linked to the mixing behaviour of turbulent flows, and will be discussed in greater detail in section 7.5.

As is customary in turbulence modeling, the turbulent transport term is often approximated by an expression that belongs to a gradient–diffusion process. When this expression is combined with the molecular diffusion term, the following approximation is found:

\[ \frac{1}{2} \langle \rho \left( \mathbf{v}'' \cdot \mathbf{v}'' \right) \rangle = \left( \mu + \mu_t / \sigma_k \right) \nabla \{ k \}, \quad (7.17) \]

where \( \sigma_k \) is a closure coefficient. Inserting the modeled expressions and approximations presented in this section into the averaged flow equations (7.10a)–(7.10c) gives:

\begin{align*}
\frac{\partial}{\partial t} \langle \rho \rangle + \nabla \cdot (\langle \rho \rangle \{ \mathbf{v} \}) &= 0, \\
\frac{\partial}{\partial t} (\langle \rho \rangle \{ \mathbf{v} \}) + \nabla \cdot (\langle \rho \rangle \{ \mathbf{v} \} \{ \mathbf{v} \}) &= - \nabla \langle p \rangle + \nabla \cdot (\langle S \rangle + \mathbf{T}), \\
\frac{\partial}{\partial t} (\langle \rho \rangle E) + \nabla \cdot (\langle \rho \rangle \{ \mathbf{v} \} H) &= \nabla \cdot (- \langle \mathbf{q} \rangle - \mathbf{q}_t - \left( \mu + \mu_t / \sigma_k \right) \nabla \{ k \}) + \\
&\quad \nabla \cdot (\{ \mathbf{v} \} \cdot (\langle S \rangle + \mathbf{T})),
\end{align*}

where \( E \) and \( H \) can now be rewritten into:

\begin{align*}
E &= \{ u \} + \frac{1}{2} \{ \mathbf{v} \} \cdot \{ \mathbf{v} \} + \{ k \}, \\
H &= \{ h \} + \frac{1}{2} \{ \mathbf{v} \} \cdot \{ \mathbf{v} \} + \{ k \}.
\end{align*}

All of the five terms containing fluctuating variables have been replaced by their modeled expressions. At first sight this does not seem to be much of an improvement, since the unknown terms have been replaced by other unknowns, such as the turbulence viscosity \( \mu_t \) and the averaged turbulence kinetic energy \( \{ k \} \). Therefore, despite the closure approximations, the system of averaged flow equations still remains unclosed. This problem will be addressed in the next section, where the introduction of eddy viscosity turbulence modeling will provide the necessary tools to obtain full closure.
7.3 Eddy Viscosity Turbulence Modeling

Eddy viscosity models are based on the idea that the additional turbulent stresses can be modeled by an additional viscosity, called the turbulence viscosity or eddy viscosity. It is however important to keep in mind that this turbulence viscosity is not a property of the fluid, but rather a consequence of the local state of turbulence. The value of the turbulence viscosity is usually calculated on the basis of a semi-empirical relationship that is based on dimensional analysis. In this thesis the value of the turbulence viscosity is calculated from the values of the turbulence kinetic energy \( k \) and the turbulence viscous dissipation \( \varepsilon \). The values of both \( k \) and \( \varepsilon \) are supplied by the well-known two-equation \( k-\varepsilon \) turbulence model, and will be examined in more detail in the following subsections.

7.3.1 The \( k-\varepsilon \) Equation

Since almost all two-equation turbulence models contain the turbulence kinetic energy equation, it is worthwhile to examine this equation more closely. The kinetic energy equation can be obtained by multiplying the instantaneous momentum equation by \( k \) and average, which yields:

\[
\frac{\partial}{\partial t} \left( \langle \rho \{ k \} \rangle \right) + \nabla \cdot \left( \langle \rho \{ \mathbf{v} \} \{ k \} \rangle \right) = \mathbb{T} : \nabla \{ \mathbf{v} \} - \langle \mathbb{S} : \nabla \mathbf{v}'' \rangle + \nabla \cdot \left( \langle \mathbb{S} \cdot \mathbf{v}'' \rangle - \langle \rho \mathbf{v}'' k \rangle - \langle p' \mathbf{v} \rangle \right) - \langle \mathbf{v}'' \rangle \cdot \nabla \langle p \rangle + \langle p' \nabla \cdot \mathbf{v}'' \rangle. \quad (7.21)
\]

The derivation of this equation is omitted here, but a detailed derivation may be found in [17]. Each of the terms in equation (7.21) represents a physical process, which are listed below:

\[
\begin{align*}
\mathbb{T} : \nabla \{ \mathbf{v} \} & : \text{Turbulence kinetic energy production}, \\
\langle \mathbb{S} : \nabla \mathbf{v}'' \rangle & : \text{Dissipation of turbulence kinetic energy}, \\
\langle \mathbb{S} \cdot \mathbf{v}'' \rangle & : \text{Molecular diffusion of turbulence kinetic energy}, \\
\langle \rho \mathbf{v}'' k \rangle & : \text{Turbulent transport of turbulence kinetic energy}, \\
\langle p' \mathbf{v} \rangle & : \text{Pressure diffusion}, \\
\langle \mathbf{v}'' \rangle \cdot \nabla \langle p \rangle & : \text{Pressure work}, \\
\langle p' \nabla \cdot \mathbf{v}'' \rangle & : \text{Pressure dilatation}.
\end{align*}
\]

As was the case with the averaged flow equations, the Reynolds decomposition during its derivation has introduced a large number of additional unknowns. Therefore, also some form of turbulence modeling is necessary to be able to solve the turbulence
kinetic energy equation.

Let us start by substituting the Boussinesq approximation from equation (7.13) into the turbulence kinetic energy production term:

\[
P_k = \mathbb{T} : \nabla \{\mathbf{v}\}
\]

\[
= 2\mu_t (\mathbb{E} : \nabla \{\mathbf{v}\}) - \frac{2}{3} \mu_t (\nabla \cdot \{\mathbf{v}\}) \cdot \nabla \{\mathbf{v}\} - \frac{2}{3} \langle \rho \rangle k (\mathbb{I} : \nabla \{\mathbf{v}\})
\]

\[
= 2\mu_t (\mathbb{E} + \mathbb{E}_A) - \frac{2}{3} \mu_t (\nabla \cdot \{\mathbf{v}\})^2 - \frac{2}{3} \langle \rho \rangle k (\nabla \cdot \{\mathbf{v}\})
\]

\[
= 2\mu_t (\mathbb{E} : \mathbb{E}) - \frac{2}{3} \mu_t (\nabla \cdot \{\mathbf{v}\})^2 - \frac{2}{3} \langle \rho \rangle k (\nabla \cdot \{\mathbf{v}\}),
\] (7.22)

where \(\mathbb{E}_A\) is the anti-symmetric rate-of-strain tensor. Also notice that this expression simplifies considerably for incompressible flows, since the divergence of the velocity field is then equal to zero.

Besides the production of turbulence kinetic energy, there is also a term that describes its destruction. As mentioned earlier during the discussion of Richardson’s energy cascade, turbulence kinetic energy is transfered from the larger eddies to smaller ones, until the eddy becomes so small that the turbulence kinetic energy is converted into heat by viscous dissipation. It makes thus sense to model the destruction of turbulence kinetic energy as follows:

\[
\langle \mathbb{S} : \nabla \mathbf{v}'' \rangle = \langle \rho \rangle \{\varepsilon\},
\] (7.23)

where \(\{\varepsilon\}\) represents the averaged turbulence viscous dissipation. At this moment it is not yet clear how to obtain a value for \(\{\varepsilon\}\), but this issue will be addressed shortly in the next subsection during the discussion of the \(k-\varepsilon\) model.

Fortunately, the terms for the molecular diffusion and turbulent transport of turbulence kinetic energy are exactly the same as those already encountered in equation (7.17). Therefore, little effort is required to find the following expression:

\[
\langle \mathbf{v}'' \cdot \mathbb{S} \rangle - \langle \rho \mathbf{v}'' k \rangle = \left(\mu + \frac{\mu_t}{\sigma_k}\right) \nabla \{k\}.
\] (7.24)

Unfortunately not much is known about the remaining three terms, and as a result these pressure related terms are often simply ignored. As Wilcox [17] points out, this can lead to significant errors given the influential role that pressure plays in a compressible flow. When the pressure related terms are ignored, the following equation for the turbulence kinetic energy is found after substitution of the remaining modeled expressions:

\[
\frac{\partial}{\partial t} (\langle \rho \rangle \{k\}) + \nabla \cdot (\langle \rho \rangle \{k\} \{\mathbf{v}\}) = \nabla \cdot ((\mu + \mu_t/\sigma_k) \nabla \{k\}) + P_k - \langle \rho \rangle \{\varepsilon\}.
\] (7.25)
7.3.2 The \( k-\varepsilon \) Model

The most well-known and most widely validated two-equation turbulence model today is the \( k-\varepsilon \) model. Both the turbulence kinetic energy \( \{k\} \) and the turbulence viscous dissipation \( \{\varepsilon\} \) can be calculated with the following set of differential equations:

\[
\frac{\partial}{\partial t} (\langle \rho \rangle \{k\}) + \nabla \cdot (\langle \rho \rangle \{k\} \{v\}) = \nabla \cdot ((\mu + \mu_t/\sigma_k) \nabla \{k\}) + \mu_t G - \langle \rho \rangle \{\varepsilon\}, \quad (7.26)
\]

\[
\frac{\partial}{\partial t} (\langle \rho \rangle \{\varepsilon\}) + \nabla \cdot (\langle \rho \rangle \{\varepsilon\} \{v\}) = \nabla \cdot ((\mu + \mu_t/\sigma_\varepsilon) \nabla \{\varepsilon\}) + \mu_t G C_{1\varepsilon} \frac{\{\varepsilon\}}{\{k\}} - C_{2\varepsilon} \langle \rho \rangle \frac{\{\varepsilon\}^2}{\{k\}}, \quad (7.27)
\]

where \( G \) is the turbulence production rate defined by:

\[
P_k = \mu_t G. \quad (7.28)
\]

Similar to the derivation of the \( k \)-equation, the \( \varepsilon \)-equation is derived by multiplying the instantaneous momentum equation with \( \varepsilon \) and average. The derivation of the \( \varepsilon \)-equation is also lengthy and complicated, and is therefore not included in this thesis. Instead, the really interested reader is encouraged to consult [14].

The expression for the turbulence viscosity \( \mu_t \) is a semi-empirical relationship that is derived on the basis of dimensional arguments and experimental data:

\[
\mu_t = \langle \rho \rangle C_\mu \frac{[k]^2}{\{\varepsilon\}}. \quad (7.29)
\]

What remains are the empirical constants in equations (7.26), (7.27) and (7.29), which were found by comprehensive data fitting for a wide variety of flows [11]:

\[
C_\mu = 0.09, \quad \sigma_k = 1.00, \quad \sigma_\varepsilon = 1.30, \quad C_{1\varepsilon} = 1.44, \quad C_{2\varepsilon} = 1.92. \quad (7.30)
\]

With the determination of these last empirical constants, the system of averaged flow equations is now fully closed.

7.4 Near Wall Behaviour

As already mentioned in the introduction, walls play an important role in the production of turbulence. In this section the near wall behaviour of turbulent flows
Figure 7.6: Schematic depiction of the formation of a turbulent boundary layer.

is therefore studied more closely. Of particular interest is the turbulent boundary layer that exists in the very narrow region along the wall, which will be discussed in greater detail in the next subsection. Another consequence of the presence of walls is that the $k$-$\varepsilon$ model needs to be modified to cope with low Reynolds number flows in the vicinity of walls. This leads to the introduction of so-called wall damping functions, which is also discussed in greater detail in one of the following subsections.

7.4.1 Turbulent Boundary Layers

One of the important discoveries in the field of fluid dynamics was the discovery of the boundary layer in 1904 by Prandtl. He found that flows can be subdivided into a very thin viscous boundary layer near the wall, and an inviscid outer region further away from the wall. The complete flow field can then be solved by coupling the boundary layer and inviscid outer region together.

A good starting point to analyze the turbulent boundary layer is the stationary momentum equation for incompressible Newtonian fluids:

$$\rho \left( v \cdot \nabla v \right) = -\nabla p + \mu \nabla^2 v, \quad (7.31)$$

where the viscosity $\mu$ is assumed to be constant. Equation (7.31) is then expanded for a two-dimensional cartesian coordinate system, which yields:

$$\rho \left( u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} \right) = -\frac{\partial p}{\partial x} + \mu \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), \quad (7.32)$$

where $(x, y)$ is the spatial coordinate and $(u, v)$ the velocity vector. The averaged momentum equation can be obtained by taking the ensemble average of both sides of equation (7.32). After a fair amount of tedious algebra, the averaged momentum
7.4. Near Wall Behaviour

equation takes the following form:
\[
\rho \left( \{u\} \frac{\partial \{u\}}{\partial x} + \{v\} \frac{\partial \{u\}}{\partial y} \right) = -\frac{\partial \langle p \rangle}{\partial x} \\
+ \mu \left( \frac{\partial^2 \{u\}}{\partial x^2} + \frac{\partial^2 \{u\}}{\partial y^2} \right) - \rho \left( \frac{\partial}{\partial x} \langle u''u'' \rangle + \frac{\partial}{\partial y} \langle u''v'' \rangle \right). \tag{7.33}
\]

Equation (7.33) can be reduced considerably by using scale analysis:
\[
x \sim L, \quad y \sim \delta, \quad u' \sim v_t, \quad v' \sim v_t, \tag{7.34}
\]
where it should be noted that the turbulent boundary layer thickness \( \delta \) is much smaller than the typical dimension scale \( L \) in the streamwise direction. This leads to the following estimations:
\[
\frac{\partial}{\partial x} \langle u''u'' \rangle \ll \frac{\partial}{\partial y} \langle u''v'' \rangle, \tag{7.35}
\]
\[
\frac{\partial^2 \{u\}}{\partial x^2} \ll \frac{\partial^2 \{u\}}{\partial y^2}. \tag{7.36}
\]

The reduced momentum equation of a turbulent boundary layer now takes the following form:
\[
\langle \rho \rangle \left( \{u\} \frac{\partial \{u\}}{\partial x} + \{v\} \frac{\partial \{u\}}{\partial y} \right) = -\frac{\partial \langle p \rangle}{\partial x} + \frac{\partial}{\partial y} \left( \mu \frac{\partial \{u\}}{\partial y} - \langle \rho u''v'' \rangle \right), \tag{7.37}
\]
which shows that the shear stress \( \tau \) is the sum of a viscous and a turbulent component:
\[
\tau = \mu \frac{\partial \{u\}}{\partial y} - \langle \rho u''v'' \rangle. \tag{7.38}
\]

**Viscous Sublayer**

Even in fully developed turbulent flows there is always a very thin layer near the wall that remains laminar. In this so-called **viscous sublayer** the effects of turbulence are not present, and the expression for the shear stress contains only a viscous component. The flow inside the viscous sublayer is considered to be shear–driven, and has a velocity with a streamwise component only\(^2\). Under these circumstances

\(^2\)This type of flow is also known as a shear–driven Couette flow.
the averaged momentum equation can be rewritten into:

\[
\frac{\partial^2 \{u\}}{\partial y^2} = 0.
\]

The shear stress is constant throughout the entire viscous sublayer, which can be seen after multiplying equation (7.39) by \( \mu \) and integrating it once:

\[
\mu \frac{\partial \{u\}}{\partial y} = \text{constant} = \tau_w,
\]

where \( \tau_w \) is the shear stress at the wall. Applying the no–slip condition and solving for \( \{u\} \) yields the following linear velocity profile:

\[
\{u\} = \left( \frac{\tau_w}{\mu} \right) y
\]

This can be expressed in a more convenient form by introducing the dimensionless velocity \( u^+ \) and wall distance \( y^+ \):

\[
u^+ = \frac{u}{u_\tau},
\]

\[
 y^+ = \frac{\rho y u_\tau}{\mu},
\]

where the friction velocity \( u_\tau \) is given by:

\[
u_\tau = \sqrt{\frac{\tau_w}{\rho}}.
\]

Combining equations (7.41) to (7.43) then gives the following expression for the velocity profile within the viscous sublayer:

\[
u^+ = v^+,
\]

which is only valid for \( y^+ \lesssim 5 \).

**Inner Layer**

Outside the viscous sublayer the remaining part of the turbulent boundary layer can be divided into an *inner layer* and *outer layer*. The inner layer lies closest to the
7.4. Near Wall Behaviour

wall, and is formally characterized by:

$$\eta \ll 1, \quad \eta = \frac{y}{\delta}. \quad (7.46)$$

Inside the inner layer the shear stress is a combination of viscous and turbulence stresses. It is therefore assumed that the velocity profile depends on the density $\rho$, viscosity $\mu$, wall distance $y$, wall shear stress $\tau_w$, and velocity $\{u\}$:

$$\{u\} = f(\rho, \mu, \{u\}, y, \tau_w). \quad (7.47)$$

Using dimensional analysis, the above equation can be rewritten in terms of two independent dimensionless parameters:

$$\frac{u}{u_\tau} = f\left(\frac{\rho_y u_\tau}{\mu}\right), \quad (7.48)$$

where $f$ is a function yet to be determined. Plugging in equations (7.42) and (7.43) then yields the following velocity profile:

$$u^+ = f(y^+), \quad (7.49)$$

which is commonly known as the law of the wall.

Outer Layer

The outer layer lies further away from the wall than the inner layer, and is characterized by $y^+ \gg 1$. The shear stress within the outer layer is determined almost exclusively by turbulence effects. This means that the velocity profile does not depend on the molecular viscosity $\mu$, but rather on the properties of the external flow. In this case the boundary layer thickness $\delta$ becomes an important parameter, since the position relative to the border of the boundary layer determines the amount of influence that the external flow has. Therefore, it is presumed that velocity profile within the outer layer answers to the following relationship:

$$\{u\} = g(\rho, \delta, \{u\}, y, \tau_w) \quad (7.50)$$

where $g$ is a function yet to be determined. Analogous to the inner layer, this expression can be rewritten in a simpler form with the use of dimensional analysis:

$$\frac{\{u_e\} - \{u\}}{u_\tau} = g\left(\frac{y}{\delta}\right), \quad (7.51)$$
where \( \{ u_e \} \) is the external flow velocity at the border of the turbulent boundary layer. Substituting equations (7.42) and (7.43) into equation (7.51) gives the following velocity profile:

\[
 u_e^+ - u^+ = g(\eta),
\]

which is known as the velocity defect law.

**Overlap Layer**

The inner and outer layer within the turbulent boundary layer are not clearly separated. As a consequence, there exists an overlap layer in which both the law of the wall and the velocity defect law are valid. Summing up equations (7.49) and (7.52) then gives the following relationship for the overlap layer:

\[
 u_e^+ = f(y^+) + g(\eta),
\]

which can be rewritten into:

\[
 y \frac{\partial u_e^+}{\partial y} = y \frac{\partial f}{\partial y^+} \cdot \frac{\partial y^+}{\partial y} + y \frac{\partial g}{\partial \eta} \cdot \frac{\partial \eta}{\partial y},
\]

\[
 0 = y \frac{\partial f}{\partial y^+} \cdot \frac{\rho u_\tau}{\mu} + y \frac{\partial g}{\partial \eta} \cdot \frac{1}{\delta},
\]

\[
 0 = y^+ \frac{\partial f}{\partial y^+} + \eta \frac{\partial g}{\partial \eta},
\]

This relationship must hold for each value of \( y^+ \) and \( \eta \), which means that:

\[
 y \frac{\partial f}{\partial y^+} = -\eta \frac{\partial g}{\partial \eta} = \text{constant} \tag{7.57}
\]

\[
 = \frac{1}{\kappa}, \tag{7.58}
\]

where \( \kappa \) is the \textit{von Kármán constant}. Solving for \( f \) and \( g \) yields:

\[
 f(y^+) = \frac{1}{\kappa} \ln(y^+) + A, \tag{7.59}
\]

\[
 g(\eta) = -\frac{1}{\kappa} \ln(\eta) + B, \tag{7.60}
\]
where $A$ and $B$ are integration constants. Plugging these solutions back into equations (7.49) and (7.52) gives the following logarithmic velocity profiles:

$$u^+ = \frac{1}{\kappa} \ln (y^+) + A,$$

(7.61)

$$u_e^+ - u^+ = -\frac{1}{\kappa} \ln (\eta) + B,$$

(7.62)

which are valid for $30 \lesssim y^+ \lesssim 500$. Because of the logarithmic behavior, the overlap layer is often referred to as the log layer.

### 7.4.2 Wall Damping Functions

Experimental evidence has shown that the $k-\varepsilon$ model derived in section 7.3.2 is unable to accurately predict low Reynolds number flows near walls. A common approach to correct this shortcoming of the $k-\varepsilon$ model, is to introduce wall damping functions. In this case the $k-\varepsilon$ model takes the following form:

$$\frac{\partial}{\partial t} \langle \rho \{k\} \rangle + \nabla \cdot \langle \rho \{k\} \{v\} \rangle = \nabla \cdot ((\mu + \mu_t/\sigma_k) \nabla \{k\}) + \mu_t G - \langle \rho \{\tilde{\varepsilon}\} \rangle,$$

(7.63)

$$\frac{\partial}{\partial t} \langle \rho \{\tilde{\varepsilon}\} \rangle + \nabla \cdot \langle \rho \{\tilde{\varepsilon}\} \{v\} \rangle = \nabla \cdot ((\mu + \mu_t/\sigma_{\tilde{\varepsilon}}) \nabla \{\tilde{\varepsilon}\}) + \mu_t GC_1 f_1 \frac{\{\tilde{\varepsilon}\}}{\{k\}} - C_2 \varepsilon \langle \rho \{\tilde{\varepsilon}\} \rangle f_2 \frac{\{\tilde{\varepsilon}\}^2}{\{k\}} + E,$$

(7.64)

where $\varepsilon = \varepsilon_0 + \tilde{\varepsilon}$. The turbulence viscosity is also corrected with a wall damping function and is given by:

$$\mu_t = \langle \rho \rangle C_\mu f_\mu \frac{\{k\}^2}{\{\tilde{\varepsilon}\}}.$$

(7.65)
The wall damping functions $f_\mu$, $f_1$, $f_2$, $E$ and $\varepsilon_0$ can be calculated in various ways. The Chien model for instance, is known for its favourable numerical properties:

\[
\begin{align*}
    f_\mu &= e^{-0.0115y^+}, \\
    f_1 &= 1, \\
    f_2 &= 1 - 0.22e^{-(Re_t/6)^2}, \\
    \varepsilon_0 &= \frac{2\mu}{\rho} \left( \frac{k}{y^2} \right), \\
    E &= -\frac{2\mu}{\rho} \left( \frac{\overline{\varepsilon}}{y^2} \right) e^{-y^+/2}, \\
    C_\mu &= 0.09, \quad \sigma_k = 1.00, \quad \sigma_\varepsilon = 1.30, \quad C_{1\varepsilon} = 1.35, \quad C_{2\varepsilon} = 1.80,
\end{align*}
\]

where the turbulent Reynolds number is given by:

\[
Re_t = \frac{\langle \rho \rangle \{k\}^2}{\{\varepsilon\} \mu}.
\]

Another wall damping model that is more appropriate for flows that involve detachment and reattachment is the Launder–Sharma model, which takes the following form:

\[
\begin{align*}
    f_\mu &= e^{-3.4/(1+Re_t/50)^2}, \\
    f_1 &= 1, \\
    f_2 &= 1 - 0.3e^{-Re_t^2}, \\
    \varepsilon_0 &= \frac{2\mu}{\rho} \left( \frac{\partial \sqrt{k}}{\partial y} \right)^2, \\
    E &= \frac{2\mu\mu_t}{\rho^2} \left( \frac{\partial^2 U}{\partial y^2} \right)^2, \\
    C_\mu &= 0.09, \quad \sigma_k = 1.00, \quad \sigma_\varepsilon = 1.30, \quad C_{1\varepsilon} = 1.44, \quad C_{2\varepsilon} = 1.92.
\end{align*}
\]

Besides these two examples, there are a lot of other wall damping models that can be found in literature, for example Lam–Bremhorst or Jones–Launder.

### 7.5 Mixing in Turbulent Flows

One of the most important aspects of turbulent flows, that has been overlooked until now, is its influence on mixing. Turbulent flows can dramatically enhance
Mixing in Turbulent Flows

Diffusing patch

Patch after diffusion and transport

Figure 7.7: Mixing of a patch of dye in a turbulent flow. The larger eddies simply translate and distort the entire patch, while the smaller eddies distort its internal structure.

mixing behaviour, which makes it important for flows where multiple species are involved. The mechanism by which the mixing occurs is depicted schematically in figure 7.7. As discussed in section 7.1, turbulent flows contain eddies with a wide range of sizes. Some of these eddies are much larger than the patch of dye shown in figure 7.7, causing the patch to simply move from one position to the other. The smaller eddies however, continuously deform the internal structure of the patch. This creates steep density gradients, which are then smoothed out by molecular diffusion. The combination of both processes is what we perceive as turbulent mixing.

To put this in a more formal context, one has to start with the instantaneous convection–diffusion equation that describes the local mass fraction \( Y_i \) of species \( i \):

\[
\frac{\partial (\rho Y_i)}{\partial t} + \nabla \cdot (\rho \mathbf{v} Y_i) - \nabla \cdot (\rho D_i \nabla Y_i) = S_i, \tag{7.69}
\]

where \( D_i \) is the molecular diffusion coefficient, and \( S_i \) is a source term that describes the production of species \( i \). The averaged convection–diffusion equation can be readily obtained by taking the ensemble average of both sides of the instantaneous equation:

\[
\frac{\partial \langle (\rho Y_i) \rangle}{\partial t} + \nabla \cdot (\langle \rho \rangle \{ \mathbf{v} \} \{ Y_i \}) - \nabla \cdot \mathbf{J}_i = \{ S_i \}, \tag{7.70}
\]

where the diffusion flux \( \mathbf{J}_i \) is given by:

\[
\mathbf{J}_i = \langle \rho \rangle \{ D_i \} \nabla \{ Y_i \} + \langle \rho \rangle \{ D_i'' \nabla Y_i'' \} - \langle \rho \rangle \{ \mathbf{v}'' Y_i'' \}. \tag{7.71}
\]
In a turbulent flow the velocity and mass fraction fluctuations are strongly correlated, and the above expression for the diffusion flux will be dominated by the last term $\langle \rho \rangle \{ \mathbf{v''} Y_i'' \}$. Therefore, the other two terms are only of minor importance, and may be neglected for turbulent flows. For laminar flows on the other hand, the diffusion flux is determined solely by the first term $\langle \rho \rangle \{ D_i \} \nabla \{ Y_i \}$, since the other terms vanish in the absence of turbulence fluctuations. This means that only the middle term $\langle \rho \rangle \{ D_i'' \nabla Y_i'' \}$ of equation (7.71) may be neglected, which results in the following expression for the diffusion flux:

$$
J_i = \langle \rho \rangle \{ D_i \} \nabla \{ Y_i \} - \langle \rho \rangle \{ \mathbf{v''} Y_i'' \}.
$$

(7.72)

The careful observer may have noticed, that even though the expression for the diffusion flux has simplified, it still retains the ability to describe diffusion processes in both laminar and turbulent flows.

In its current form equation (7.72) still contains unknown quantities, which need to be modeled in order to solve the averaged convection–diffusion equation. This is often done by interpreting the unknown term as a Fickian diffusion process:

$$
\{ \mathbf{v''} Y_i'' \} = -D_i \nabla \{ Y_i \},
$$

(7.73)

where $D_i$ is called the turbulence diffusion coefficient. Much like the turbulence viscosity, the turbulence diffusion coefficient is a property of the fluid flow, and not a property of the fluid itself. In most cases empirical data is used to determine the value of the turbulence diffusion coefficient. When no empirical data is available, one may also use the following relationship:

$$
D_i = \frac{1}{\langle \rho \rangle} \frac{\mu_t}{S_{ct}},
$$

(7.74)

where the turbulence Schmidt number $S_{ct}$ is defined as the ratio between turbulence viscosity and turbulence diffusivity of mass:

$$
S_{ct} = \frac{\mu_t}{\langle \rho \rangle D_i}.
$$

(7.75)

The value of $S_{ct}$ depends on the type of flow involved, but experimental data shows that in most cases it is approximately equal to unity [6, 13, 8]. Together with equations (7.73) and (7.74) the expression for the diffusion flux now takes the following form:

$$
J_i = \left( \langle \rho \rangle \{ D_i \} + \frac{\mu_t}{S_{ct}} \right) \nabla \{ Y_i \}.
$$

(7.76)
7.6 Numerical Implementation

After a considerable amount of theoretical groundwork, it is now time to translate the theory of the previous sections into practical numerical schemes. As this section will show, the numerical implementation of the averaged flow equations can be achieved in a rather straightforward manner by replacing the flow variables with a set of their effective counterparts. The $k$–$\varepsilon$ model however, has a reputation of being more problematic. Especially at the beginning of the simulation, the values of $k$ and $\varepsilon$ are prone to wild oscillations, which can lead to non-physical negative values. These issues, as well as others, are also addressed in the second part of this section.

7.6.1 Averaged Flow Equations

Fortunately, the averaged flow equations have the same basic structure as their instantaneous counterparts. Consequently, an existing CFD code that solves the instantaneous flow equations can be relatively easily modified to solve the averaged flow equations. This can be readily achieved by introducing the following collection of effective variables:

$$p_{\text{eff}} = \langle p \rangle + \frac{2}{3} \langle p \rangle k,$$

$$S_{\text{eff}} = 2 (\mu + \mu_t) \left( \langle \| E \| \rangle - \frac{1}{3} (\nabla \cdot \langle v \rangle) \right),$$

$$q_{\text{eff}} = \langle q \rangle + \langle q \rangle + (\mu + \mu_t / \sigma_k) \nabla k,$$

$$e_{\text{eff}} = \{ e \} + k,$$

$$h_{\text{eff}} = \{ h \} + k.$$ (7.77-7.81)

Substituting these expressions into the averaged flow equations (7.18a)–(7.18c) gives:

$$\frac{\partial}{\partial t} \langle p \rangle + \nabla \cdot (\langle p \rangle \{ v \}) = 0,$$ (7.82a)

$$\frac{\partial}{\partial t} (\langle p \rangle \{ v \}) + \nabla \cdot (\langle p \rangle \{ v \} \{ v \}) = - \nabla p_{\text{eff}} + \nabla \cdot S_{\text{eff}},$$ (7.82b)

$$\frac{\partial}{\partial t} (\langle p \rangle \{ e \}) + \nabla \cdot (\langle p \rangle \{ h_{\text{eff}} + \frac{1}{2} \{ v \} \cdot \{ v \} \} \{ v \}) = - \nabla \cdot q_{\text{eff}} +$$

$$\nabla \cdot (\{ v \} \cdot S_{\text{eff}}).$$ (7.82c)

A quick comparison shows that this system of equations is mathematically identical to the system of instantaneous flow equations. It is thus possible to solve the averaged flow equations with an existing CFD code, merely by replacing the appropriate variables by their effective counterparts. However, one should always keep...
in mind that the modified CFD code solves the *averaged* flow field, rather than the instantaneous flow field.

### 7.6.2 The \( k-\varepsilon \) Model

The Plasimo framework possesses a dedicated infrastructure for solving diffusion–convection equations of the following form:

\[
\frac{\partial}{\partial t} (\rho \phi) + \nabla \cdot (\rho \mathbf{v} \phi) = \nabla \cdot (\Gamma \nabla \phi) + S^{(C)} + S^{(L)} \phi, \tag{7.83}
\]

where \( \phi \) is a transport variable, \( \Gamma \) the diffusion coefficient, and \( S^{(C)} + S^{(L)} \phi \) the source term. Recasting the \( k-\varepsilon \) model equations (7.63) and (7.64) in the same form as equation (7.83) yields:

\[
\Gamma_k = \mu + \mu_t/\sigma_k, \tag{7.84}
\]
\[
\Gamma_\varepsilon = \mu + \mu_t/\sigma_\varepsilon. \tag{7.85}
\]

The following source term linearization provides good numerical stability [3]:

\[
S^{(C)}_k = \mu_t G - \rho \varepsilon_0, \tag{7.86}
\]
\[
S^{(L)}_k = -\rho^2 C_\mu f_\mu k^2/\mu_t, \tag{7.87}
\]
\[
S^{(C)}_\varepsilon = \rho C_\mu C_{1\varepsilon} f_\mu f_1 k G + E, \tag{7.88}
\]
\[
S^{(L)}_\varepsilon = -\rho^2 C_\mu C_{2\varepsilon} f_\mu f_2 k^2/\mu_t. \tag{7.89}
\]

Unfortunately, there is no a priori guarantee that the values of the turbulence kinetic energy and turbulence viscous dissipation will remain positive at all grid points throughout the entire numerical simulation. This can inadvertently lead to non-physical negative values of the turbulence viscosity \( \mu_t \) and the diffusion coefficients \( \Gamma_k \) and \( \Gamma_\varepsilon \). It is therefore a good practice to limit the value of the turbulence viscosity as follows:

\[
10^{-5} \leq \frac{\mu_t}{\mu} \leq 10^5, \tag{7.90}
\]

where \( \mu_t/\mu \) is called the *Turbulence Viscosity Ratio* (TVR). By limiting the value of the turbulence viscosity as proposed in equation (7.90), the correct sign of the diffusion coefficients is ensured without the loss of any physical significance. Furthermore, it also improves the numerical stability of the source term calculation, since it effectively removes the possibility of a division by zero.
The source terms of the $k-\varepsilon$ model equations are also not immune to the erratic behaviour of $k$ and $\varepsilon$. To ensure the negativity of the linear source terms the following limits are imposed:

\[
S_k^{(L)} = \min \left[ S_k^{(L)}, 0 \right],
\]
\[
S_{\varepsilon}^{(L)} = \min \left[ S_{\varepsilon}^{(L)}, 0 \right].
\]

Rather than limiting the values of $k$ and $\varepsilon$ directly, the values of the derived source terms are limited, and $k$ and $\varepsilon$ are allowed to take on any value possible. At convergence however, the values of $k$ and $\varepsilon$ should have a physically correct non-negative value.

### 7.7 Validation Experiments

The $k-\varepsilon$ model has been implemented into the Plasimo framework, but it needs to be validated before it can be used. During validation the newly added turbulence code is tested by solving relatively simple problems and comparing the outcome of the simulation with the available experimental data. To test the implementation of the $k-\varepsilon$ model, a number of test cases have been evaluated, viz. turbulent pipe flow, turbulent channel flow, and turbulent flow over a backward facing step.

#### 7.7.1 Pipe Flow

Because of its engineering importance, the behaviour of turbulent flows through cylindrical pipes has been studied extensively. The experimental data obtained from these studies provides an excellent reference to which the results of the $k-\varepsilon$ turbulence model can be compared. Therefore, it comes as no surprise that the first validation experiment involves a fully turbulent flow through a cylindrical pipe.

To test the performance of the $k-\varepsilon$ turbulence model, the results of the numerical simulations were compared to the results of wind tunnel experiments performed by Laufer [10]. In his experiments, Laufer studied the behaviour of fully developed turbulent flows ($Re_D \geq 50,000$) through a simple 10 inch cylindrical pipe. This geometry is easily recreated by employing a cylindrical coordinate system for the simulation. To resolve the viscous sublayer, the computational grid has a high grid density near the wall of the pipe. The correct physical near-wall behaviour of the flow is also ensured by the use of Chien wall damping functions, as was discussed previously in section 7.4.2. The flow through the pipe is driven by an axial pressure gradient, which can be created by specifying the pressure at both the inlet and outlet of the pipe. As is customary, the no-slip condition applies to the wall of the pipe.
∀x ∈ δΩ_{inlet}: \quad \nabla v_z \cdot n = 0, \quad v_r = 0, \\
\nabla k \cdot n = 0, \quad \nabla \tilde{\epsilon} \cdot n = 0, \\
p_{\text{eff}} = p_{\text{inlet}}, \quad \nabla^2 p_{\text{eff}} \cdot n = 0, \quad \nabla \tilde{\epsilon} = 0, \\

∀x ∈ δΩ_{outlet}: \quad \nabla v_z \cdot n = 0, \quad v_r = 0, \\
\nabla k \cdot n = 0, \quad \nabla \tilde{\epsilon} \cdot n = 0, \quad \nabla \tilde{\epsilon} = 0, \\
p_{\text{eff}} = p_{\text{outlet}}, \quad \nabla^2 p_{\text{eff}} \cdot n = 0, \quad \nabla v_z = 0, \quad v_r = 0, \\
\n∀x ∈ δΩ_{wall}: \quad v = 0, \\
\quad k = 0, \quad \nabla \tilde{\epsilon} = 0, \\
\n∀x ∈ δΩ_{axis}: \quad \nabla v_z \cdot n = 0, \quad v_r = 0, \\
\nabla k \cdot n = 0, \quad \nabla \tilde{\epsilon} \cdot n = 0, \quad \nabla \tilde{\epsilon} = 0, \\

Table 7.1: Summary of the boundary conditions used for the numerical simulation of fully developed turbulent flows through a cylindrical pipe.

pipe. The turbulence kinetic energy and turbulence viscous dissipation, as well as the velocity, are governed by homogeneous Neumann boundary conditions at the inlet and outlet, because the flow is considered to be fully developed. Denoting the simulation domain as Ω, the boundary conditions on its boundary δΩ = δΩ_{inlet} ∪ δΩ_{outlet} ∪ δΩ_{wall} ∪ δΩ_{axis} may be summarized as listed in table 7.1. Furthermore, the flow through the pipe is assumed to be incompressible, which is usually a good approximation for flows with Mach numbers below 0.3.

The results of the simulations are shown in figure 7.8, which shows that the calculated velocity profiles are a close match to the experimental data. As expected, the calculated velocity profile is severely flattened, which is caused by the increased turbulence viscosity near the center of the pipe. As figure 7.9 clearly shows, the turbulence viscosity reaches values that are several hundreds of times larger than the molecular viscosity. But the increased turbulence viscosity does not only change the shape of the velocity profile, but also the maximum velocity at which the air can flow through the pipe. This behaviour can be understood by examining the well-known parabolic velocity profile belonging to a laminar pipe flow:

\[ v = -\frac{1}{4\mu} \frac{\partial p}{\partial z} \left(1 - \frac{r^2}{R^2}\right), \]  

(7.93)
Figure 7.8: Simulated velocity profiles generated by Plasimo and experimental velocity profiles measured by Laufer [10] of fully developed turbulent flows through a 10 inch cylindrical pipe.
where \( r \) represent the radial coordinate, \( R \) the radius of the pipe, \( \mu \) the molecular viscosity, and \( -\partial p/\partial z \) the axial pressure gradient. Keeping in mind equation (7.93), one can conclude that the additional turbulence viscosity reduces the maximum velocity by a factor that is approximately equal to the turbulence viscosity ratio. This behaviour was also successfully predicted by the \( k-\varepsilon \) turbulence model.

### 7.7.2 Channel Flow

The second validation experiment involves a fully developed turbulent flow through a 5 inch two-dimensional channel. For this simulation experimental data from Laufer [9] was used to validate the implementation of the \( k-\varepsilon \) turbulence model. In this case a cartesian coordinate system was used to model the geometry of the channel. The grid density near the walls was also increased, in order to resolve the steep gradients within the viscous sublayer. The correct near-wall behaviour of the \( k-\varepsilon \) model was ensured by the use of Chien wall damping functions. For similar reasons as those discussed in the previous subsection, the flow was also assumed to be incompressible.

The boundary conditions on the boundary \( \delta \Omega = \delta \Omega_{\text{inlet}} \cup \delta \Omega_{\text{outlet}} \cup \delta \Omega_{\text{wall}} \) of the simulation domain \( \Omega \) may be found in table 7.2. These boundary conditions are...
7.7. Validation Experiments

\forall x \in \delta \Omega_{\text{inlet}}: \quad \nabla v_x \cdot n = 0, \quad v_y = 0, \\
\quad \nabla k \cdot n = 0, \quad \nabla \tilde{\varepsilon} \cdot n = 0, \\
\quad p_{\text{eff}} = p_{\text{inlet}}, \\

\forall x \in \delta \Omega_{\text{outlet}}: \quad \nabla v_x \cdot n = 0, \quad v_y = 0, \\
\quad \nabla k \cdot n = 0, \quad \nabla \tilde{\varepsilon} \cdot n = 0, \\
\quad p_{\text{eff}} = p_{\text{outlet}}, \\

\forall x \in \delta \Omega_{\text{wall}}: \quad v_x = 0, \quad v_y = 0, \\
\quad k = 0, \quad \nabla \tilde{\varepsilon} = 0, \\
\quad \nabla^2 p_{\text{eff}} \cdot n = 0.

Table 7.2: Summary of the boundary conditions used for the numerical simulation of fully developed turbulent flows through a channel.

similar to those used for the pipe flow simulations in the previous subsection, and show that the flow is driven by a difference in pressure between the inlet and outlet of the channel. Most of the other flow quantities, such as the turbulence kinetic energy, turbulence viscous dissipation and the velocity, are governed by homogeneous Neumann boundary conditions at the inlet and outlet, since the flow is assumed to be fully developed.

Figures 7.10 and 7.11 show that the results of the simulations are in agreement with the experimental data. As was the case with the fully developed turbulent pipe flows in the previous subsection, the velocity profiles are also severely flattened due to the increased turbulence viscosity near the center of the channel. In this case the turbulent boundary layer was also investigated, and it was found to behave as predicted by equations (7.45) and (7.61). As figure 7.12 shows, the dimensionless velocity \( v^+ \) scales linearly with the dimensionless wall distance \( y^+ \) within the viscous sublayer. Outside the viscous sublayer, \( v^+ \) is shown to scale logarithmically with \( y^+ \) in what is fittingly known as the log layer.

7.7.3 Backward Facing Step

The backward facing step problem was first introduced at the 1980–81 AFOSR–HTTM Stanford Conference on Complex Turbulent Flows, as a method to evaluate the performance of turbulence models for more complex flows that involve separation, recirculation and reattachment. Since then it has become one of the most popular and widely used validation methods for turbulence models, due to its simple geometry and the availability of good experimental data.

The geometry of the backward facing step is characterized by a step–wise increase
Figure 7.10: Simulated velocity profiles generated by Plasimo and experimental velocity profiles measured by Laufer [9] of fully developed turbulent flows through a 5 inch two-dimensional channel.
in cross-sectional area within a two-dimensional channel. The height of the channel and the step itself are not fixed, and may vary from experiment to experiment. In this particular case the geometry was chosen identical to the geometry used by Driver & Seegmiller [4]. As figure 7.13 shows, it is characterized by a channel height of $9h$, where the step height $h$ is equal to 1.27 cm. The central inlet velocity is equal to 44.2 m/s, which results in a step based Reynolds number $Re_h$ of $3.78 \times 10^4$ for air.

Figure 7.13 also shows that the flow over the backward facing step is characterized by three important features: separation, recirculation and reattachment. When the flow passes over the backward facing step, the boundary layer separates at the corner point of the step. The resulting free shear layer is what drives the recirculation zone directly behind the step, after which it reattaches to the channel wall again. The distance at which this happens is called the reattachment length $x_r$, and it serves as a good indicator for the performance of the turbulence model. Various experiments [7, 5] have shown that for fully turbulent flows ($Re_h > 6600$) the reattachment length has a relatively constant value of $(7.0 \pm 0.5) h$.

The cartesian grid needed to simulate such a fully turbulent flow over a backward facing step should not only provide a sufficiently high grid density near the walls, but also at the recirculation and reattachment zone. This was achieved by the use of grid stretching functions. In this case the Launder–Sharma wall damping
Figure 7.12: Semi-logarithmic plot generated by Plasimo of the dimensionless wall distance $y^+$ versus the dimensionless velocity $v^+$ of fully developed turbulent flows through a 5 inch two-dimensional channel.

Figure 7.13: Backward facing step as used by Driver & Seegmiller [4].
7.7. Validation Experiments

∀x ∈ δΩ_{inlet}: \begin{align*}
v_x &= v_{inlet}(x), \quad v_y = 0, \\
k &= k_{inlet}(x), \quad \varepsilon = \varepsilon_{inlet}(x), \\
\Delta p_{\text{eff}} \cdot \mathbf{n} &= 0.
\end{align*}

∀x ∈ δΩ_{outlet}: \begin{align*}
\nabla v_x \cdot \mathbf{n} &= 0, \quad v_y = 0, \\
\nabla k \cdot \mathbf{n} &= 0, \quad \nabla \varepsilon \cdot \mathbf{n} = 0, \\
p_{\text{eff}} &= p_{\text{outlet}}.
\end{align*}

∀x ∈ δΩ_{wall}: \begin{align*}
v_x &= 0, \quad v_y = 0, \\
k &= 0, \quad \varepsilon = 0, \\
\nabla^2 p_{\text{eff}} \cdot \mathbf{n} &= 0.
\end{align*}

Table 7.3: Summary of the boundary conditions used for the numerical simulation of fully developed turbulent flows over a backward facing step.

functions were used, since Chien wall damping functions depend too heavily on the dimensionless wall distance \(y^+\), which goes to zero at the point of reattachment. Similar to the previous two validation experiments, the flow was also considered to be incompressible.

Let \(\Omega\) be the domain on which the flow equations need to be solved, then the boundary conditions on its boundary \(\delta\Omega = \delta\Omega_{\text{inlet}} \cup \delta\Omega_{\text{outlet}} \cup \delta\Omega_{\text{wall}}\) may be summarized as listed in table 7.3. This arrangement of boundary conditions dictates the velocity and turbulence profiles at the inlet of the simulation domain. At the outlet it is assumed that the velocity and turbulence profiles are fully developed, and may be described by a homogeneous Neumann boundary condition. The effective pressure that drives the flow is anchored at the outlet at a constant value of \(p_{\text{outlet}} = 10^5\) Pa. At the inlet the effective pressure is simply extrapolated to the boundary, thereby allowing the pressure distribution to take on any form necessary to facilitate the prescribed velocity and turbulence profiles at the inlet.

As was the case with the previous validation experiments, Plasimo was used again to calculate the flow field and turbulence variables. Figure 7.15 shows the calculated pressure distribution, turbulence viscosity, turbulence kinetic energy and the turbulence viscous dissipation distributions. Especially the pressure distribution warrants a closer look, since the overall pressure gradient is in the opposite direction of what one would expect. The pressure at the outlet is higher than the pressure at the inlet because the outlet falls into a region where the flow stagnates. When the flow passes over the step, it is suddenly allowed to expand and accelerates as a result. But as the flow continues, it begins to stagnate due to the body of air that is already in front of the flow. The increased pressure at the outlet should therefore
Figure 7.14: Simulation results generated by Plasimo of the fully turbulent developed turbulent flow over a backward facing step.
be interpreted as a stagnation pressure. Placing the outlet boundary further downstream may illustrate this more clearly, but this is usually not done because of the increased computational cost due to the larger grid size.

Figure 7.15: Streamlines generated by Plasimo of the fully developed turbulent flow over a backward facing step.

More simulation results may be found in figure 7.15, which shows the calculated streamlines. As expected, the detailed view of the step region clearly shows that there is a recirculation zone present directly behind the step. The reattachment length \( x_r \) may be determined by examining the tangential velocity profile at the first wall adjacent row of cells at the bottom of the channel. The result is displayed in figure 7.16, which clearly shows that the reattachment length is equal to 6.1 \( h \). At a first glance this may seem as a bad result, since the experimental data [7, 5] predicts a reattachment length of (7.0 ± 0.5) \( h \). The \( k-\varepsilon \) model however, is known to under-predict the value of the reattachment length. In literature other researchers [12, 15] have reported reattachment lengths of (5.8 – 6.1) \( h \), which is consistent with the result of this validation experiment.
Figure 7.16: Tangential velocity at the first wall adjacent row of cells at the bottom of the channel.

7.8 Conclusions

The simulation results presented in this chapter show that the implemented $k-\varepsilon$ model is performing as expected. Both the turbulent pipe and channel flow experiments produce the correct turbulent velocity profile as well as the expected behaviour in the near wall region. Also the more complicated backward facing step geometry is successfully simulated, which leads to the conclusion that the newly implemented turbulence model is performing as expected. It should be noted however, that the validation experiments presented here involve incompressible flows only. For a more complete validation it is also necessary to perform more validation experiments with incompressible flows.

References


Here at the end of my thesis, I would like to take a minute and thank all the people that have helped and supported me during the past four years. First of all, I would like to thank my PhD advisor Joost van der Mullen and my ASML supervisor Luigi Scaccabarozzi for their help and support. Many thanks also to Luc Steevens and John de Kuster from ASML, who were always able to supply me with the right equipment. Furthermore, I am grateful for the support I received from the people in the EPG group at the Eindhoven University of Technology. And finally, I would like to thank my family for their years of unconditional help and support.
Index

$k\varepsilon$ model, 123
AC Stark shift, 19
Aerosol spraying, 91
Averaged flow equations, 118, 120
Backward facing step, 139
Boussinesq approximation, 119
Capillary force, 74
Cascade ionization, 21
Chapman–Jouget theory, 51
Contact mechanics, 70
Contamination control, 8
Convection–diffusion equation, 131
Coulomb interaction, 62
Coulombic image force, 76
Covalent bonding, 66
Debye interaction, 65
Depth of focus, 4
Dimensionless velocity $v^+$, 126
Dimensionless wall distance $y^+$, 126
Dispersion interaction, 65
Double patterning, 5
Eddy viscosity, 119
Eddy viscosity model, 121
Electric double layer force, 75
Electron density, 14
Electron temperature, 14
Electronegativity, 64
Ensemble averaging, 117
Euler equations, 23
Extreme ultraviolet lithography, 6
Favre averaging, 117
Fermat’s principle, 40
Fourier heat conduction, 116
Friction velocity, 126
Fundamental forces of nature, 62
Gas dynamics, 21
Gladstone–Dale relationship, 42
Hamaker constant, 68
Hydrogen bonding, 64
Immersion lithography, 4
Impact ionization, 21
Induced dipole moment, 65
Induction interaction, 65
Instantaneous flow equations, 115
Integrated circuit, 2
Inverse Bremsstrahlung, 20
Ionic bond, 62
Keesom interaction, 64
Keldysh parameter, 19
Kelvin equation, 72

Laplace pressure, 74
Laser induced breakdown, 13
Law of the wall, 127
Log layer, 129
London interaction, 65

Mach number, 22
Mean free path length, 79
Microlithography, 3
Moore’s Law, 3
Multi photon ionization, 17
Multiple patterning, 5

Navier–Stokes system, 115
Normal shock waves, 23

Optical lithography, 3
Orientation interaction, 64

Péclet number, 22
Particle removal efficiency, 93
Photomask, 3
Photoresist, 4
Polar molecule, 64
Polarizability, 65

Rankine–Hugoniot equation, 25
Rankine–Hugoniot relations, 25
Resolution, 4
Reynolds averaging, 117
Reynolds decomposition, 116
Reynolds number, 22
Reynolds stress tensor, 119

Schlieren illuminator, 48
Schlieren photography, 44
Shadowgraphy, 42
Shock adiabat, 25
Shock relations, 25
Shock strength, 35
Shock wave, 21
Shomate equation, 55
Speed of sound, 22
spin coating, 91
Subsonic flow, 22
Supersonic flow, 22

Tunneling effect, 17, 19
Turbulence diffusion coefficient, 132
Turbulence heat flux, 119
Turbulence kinetic energy, 119, 123
Turbulence production rate, 123
Turbulence Schmidt number, 132
Turbulence viscosity, 119
Turbulence viscous dissipation, 123
Turbulent boundary layer, 124
Turbulent channel flow, 138
Turbulent mixing, 130
Turbulent pipe flow, 135

van der Waals force, 66
Velocity defect law, 128
viscous sublayer, 125

Z–type Schlieren system, 48
15 August 1979

Born in Zeven, Germany

1991 – 1997

Pre-university education at the Grotius College, Heerlen, the Netherlands

1997 – 2008

M.Sc. in Applied Physics, Eindhoven University of Technology, Eindhoven, the Netherlands

2008–2012

Ph.D. project at the Group Elementary Processes in Gas Discharges, Department of Applied Physics, Eindhoven University of Technology, Eindhoven, the Netherlands. Subject: Laser–induced shock wave cleaning of EUV photomasks.