Delft University of Technology

TECHNISCHE
UNIVERSITEIT
DELFT

Refrigeration With And Of Solid State Devices

Author:
Amar Ashok JITURI

GRADUATE SCHOOL OF ELECTRICAL ENGINEERING, MATHEMATICS AND COMPUTER SCIENCE,
AND
FACULTY OF APPLIED SCIENCES
DELFT UNIVERSITY OF TECHNOLOGY,
DELFT,
NETHERLANDS

Thesis submitted as partial requirement for the Masters Degrees of
Electrical Engineering (Microelectronics)
and
Nanoscience

`j.amarashok-1@student.tudelft.nl`

Supervisor:

Prof. Dr. Ir. T.M KLAPWIJK
Prof. Dr. P.M SARRO
Asst. Prof. Dr. Jaap HOEKSTRA
Ir. Nathan VERCRUYSSEN

August 14, 2009

Abstract

Solid state coolers are reliable, cheap and easy to scale down to microscopic scales, hence very promising for on chip refrigeration in microelectronics. The previous decade has shown tremendous interest for applications both in science and industry. In this thesis we present an overview on the state of the art possibilities and applications of solid state cooling in Microelectronic industry. A study based on the different types of coolers, their applicability and limitations. This includes the operating temperature ranges, cooling power and efficiency. Physical and practical advantages and drawbacks of different systems are analysed.

Also presented is a brief literature on cryogenic NIS coolers, their operating principles,efficiencies, and limitations are touched upon.

In that respect, the problems limiting the efficiency of these refrigerators is the excessive heating, of the superconducting leads. The diffusion of hot quasi-particles is of critical importance. The physical processes behind are of strong fundamental interest, especially at low temperatures.

To address the above problem of heating, we conduct an experiment which allows us to study the diffusion and energy relaxation of quasiparticles in a superconducting wire which aids in the design and optimization of the NIS coolers. An attempt has been made identify the relevant physical processes involved. This involves study of nonequilibrium superconductivity, tunnelling, diffusion and relaxation processes.

We present the the basic physics, fabrication technology and experimental techniques.

Contents

1 **Introduction** **3**
  1.1 Motivation 4
  1.2 Overview 5

2 **Principles of Thermometry and Refrigeration with Solid State Devices** **7**
  2.1 Superconducting Tunnel Structures 9
    2.1.1 SINIS Structures 10
  2.2 Limits of Cryo Cooling 12

3 **Device Fabrication** **14**
  3.1 Photoresist Recipe 15
  3.2 Lithography and Development 17
  3.3 Metal Deposition and Oxidation 20
  3.4 Liftoff 22

4 **Device Characterization** **24**
  4.1 Criteria measured at Room temperature 24
    4.1.1 Material Selection 24
    4.1.2 Tunnel Barrier Resistance 24
  4.2 Criteria measured at Low temperature 25
  4.3 Measurements of Device Characteristics 27
  4.4 Device Layout 27

5 **Measurements** **30**
  5.1 Resistance Measurements 30
  5.2 Measurement Setup 31
    5.2.1 Current/Voltage Module 32
    5.2.2 Voltage/Current Source 33
    5.2.3 Voltage Measurement 33
    5.2.4 Current Measurement 33
    5.2.5 PicoWatt Resistance Bridge 33
    5.2.6 Dilution Fridge 34
  5.3 Current-Voltage Measurements 34
    5.3.1 Current Biased Injector Junction 36
    5.3.2 Voltage Biased Detector Junction 37

6 Theory **39**
6.1 Length Scales 39
6.2 Distribution Function 40
6.3 Superconductivity 43
6.4 Quasiparticle Non-equilibrium 46
6.5 Electron Tunneling 49
  6.5.1 Superconductor-Superconductor Tunneling 50

7 Results and Analysis **52**
7.1 Nonequilibrium Effects and Energy Gap deviations 54
7.2 Effect of Josephson Current 57
7.3 Detection and Measurement of Charge Imbalance 60

8 Thermal Management (Cooling) Techniques and Solutions of Solid State Devices **72**
8.1 Materials - Thermal Conductivity 75
  8.1.1 Traditional Materials 75
  8.1.2 Ultrahigh Thermal Conductivity Materials 76
8.2 Thermal Properties at Nanoscale 80
  8.2.1 Phonons 80
8.3 Thermal Management-Techniques Solutions 83
  8.3.1 Fans 83
  8.3.2 MEMS 83
  8.3.3 Refrigeration 84
  8.3.4 Fluidics 84
8.4 Passive Cooling 85
  8.4.1 Thermal Interface Materials 86
  8.4.2 Heat Pipes 91
  8.4.3 Heat Sinks 96
  8.4.4 Vapor Chambers 102
8.5 Active Cooling 104
  8.5.1 Micropumps 105
  8.5.2 Electroosmotic Pumps 112
  8.5.3 Microchannels 115
  8.5.4 Thermoelectric Coolers 120
  8.5.5 Conclusions 125

9 Conclusions **132**

Chapter 1

Introduction

Temperature and power dissipation in Microelectronic VLSI circuits, packages and systems play an important role in the functionality, performance and reliability. Since the electrical properties of these devices are directly dependent on the temperature and power, higher operating temperatures and power dissipation can lead to thermal drift and performance degradation. Variation of the electrical resistance of the device or the package is an important issue which normally increases with the temperature and with power. Fig 1.1 gives the power distribution for Intel microprocessors up to a Pentium II-300 which uses approximately 40W of power. Also the usage can be seen to increase exponentially over the years. In just a decade's time, the power has increased from about 2W for the 386 to 40W for the Pentium II 30. The second chart displays the power consumption of the microprocessor over the last 25 years to have doubled every 3 years!.

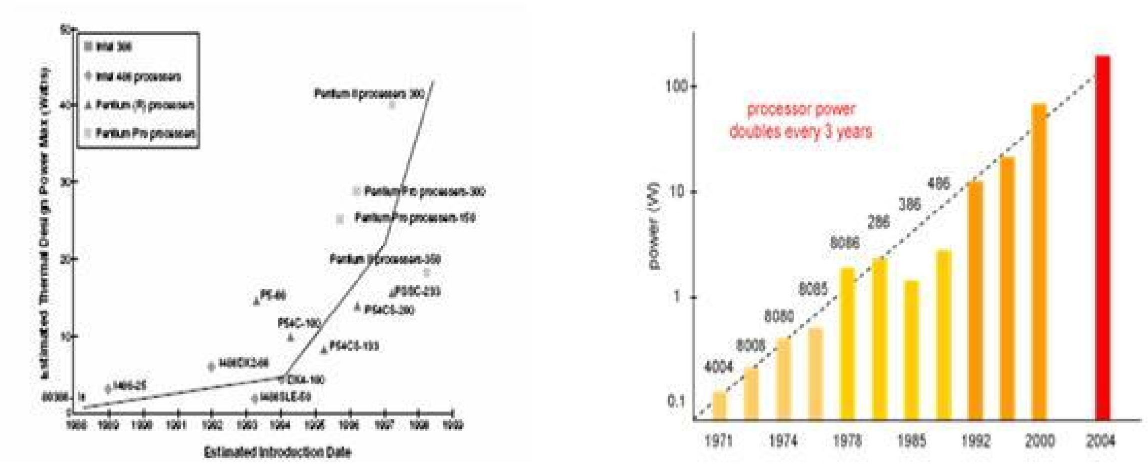
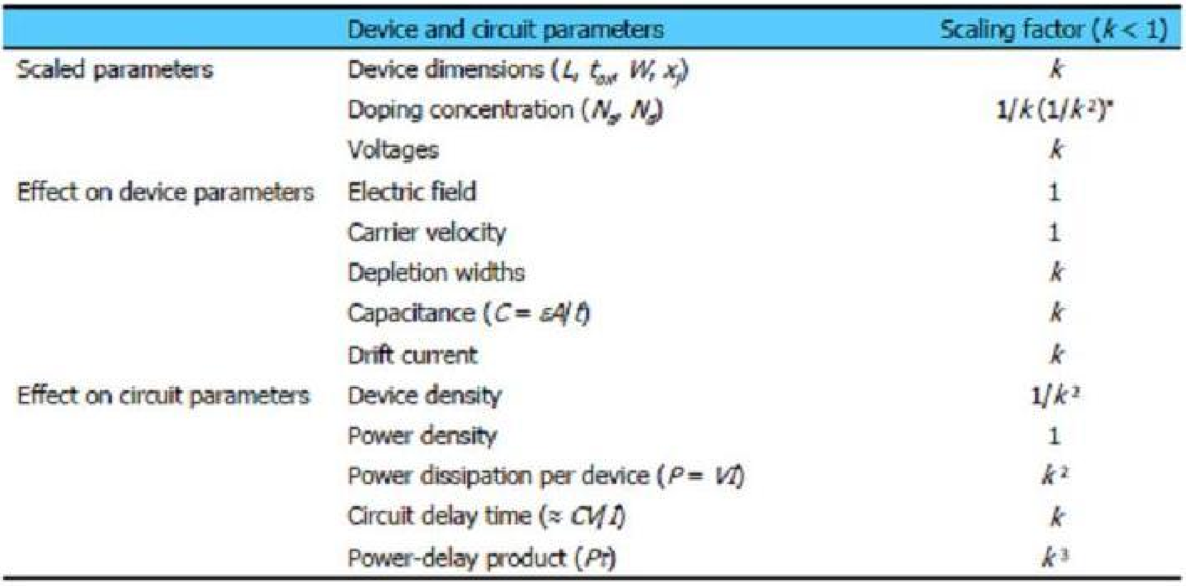


Figure 1.1: Power dissipation in Intel's microprocessor family over the years [28]

The general trend involved in shrinking the device dimensions is the so called "Constant Field Scaling" which means, that the device dimensions and voltages are scaled such that the electric fields remain constant in order to ensure the reliability is not compromised. Fig 1.2 shows the variation of device parameters in constant field scaling Although, the voltage is supposed to be reduced by the scaling factor, in reality is kept constant. This



	Device and circuit parameters	Scaling factor ($k < 1$)
Scaled parameters	Device dimensions (L, t_{ox}, W, x_j)	k
	Doping concentration (N_a, N_d)	$1/k\ (1/k^2)^*$
	Voltages	k
Effect on device parameters	Electric field	1
	Carrier velocity	1
	Depletion widths	k
	Capacitance ($C = \varepsilon A/t$)	k
	Drift current	k
Effect on circuit parameters	Device density	$1/k^2$
	Power density	1
	Power dissipation per device ($P = VI$)	k^2
	Circuit delay time ($\approx CV/I$)	k
	Power-delay product ($P\tau$)	k^3

Figure 1.2: Summary of constant field device scaling
[162]

increases the power dissipation per device and contributes directly to the heat flux. This has created major problems for designers working in the areas of thermal management to come up with new and innovative techniques of heat dissipation in keeping with the pace of the semiconductor industry.

On the other hand, there are solid state electronic devices that operate at milli-Kelvin temperatures. These devices also dissipate heat but, the usual run of the mill method used for cooling the microelectronic IC devices are practically useless. For these devices knowledge of quantum mechanics is essential in order to employ efficient heat dissipation techniques. The reason being, at such low temperatures, the electron temperature and the lattice temperature can be different and due to the low interaction between the electrons and phonons, the hot electrons do not readily transfer their energy to the lattice.

1.1 Motivation

It is clear that managing heat dissipation is an important task for both the general integrated circuits and the highly sophisticated croygenic devices. In this thesis we take a close and detailed look at the state of thermal management techniques employed in the Microelectronic Industry for cooling of solid state integrated circuits. Divided as active and passive coolers, a review of different types of implementations is presented with their advantages, disadvantages, efficiency and limitations.
For the cryogenic devices, we study refrigeration techniques using tunnel junctions where the junction is made between a normal metal and a superconducting metal. Such junctions are popularly known as NIS junctions and we take a look at their operation and some implementations, variations, efficiency and limitations.

We carry out an experiment with a tunnel junction having superconductor metal on either sides. Such junctions are popularly known as SIS junctions. In this experiment we induce non-equilibrium in a superconducting wire and study how this non-equilibrium diffuses and decays. The motivation for this experiment arises from the fact that this non equilibrium is one of the major factors that affects the performance of the cryogenic refrigeration devices.

We utilize the concept of superconducting tunneling since it has been proven to be extremely useful for the creation and detection of non equilibrium states. Another facet of non equilibrium is the concept of charge imbalance, that is an exchange of charge between the superconducting pairs and the quasiparticles. Many non equilibrium states exhibit charge imbalance and the relaxation of this imbalance back to equilibrium determines the properties of these states. The concept of charge imbalance is particularly useful in describing properties of normal-superconductor metal interfaces, Josephson devices, thermoelectric effects etc. In one of the reviews, Giazotto et al [13] has enumerated a number of applications relating to thermometry and refrigeration. The review mainly concentrates on applications exploiting the non-equilibrium phenomena like in thermometry, where the use of SIS and NIS superconducting tunnel junctions is explained. Other applications mentioned are Thermal Detectors like Hot-electron Bolometers and Electronic Refrigeration

1.2 Overview

The next chapters are arranged as follows:

- In Chapter 2 the refrigeration of cryogenic devices using NIS and SINIS tunnel junctions is explained.
- Chapter 3 gives details on the fabrication of the devices used for the experiments in this thesis. It describes the initial assumptions made to come up with a test structure. Fabrication steps are described in complete detail which include photoresist application, patterning by electron-bean lithography, metal deposition by evaporation, and development.
- Chapter 4 explains device characterization. It includes deciding the right material, calculating the right dimensions, calculating the right tunnel barrier resistances, the final layout of the design structure and the criteria for device at low temperatures are all described in detail.
- Chapter 5 provides the measurements and the experimental setup
- Chapter 6 presents the theory relevant for the experiments performed.
- Chapter 7 explanations of the results for the study of the non-equilibrium state by injection of large number of quasiparticles from an external current source. The experimental results and the theory are both analyzed and explained.
- Chapter 8 gives a detailed overview of the cooling mechanisms of solid state devices present in the microelectronic industry.

This thesis can be divided into three parts

- Part A which is the common part included chapters 2 to 5
- Part B which is specific to the Physics/Nanoscience section includes chapters 6 and 7
- Part C which is specific to the Microelectronics section is chapter 8 which is a review on current and existing thermal management techniques for microelectronic integrated circuits and packaging.

Chapter 2

Principles of Thermometry and Refrigeration with Solid State Devices

The discovery of thermoelectric cooling dates back to more than 170 years to Peltier. The last few decades have seen phenomenal progress in developing practical thermoelectric refrigerators for scientific and industrial purposes. This interest has been driven primarily by the semiconductor industry where minimizing the dimensions of the Integrated Circuit at the same time increasing the density has led to exponential increase in power and heat. IC cooling techniques are discussed in detail in the second part of this thesis. In this chapter an attempt has been made to describe the principles and properties of electronic refrigerators for low temperature applications mainly in the sub-kelvin range.

Solid state coolers have much lower efficiency in comparison to the traditional coolers used in the industry. On the other hand, they are far more reliable, cheap and most importantly they can be integrated with the device under consideration and can be scaled down the dimensions of the device itself. This property of solid state coolers makes then very attractive as on-chip coolers with various kinds of micro and nano-devices.

A schematic of a typical setup is as shown in 2.1. The main object is a diffusive metal connected to two large reservoirs on either side where electrons can thermalize quickly. Due to the diffusive nature of the wire, the electrons in the wire experience inelastic collisions. With an applied voltage bias, the energy distribution of of the electrons depends on these collisions, the temperature of the system and on the state of the material whether it is superconducting, semiconducting or a normal metal. This energy can be described by the Fermi function as:

$$f(E) = \frac{1}{exp[(E-\mu)/k_BT]+1} \tag{2.1}$$

where T is the electron temperature, and μ is the potential. As mentioned in section 6.1 "Length Scales", we concentrate on wires that are in the diffusive limit. this implies the Fermi wavelength λ_f and the inelastic scattering length l_e are much less than the length of the wire. In this regime the electron distribution is well defined. The reservoirs or the

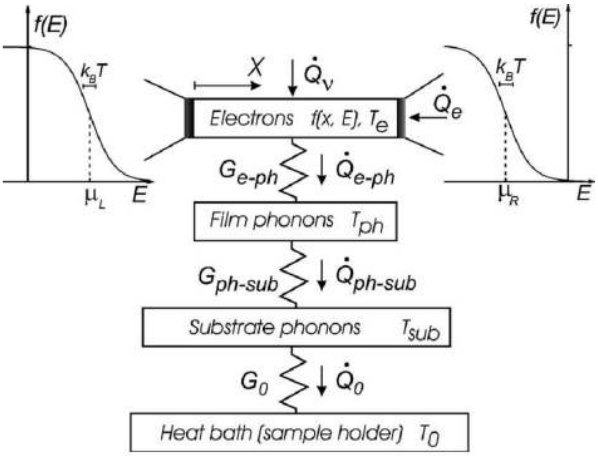


Figure 2.1: Schematic of a typical setup
[13]

contacts are also assumed to be quite large such that the capacitance C of the contacts is large and hence the charging energy $E_c = e^2/2C$ is much smaller and hence can be ignored.

Consider a generic thermoelectric refrigerator as shown in Fig 2.2 In the center is the

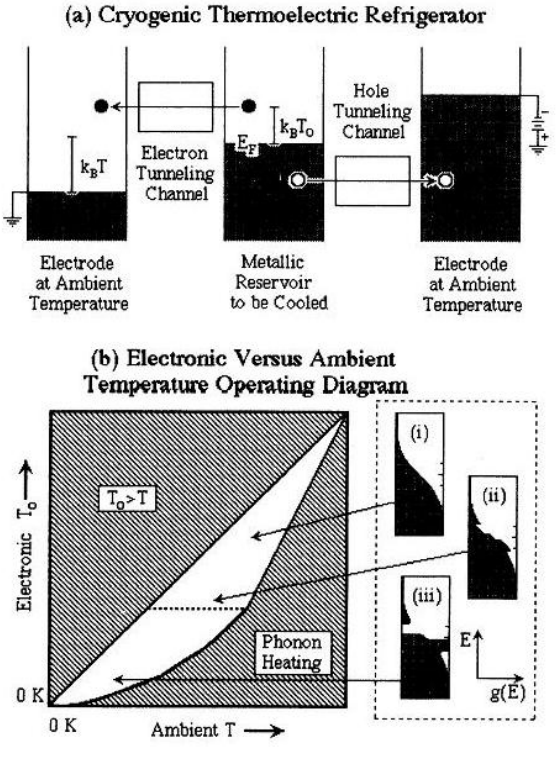


Figure 2.2: (a) Hypothetical cryogenic refrigerator (b) Operating region of the refrigerator
[155]

metallic island which is to be cooled. Electrons and holes tunnel into the island depending on the bias. If these excited electrons and holes are in some way extracted from the metallic island, it is possible to cool the metal below the ambient temperature.

The cooling power of this refrigerator is proportional to ${T_o}^2$ since each pair of excited electron and hole removes energy of T_o. At low temperatures, the electrons absorb heat proportional to $(T^5 - T_o^5)$ where T is the ambient temperature and T_o is the electron temperature. Thus at low temperatures the refrigerator is very adept at cooling the electrons. In a diffusive metal, at low temperatures, electron-electron scattering limits the efficacy of the refrigerator.

Fig 2.2 (b) shows the temperature range where the solid state refrigerator works best. If the ambient temperature is too high, phonon heating causes the refrigerator not to perform. If the electron temperature T_o if higher than the ambient temperature, there is heating and again the refrigerator is unable to cool. Thus in between these two extremities, there lies a region where the operation of the device is able to function as a refrigerator.

Inset shows the electron distribution at particular operating regions and the dotted line describes electrons in thermal equilibrium.

2.1 Superconducting Tunnel Structures

Consider a Superconductor-Insulator-Normal Metal (NIS) tunnel structure wherein a large superconducting metal reservoir connects to a large normal metal reservoir through a tunnel barrier. A voltage bias V is applied over the system and the resulting Energy diagram is shown in Fig. 2.3 The physical mechanism that determines the cooling is such a system is simple: In the presence of the superconductor, quasiparticles with energies

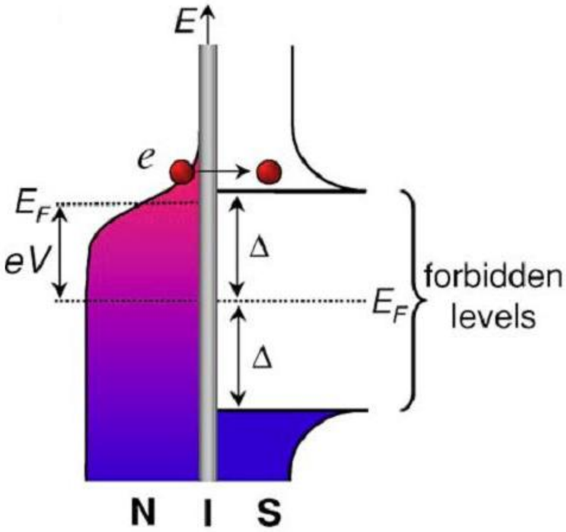


Figure 2.3: Energy Band Diagram of a NIS Tunnel Junction.
[13]

$E < \triangle$ are unable to tunnel into the superconductor due to the presence of the forbidden energy gap. On the other hand, quasiparticles with energies $E > \triangle$ easily traverse through the tunnel barrier into the superconductor. This so called selective tunneling enables the removal of the hot quasiparticles from the normal metal. Hence the NIS junction behaves as a electron cooler.

For the NIS tunnel junctions containing a diffusive wire connected to two reservoirs the quasiparticles charge and heat currents are given by

$$I = \frac{1}{eR_T} \int dE N_L(\hat{E}) N_R(E)[f_L(E) - f_R(E)] \tag{2.2}$$

$$\grave{O} = \frac{1}{e^2 R_T} \int dE \hat{E} N_L(\hat{E}) N_R(E)[f_L(E) - f_R(E)] \tag{2.3}$$

In the above equations, $\hat{E} = E$ - eV, $R_T = 1/G_N$, $N_{L/R}(E) = v(E)/v_F$ is the reduced density of states $N(E) = 1$ for a normal metal and $= N_S(E)$ for a superconducting material. The resulting charge current from the above equation is sensitive to temperature and hence tunnel junctions find applications in thermometry, radiation detection and cooling.

The general principle of thermoelectric cooling is that the energy transfer is related to the quasiparticle current. Under suitable conditions, if the heat is transferred from a cold region to the hot region of the system, the system acts as a refrigerator. Giazotto [13] refers to refrigeration as a process wherein the temperature of the system is lowered in comparison to the ambient while cooling just means removing heat. The efficiency of a cooler is given by its Coefficient of Performance (COP) (η), i.e COP is the ratio of the ratio of the cooling power (Ó) to the total input power(P_{total})

$$\eta = \frac{\acute{O}}{P_{total}} \tag{2.4}$$

The COP or enhancement of cooling power can be achieved by connecting several coolers in parallel.At sub-millikelvin temperatures,superconducting materials can be exploited efficiently. They can be used as peltier coolers because of their low thermal conductivity and zero electrical resistance their energy gap can be used in NIS coolers where the property of energy dependent electron tunneling is exploited. Following, we shall take a look at a few implementations of superconducting electron refrigerators where superconductors and tunnel junctions are used as refrigeration devices.

2.1.1 SINIS Structures

Utilizing simple NIS tunnel junctions Nahum [22] et. al presented a novel electron microrefrigerator that could cool conduction electrons below the lattice temperature. Later, Levio et al [23] presented a series implementation of two NIS junctions forming a SINIS structure as an efficient Peltier cooler. In their experiment a temperature reduction of 200mK is reported.

The energy band diagram of a SINIS tunnel junction is as shown in Fig. 2.4. A SINIS

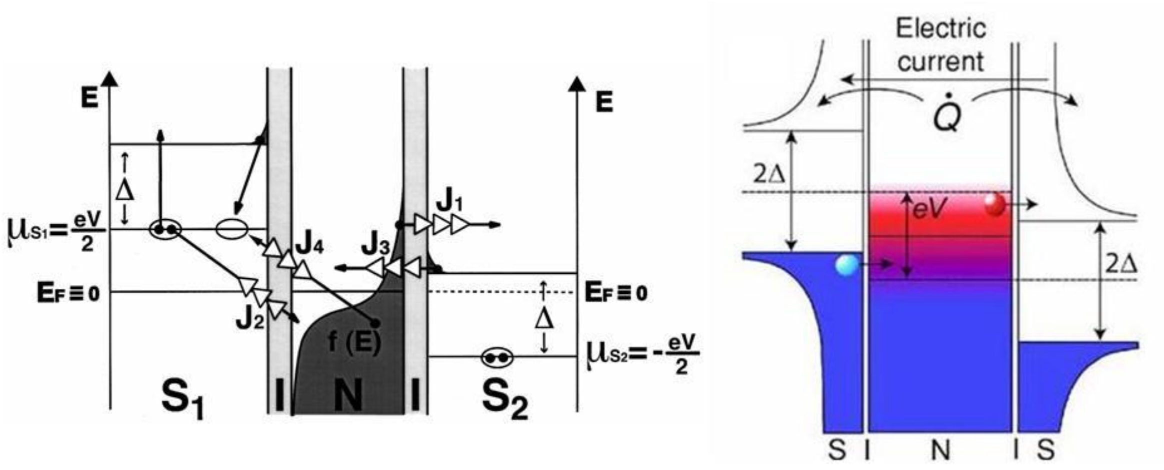


Figure 2.4: Energy Band Diagram of a SINIS Structure with different tunneling processes. [13] [25]

tunnel junction structure is able to function as a refrigerator because the cooling power $\acute{O}$ is a symmetric function of V. Hence with V fixed at a particular value, electrons with energies above Δ are extracted into N region from one tunnel junction and from the other junction holes get into the N region with energies less than -Δ.

- When $V_{ref} = 0$ the refrigerator is off. For V_{ref} biased slightly lower than the gap voltage V_g =$2\Delta/e$ electrons with energy larger than the Fermi Energy F_E tunnel to the superconductor S_2 and there is electron cooling. Both, charge and heat flow in the same direction (J_1).
- On the other superconductor S_1 due to cooper pair breaking, quasiparticles are injected in the normal metal below the Fermi Energy E_F. In this case charge and energy flow in opposite directions (J_2).

Since heat flows from the normal metal to the superconductors, both the processes J_1 and J_2 decrease the average energy of the metal electrons. Two other processes, J_3 and J_4 relate to the population of quasiparticles in the two superconductors. When the thermal energy K_T is less than the energy gap Δ of the superconductor both these processes can be neglected.

In general, in a SINIS refrigerator the final electron temperature is due to several forces that drive power into the system. Since these refrigerators are in the diffusive limit with strong inelastic interactions, the system is driven into a quasiequillibrium state which is described by the Fermi function at a temperature T. This temperature is the electron temperature and can be different from the lattice or bath temperature. In such a scenario, the cooling power can be given by equation 2.5[13] where 2 is for two identical NIS junctions and $T_{e,S} = T_{ph}$.

$$\acute{O}_{refr} = 2\acute{O}(V/2; T_{e,N}, T_{E,S}) - \acute{O}_{e-ph}(T_{e,N}, T_{ph}) \tag{2.5}$$

SINIS refrigerators have found real world implementations. In some implementations Clark and Miller [24] have demonstrated a solid state onchip refrigerator that is able to reach temperatures of 100mK and used it to cool bulk material with no electrical connections to the refrigerating elements. In another implementation by Leoni et al [25] claim to have cooled down electrons in a $2\mu\text{m}^3$ copper strip by using a on chip microcooler which is a SINIS structure.

2.2 Limits of Cryo Cooling

The above mentioned NIS and the symmetric SINIS refrigerators are characterized by their cooling power, efficiency of performance and operating temperature. To be able to define the term ”Operating Temperature” it is necessary that the energy relaxation within the system is much faster than any other process that are able to transfer heat flux between the system and its surrounding environment. If this condition fails then the energy distribution of the particles in the system is non thermal and the concept of operating temperature does not hold good. When cryogenic refrigerators are scaled down to sub-micron sizes, this nonthermal energy distribution starts playing a dominant role at low temperatures. This is because, at low temperatures, the lattice can be considered to be hypothetically non-existent. Energy relaxation happens due to the inelastic electron-electron (e-e) interaction. Depending upon the rate of this interaction one can define the state of the system as being in equilibrium, quasi-equilibrium or non-equilibrium.

In their paper [159] Pekola et al present their findings on the limitations of refrigeration using NIS junctions and suggest that at low temperatures typically less than 100mK, the slow e-e interaction restricts the use of the concept of a temperature and the non zero DOS in the normal metal within the gap of the superconductor gives rise to anomalous heating which ultimately limits the minimum temperature that is reached. Jochum et al [160] modeled the power flow in NIS junctions and present their findings based not only on the operating temperature but also on the size of the junctions. They argue that junctions with larger sizes provide the tunneled electrons with a much longer time to diffuse out. This gives rise to two processes, one, the electrons will tunnel back to the normal metal which is absolutely useless in the cooling process and two, quasiparticles may recombine with the emission of a phonon which is consumed by the normal metal. Efficient extraction of quasiparticles that tunnel into the superconducting material on the other side of the tunnel junction helps greatly in improving the efficiency of these coolers. Placing quasiparticle traps is one way of improving the efficiently [161]

The experiment described in this thesis aims precisely at understanding the quasiparticles energy relaxation processes. Fig 2.5 shows the schematic of the experimental device. Using a SIS tunnel junction, we inject high energy quasiparticles using the large junction at one end. We probe these injected quasiparticles at varying distances from the injection point and study the process of relaxation and recombination. By measuring the $I - V$ characteristics of these probes we can study the non equilibrium properties. We can conclude whether it is the energy relaxation that is playing a dominant role or it is the recombination of quasiparticles that is having a major effect.

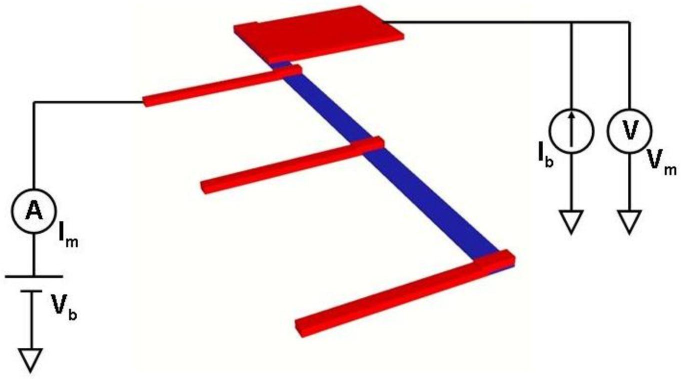


Figure 2.5: Schematic showing the SIS junction with the large junction used to inject quasiparticles and the probes at various distances used to measure non-equilibrium

Chapter 3

Device Fabrication

In this chapter we take a look at procedure involved in fabricating tunnel junctions. Device fabrication is a tedious task with a number of fabrication issues to be taken into account. Important steps in this regard are

1. Designing with a proper recipe for the photoresist
2. Determining the correct exposure for E-beam lithography
3. Development of the exposed photoresist
4. Deposition of the metal, insitu oxidation, and shadow evaporation to obtain tunnel barriers
5. Liftoff of the deposited metal to obtain the final device.

In order to try out the fabrication process, it was necessary to have a device layout such that it would not only help us in determining the correct requirements for fabrication but it would also help us to come up with the dimensions required to make the actual device used in the experiment. The following design was chosen as a test layout. The above test Fig. 3.1 structure is designed in *Autocad*®.The wire width is 400nm. The distance between the probes and the wire is 300 nm. The width of the probes was varied from 4 μm to 200nm. As seen, the most important parts of this layout are the ones in red which consist of the wire, the probes which are seen to vary in their width. The orange/yellow part seen is the area where we need the undercut[1]. The advantages this test layout offers us are as follows:

- Helps in designing with the proper recipe for the double layer photoresist
- Helps in determining the correct E-Beam dose required for proper exposure as well as obtaining the undercut.
- With the probes varying in their width, they automatically have different resistances after the oxidation which helps us in determining the correct oxidation pressure and time.

We shall now take a look at each of the above mentioned fabrication steps in detail

---

[1] Undercut is the area defined where the bottom layer of the double layered photo resist is developed while maintaining the top layer, creating a sort of a bridge under which the probes will overlap the wire after the second evaporation

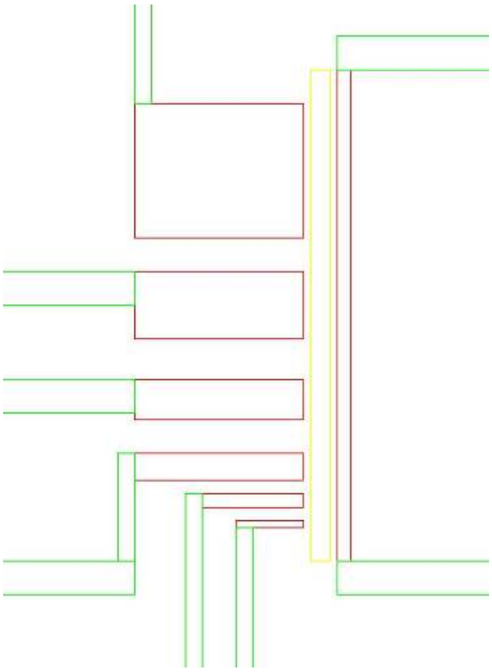


Figure 3.1: Test Structure schematic. Green → connections to pads. Red → varying widths of probes and the wire over which the tunnel barrier is to be made. Yellow → bridge for the undercut.

3.1 Photoresist Recipe

The substrate used is Silicon, with approximately 250nm of thermally grown oxide. As a reference the initial recipe used was that used earlier [6] which is a bilayer of a combination of PMMA-MMA[2] which provided a good undercut. The recipe yielded a total resist thickness of approximately 1μm. A total undercut of 300nm (distance between the wire and probes) was required in order that the probes move a distance of 700 nm to overlap the wire completely and there is no overlap of the wire with itself.

Although the above recipe has been used quite extensively to obtain good undercuts, it did not provide the required 300nm of undercut as seen in Fig. 3.2[3].

In order to overcome this problem, a ghost pattern was utilized which was exposed with a slightly lower e-beam strength. The idea behind this was the fact that the bottom layer of the bi-layer resist is very sensitive and hence with a low strength of the e-beam over this area *(yellow area of Fig. 3.1)*, the bottom layer is exposed while the top layer stays thus providing the undercut. Unfortunately this technique did not work too. The failure of this PMMA-MMA bilayer resist can be due to the following:

- The resists used were quite old (5 years since the expiry date on the storage)
- The area between the wires and the probes was exposed with a small beam current. While in the case of [6], the probes were in close proximity of the pattern which

---

[2] Poly Methyl Methacrylate - Methyl Methacrylate

[3] If there is an undercut, the wires and probes should overlap which is not seen

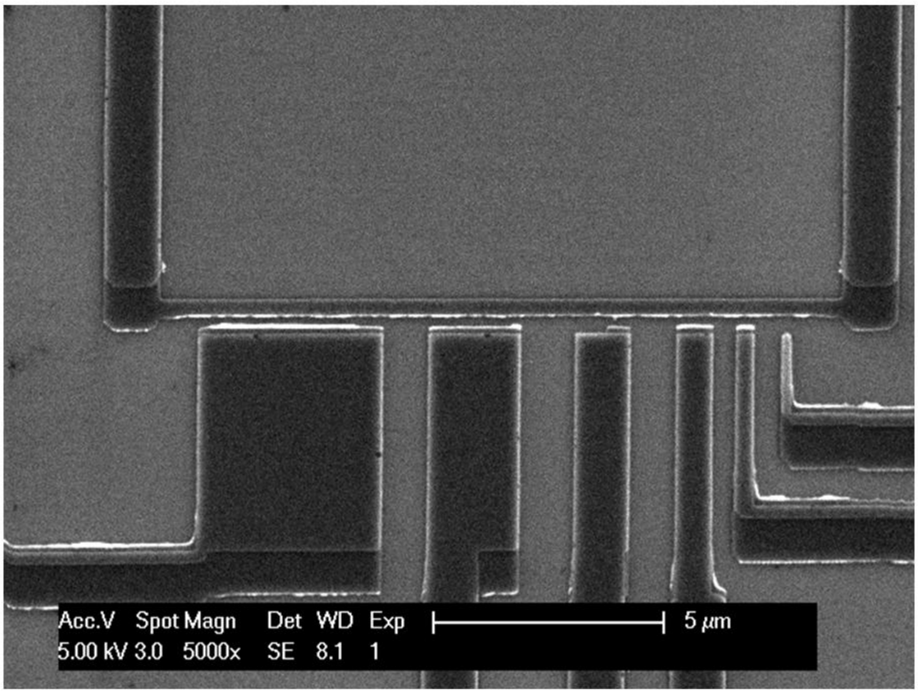


Figure 3.2: SEM image of the test structure developed using the PMMA-MMA bilayer of photoresist showing not enough undercut (no overlap of the probes and wires)

was exposed with a large beam current thus raising the question of proximity effect, causing the large beam to give an undercut under the small probes.

- Another reason for the absence of the undercut might be the energy of the E-beam used. A 100KeV E-beam was used to expose the design. This high energy beam also gives a high directionality to the electrons. As seen in Fig 3.3 the proximity affect is higher for the electron beam with 20KeV energy than it is for the 50KeV beam. It will be even less for a e-beam with a 100KeV as used in our experiments.

Due to time limitations, the exact cause for the behavior was not investigated. In due course, while following this cause [7], it was learnt that with small beam currents, the maximum undercut seen was of the order of a 100 nm. In order to obtain large undercuts in areas that are exposed with a small beam strength, deep UV flood exposure could be utilized for the bottom layer of the bilayer resist recipe.

The recipe of PMMA-MMA bilayer was changed and a different bilayer recipe was used with the bottom layer being PGMI in 7% SF and the top layer of PMMA 950K in 4% Anisol. The resist coating steps are outlined below:

1. On a dry clean wafer, spin coat a layer of PGMI-SF7 at 2250 RPM[4] and bake at 200°C for 5 mins.
2. Repeat the above step to obtain a double layer of PGMI-SF7.

---

[4] 5 seconds at 500 RPM[5] and 55 secs at 2250 RPM

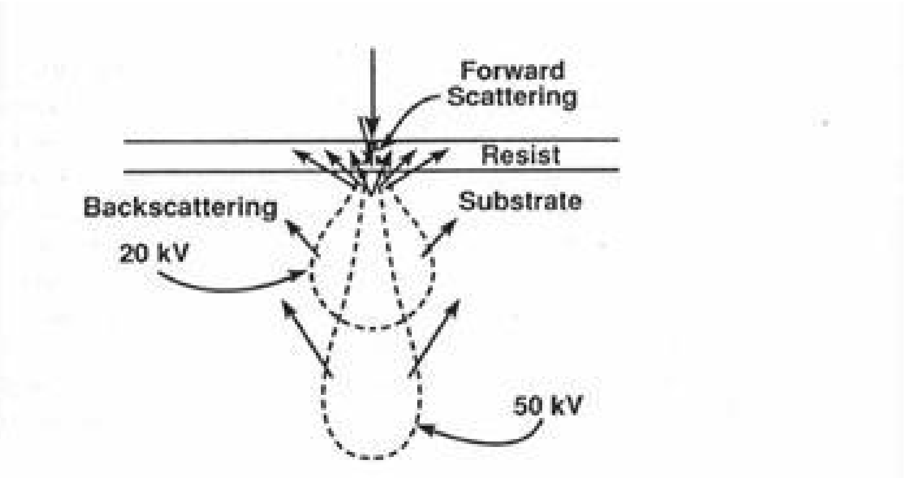


Figure 3.3: Proximity effect realized by e-beam strengths of 20KeV and 100KeV

3. Spin coat a layer of PMMA 950K in 4% Anisol at 6000 RPM[6] and bake at 175°C for 15 mins.

A single layer of PGMI-SF7 spun at 2250 RPM gives a thickness of around 420nm to 450nm and PMMA spun at 6000 RPM gives a thickness of about 120nm to 150nm. In order not to change the design, a double layer of the PGMI resist was used with PMMA on top to obtain the required 1μm thickness of the resist.

With this receipe the required undercut was obtained as seen in Fig. 3.4 which is clear from Fig 3.2 where it is clear that the probes and the wire do not overlap.

3.2 Lithography and Development

Once the resist spinning step is completed, the next step is to pattern the design on the wafer using Photo Lithography. The design created in *Autocad*® is saved as a .dxf file. This dxf file is converted into a format recognized by the E-Beam system using a software known as Layout Beamer. Each color in the *Autocad*® .dxf file is exported as a separate layer by the Layout Beamer with specific beam step sizes and acceleration voltage and is stored as a .gpf file. The gpf file is included in the script file which the E-Beam system executes to write the pattern on the wafer.

The executable script is the .layout file. This file contains all the details needed for the E-Beam system to be able to write the pattern on the wafer[7]. The most important parts in the .layout file are

- Name and number of the holder on which the wafer is mounted
- Coordinates of the wafer
- Size of the pattern

---

[6] 5 seconds at 500 RPM and 55 secs at 6000 RPM

[7] Refer to Appendix B for the structure and syntax of the .layout file used

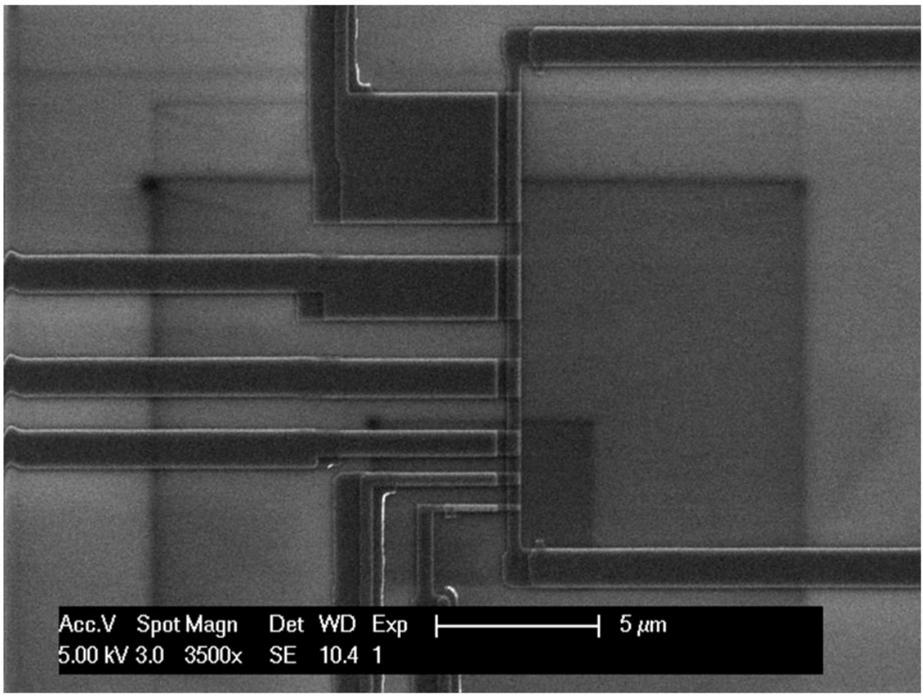


Figure 3.4: SEM image of the test structure developed using PGMI-PMMA recipe

- Total number of patterns
- E-Beam details
- Dose used

The flowchart of 3.5 summarizes the steps involved starting form design to finally obtaining an exposed sample from the E-Beam system.

The design is exported as two seperate entities to be patterened using the Eletron Beam Pattern Generator (*EBPG*). The probes and the wire are exported as one entity and were exposed with a smaller beam current and the rest of the connections are exported as another separate entity and exposed with a higher beam current. Calculating the correct beam strength and the dosage is important for proper exposure of the resist. This is explained below:

- The **Dose** is chosen by exposing the pattern to a variety of doses for a particular beam strength. The units are $\mu C cm^{-2}$. For our case a dose of 1200 was used which results in $6.5 * 10^{21}$ electrons per cm^2.
- The acceleration voltage is 100KeV.
- The Beam Step Size (BSS) is specified when exporting the files for the EBPG. For the probes and the wire, the BSS is set to 50nm and for the other connections is set to 100nm.

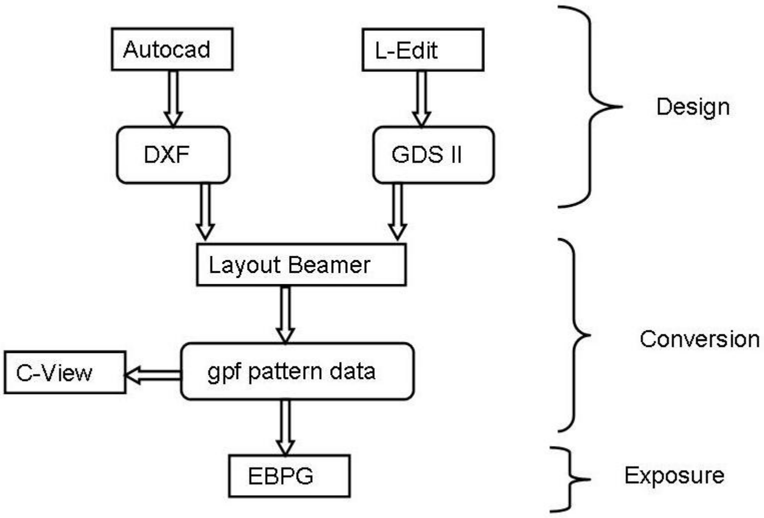


Figure 3.5: Flowchart for EBPG flow.

- Based on the BSS, beams are selected from the list wherein 1.5 - 1.2 times the spot size is approximately equal to the BSS. From the list of beams with their estimated spot size, the match is made by defocussing the beam. The relation is

  $$d^2_{calc} = d^2_{est} + d^2_{defocus} \tag{3.1}$$

  Defocussing is in discrete steps of 20nm/bit. In our case the defocussing was set to 2. With this the spot size comes out as 45nm for the smaller beam and 178nm for the larger structures.

- The above selected beams have different currents associated with them. The strength determines the Beam Step Frequency (BSF) which also depends upon the dosage and is given by

  $$\begin{align} \mathit{BSF\ (MHz)} &= \frac{1}{T(\mu s)} \tag{3.2} \\ &= 0.1 * \frac{I_{spot}(nA)}{Q(\mu C/cm^2) * BSS(\mu m)^2} \tag{3.3} \end{align}$$

  This BSF is required to be from 2 to 10 MHz since this determines how long the beam is active at a particular area.

- From the above calculations, the final beams chosen for exposing the small and the large structures are respectively:

  1. $28_{nm}\_26_{nA}\_400_{\mu M}\_100KeV$
  2. $175_{nm}\_317_{nA}\_400_{\mu M}\_100KeV$

  The measured spot size for the beams are 20nm and 174nm respectively.

After the design pattern is exposed by the E-Beam, wet etching is used to develop the exposed resist. PMMA is sensitive to E-BEAM exposure and is a positive resist[8]. The development of the exposed sample is explained below:

- Exposed PMMA is developed by etching it in a solution of MIBK[9] diluted with IPA in the ratio 1:3 respectively. The sample is dipped in the solution and stirred for 60 seconds to completely etch out the top layer PMMA. The sample is then dipped in IPA[10] for another 60 seconds to completely take off the MIBK solution and carefully dried with nitrogen blow drier.
- Unlike PMMA and other photoresists, PGMI is not affected by the e-beam exposure. PGMI is also not dissolved in the diluted solution of MIBK. Hence in order to dissolve PGMI, the sample after treatment of MIBK and IPA is dipped in a solution of MF321 and stirred for 15 seconds[11]. This stirring of 15 seconds is just enough to obtain a undercut of 300nm.
- The sample is then immediately dipped in de-ionized water and stirred for about 60 seconds again to take off the effects of MF321[12].
- As a final step, the sample is dipped in clean IPA again and dried carefully and inspected under optical microscope.

3.3 Metal Deposition and Oxidation

Once the mask is developed and inspected, the sample is carefully loaded into the evaporator to deposit the metal, oxidize the first layer and deposit the second layer on top at an angle in order to obtain tunnel barriers. For simplicity a single angle deposition is done for the second layer which makes the tunnel barriers.

The metal deposition in done in an HV[13] chamber with pressures of low 10^{-7} mbar and at a rate of $1.2\AA$ to $1.5\AA$ per sec. The second step of deposition after oxidation takes place at an angle. The angle of the second deposition is determined based on the following two factors as shown in Fig. 3.8

1. The thickness of the resist $\boldsymbol{h}$
2. The distance the second deposition has to move in order to overlap the first layer $\boldsymbol{d}$

The angle required is then calculated by using the relation

$$\boldsymbol{d} = \boldsymbol{h} \tan(\theta) \tag{3.4}$$

---

[8] A positive resist is one where the portion of the photoresist that is exposed is soluble to the photoresist developer. The portion of the photoresist that is unexposed remains insoluble to the photoresist developer

[9] Methyl Isobutyl Ketone

[10] 2-Propanol

[11] Care should be exercised here because PGMI dissolves rapidly in MF321 giving large undercuts and causing potentially useless structures

[12] IPA can be used in place of deionized water, but the results are not good.

[13] High Vacuum

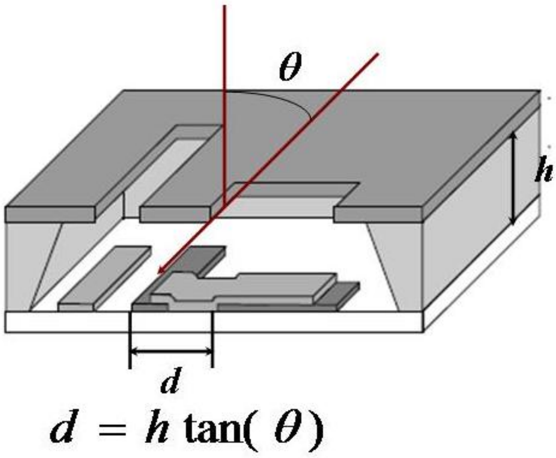


Figure 3.6: Principle of shadow evaporation.

In our case, the thickness of the resist ***h*** is about 1μm and the distance ***d*** is 700nm and the resulting θ is 35.

The complete process of evporation and deposition of metal is described below

- After the sample is loaded into the loadlock of evaporator, it is isolated from the evaporation chamber. The e-beam is switched on and current increased gradually until a steady rate of evaporation is seen typically 1.2nm to 1.5nm per second as mentioned earlier.
- The sample is then brought into the evaporation chamber and metal is deposited on it. The first evaporation step deposits 40nm of metal.
- The sample is then pulled out of the evaporation chamber into the loadlock and isolated. Oxygen is then pumped into the loadlock at the rate of 50μbar for a set duration of time. Pressures of the loadlock, evaporation chamber are monitored.
- Once the oxidation is done, the sample is then put back into the evaporation chamber. This time the second deposition takes place at an angle determined earlier. The second deposition is made for a thickness of 80nm of the metal.
- The sample is now ready to be taken out of the evaporator for further processing.

Oxidation deserves some detailed explanation. With the different widths of the probes, the overlap of them with the wire gives different overlap areas. By varying the pressure of oxygen and the oxidation time, the oxide thickness is varied which in turn influences the resistance of the tunnel junction area. For a given oxidation pressure and time, the resistance increases linearly with the area of the tunnel junction. A larger the area of the junction, lower is the resistance. By experimenting with different pressure and oxidation times, for our tunnel junctions, oxidation time of 45 mins at a pressure of 50μbar yielded a RnA[14] value of 2.4K$\Omega\mu^2$

---

[14] Resistance of the junction multiplied by its area

3.4 Liftoff

The final stage in the fabrication process is the Liftoff. The deposited metal has to be taken off from the unwanted places, namely from the areas where the resist is not exposed and developed. Liftoff in this process where the underlying resist is PGMI is done by placing the sample in a solution of PRS 3000 at 60°C. If time is not a constraint, it is best advised to leave it for atleast 6 to 8 hours without the ultrasonic. Also a slight ultrasonic may be put on and the liftoff then takes about 4 hours. It has been observed that the structures are developed better without the ultrasonic although it takes some extra time.

The sample now is complete of all the clean room fabrication steps. It is inspected under SEM[15] to determine those structures which are not damaged. Then the sample is diced and the individual structures are glued on ceramic holders and bonded for measurements.

The final structure after the liftoff is as shown below:

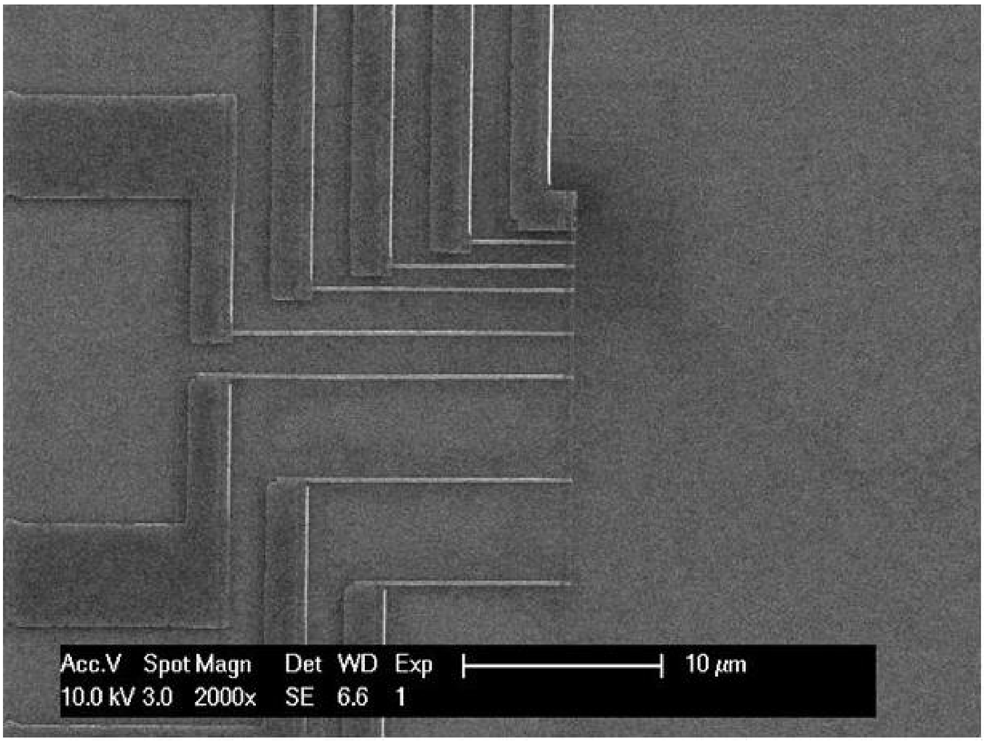


Figure 3.7: SEM image of the final device.

---

[15] Scanning Electron Microscope

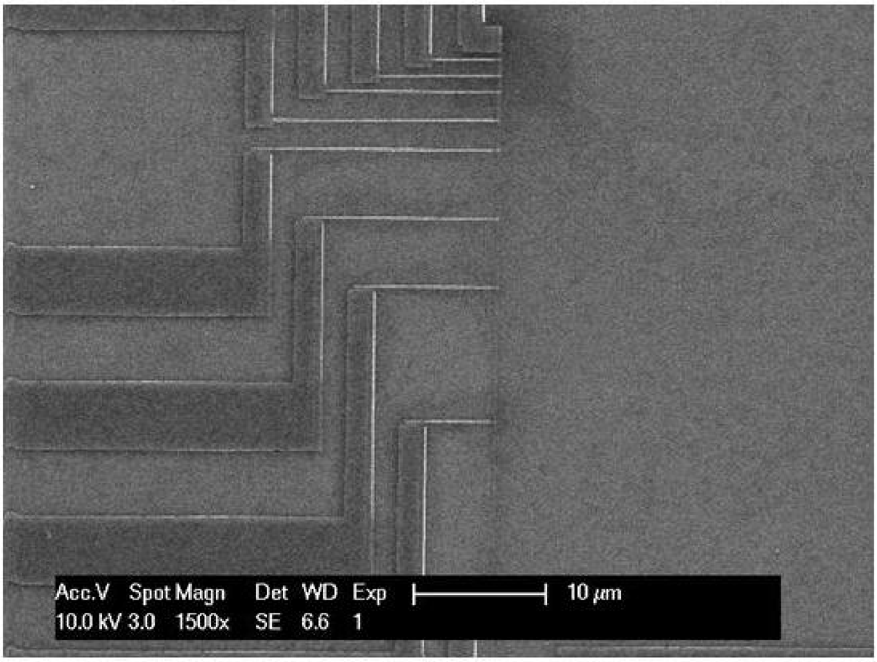


Figure 3.8: SEM image of the final device

Chapter 4

Device Characterization

The initial sample or the structure is fabricated with certain assumptions. The probes were of different widths in order to fine tune them or select the width that best fits the theoretical model. Oxidation as mentioned in the previous chapter is also an important criteria in characterizing the device. As we seek to probe non equilibrium properties, there has to be a source that induces this non equilibrium which is a clean contact to the wire, and through the tunnel junctions this non-equilibrium is extracted and studied. Hence the fabricated sample structure has to pass a number of criteria to be a suitable candidate for final measurements.

In this chapter the criteria for characterizing a device which is deemed suitable for measurements are presented.First the material properties that determine the choice of a particular material are presented, next in line are the characteristics a fabricated device has to satisfy at room temperature are verified and then finally the behavior a good device should present when cooled down to low temperature are presented.

4.1 Criteria measured at Room temperature

4.1.1 Material Selection

Aluminium was the material of choice for the superconducting metal. Several reasons go in favor of this choice. Aluminium oxide (AlO_x) is arguably the most wide spread insulator used to fabricate metallic tunnel junctions because it can be easily and reliably grown. Other factors that favor aluminium as a material of choice are:

1. Aluminium is a type I superconductor.
2. The electron electron interaction length l_{ee} is about 1μm (mK temperatures - order of magnitude) which helps us in probing the nonequilibrium properties of the tunnel junctions.

4.1.2 Tunnel Barrier Resistance

In situ vacuum oxidation of aluminium and deposition of the second layer gives a sandwich of metal-insulator-metal junction. The insulator is the natural oxide of aluminium. As mentioned earlier, the pressure and duration determine this oxidation and the resulting barrier resistance. For a fixed temperature and pressure, the resistance of the junction scales linearly with the area. For the experiments performed, it was desired to obtain an RnA value of more than 1K$\Omega\mu^2$. A thick oxide barrier would definitely give higher resistance and a thinner would give a lower resistance. The problem with a thin barrier is that there might be pinholes which would make the tunnel junction useless.

For the tunnel barriers used in the experiments mentioned in this thesis, the oxidation was carried out for 45 mins at pressure of 50μbar. The resulting resistances obtained for the injector and the detector junctions are 1.5 KΩs and 15 KΩs. The RnA value for the injector junction is 1.2K$\Omega\mu^2$ and for the detector junction is 2.4K$\Omega\mu^2$

4.2 Criteria measured at Low temperature

The criteria that determine whether a tunnel junction is suitable for measurements or not at low temperatures were first put forth by J.M Rowell [11] and are known as the Rowell Criteria

- After cooling the junctions to nitrogen temperature[1], the first test is to measure the resistance. If the resistance of the junction is not inversely proportional to the area ±20%, the barrier cannot be trusted.
- On cooling further to helium temperatures, the junction resistance should show a slight increase. Any decrease in resistance qualifies the junction to be discarded from further considerations.
  Resistance of the junctions when cooling to helium temperatures was measured and it indeed showed increase in comparison to the resistance measured at room temperature.For the three tunnel probes the resistance at room temperature was 11.2KΩ 12.5KΩ and 9KΩ. At helium temperature, these resistances increased to 12KΩ, 13KΩ and 9.5KΩ respectively.
- Observation of the superconducting gap. If the experiment does not show the superconducting gap, none of the effects of the tunnel junctions mentioned above are to be believed.
  The superconducting gap was observed on further cooling below the theoretical value of T_c of $1.2K$
- If the observed conductance at zero bias is not less than 10^{-1} of that measured at normal state, the observed superconducting gap is not right.
- For junctions which have superconducting metals on either side, the negative resistance region should be observable. The cusp at $\Delta_1 - \Delta_2$ should be well defined and the current onset at $\Delta_1 + \Delta_2$ should be sharp.

Once the device under consideration passes the above criteria, the first measurements to be done are for T_c. Thin film T_c can be determined resistively or by finding the temperature where the junction conductance at zero bias first begins to decrease. This

---

[1] 77K

method gives the uniformity of the film and its transition.

The next measurement is that of the gap. Since no junction is ideal, and with an infinite cusp at $\Delta_1 - \Delta_2$ and a discontinuous rise at $\Delta_1 + \Delta_2$, a smeared characteristic makes it difficult to pick the right point. Rowell [11] has suggested a way to pick this point which is explained as below:

- Find the voltage where the current is half the value extrapolated from voltages greater than 2Δ.
- The size of the current jump should be $\pi/4$ times the normal current at that voltage.

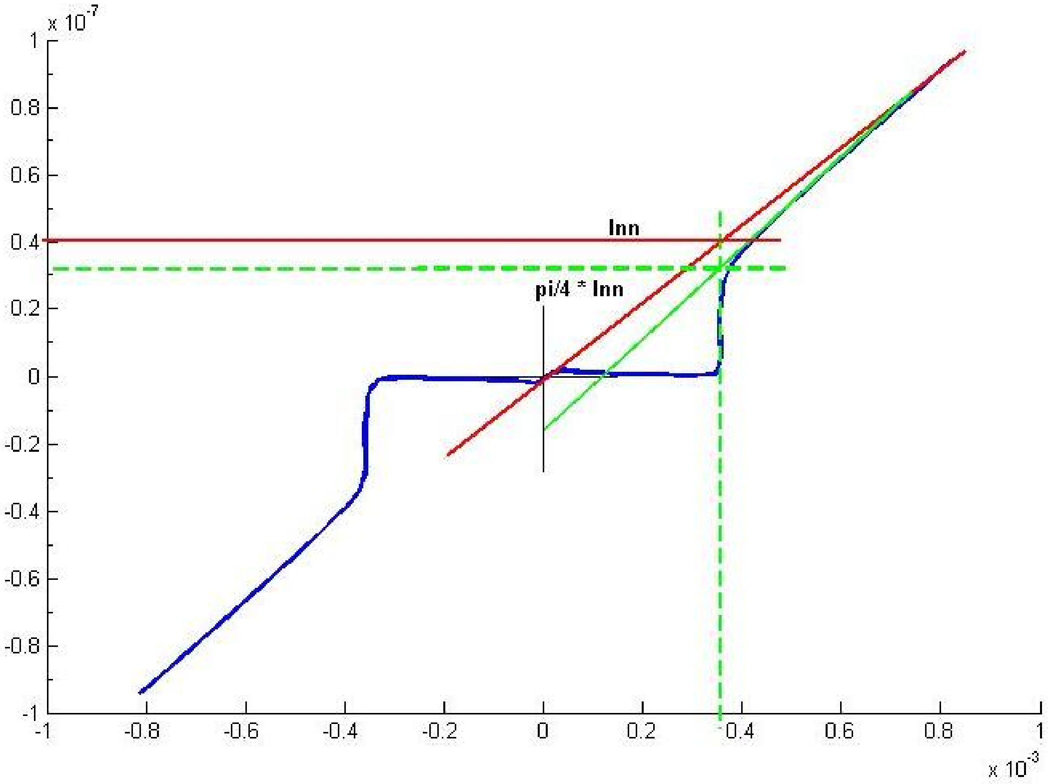


Figure 4.1: Schematic to calculate the energy gap using the Rowell Criteria.

The superconducting gap value was calculated using the procedure underlined above. Fig. 4.1 shows the calculation of the energy gap Δ using the Rowell criteria mentioned above. For the device measured, I_{nn} is 0.443, so π by four gives 0.3443 which is the value for 2Δ

The criteria mentioned above pertain to the performance or behavior of the device. There ia another set of criteria which can be termed as design criteria which mainly consists of the dimensions, in particular the length scales. These are presented below:

- The superconducting wire is a quasi 1 dimensional diffusive system.

- A clean contact on one side which would serve as the ground potential and a tunnel contact of much larger width than the probes placed on the other end which is used to drive the wire out of equilibrium

- The probes are placed on the wire for extracting and probing this non equilibrium.The position of these probes is depends on the calculated length scales. Although not all of them are important, the probes closest to the large tunnel contact that drives the wire out of equilibrium are important. The distance of the first probe is calculated based on the diffusion constant D which is determined from the Einstein Relation
$$D = \frac{1}{N(0)e^2\rho_o} \tag{4.1}$$
where $N(0) = 2.2 \times 10^{47} J^{-1}m^{-3}$ [9] is the aluminium density of states at the Fermi Level ρ_o is the impurity resistivity given by
$$\rho_o = \frac{\rho_{ph}}{RRR - 1} \tag{4.2}$$
and the phonon resistivity ρ_{ph} of aluminium is taken to be $2.7\mu\Omega$cm [12]

- Using the values obtained for the diffusion constant D and the inelastic collision length $l_e = 3.65 \times 10^7 \Omega^{-1}m^{-1}$ [4] we can calculate the diffusion time τ_e.

4.3 Measurements of Device Characteristics

The DC parameters for Aluminium are measured as below:

- The resistance of the Hall bar on cooling to helium temperature was measured to be 400Ω. Assuming the resistance of the wires did not vary much, the actual resistance if the hall bar is 20Ω. This results in a RRR Value of 4 [2].

- Using the values for RRR, and phonon resistivity ρ_{ph} from equation 4.2 and $N(0) = 2.2 \times 10^{47} J^{-1}m^{-3}$ [9], we calculate the Diffusion Constant D using equation 7.3 which comes out to 209 cm^2s^{-1}.

- The Coherence Length (ξ) can be calculated using the relation
$$\xi = \sqrt{\frac{\hbar D}{2\pi k_b T_c}} \tag{4.3}$$
The coherence length is 184nm.

4.4 Device Layout

From the above discussion, the layout of the design suitable for measurements in the experiments is come up as shown in Fig 4.2 and Fig 4.3.

---

[2] Although the variation of the wire resistance was not taken into account, the RRR value fits from previous experiments and trials.

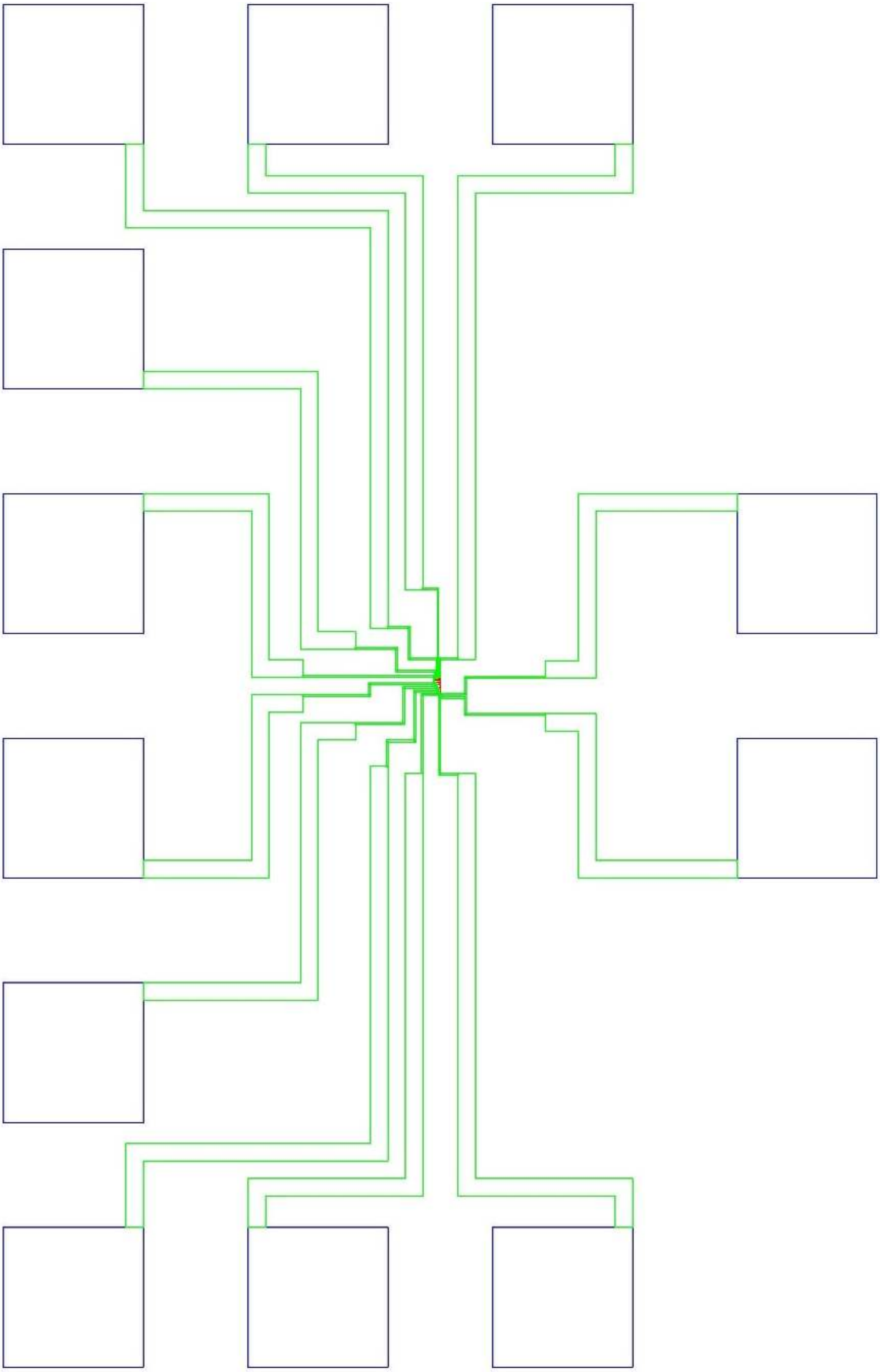


Figure 4.2: Final design showing the complete structure.

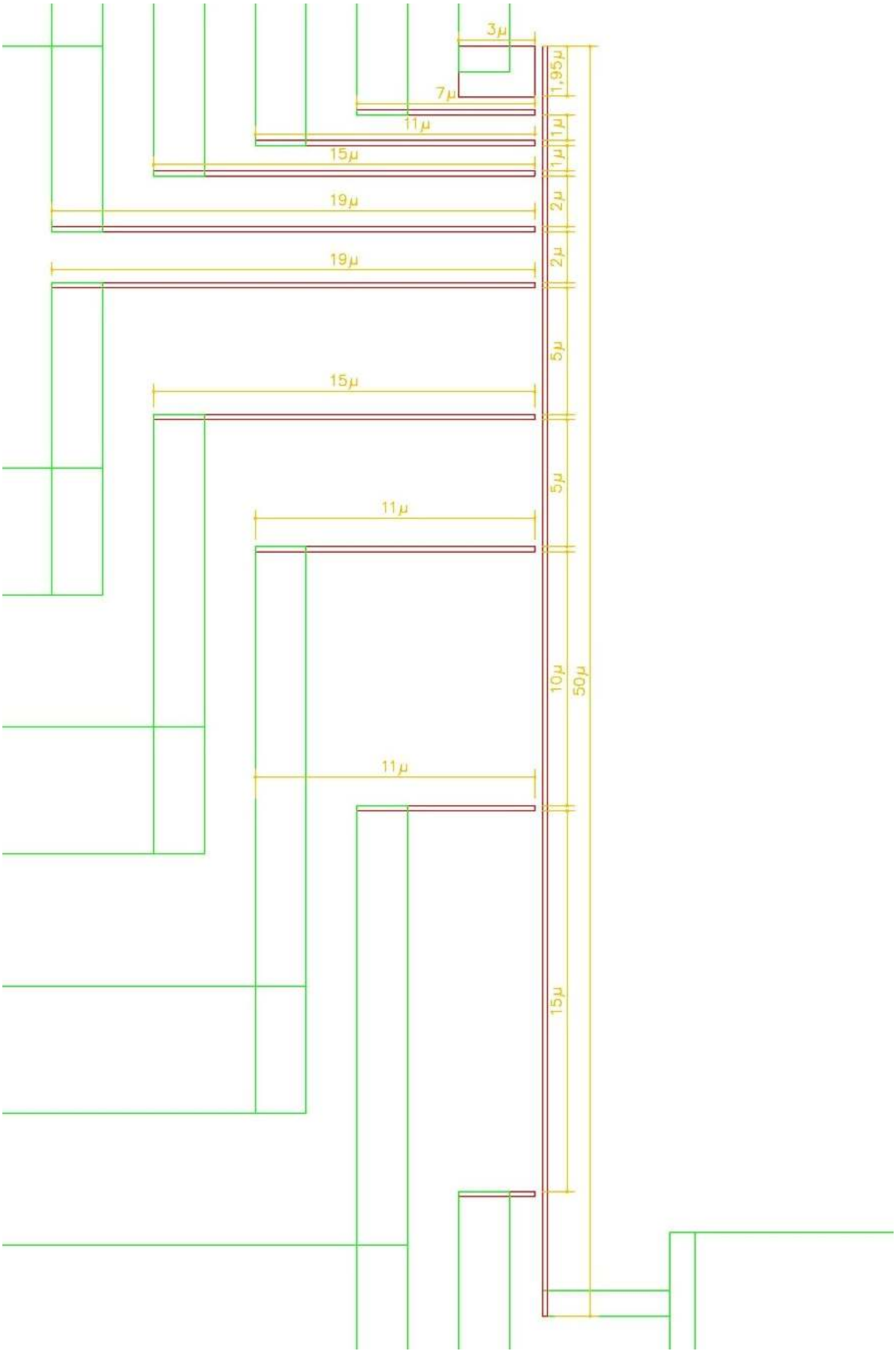


Figure 4.3: Final design showing details of the probes with measurements.

Chapter 5

Measurements

In this chapter we will take a look at the measurement setup. Since our measurements happen to be all two point, as a first step, we shall take a look at the difference between a 2 point and 4 point resistance measurement methods. Next presented is the measurement setup in which consists among other things, The IVVI Rack, Dilution Fridge that are explained. Lastly we take a brief look at the IV characteristics of the injector junction and the detector junction and the non local behavior.

5.1 Resistance Measurements

Resistance measurement are usually made with just 2 point measurements. The disadvantage of a two point measurement is that the lead resistance gets added to the total resistance measured. Hence when accurate values of resistance are needed a four point resistance is done. In principle using the four point measurement technique one can eliminate the lead resistance. In this section we shall take a look at $TwoPoint$ and $FourPoint$ measurement techniques. Fig 5.1 shows these two scenarios where a voltmeter and a current source are connected. The two current leads comprise a two-wire current source that circulates current through the resistance under test ($Green$). Two potential leads provide a two-wire voltage measurement circuit that measures the voltage drop across the resistance under test (Red).
The resistance with the four point method can be calculated as follows:

$$V_1 = R_1 * I_b \tag{5.1}$$
$$V_2 = R_2 * I_b \tag{5.2}$$
$$V_3 = R_3 * I_m \tag{5.3}$$
$$V_4 = R_4 * I_m \tag{5.4}$$
$$V_t = R_t * I_b + R_t * I_m \tag{5.5}$$

Hence we have

$$V_m = R_4 * I_m + R_3 * I_m + R_t * I_b + R_t * I_m \tag{5.6}$$
$$= (R_4 + R_3 + R_t) * I_m + R_t * I_b \tag{5.7}$$

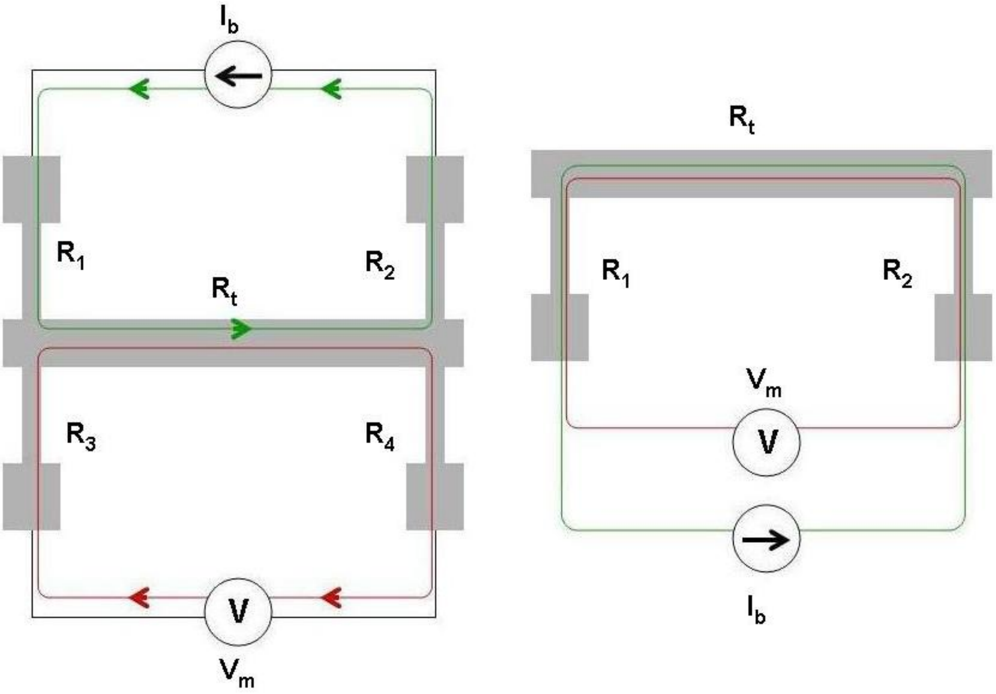


Figure 5.1: Two point and Four point resistance measurement techniques

I_m is negligible because the voltmeter does not put in current.

$$V_m = R_t * I_b \tag{5.8}$$

Thus using the four point measurement method, the lead resistance effectively cancels out.

Using similar argument presented above it is easy to see that using the two point method includes the lead resistance.

5.2 Measurement Setup

The measurement setup can be divided into two

1. The IVVI Rack which consists of the QT-Delft Box which contains the current and voltage sources and current and voltmeters used in the experiments.
2. The "Picowatt Resistance Bridge which is used to monitor the temperature of the dilution fridge.
3. A Microsoft Windows based PC
4. A DAC Card to interface the PC
5. A Lockin Amplifier which can be used as an IV source/meter
6. The Dilution fridge

Fig 5.2 shown the voltage and current modules from the ”Delft QT Box which are utilized in our experiment. In Fig 5.2 the source $S4c$ is used as the current source to bias the injector junction and $S3b$ is used to bias the detector junction. $M2m$ is the voltmeter across the injector junction and $M1b$ is the current meter at the detector junction.

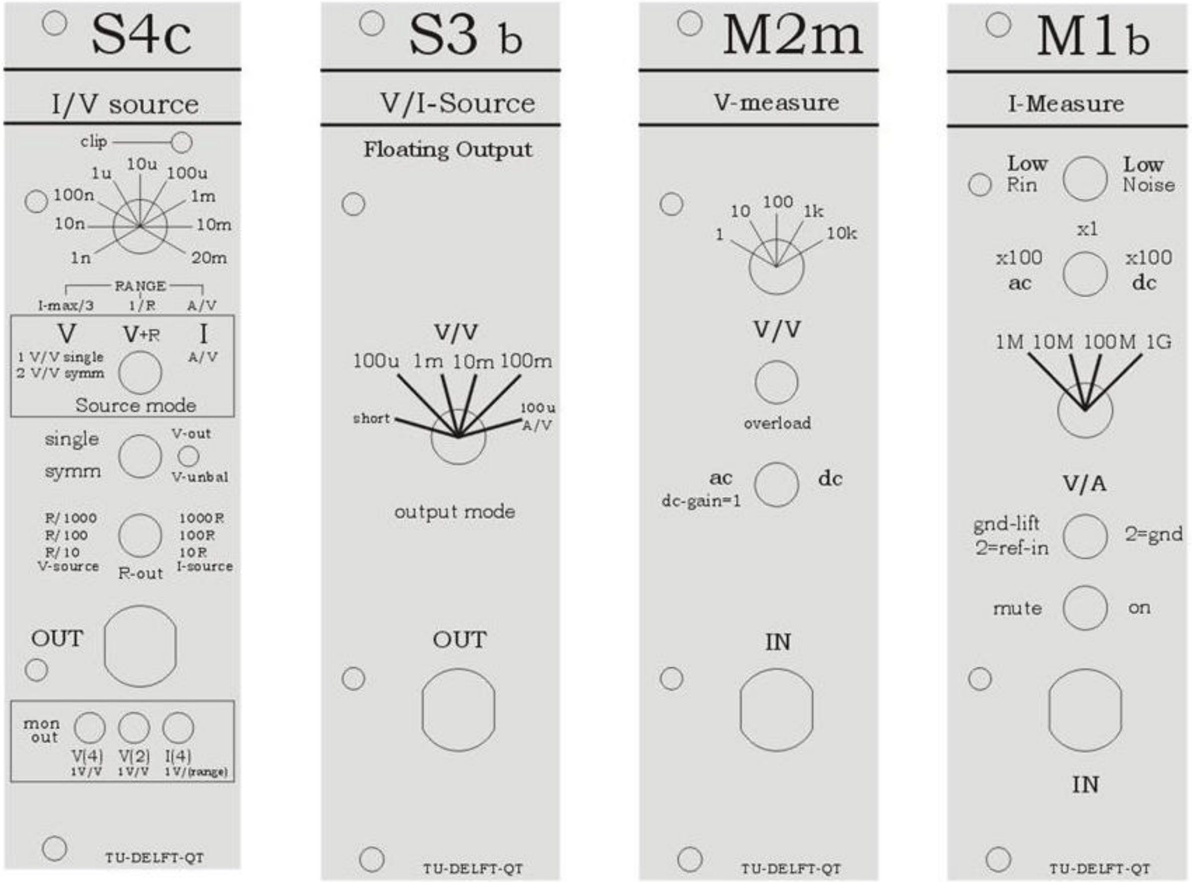


Figure 5.2: Schematic of the voltage and current sources and meters used in the measurements.

5.2.1 Current/Voltage Module

The first block of Fig. 5.2, $S4c$ is a versatile source for V-I or I-V measurements. In our measurements we utilize it as the current source across the injector junction. The *Sourcemode* switch sets the module operating as:

- V a voltage source with the RANGE switch setting the current limit to approx 3x(1nA..20mA)
- $V + R$ a voltage source with RANGE setting the output resistance (1Gohm..50ohm)
- I a current source with RANGE setting the current range (1nA..20m A/V)

The $R - out$ switch makes it possible to set the output resistance of the selected V or I source (R is set by 1/RANGE). The R/1000 1000R, being the usual setting, gives the highest accuracy. The R/10 10R gives the lowest noise. Highest accuracy setting were used in the experimental setup. The output resistance of the current source with the settings result in 1GΩ's.

5.2.2 Voltage/Current Source

The second module in Fig 5.2 *S3b* is designed to be a very low noise bias source for VI/IV measurements and is utilized as a Voltage source across the detector junction. The output of the source is fully isolated i.e floating source, energy is send via an optical path. The switch *outputmode* sets the full scale (@1V control) voltage/current. The short setting is there to check for offset/noise.

The voltage source has an output resistance (<1% error) that is 1kΩ@range-100mV 100Ω@10mV 10Ω@1mV 1Ω@100uV. For our measurements, the 1mV range was used which would give us an output resistance of 10Ω.

5.2.3 Voltage Measurement

The third module in Fig 5.2 the general voltage amplifier which is utilized as the Voltmeter across the injector junction. It is a differential voltage amplifier with the input stage of the amplifier being formed by cmos FET-opamps.For our experimental setup, 1V and DC settings were used.

5.2.4 Current Measurement

The last module in Fig 5.2 is the current meter used at the detector junction.The output resistance of this meter is not fixed but it depends on the user settings. For our experimental setup *LowRin* was selected and the V/A was set to 1M. The postgain switch is set to $*100dc$ giving a max total output of 100G V/A.

The ground-lift switch $gnd - lift$ solves problems with thermovoltages that are usually present in cryogenic setups. We set the ground lift to "gnd" which implies the ground pin is connected to the I-V rack.

The input resistance is a result of the above settings. It is 2kΩ + 1E-3 * [V/A] for $low - Noise$ and 2 kΩ +1E-4 * [V/A] for [$low - Rin$]. For our measurements, we used the *LowRin* and $1MV/A$ which gives us an internal resistance of 2100Ω's.

With the above settings for the I-V measurements, the device is mounted on the sample holder and cooled in the dilution fridge to a temperature of 400 mK, far below the critical temperature T_c of Aluminium to suppress the quasiparticle current due to thermal excitation which masks the detection of the excess current.

5.2.5 PicoWatt Resistance Bridge

Temperature of the dilution fridge is monitored using the "Picowatt-Resistance Bridge" which has temperature sensitive resistances that monitor the temperature of the sample holder, mixing chamber etc. Based on the readout of the particular resistance values from the Resistance bridge, one can determine the temperature of the device at which the measurements are made.

5.2.6 Dilution Fridge

Measurements are done using the Oxford Kelvinox dilution fridge. The sample is glued to the sample carrier and mounted on the sample holder. This sample holder in the dilution fridge is connected to the mixing chamber. The device is lowered in the dilution fridge by enclosing it in a copper box which is connected to the mixing chamber. The connections from the sample to the outer world are filtered using copper powder filters and the external electric noise is eliminated using low pass filters. A schematic of the dilution fridge and the electrical wiring in the dilution fridge for external connections is shown in Fig 5.3 and Fig 5.4

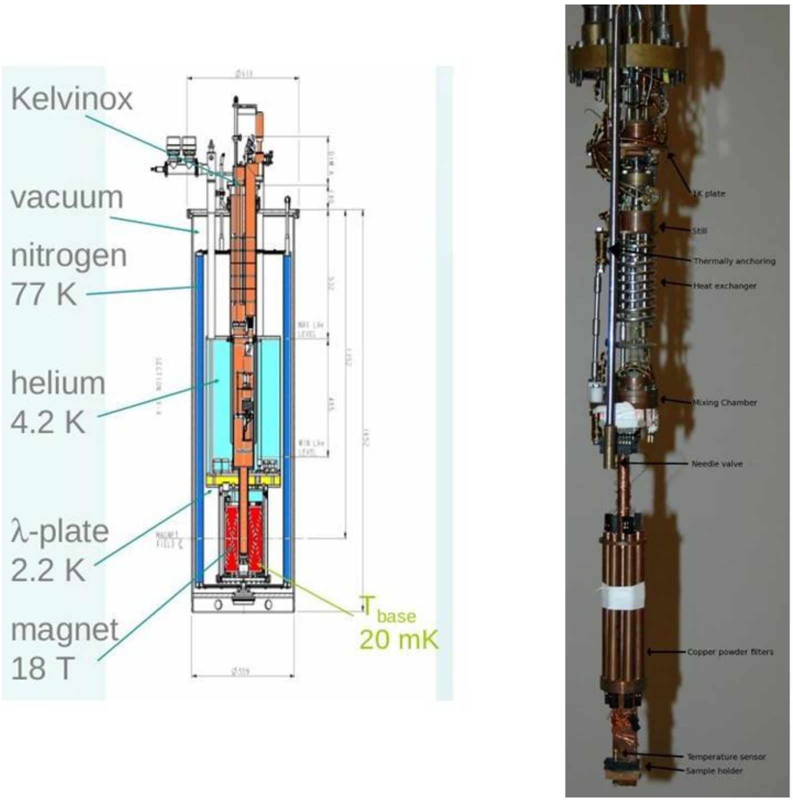


Figure 5.3: Schematic of a dilution fridge (Left) Photograph of the dilution fridge used for measurement.

[6]

5.3 Current-Voltage Measurements

Fig 5.5 shows the schematic of the device measured. The injector junction has an area of 2μ by 400nm and a resistance of 1.5KΩ. The three detector junctions are of areas 400nm by 400nm each with a resistance of 15KΩ. The 3 detector junctions are at a distance of

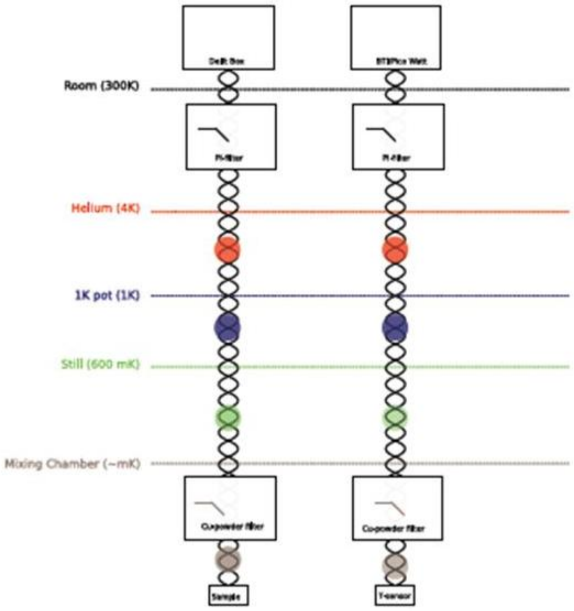


Figure 5.4: Schematic of the electrical wirings in a dilution fridge
[6]

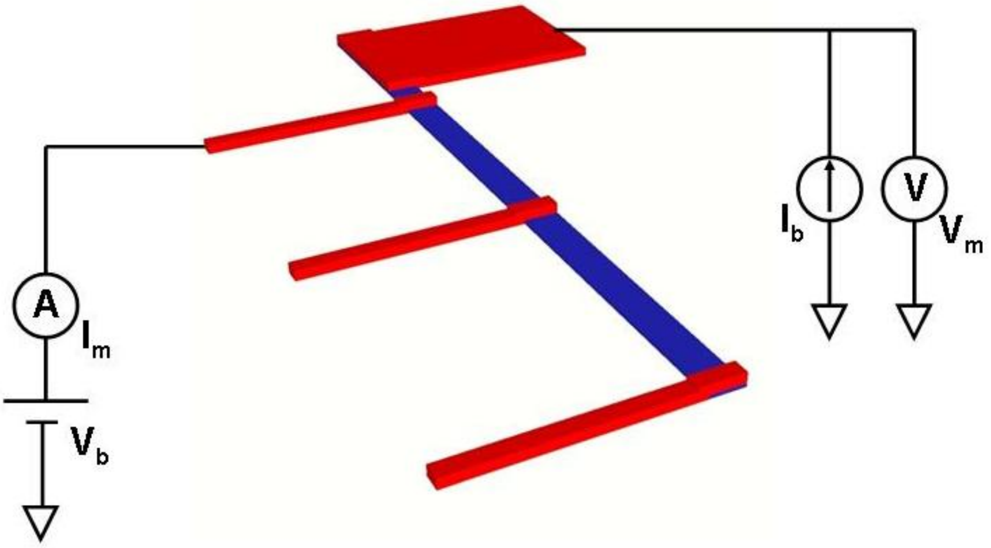


Figure 5.5: Schematic of the device structure showing the injector junction as well as the 3 detector junctions with the respective sources and measurement connections.

500nm, 28μm and 43μm. The circuit connections are as shown in Fig. 5.5. The voltage and current sources V_b, I_b, and the voltmeter and current meter V_m, I_m connections are made respectively. The voltage in the voltmeter is swept from ± 1 mVolt and the current in the current source is swept from ± 1 μ Amps. We shall now take a look at the IV characteristics of the current biased injector junction and the voltage biased detector junction.

5.3.1 Current Biased Injector Junction

Fig. 5.6 shows the IV characteristics of the injector junction. For the characteristics, the voltage V_b at the detector is fixed at a certain value $(0 in Fig 5.6)$ and the bias current I_b is swept from $-1\mu A$ to $+1\mu A$. From the IV Characteristics we can see that there is no voltage across the injector until the current across the junction reaches a certain value known as the critical current I_c. The initial current which rises from zero to I_c without any voltage, is a dissipationless current known as the Josephson Current. As soon as the current across the junction exceeds this limiting value, the junction switches into normal mode. This transition is seen as a sudden rise in the voltage measured across the junction. For further increase in the voltage sweep, the IV characteristics show a normal behavior. The critical current $I_c$measured for the injector junction is 103nA

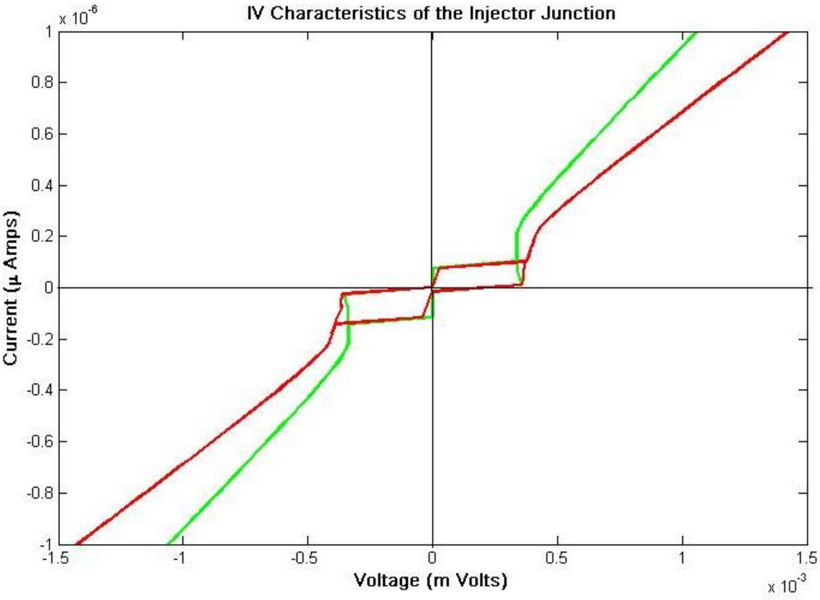


Figure 5.6: IV characteristics of the injector junction. Green curve is after taking into account the internal resistance of the current source, Red is the original measured.

Hence with a current biased SIS junction it is clear that the sub gap measurements are not possible due to the above behavior. If one has to study the behavior of the junction in the subgap region, one needs to bias the junction with a voltage source which is what is done for the detector junction.

As seen in the IV plot, the *Green* curves are plotted by taking the internal resistance into account. Total internal resistance calculated is about 375 Ω's. This is 10 Ω's from the voltage source plus 365 Ω's of the wires of the setup. The output resistance of the current source is 1G Ω's and drop across the current source can be neglected. Hence the measured voltage can be considered due to the full flow of current across the injector junction.

5.3.2 Voltage Biased Detector Junction

For the IV characteristics of the detector junction the voltage V_b across the detector is swept from ± 1mV and the current I_b at the injector is fixed at some value. The resulting IV curve of the detector junction is as shown in Fig. 5.7

As seen, the internal resistance of the sources definitely effects the IV curve of both the injector as well as the detector junctions and manifests itself contributing as a series resistance. This is seen as a slope in the *Red* curve when the current transitions to the normal state. This internal resistance however does not affect the gap of the superconducting aluminium in either case.

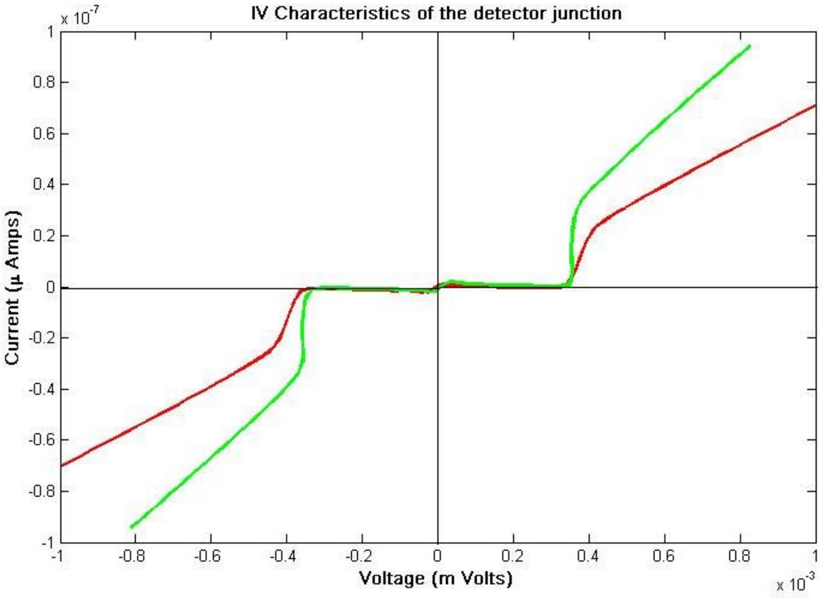


Figure 5.7: IV characteristics of the detector junction. Green curve is after taking into account the internal resistance of the voltage source, Red is the original measured.

Quasiparticles are injected into the superconducting wire and they are extracted at the detector junction. Hence a non equilibrium induced at some point in the wire cause a non local voltage at the point of measurement. This is seen in the IV characteristics of the detector junction shown in Fig 5.8. From section 5.2.4 the resistance of the current meter is calculated to be 2100 Ω's + 365 Ω's of the wires connecting the meter to the setup. When the drop due to the resistance of the current meter is subtracted, the *Green* IV curve is obtained. The energy gap measured with the above measurement is 0.346 m Volts which is the value of 2Δ.

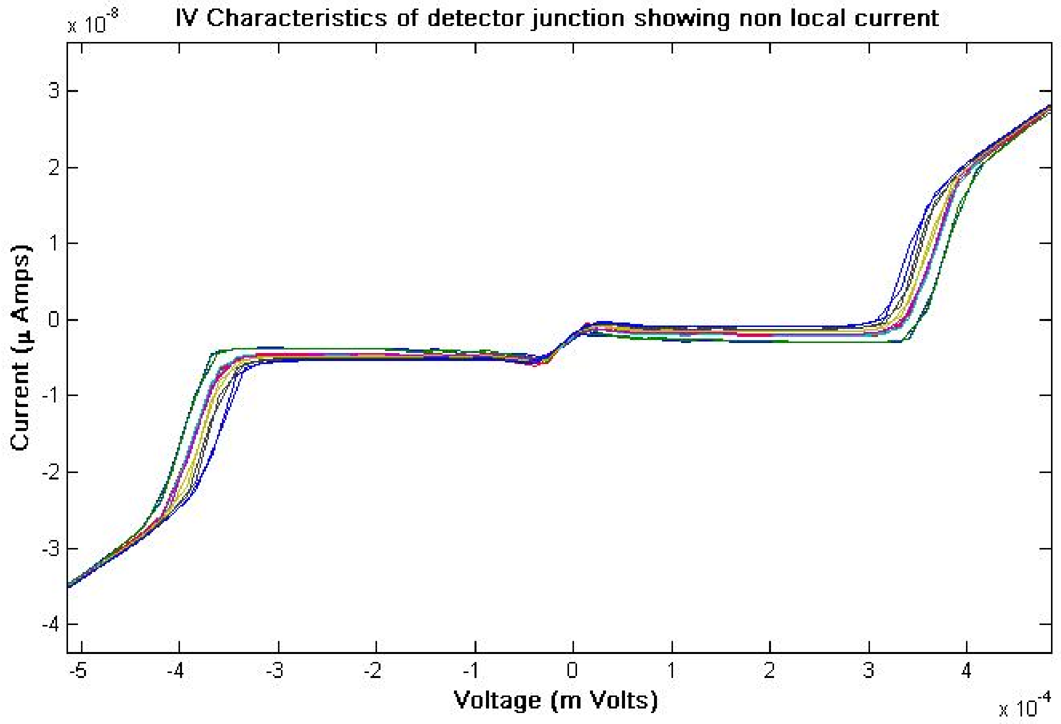


Figure 5.8: IV characteristics of the detector junctionshowing the non local behavior for different injector currents.

Chapter 6

Theory

In this chapter we shall present a a short review of the fundamentals of quantum mechanics, a brief overview of superconductivity and the theory of electron tunneling and charge imbalance due to tunneling.

6.1 Length Scales

Length scales determine the dimensionality of the system. In the three directions in space, the length can be different. For systems where electron behavior is wave like, it is important to know that there is an energy dependent wavelength associated with them known as de Broglie wavelength. For an electron with a mass m traveling with a velocity v and having a kinetic energy E_k, the de Broglie wavelength λ is given by [1]

$$E_k = \frac{mv^2}{2} = \frac{p^2}{2m} \tag{6.1}$$

where:

p is the momentum given by

$$p = mv = \hbar k = \frac{h}{\lambda} \tag{6.2}$$

In thermal equillibrium or in the ground state, all energy states below the Fermi Energy E_f are occupied and all energy states above are empty. Thus, the Fermi Energy is the state which has the maximum kinetic energy in the system. For this energy E_f equation 6.1 and 6.2 provide us the velocity v_f and momentum p_f. The Fermi wavelength λ_f is given by[1]

$$\lambda_f = \frac{h}{\sqrt{2mE_f}} \tag{6.3}$$

In real systems, electron transport is hardly ideal. Lattice irregularities, impurities and surface boundaries cause electrons to scatter. This scattering induces what is known as a mean free path, which can be defined as the length that an electron can travel before it encounters a scattering object. Scattering can be elastic or inelastic.

Elastic scattering occurs due the presence of static faults in the system, like, impurities, surfaces, lattice imperfections. In elastic scattering, the energy of the electron before and after the scattering is preserved. The direction of the electron traveling is however changed. The average distance an electron can hence travel before it experiences an elastic scattering is known as the elastic mean free path or elastic scattering length l_e.

Inelastic scattering also known as non-stationary or time-dependent scattering [1] is primarily due to the electron scattering via phonons (electron-phonon interaction) and collisions with other electrons (electron-electron interaction). It is termed inelastic, because an electron loses a part or all of its energy with such a collision. In such an event, neither the energy of the electron nor its wave vector are preserved. As in the case of elastic scattering, the inelastic mean free path, is the average length an electron travels before it experiences an inelastic collision. Since this type of a collision results in exchange of energy, the inelastic scattering length is also known as energy relaxation length l_ε.[1]

There is a crucial difference in regards to quantum mechanical consequences arising due to the elastic and inelastic scattering which relates to the phase of the traveling electron wave and is known as the phase coherence length.

From the above definitions it is clear that an elastic scattering preserves the phase of the electron, which means, any electron starting at a particular position will experience the same scattering sites and the phase will evolve in a deterministic way. Hence elastic scattering does not randomize the phase of the electron.

In contrast, inelastic scattering is non deterministic or it is statistical, which means, any electron starting in a particular position may not experience the same scattering if it is made to re-travel the same path again. Since the inelastic scattering mechanisms are statistical, the phase of a traveling electron is not predictable. With many such scattering, the phase ultimately becomes random.

Hence, a third length scale can be defined which is known as the phase coherence length l_φ [1]. It is the length an electron can travel before its phase is randomized.

It is evident that phase coherence length is important because it determines the degree to which the electron experiences process where the phase is important. Since phase breaking and energy exchange are clearly related (both resulting from inelastic scattering), this is an important length scale that has to be kept in mind.

Table 6.1: Comparison of different length scales in transport regimes.

Diffusive	Classical	$\lambda_f, l_e \ll L, l_\varphi < l_e$
	Quantum	$\lambda_f, l_e \ll L, l_\varphi > l_e$
Ballistic	Classical	$\lambda_f \ll L < l_e, l_\varphi$
	Quantum	$\lambda_f \approx L < l_e, l_\varphi$

6.2 Distribution Function

Electrons in a solid occupy well defined states as defined by the Pauli exclusion principle which states that no two electrons may occupy the same energy state. In a solid, the electrons start occupying energy levels from the lowest and fill up to the highest. This highest level that is filled is known as the Fermi Energy. Among the energy levels lie states and the occupation of these states is given by a distribution function.

The distribution function denoted by $f(E)$ is the probability that a state of energy E is occupied by a particle. Three distinctly different distribution functions are found in physics. The classical Maxwell-Boltzmann distribution and the Bose-Einstein and Fermi-Dirac distribution which are quantum mechanical.

For our study we are interested in the Fermi-Dirac distribution. The Fermi-Dirac distribution applies to fermions, particles with half-integer spin which must obey the Pauli exclusion principle.The significance of the distribution is seen at T=0. At absolute zero, the probability of occupation is = 1 for energies less than the Fermi energy and zero for energies greater than the Fermi energy E_F. We picture all the levels up to the Fermi energy as filled, but no particle has a greater energy. This is entirely consistent with the Pauli exclusion principle where each quantum state can have one but only one particle. The Fermi-Dirac distribution function is given by

$$f(F_E, T) = \frac{1}{e^{(E-E_F)/kT} + 1} \tag{6.4}$$

From the above equation it is clear that the distribution function depends upon the temperature. The function changes with the increase in temperature as shown below in Fig 6.2 The temperature T defined above refers to the temperature of the electrons. For a general definition, the electron temperature and the lattice or phonon temperature are considered the same and the inelastic interactions between the electrons and phonons is the primary energy relaxation mechanism. Under such a condition the distribution is said to be in equilibrium. When temperature of the system is reduced, the electron-electron interaction becomes more prominent while the electron-phonon interaction starts decreasing. In general, the electron temperature may not be equal to the lattice temperature, it is higher due to high energy quasiparticles. The distribution under such conditions is

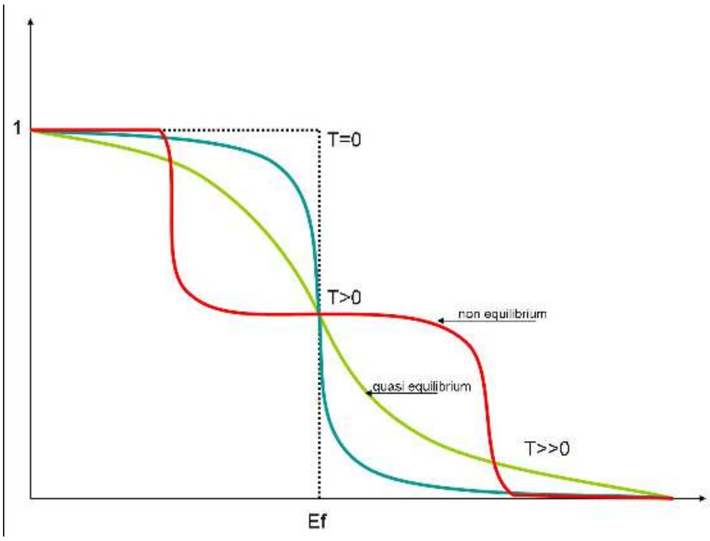


Figure 6.1: Fermi-Dirac distribution showing distribution for T = 0, T> 0 and T >> 0 Quasi-Equilibrium and non-equilibrium

known to be in quasi-equilibrium. Finally, on further reducing the system temperature, the lattice vibrations can be completely neglected. The only means of energy relaxation are due to the inelastic electron-electron interactions. Under such conditions, it is not appropriate to define a common temperature for the system since the electron temperature is not the same all over. The system is then said to be in non-equilibrium and the the energy distribution cannot be described by equation 6.4.

Pothier [163] presents a novel way using NIS tunnel junctions in which the energy distribution of quasiparticles can be measured. Fig 6.2 shows the experimental setup. The main idea behind this experiment was to exploit the property of the distribution function to have different shapes depending on the amount of inelastic collision a quasiparticle undergoes while diffusing through the wire. The distribution function $f(x, E)$ at a distance x and at an energy E determines how much energy the quasiparticles lose during their diffusion time τ_D.

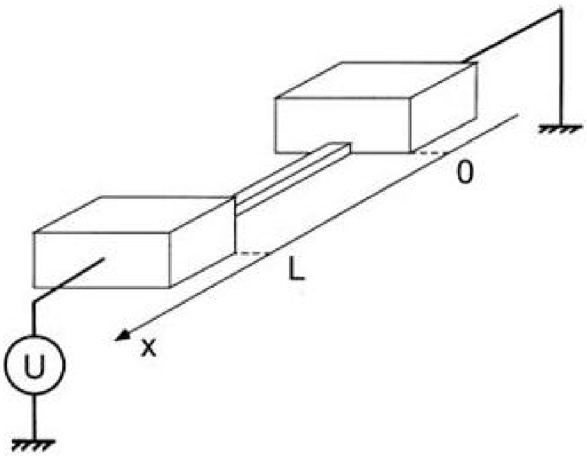


Figure 6.2: Experimental layout: diffusive wire of length L connected to the two reservoirs biased at potentials 0 and U

[163]

In this regard three limiting regimes can be thought of which explain how the properties of the distribution function depend upon the inelastic length scales mentioned in section 6.1:

- No inelastic scattering
- Strong quasiparticle scattering
- Strong phonon scattering

In the absence of any inelastic scatterings, the quasiparticles conserve their energy along the wire. In this case the distribution function $f(x, E)$ reflects the probability to find a particle at position x with energy E and it obeys the Boltzmann diffusion equation

$$\frac{\partial f(x,E)}{\partial t} - D\frac{\partial^2 f(x,E)}{\partial x^2} = 0 \tag{6.5}$$

The solution of the above equation given the boundary conditions of the wire is

$$f(x,E) = (1 - \frac{x}{L})f_T(E) + \frac{x}{L}f_T(E + eU) \tag{6.6}$$

where $f_T(E) = \frac{1}{1+e^{(E/k_BT)}}$ is the Fermi-Dirac function at temperature T. If $k_BT < eU$, the distribution function has a step at $f(x,E) = \frac{x}{L}$ for $-eU < E < 0$. The plot of the function given by equation 6.6 is shown in Fig 6.2.

In the case where there is no phonon scattering but the inelastic quasiparticle-quasiparticle scattering is the dominant mechanism of energy relaxation, a local thermal equilibrium is reached. The distribution function is a Fermi-Dirac function given by

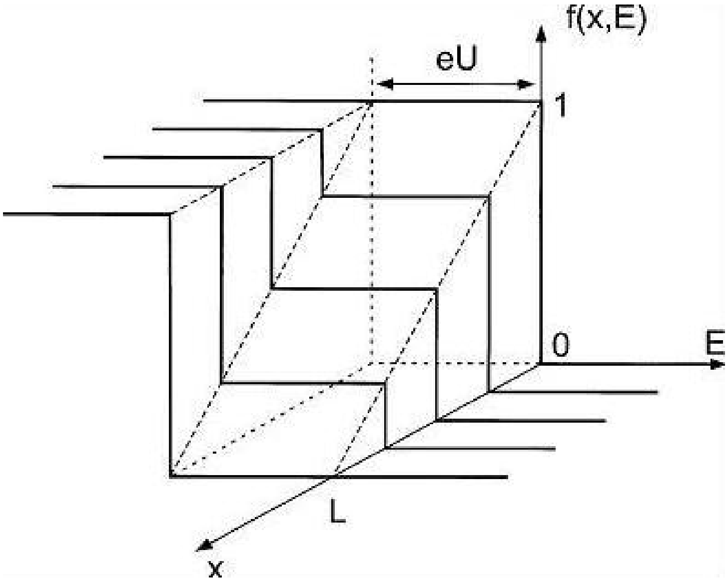


Figure 6.3: Distribution function $f(x,E)$ as a function of position x and energy E for non-interacting quasiparticles

[163]

$$f(x,E) = f_{Te(x)}(E - \mu(x)) \tag{6.7}$$

where $\mu(x) = -eU\frac{x}{L}$ is the local electrochemical potential and $T_e(x)$ is the electron temperature. The spatial dependence of the distribution function is shown in Fig 6.2.

The last variation is when there is a strong phonon-quasiparticle inelastic scattering wherein the electrons thermalize through the phonon interaction. The distribution function in this case is given by

$$f(x,E) = f_T(E - \mu(x)) \tag{6.8}$$

where $\mu(x) = -eU\frac{x}{L}$. the spatial dependence of the distribution is as shown in Fig 6.2.

6.3 Superconductivity

In 1911, Kamerlingh Onnes observed that the electrical resistance of Mercury suddenly vanished. Onnes stated `Mercury has passed into a new state, which on account of`

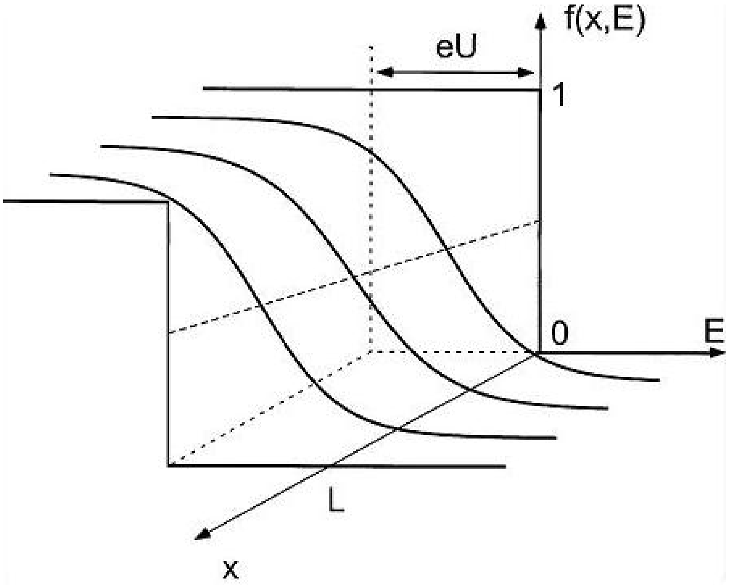


Figure 6.4: Distribution function $f(x, E)$ as a function of position x and energy E for strong quasiparticle scattering but negligible phonon scattering
[163]

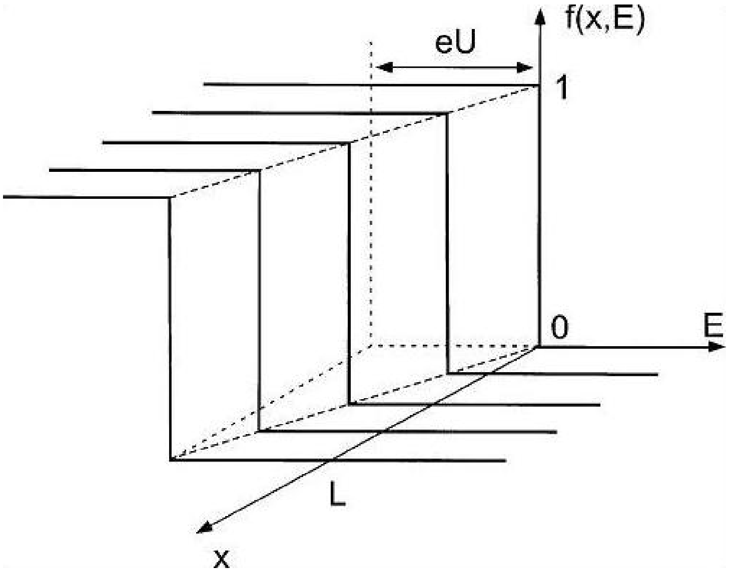


Figure 6.5: Distribution function $f(x, E)$ as a function of position x and energy E for strong phonon scattering
[163]

`its extraordinary electrical properties may be called the superconductive state.` [137] Thus evolved a new branch of physics known as "Superconductivity"

Superconductivity is a phenomenon occurring in certain metals when they are cooled below a certain temperature which is known as "Critical Temperature". Superconductivity is characterized by two main features

- Zero Resistivity
  $\rho = 0$ for all $T < T_c$

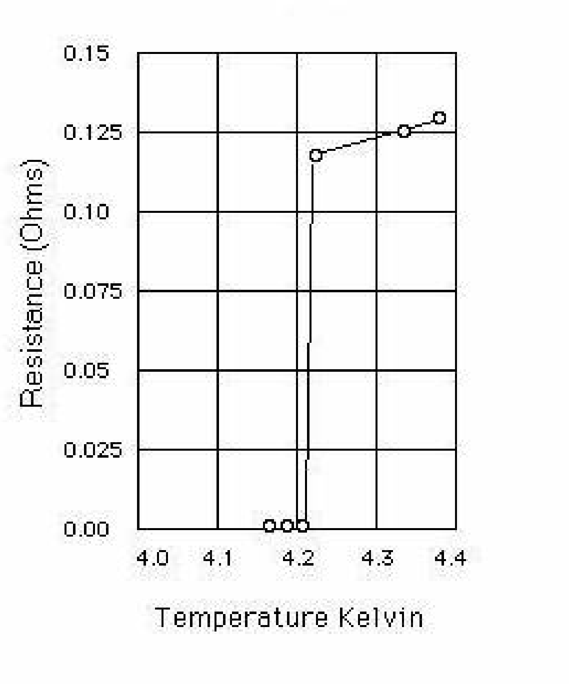


Figure 6.6: Temperature dependence of electrical resistivity

- No Magnetic Induction or Perfect Diamagnetism
  $B = 0$ inside the superconductor

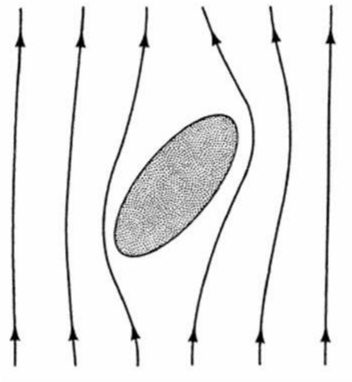


Figure 6.7: Weak external magnetic field expelled from the superconductor [2]

In 1957, the quantum theory of superconductivity was proposed by Bardeen, Cooper and Schriefer [3]. The specific accomplishments of this theory in brief are

1. The attractive interaction between electrons is a consequence of the indirect interaction between the electron-lattice-electron leading to the formation of "Cooper Pairs"

2. An attractive interaction between electrons leads to the separation of the ground state from the excited state leading to a formation of an energy gap. The critical field, thermal properties are a consequence of this energy gap
3. The penetration depth λ and coherence length ξ are the result of the natural consequence of the BCS theory.[4]

The minimum energy required to break the Cooper pair is given by $E_g = 2\Delta = 3.538k_bT_c$[5] Cooper pairs can be broken either by heating the superconductor beyond the critical temperature T_c leading to a temperature dependent quasiparticle distribution or by applying a current leading to a charge mode quasiparticle distribution. In both the cases, the resulting single electrons are called as "Quasiparticles".
The energy gap E_g mentioned above is given as below:

$$\frac{2}{V} = \sum_k \frac{1-2f_k}{E_k} = \sum_k \frac{1-2f_k}{(\Delta^2+\xi^2)^{1/2}} \tag{6.9}$$

where ξ_k is the energy of the state k w.r.t Fermi energy and f_k is the actual occupation numbers. If f_k equals $f_o(E/k_T)$ one obtains the BCS form of the Δ(T) which goes to zero as $3.07kT_c(1-T/T_c)^{1/2}$. If $f_k \neq f_o(E/k_T)$ Δ will take a different value and given by equation 6.9 The variation of the gap with temperature is as shown below in Fig 6.3

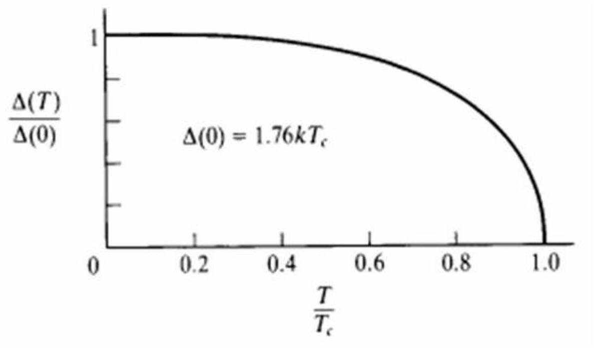


Figure 6.8: Variation of the gap with temperature
[5]

6.4 Quasiparticle Non-equilibrium

The Cooper pairs, that constitute the BCS ground state consist of a pair of electrons with equal and opposite momentum and spin. Hence a Cooper pair has twice the charge of an electron and zero momentum. When such a pair is broken, the single particle excitation energy for a particular state k is given by

$$E_k = (\Delta^2 + \xi_k^2)^{1/2} \tag{6.10}$$

where ξ_k is the energy of the state k w.r.t Fermi energy and Δ is the gap. In thermal equilibrium, these quasiparticle states are occupied with a probability given by the Fermi

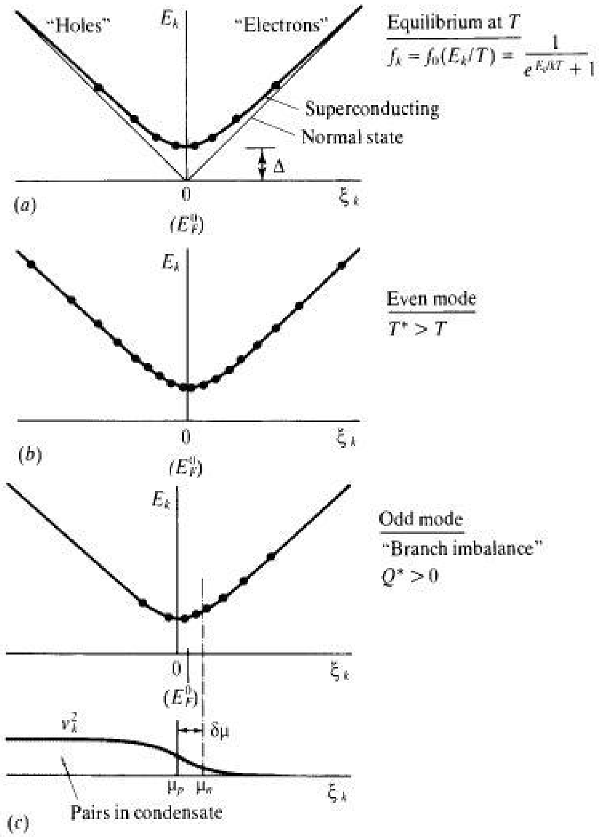


Figure 6.9: (a) Dispersion curves of excitation energies in Normal and Superconducting states. (b) Energy-Mode excitation with $T > T^*$. (c) Charge-Mode excitation showing charge imbalance with $Q > T^*$ and shift of μ_n and μ_p w.r.t equilibrium Fermi E_f

distribution function

$$f_0(E_k/kT) = [1 + e^{E_k/kT}]^{-1} \tag{6.11}$$

As the Fermi surface is traversed from outside to inside over the range $2\Delta > \xi_k \sim -2\Delta$, the nature of the excitations is given by

$$q_k = (u_k^2 - v_k^2) \tag{6.12}$$

where q_k is the charge of a quasiparticle.

The coherence factors $u_k^2(\epsilon_k)$ and $v_k^2(\epsilon_k)$ have their usual definition $u_k^2(\epsilon_k) = \frac{1}{2}(1 + \epsilon_k/E_k)$ $= v_k^2(-\epsilon_k)$ and $v_k^2(\epsilon_k) = \frac{1}{2}(1 - \epsilon_k/E_k) = u_k^2(-\epsilon_k)$. The non-equilibrium part of the quasiparticle distribution function in a superconductor can be decomposed into two orthogonal modes; one involving net quasiparticle charge and one involving only quasiparticle energy. The first gives rise to measurable potential differences between pairs and quasiparticles where there is conversion of supercurrent to normal current as in S-N interfaces and in phase-slip centers in superconducting filaments. This is also known as Charge-Mode Non-equilibrium or Branch Imbalance. The second can either involve a heating effect which can limit superconducting applications, or it can involve a cooling effect giving rise to enhanced or stimulated superconductivity. This is also known as Energy-Mode Non-equlibrium. Figuratively, the two orthogonal modes are shown in Fig. 6.9 [5] In our experiments we are interested in the od mode or Branch imbalance Q which is created by injection of quasiparticles. If the bias voltage of the tunnel junction V_{inj} is much larger than Δ/e and k_BT/e, the injected quasiparticles ar distributed in energy from Δ to eV_{inj}

but due to the polarity of the bias voltage most of these quasiparticles will be in the $k_>$ branch than the $k_<$ branch. If the bias is of the opposite polarity then these populations are also reversed with $k_<$ branch having higher population than the $k_>$ branch which implies injection is of holes rather than that of electrons.If on the other hand V_{inj} is $\sim$ to Δ then $u_k^2 \sim v_k^2$, and the branches are almost equally populated. The resulting Q will be therefore small.

The branch imbalance Q is defined as

$$Q = 2N(0) \int_{\Delta}^{\infty} N(E)(f_{k>} - f_{k<})dE \tag{6.13}$$

[156] where $N(0)$ is the density of states in the superconductor at the Fermi Surface $N(E) = E/\sqrt{(E^2 - \Delta^2)}$ is the normalized BCS density of states and ($f_{k>}$ and $f_{k<}$) are the steady state non equilibrium populations of the $k_>$ and $k_<$ branches.

As shown by Tinkham, [157] The steady state value of Q is Q_{inj} which determines how much imbalance is created by the injection current. Since the applied voltage is usually of the order of millivolts and the imbalance observed is of the order of nanovolts, the injected imbalance Q_{inj} is given by

$$Q_{inj} = \frac{G_{nn}}{e^2\Omega} \int_{\Delta}^{\infty} [f(E - eV_{inj}) - f(E + eV_{inj})]dE \tag{6.14}$$

where G_{nn} is the conductance in the normal state and f is the Fermi function. Q_{inj} differs from the injection current as shown below[156]

$$I/e = \frac{G_{nn}}{e^2\Omega} \int_{\Delta}^{\infty} N(E)[f(E - eV_{inj}) - f(E + eV_{inj})]dE \tag{6.15}$$

by a factor of $N(E)$ From equations 6.4 and 6.4 following Tinkham [156] we introduce the parameter $F(V_{inj}, T) = \Omega e Q_{inj}/I$ which characterizes the degree of imbalance of the injected population and is given by

$$F = \frac{\int_{\Delta}^{\infty} [f(E - eV_{inj}) - f(E + eV_{inj})]dE}{\int_{\Delta}^{\infty} N(E)[f(E - eV_{inj}) - f(E + eV_{inj})]dE} \tag{6.16}$$

F is dimensionless parameter and is shown to have the following limiting forms given by Tinkham [157]

$$F = \begin{cases} 2f(\Delta)/g_{ns} & eV_{inj} \ll k_BT \quad all \quad T \\ 1 & T \approx T_c \quad all \quad V_{inj} \\ [(eV_{inj} - \Delta)/(eV_{inj} + \Delta)]^{1/2} & T = 0 \quad all \quad V_{inj} \\ 1 & eV_{inj} \gg \Delta \quad all \quad T \end{cases}$$

6.5 Electron Tunneling

Giaever pioneered a technique to confirm the density of states and the temperature dependence of the energy gap as predicted by the BCS theory. The basic idea is, when two conductors are separated by a thin insulating barrier, there is a non zero exponential probability of charge transfer through quantum mechanical tunneling of electrons between the two conductors. This probability of tunneling depends upon the thickness and other details of the insulating barrier.

Tunneling can be explained conveniently by making use of the so called "Semiconductor Model" In this model, the normal metal is represented as the one with the continuous distribution of the density of states *N(O)* where energies below the Fermi level are occupied and above are empty. The superconductor on the other hand is represented as a semiconductor with the energy gap representing the forbidden barrier, the conduction band on top which is devoid of charge carriers and the valence band which is filled with charge carriers as seen in Fig. 6.10 [5].

The semiconductor model treats all tunneling transitions as horizontal i.e, they occur at constant energy after the chemical potential μ is adjusted to take into account the applied bias eV. As this method greatly simplifies computations, it is used extensively to describe the characteristics of various types of tunnel junctions. However due to the oversimplification that this method provides, it fails to explain important characteristics when superconducting materials are used to make tunnel junctions, namely the mixing of hole and electron states and charge imbalance regimes where the states inside and outside the Fermi are not in equilibrium.

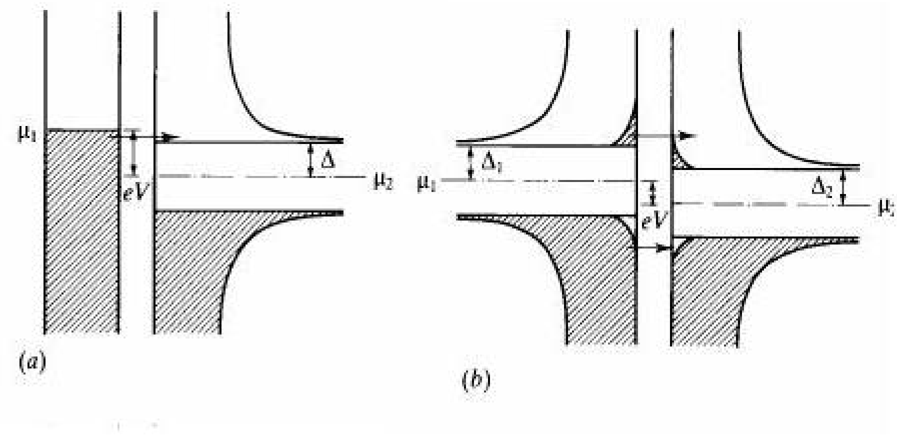


Figure 6.10: Semiconductor model description of tunneling. (a)*N-S* tunneling at $T = 0$ with bias eV slightly exceeding the energy gap Δ. (b) *S-S* tunneling at $T = 0$ with bias voltage $eV < \Delta_1 + \Delta_2$. Arrows depict tunneling.
[5]

Within the independent particle approximation, tunneling current from conductor 1 to conductor 2 can be expressed as [5]

$$I_{1\to 2} = A|T|^2 \int_{-\infty}^{\infty} N_1(E)f(E)N_2(E+eV)[1-f(E+eV)]dE \tag{6.17}$$

where:

V is the applied voltage eV is the difference in chemical potential and *N(E)* is the density of states. $N_1 f$ and $N_2(1-f)$ are the occupied and empty states. T is the tunneling matrix and A is the proportionality constant.

Using the above expresions we describe tunelling between two superconducting materials which is required for our experiments.

6.5.1 Superconductor-Superconductor Tunneling

If both the metals on either side of the tunnel barrier are superconductng the energy diagram of Fig 6.10b is relevant. In this case, equation 6.17 changes to

$$\begin{aligned} I_{ss} &= \frac{G_{nn}}{e}\int_{-\infty}^{\infty} \frac{N_{1s}(E)}{N_1(0)}\frac{N_{2s}(E+eV)}{N_2(0)}[f(E)-f(E+eV)]dE \\ &= \frac{G_{nn}}{e}\int_{-\infty}^{\infty} \frac{|E|}{[E^2-\Delta^2]^{1/2}}\frac{|E+eV|}{[(E+eV)^2-\Delta_2^2]^{1/2}}[f(E)-f(E+eV)]dE \end{aligned} \tag{6.18}$$

Fig 6.11 describes the qualitative behavior. At $T = 0$, no current can flow untill $eV =$

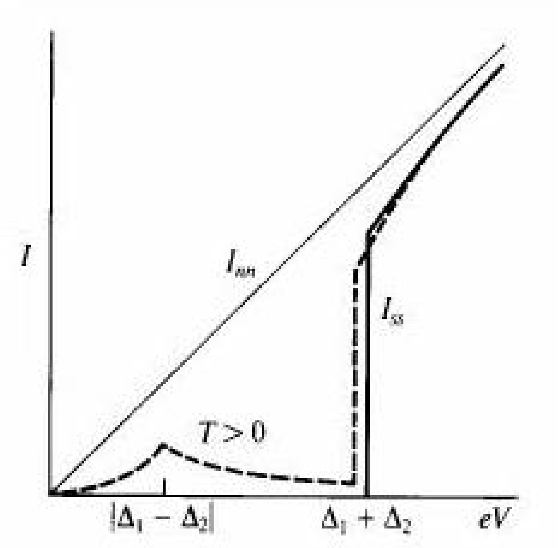


Figure 6.11: Tunneling Characteristics of Superconductor-Superconductor tunneling. Solid Lines are for T = 0 and dotted lines are for $T > 0$.
[5]

$\Delta_1 + \Delta_2$. At $eV = \Delta_1 + \Delta_2$ there is a sudden rise in the current I_{ss} due to the availability of infinite density of states at the gap edges. For $T > 0$, there is a current even at lower voltages due to the presence of thermally excited quasiparticles.

Chapter 7

Results and Analysis

In this chapter we shall take a look at the measurement setup and a brief look at the IV characteristics of the injector junction and the detector junction. We shall concentrate in the region of the energy gap for the voltage biased detector junctions because it is in this subgap where the non-equilibrium properties can be observed and studies. Fig 7.1 shows the schematic of the setup Fig 7.2 shows the IV characteristics of the injector junction. The * are the points where this junction is current biased for studying the affects of nonequilibrium on the detector junctions.

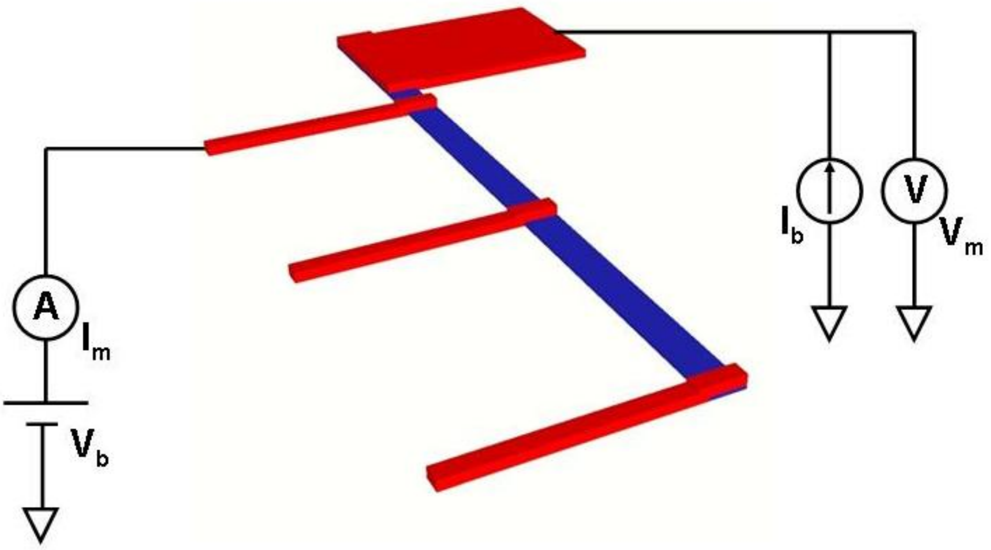


Figure 7.1: Schematic of the device structure showing the injector junction as well as the 3 detector junctions with the respective sources and measurement connections.

In the first part of the analysis of the results we shall look at how a nonequilibrium induced at the injector by forcing a current induces nonequilibrium at the detector. Second we take a look at the quasi-particle current measured at the particular detector junction and see how this current affects the energy gap and how it changes with the distance from the injector for different values of injector current. Third we study whether the dissipationless Josephson Current mentioned in section 5.3.1 also has an effect in the subgap current of the detector. Finally, we shall see the excess current manifesting itself within the subgap of the detector junction and its variation with regards to the injector current and the detector distance.

Fig 7.3 shows the quasiparticle thermal density or aluminium. The quasiparticle density

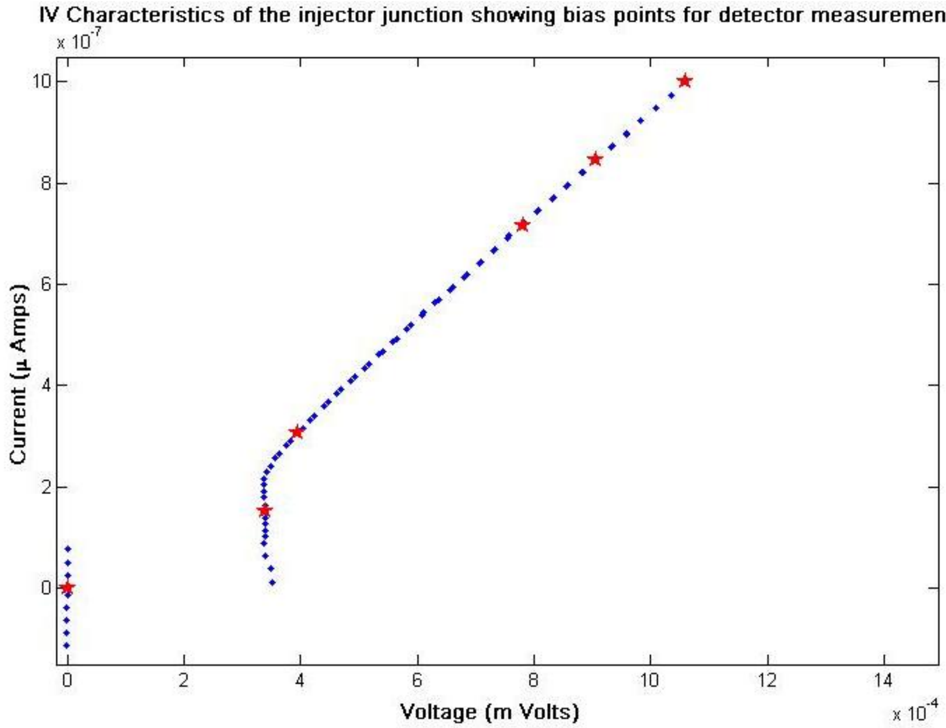


Figure 7.2: IV Characteristics of injector junction showing specific bias points where detectors are probed

for the injected current is calculated as follows:

1. The current $\frac{I}{e}$ is equal to the number of quasiparticles per second $\frac{\sharp qp}{s}$
2. The quasiparticle density is then given by $\frac{\sharp qp}{V}$ where V is the volume of the injector.
3. The recombination is equal to the injected quasiparticles is given by γn^2 where γ is the recombination rate given by

$$\gamma = (\frac{2\Delta}{k_B T_c})^3 \frac{1}{2\Delta N_0 \tau_o}[165] \tag{7.1}$$

4. Hence the density of qusiparticles can be given by

$$n = \sqrt{\frac{I/e}{V\gamma}} \tag{7.2}$$

where: τ_0 is the intrinsic relaxation rate = 100ns [156] N_0 is the thermal density of states = $3.5244 \times 10^{10} \mu m^{-3}$, 2Δ is the gap at injection current of 1μA = 0.3146 mVolts, T_c is the critical temperature of Aluminium = 1.19K, k_B is the Boltzmanns constant =

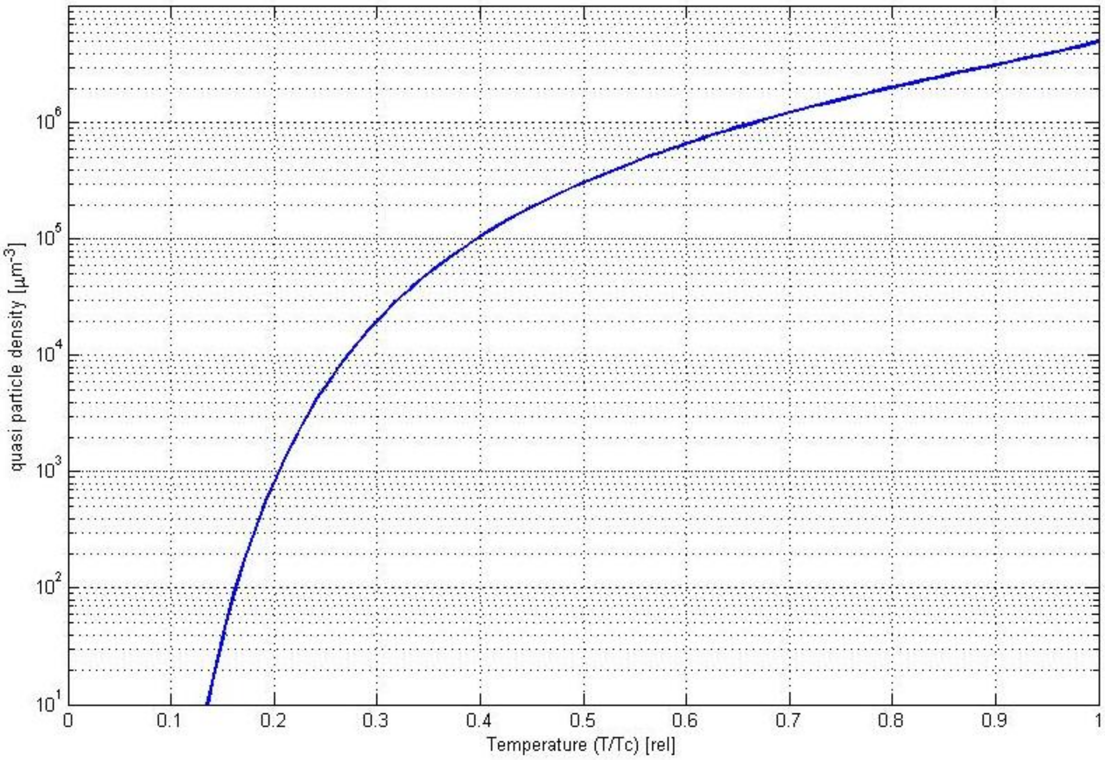


Figure 7.3: Plot showing the quasiparticle density as a function of temperature for Aluminium

$8.167 \times 10^{-5} eVK^{-1}$.
With the above calculations, the density of quasiparticles comes out to be $2.395 \times 10^6 \mu m^3$. and the calculated thermal density is $5.05 \times 10^6 \mu m^{-3}$. Thus the amount of quasiparticles injected is almost equal to the thermal density and we can safely say that we are in the high injection regime.

7.1 Nonequilibrium Effects and Energy Gap deviations

Fig 7.4, 7.5, and 7.6 show the IV characteristics of the three superconducting detector junctions connected to a superconducting wire into which quasi-particles are injected. Fig 7.4 is for detector junction 1. From the IV characteristics it is clear that for increase in the injected current, there is an increased subgap quasiparticle current at the detector. Another property visible from the IV of Fig 7.4 is that as long as the magnitude of the injected current is below the energy gap of the injector, no increase in the quasiparticle current is seen in the detector[1]
Fig 7.5, and 7.6 show the characteristics for the detectors two and three respectively. The overall IV characteristics are the same for both the junctions but there is a marked

---

[1] Overlap of *black* and *cyan*

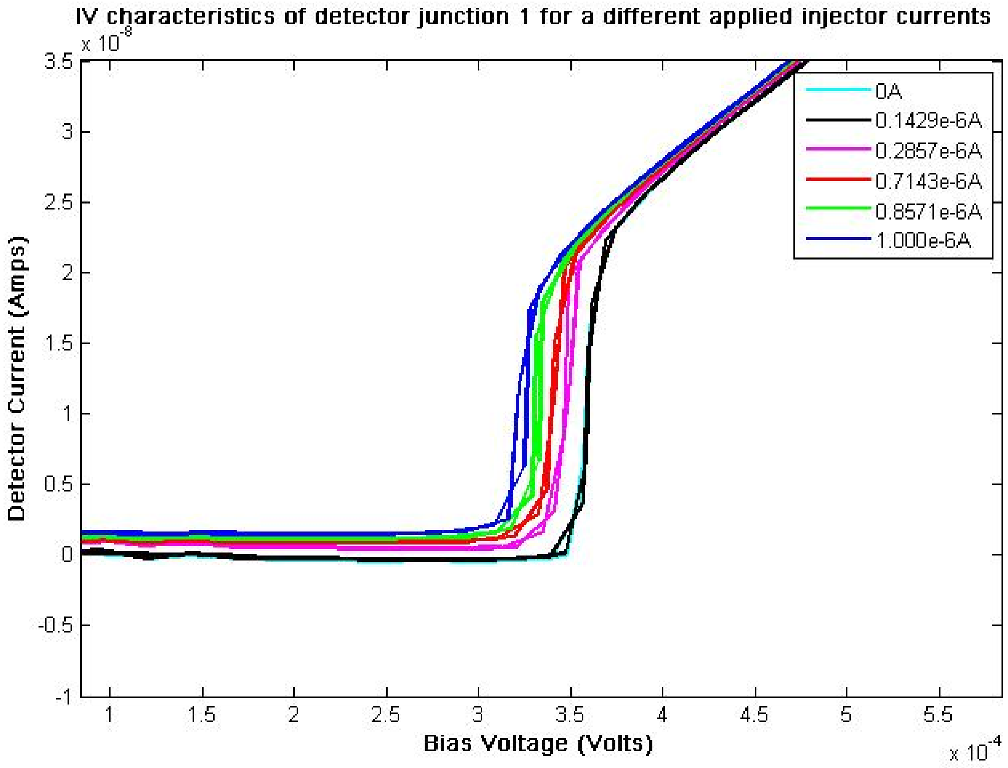


Figure 7.4: IV Characteristics of detector junction 1 for different injector currents shown in Fig 7.2

difference in the subgap current measured among the three junctions. The decrease is clearly visible. This decrease can be attributed to the processes of recombination and energy relaxation of the injected quasiparticles as they travel through the diffusive wire. Due to these, the non equilibrium at the injector which shows a strong presence at the closest detector decays off and is not very prominent at the farther detectors.

Another property observed in the above IV characteristics relates to the energy gap of the detector junctions. The injection of high energy quasi-particles results in a change of the distribution function of the superconductor which is analogous to a kind of heating of the superconductor which reduces the energy gap
As observed earlier in the case of the subgap current, when the magnitude of the injected quasi-particles is below the energy gap of the injector junction there is no reduction in the energy gap.
The following can be inferred from the IV characteristics of Fig 7.4, 7.5, 7.6

- For the quasiparticles injected with energy less than the energy gap, the observation is the same as is for the detector junction 1 where neither the energy gap nor the excess quasi-particle current change because of the Josephson current.
- When the injected quasiparticles are with higher energies, the observed change in the energy gap in junctions 2 and 3 is not as prominent as in junction 1. The affect of the quasi-particle energy on the gap is seen to decrease with the increase in the distance from the injector junction.

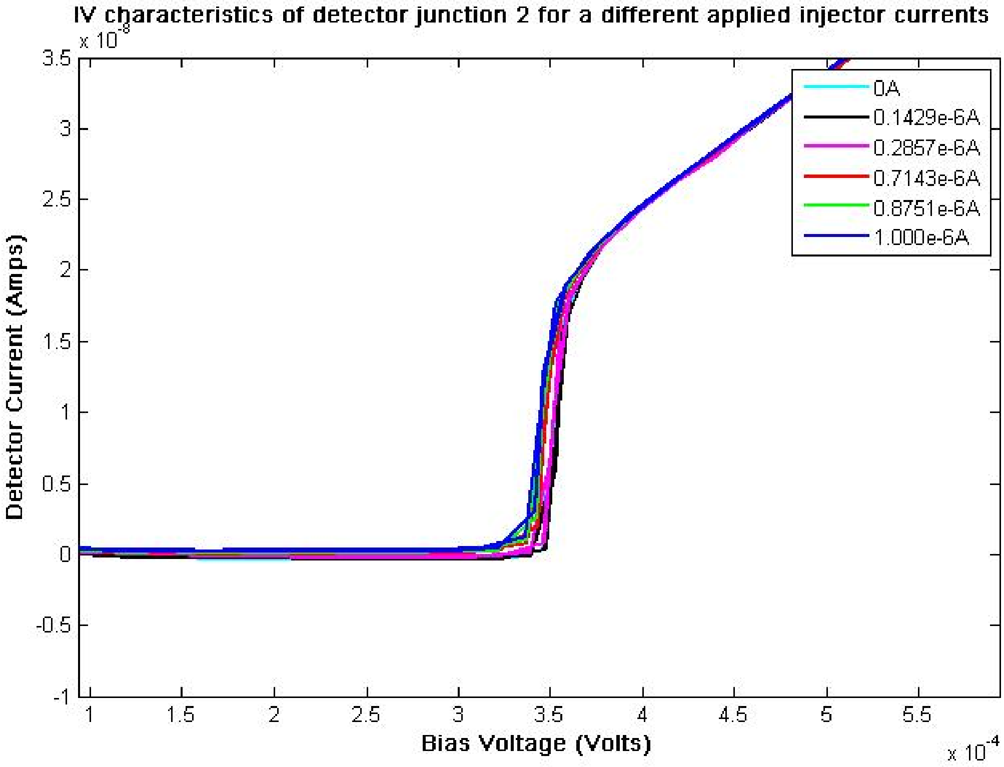


Figure 7.5: IV Characteristics of detector junction 2 for different injector currents shown in Fig 7.2

- The observed sub gap current in junctions 2 and 3 is also much less in comparison to junction 1. Again this current is seen to decrease with increase in the distance from the injector.

From the above observations the following conclusions can be drawn

1. Quasiparticles injected with high energies cause heating which affects the energy gap and also induce excess subgap current

2. When probed at different distances, it is seen that the above affects are diminished because the quasiparticles lose their energy while diffusing through the wire due to energy relaxation and recombination.

3. The diffusion or the relaxation of quasi-particle energy takes place over quite a long distance which implies that hot spots are not created by injection.

4. The phenomenon is non local since the observed affects are due to the possible transport of non equilibrium quasiparticles which show their presence over large distances from their origin.

Figure 7.7 shows the variation of the energy gap as a function of the injected current. Fig 7.8 shows the variation of the gap voltage with respect to the bias condition of the injector. The plot is calculated using equation 6.9 assuming that the detector junction is not heated ($T = T^*$. From Fig 7.8 it is clear that for none of the detectors see a change in the gap untill the injector current switches to the normal current.

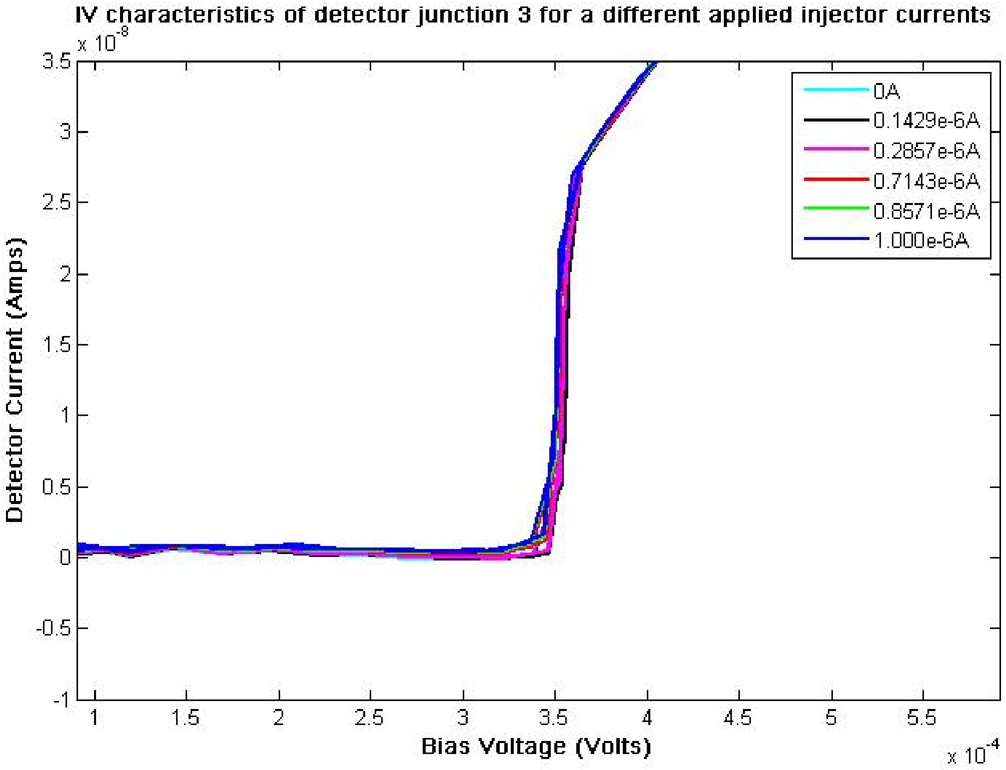


Figure 7.6: IV Characteristics of detector junction 3 for different injector currents shown in Fig 7.2

7.2 Effect of Josephson Current

In section 5.3.1 we mention about the Josephson current, a current observed at zero voltage and seen to increase till the critical current I_c is reached. We take a look now at how and if this current has any affects on the observed sub gap quasiparticle current at the detectors. Fig 7.9 shows the current-voltage characteristics of the injector junction *(lower curves with arrows)* and the excess current *(upper curves)* measured at detector junction 1. The upper curves are shifted by 0.5μ for clarity.

The excess current is calculated as follows: With a fixed voltage at the detector the injector current is varied from zero to 1μA. This current is measured at the detector. From the measured current, the value obtained for the detector current when the injector current is zero is deducted from all the other values of the measured current. The resulting current at the detector is the excess current . The current measured at the detector when the injector current is zero is due to the applied bias at the detector which is to be taken off. Hence what remains after the subtraction is only the current due to the injector and hence we term it as excess current.

As seen in Fig 7.9, increase in the Josephson current does not give rise to any quasiparticle current in the detector. As soon as the magnitude reaches I_c, the current switches to the quasiparticle branch and we observe a quasiparticle current at the detector. The sudden increase in the detector current observed at 2Δ is due to the alignment of the empty and filled DOS's on either side of the tunnel junction. As the current is further increased, we see the normal state IV characteristics.

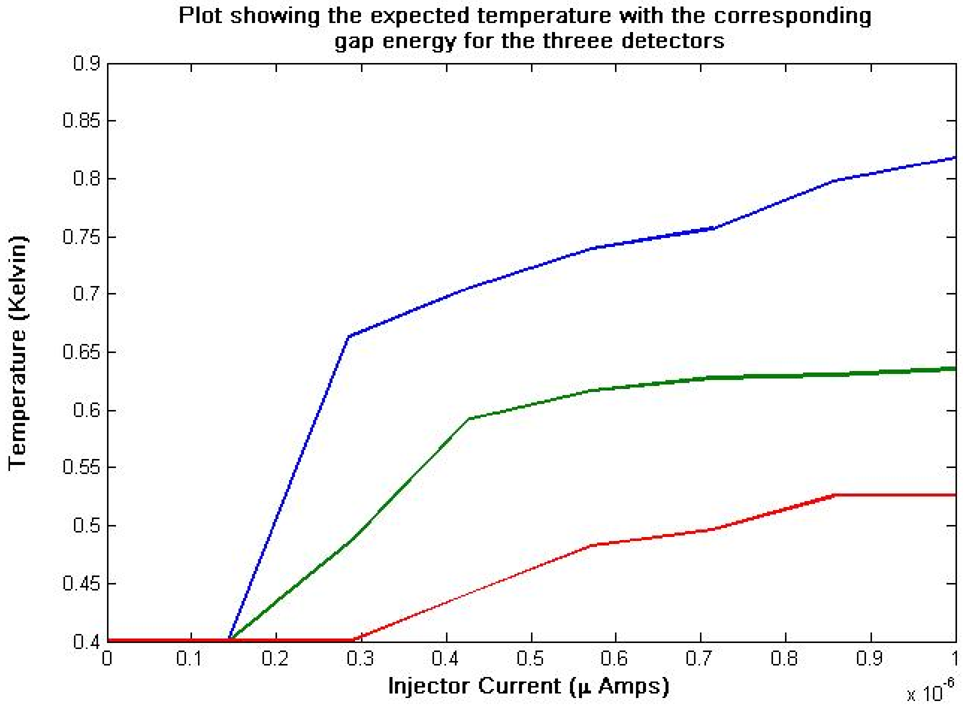


Figure 7.7: Plot showing the variation of the gap voltage at the three detectors for different injector currents

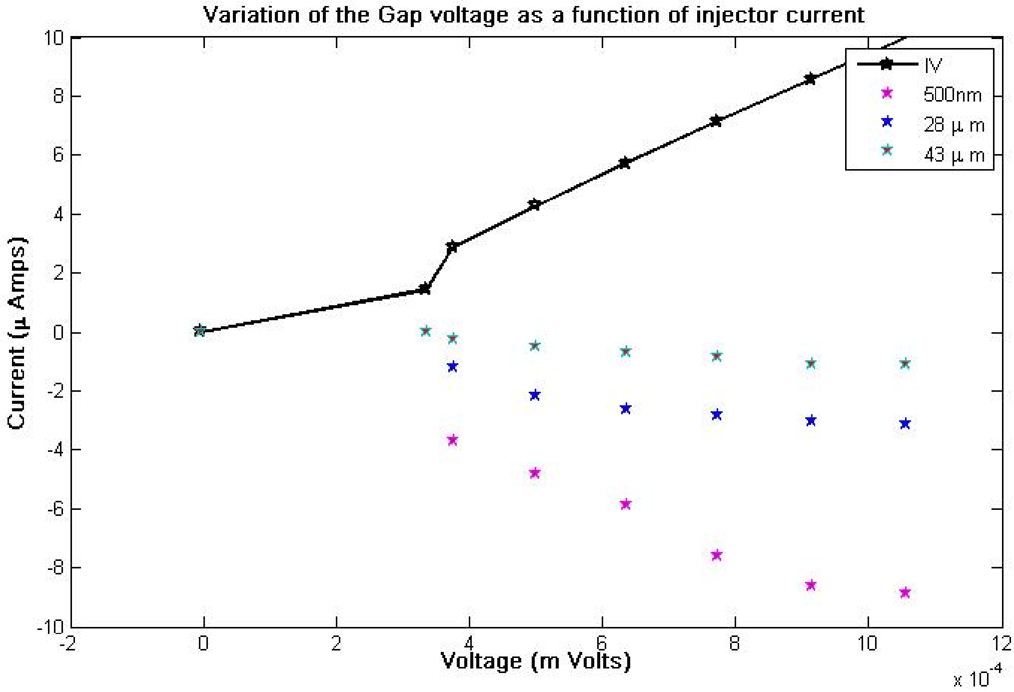


Figure 7.8: Plot showing the variation of the gap voltage as a function of the bias points along the injector IV characteristics for the three junctions

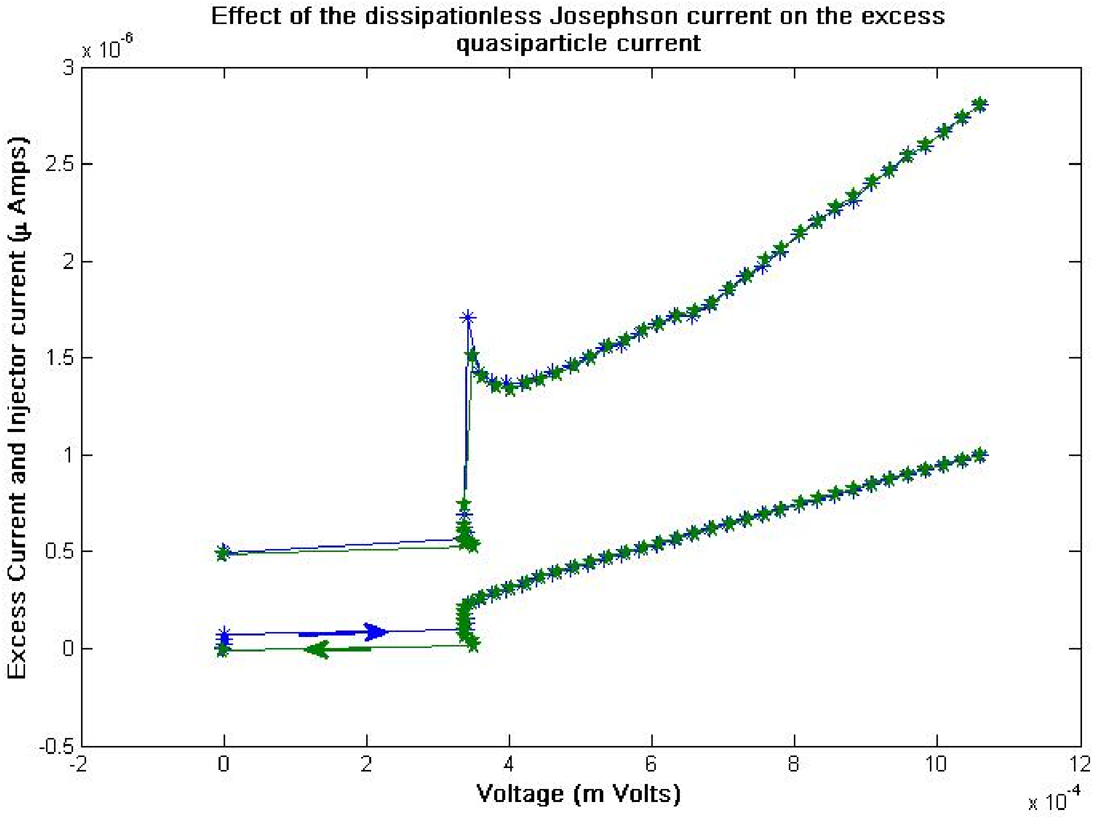


Figure 7.9: Plot showing the variation of the excess current with respect to the injector current as a function of the voltage across the injector

The green curves correspond to the case when the current across the injector is brought back to zero. It is seen that to go back to the dissipationless current rom the quasiparticle current, the magnitude of the current is lower than the one in the prior case. This is due to a phenomena known as "Washboard Potential" as explained in Tinkham [5] section 6.3. This causes the hysteresis seen in Fig 7.9.
To conclude, the Josephson current does not contribute to the quasiparticle current. The quasiparticle branch current switches to the normal state at the voltage of 2Δ and in the reverse direction, we observe a hysteresis due to the washboard potential.

In Fig. 7.9 the excess current depicted is for a particular current at the injector. This excess current varies with the injector current and as well as with the distance as shown in Fig 7.10 For all the three detector junctions, for all different values of injection current, the excess current measured at the particular detector is presented. It is seen that at the last detector which is 43μm from the injector, we still measure excess current of 1nAmp for an injector current of 1μAmp. With this observation we can conclude that nonequilibrium created at the injector does decay quite slowly and has influence over quite large distances.

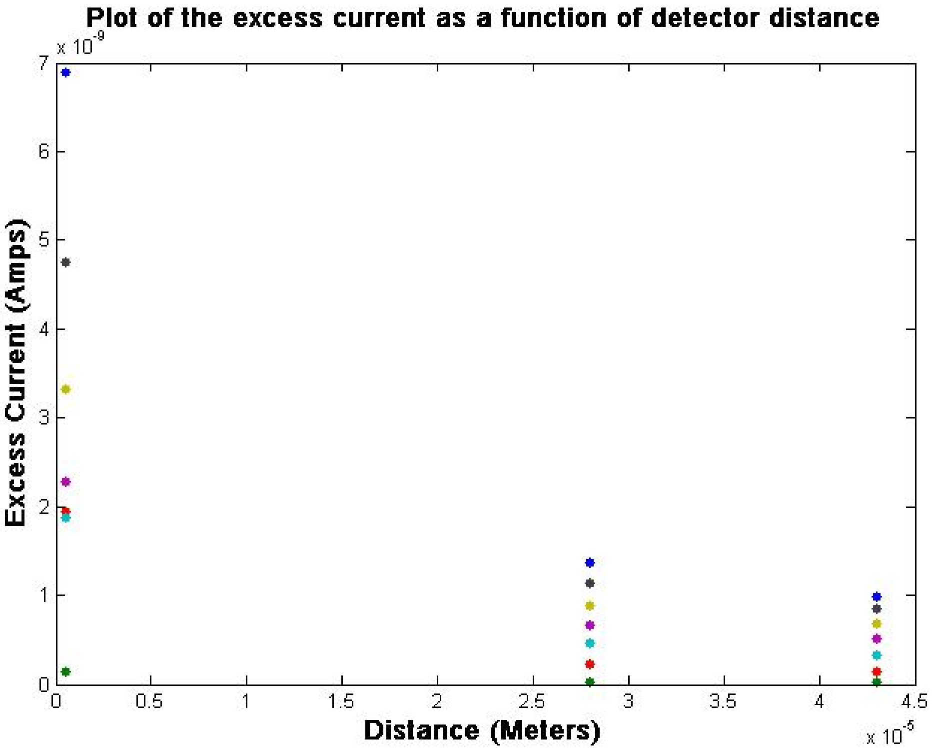


Figure 7.10: Plot showing the excess current at the three detectors for an applied bias at the detector

7.3 Detection and Measurement of Charge Imbalance

As explained in Chapter 6 section 6.4 high injection voltage from the injector much above the energy gap results in a unequal distribution of electron and hole like particles resulting in an unequal distribution of charge carriers. In this section we examine this nonequilibrium in our experiment.
We take a look at the IV characteristics of the detector junctions when the injector is fixed at positive as well as negative current polarity and the applied voltage bias across the detector is swept through a complete cycle from -1mV to +1mV.
To examine the charge non-equilibrium we consider 4 cases

- The injector current is positive and the applied voltage bias across the detector is positive
- The injector current is positive and the applied voltage bias across the detector is negative
- The injector current is negative and the applied voltage bias across the detector is positive
- The injector current is negative and the applied voltage bias across the detector is negative

For each of the above cases the variation in polarity of the injector current with respect to the polarity of the applied bias at the detector will result in either an excess of electronlike or hole like current. Consider Fig 7.11 It is observed that a with a positive injector current

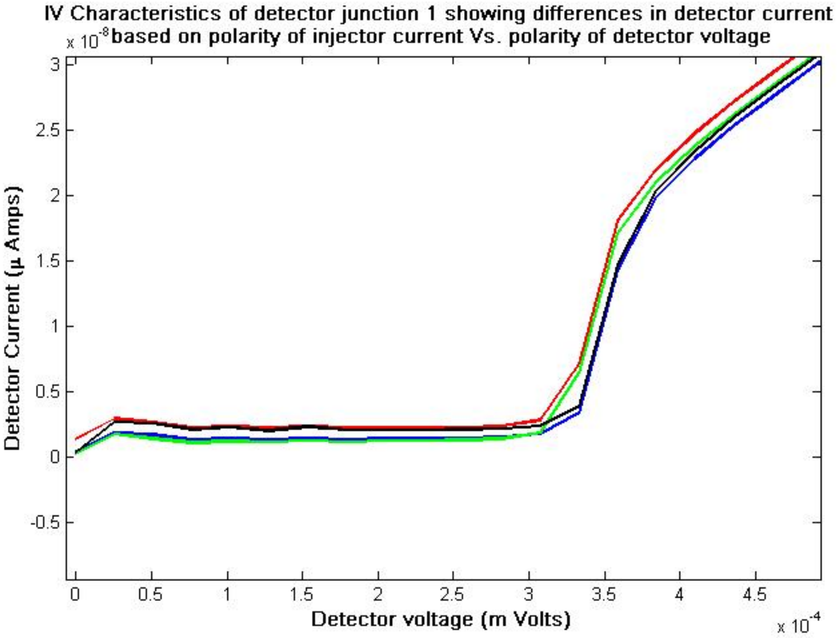


Figure 7.11: Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 1 for injection current of 1μA

and a positive voltage at the detector *Redcurve* results in a larger subgap current as compared to when the voltage at the detector is negative *Bluecurve*. Similar argument holds good when the polarity of the injector current is negative and so is the voltage at the detector *Blackcurve* and when the polarities of these two are opposite, the current is diminished *Greencurve*.

The Fig 7.12 is also for the detector junction 1 but the magnitude of the injector current is less The results are similar to the ones observed in Fig 7.11. An interesting observation in both the above Fig's is that the curves overlap each other. The current observed for the positive-positive polarities is the same as that observed for negative-negative polarities (in absolute terms). So is the case for the opposite polarities.

From the above observation we can conclude that depending on the polarity of the applied voltage bias across the detector and the polarity of the injected quasiparticle, there is a measurable potential difference across the positive and negative branches of charge distrubution. This manifests itself as charge imbalance. It should be noted here that the low injection currents have not been taken into account because one needs an energy at least greater than than Δ/e to observe a significant and measurable potential distribution The curves for the other two junctions for the same applied magnitude of the injected current are presented in Fig's 7.13 to 7.16. The observations and conclusions for these two junctions are also the same as they are for the detector 1.

The following conclusions can be drawn from the above arguments

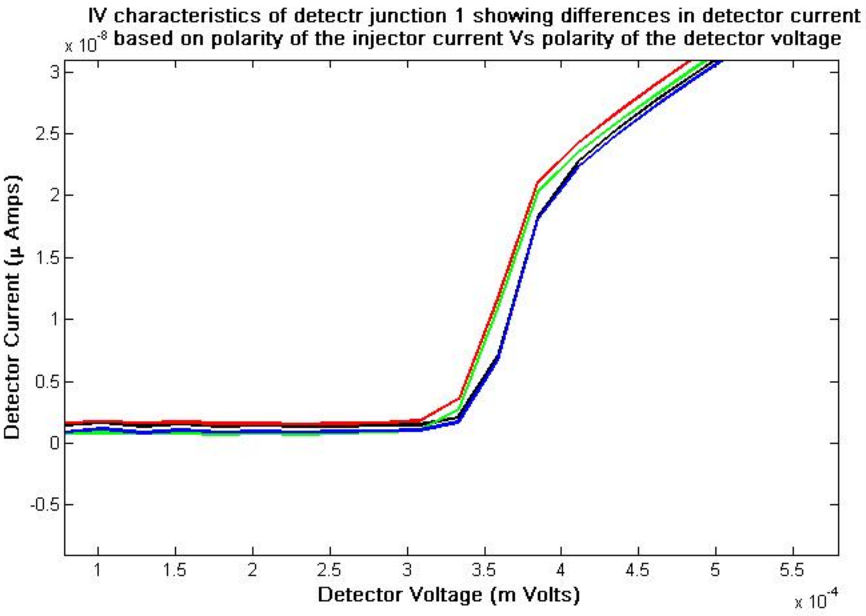


Figure 7.12: Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 1 for injection voltage of 0.85μA

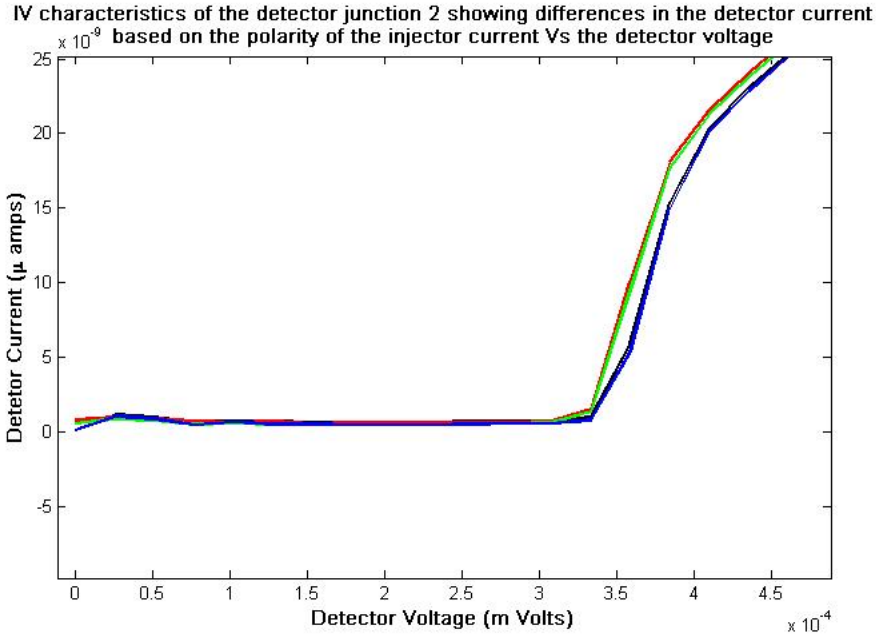


Figure 7.13: Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 2 for injection current of 1μA

- With like polarities of the injector and detector, there is excess quasiparticle current as compared to the the current when the polarities are opposite. This relates to the type of quasiparticle that is injected and the affinity it has with the detector

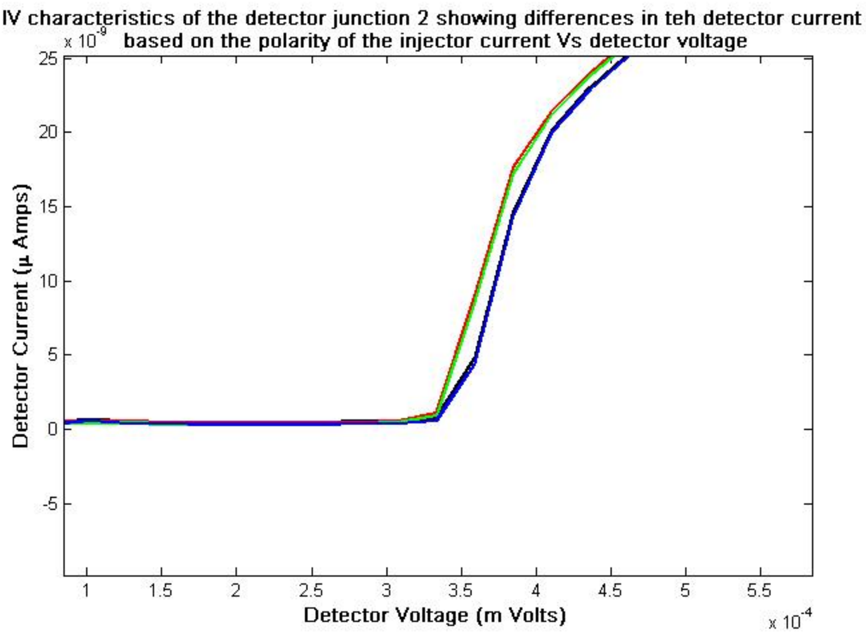


Figure 7.14: Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 2 for injection voltage of 0.85μA

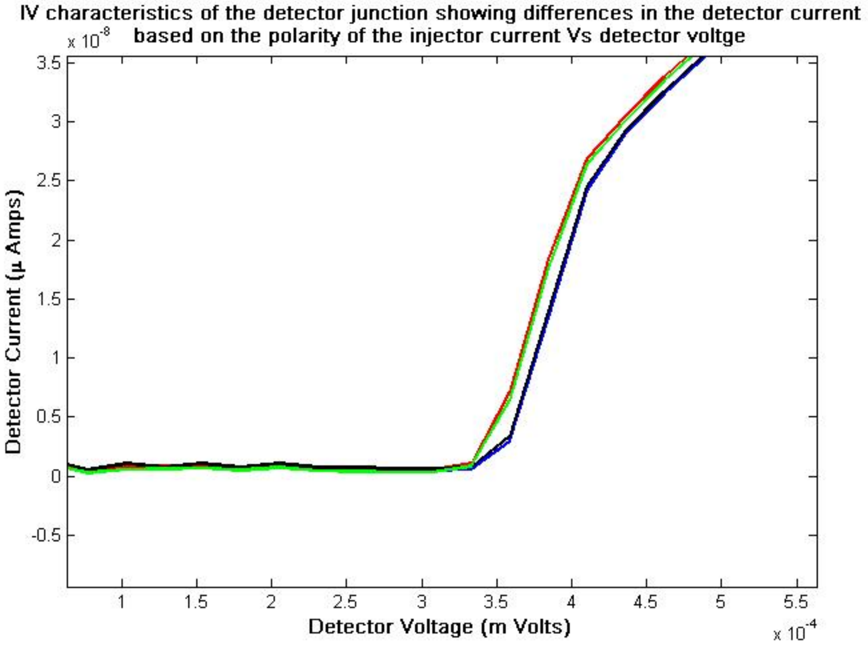


Figure 7.15: Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 3 for injection current of 1μA

polarity.

- With increasing injection energy much above the energy gap, the population of

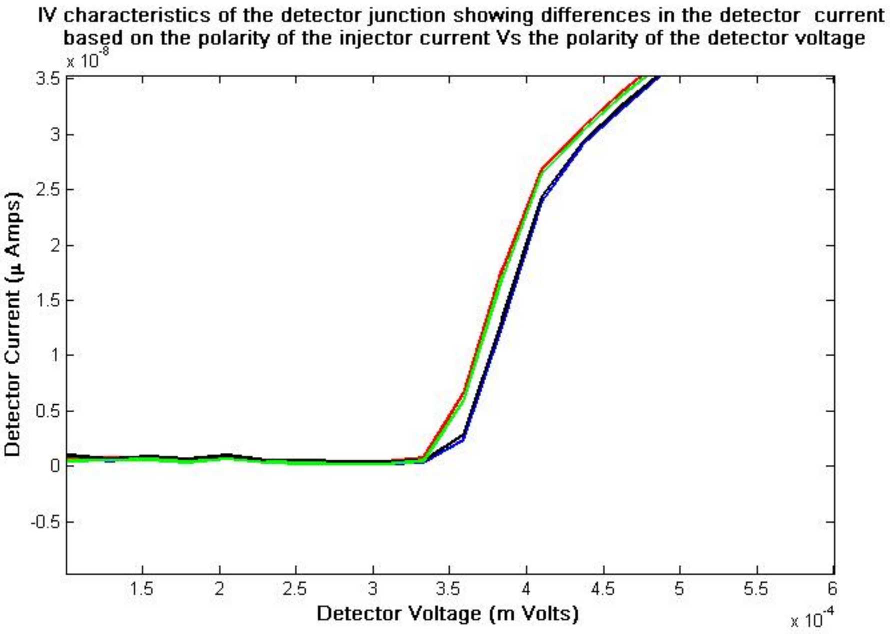


Figure 7.16: Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 3 for injection voltage of 0.85μA

> quasiparicle branches is unequal as explained in section 6.4 which results in the unequal currents observed. Hence one can conclude that this unequal current measured is due to the branch imbalance.

We shall now delve deeper and measure the degree of imbalance injected. For this we will measure the excess current as a function of injected current for different bias of the detector voltage. Since the excess current is made up of symmetric as well as asymmetric parts, we shall separate these two components. The asymmetric part of this excess current reveals the charge imbalance.

Fig 7.17, 7.18, 7.19, show the plots of the excess current against the injected current and the symmetric and asymmetric parts of the excess current. With Fig 7.17 charge imbalance is clearly visible between the two polarities of the injector current. Also seen is the rise in the excess current as the applied voltage at the detector gets closer to the energy gap Δ and then it falls off again. At the outset, it is seen that the asymmetric part shows marked behavior for the detector junction 1 which is closest to the injector. It is more dependent on the injector current at energies greater than Δ. At energies below Δ there are not many quasiparticles and hence the injector current dependency is quite low. The characteristics for the other two detectors are shown below

Fig 7.20 7.21 and 7.22 show the plots for detector junction 2. In Fig 7.20 the sharp peak observed at the energy gap Δ for the injector current has mellowed down and only a small kink is visible. This is because the injected quasiparticles have to a very large extent have lost much of their energy due to diffusion. Those that tunnel into the detectors further away from the injector do not contribute much to the branch imbalance. Another reason is that the wire being diffusive in nature, this branch imbalance decays off with time and the time taken for the quasiparticles to reach the detectors far away is much greater than

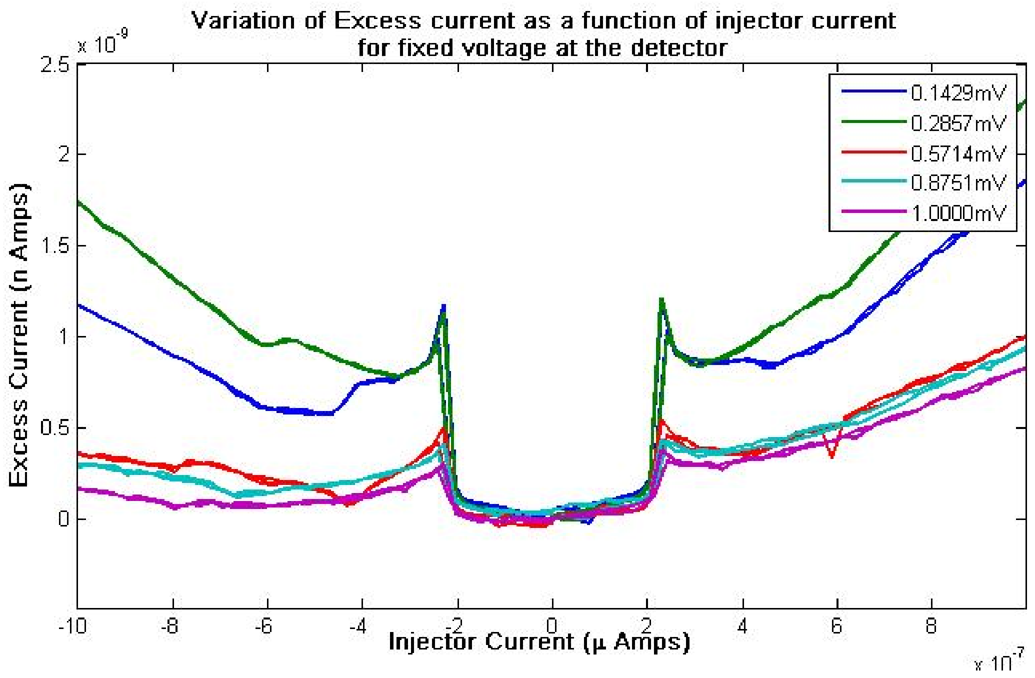


Figure 7.17: Plot showing the excess current as a function of injector current for different voltages at detector junction 1

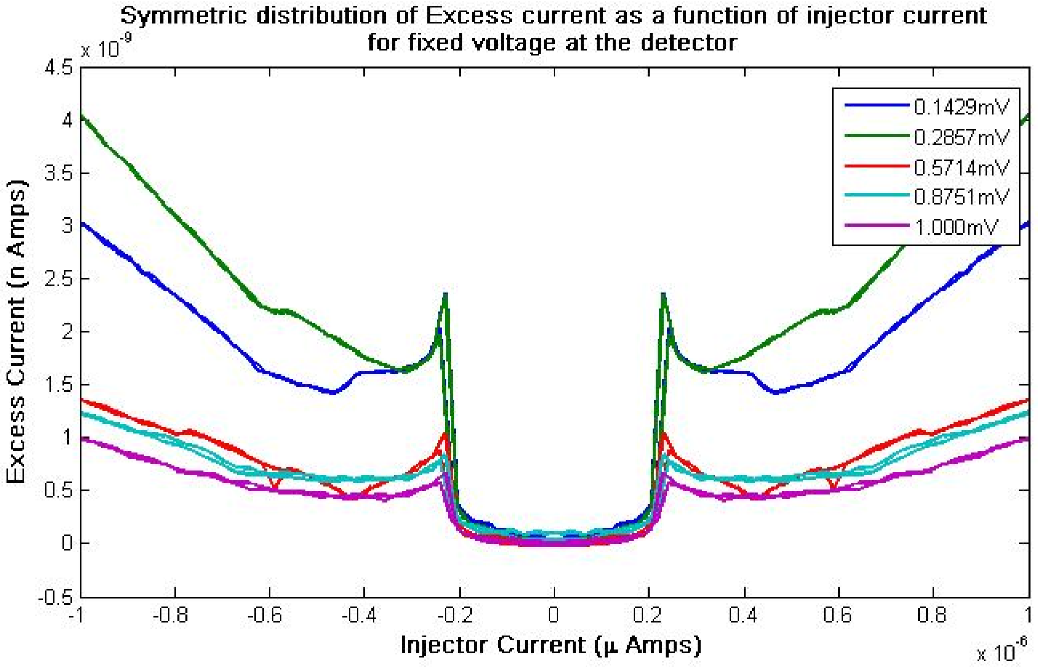


Figure 7.18: Plot showing the symmetric part of the excess current for junction 1 which is independent of the injector current

the time it takes for the branch imbalance to decay off. Lastly we take a look at the characteristics of junction 2 as shown in Fig 7.23 7.24 and 7.25 below:

Fig 7.26 shows the variation of the asymmetirc current with the distance for a fixed

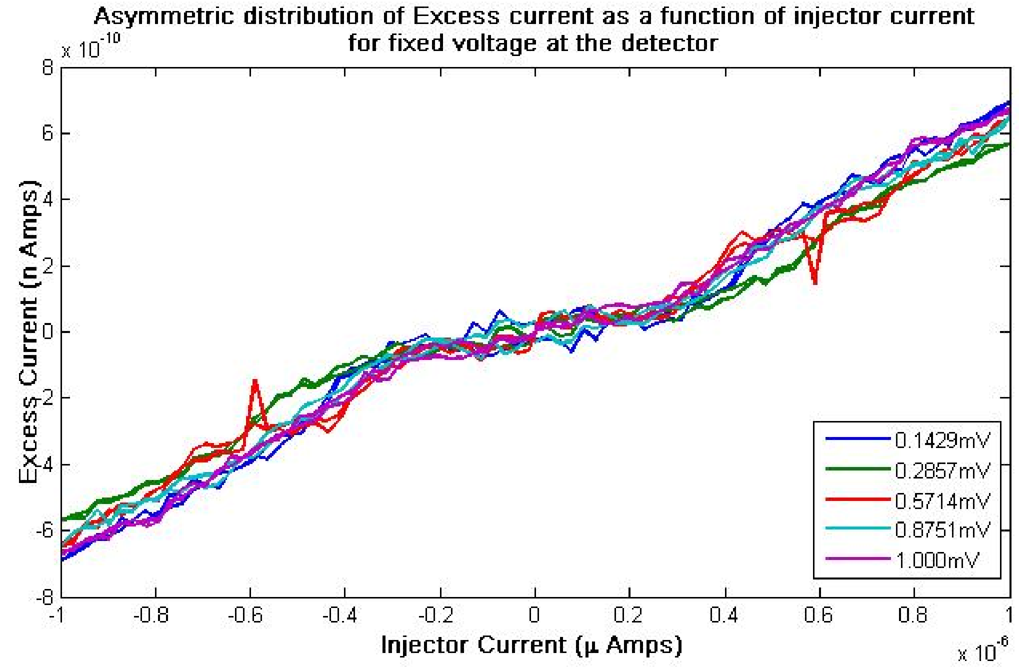


Figure 7.19: Plot showing the asymmetric part of the excess current for junction 1which is independent within the Δ region but starts becoming heavily dependent upon injector current

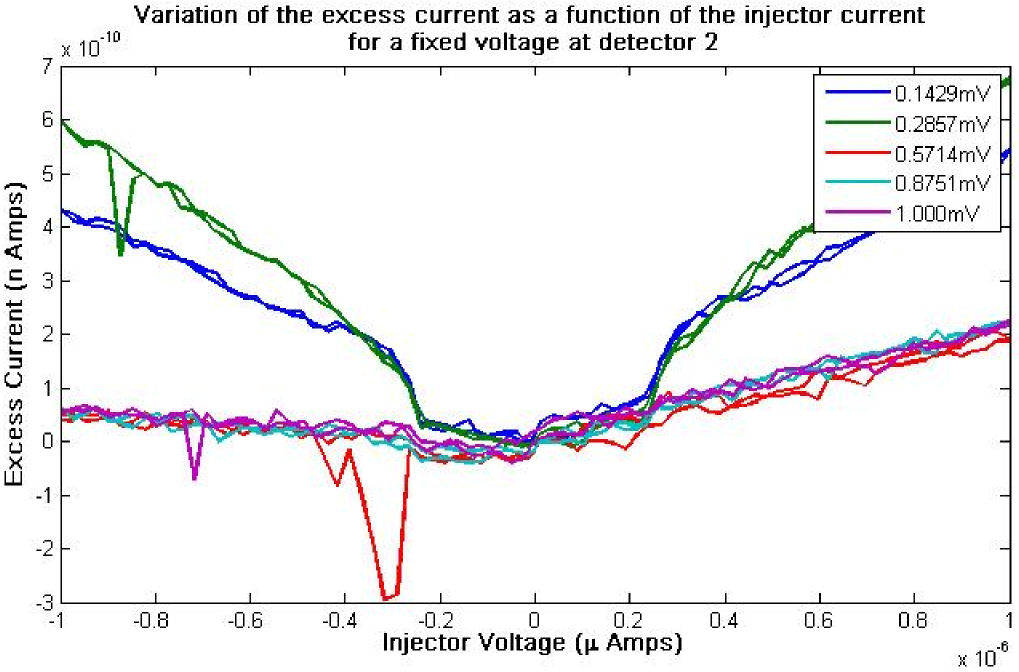


Figure 7.20: Plot showing the excess current as a function of injector current for different voltages at detector junction 2

bias at the detector.

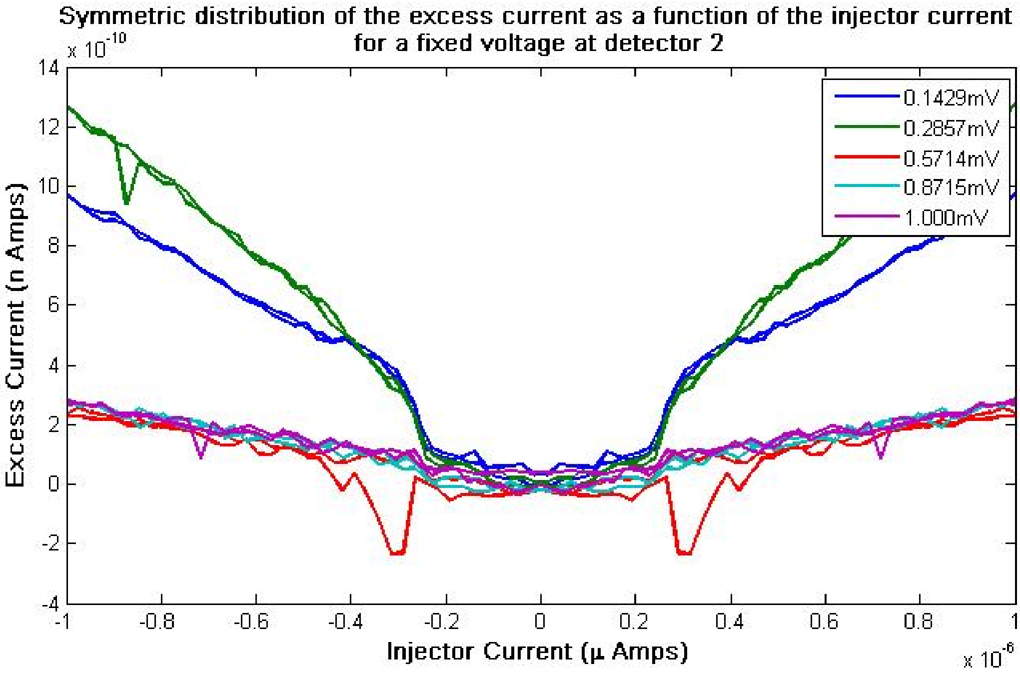


Figure 7.21: Plot showing the symmetric part of the excess current for junction 2 which is independent of the injector current

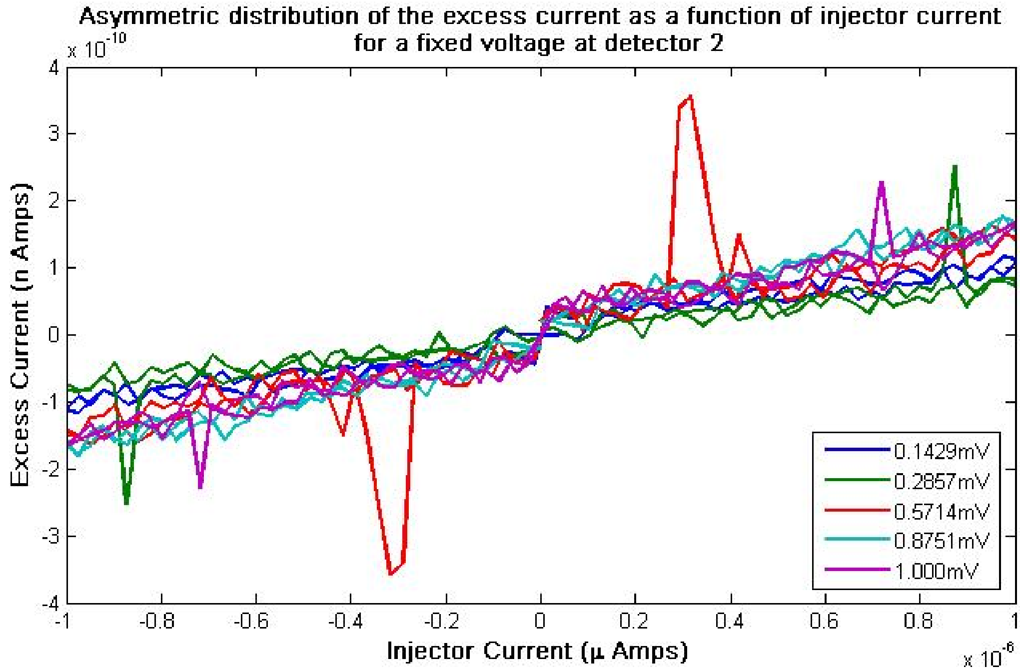


Figure 7.22: Plot showing the asymmetric part of the excess current for junction 2. Asymmetric part is dependent on injector current

The characteristics look very similar to junction 2. One similarity seen in all 3 junctions is that the excess current is maximum when the the voltage applied at the detector is close to the energy gap. This is due to the alignment of the filled and empty DOS's on

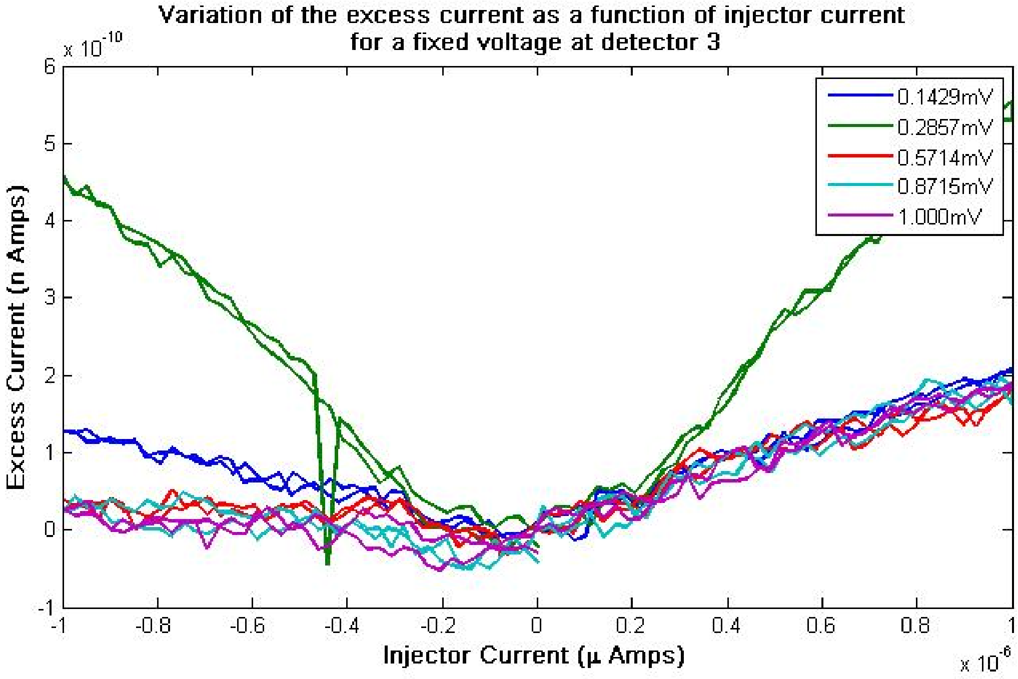


Figure 7.23: Plot showing the excess current as a function of injector current for different voltages at detector junction 3

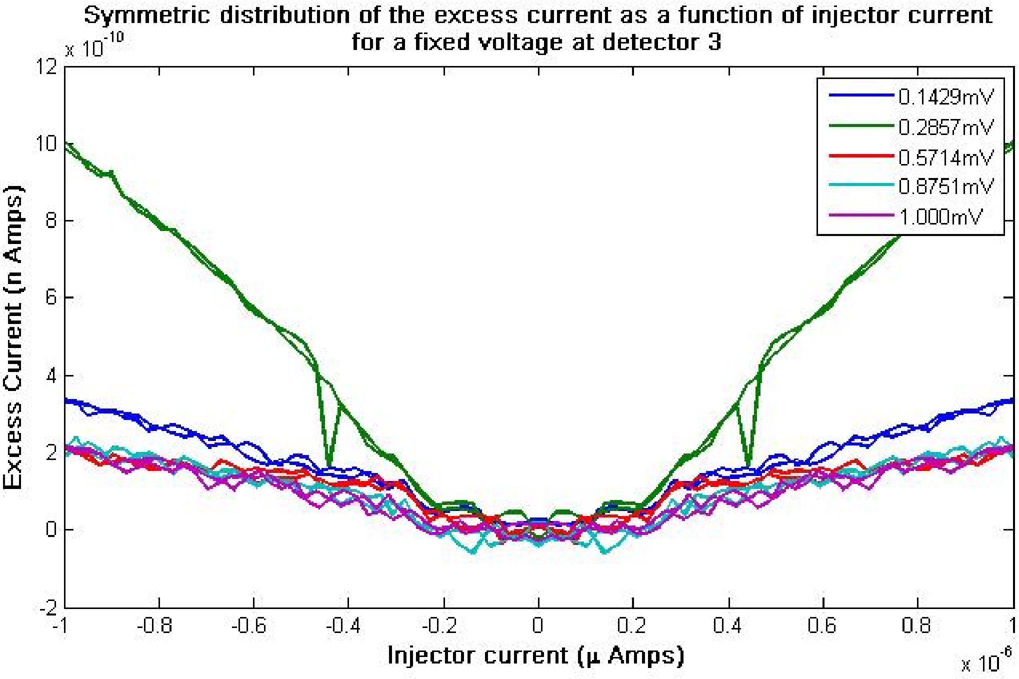


Figure 7.24: Plot showing the symmetric part of the excess current for junction 3 which is independent of the injector current

either side of the junction.

Calculation for Charge Imbalance is presented as below. The injector creates a quasiparticle imbalance in a volume Ω, this causes a current to flow through the wire which is

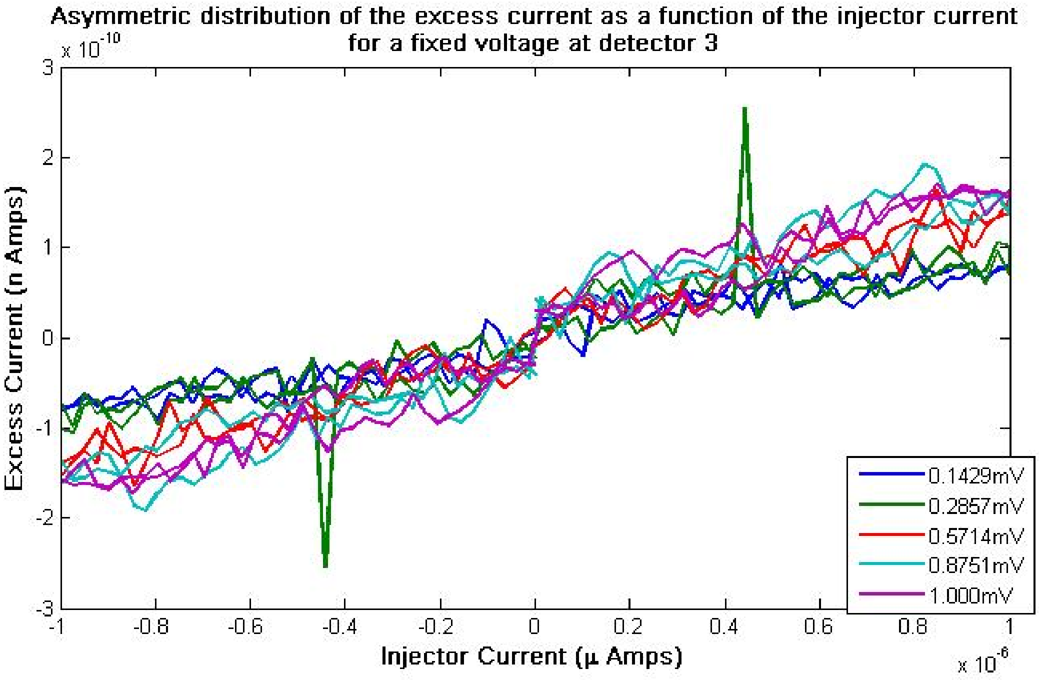


Figure 7.25: Plot showing the asymmetric part of the excess current for junction 3. Asymmetric part is dependent on injector current

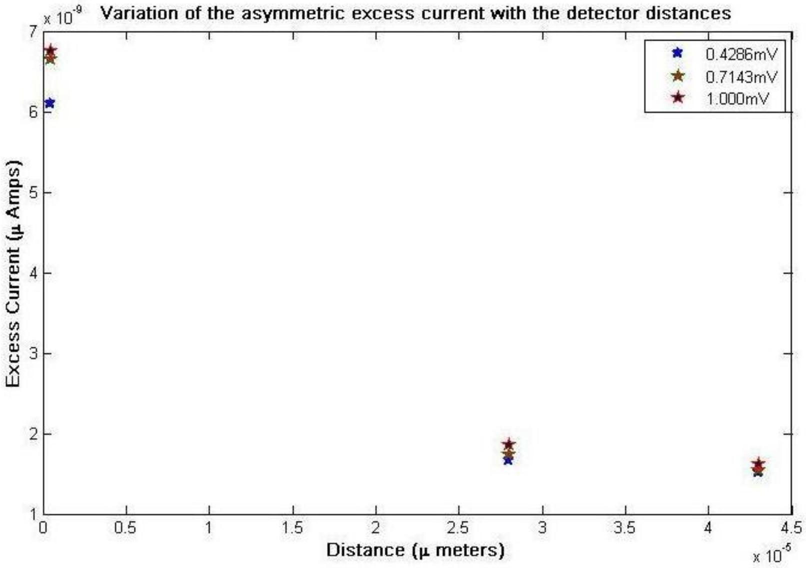


Figure 7.26: Plot showing the variation of the asymmetric part of the excess current for the three detector junctions

measured at the detector junction which is equivalent to

$$I_d(V_{bias}=0)=\frac{G_{nn}}{e}\int_0^\infty (f_{k<}-f_{k>})dE=-\frac{G_{NN}Q^*}{2N(0)e} \tag{7.3}$$

where:

$$Q^* = 2N(0)\int_{\Delta}^{\infty}(f_{k<} - f_{k>})dE \tag{7.4}$$

and G_{NN} is the normal state resistivity of the detector junction. [156]

The experimental wire in our case can be considered as a diffusive one dimensional system. Assuming constant relaxation of the injected quasiparticles, the charge imbalance Q^* can be modelled using a diffusive equation of the form

$$D\frac{d^2Q^*}{dx^2} - \frac{Q^*}{\tau_{Q^*}} = 0 \tag{7.5}$$

where:
D is the diffusion constant and τ_{Q^*} is the charge relaxation time. As seen in Fig 5.5 the injector is at one end of the wire. Hence all charge injected flows only towards the other end. If the point where this injection happens is taken as $x = 0$ at one end of the wire, and the imbalance Q^* is measured at different detectors at distances x from the injector, the solution to the equation **??** is given by

$$Q^* = (\frac{I_{inj}F}{DS})exp^{\frac{-x}{\lambda_{Q^*}}} \tag{7.6}$$

where S is the cross sectional area of the wire, F is the value from Eq: 6.4. From the limiting conditions F equals 1 in our case. λ_{Q^*} is the relaxation length which is given by $\sqrt{D\tau_{Q^*}}$.

τ_{Q^*} is determined by the relation

$$\tau_{Q^*} = \frac{4k_BT_c}{\pi\Delta(T)} \tag{7.7}$$

With the above parameters, the excess current due to the imbalance is determined by

$$\delta I = \frac{G_{nn}Q^*}{2eN(0)} \tag{7.8}$$

From the above observations regarding the charge imbalance we come to the following conclusions

- The symmetric part of the excess current is dependent of the injector voltage
- The large variations in the symmetric part of the excess current in junction 1 as compared to the other junctions are attributed to the fact that the gap of detector is most influenced by the injector current
- The peak that is seen in the excess current which gradually vanishes with the other junctions as seen in Fig's 7.18 7.21 7.24 is not clearly understood. it can be due to the alignment of the filled and empty DOS on either site of the tunnel junction. It can also due to the diffusion of the quasiparticles. More work is needed to take a deeper look into this.

- The asymmetric part remains almost constant within the gap and then it depends in an almost linear manner with the injection current
- The increase in this asymmetric current is reduced with increase in the distance from the injector due to the energy relaxation the quasiparticles undergo while diffusing through the wire.

Chapter 8

Thermal Management (Cooling) Techniques and Solutions of Solid State Devices

Temperatures and temperature gradients in Microelectronic VLSI circuits, packages and systems play an important role in the functionality, performance and reliability of electronic circuits and products. Since the electrical properties of these devices and products are temperature dependent, a higher operating temperature can lead to thermal drift and performance degradation. Variation of the electrical resistance of the device or the package is an important issue which normally increases with the temperature. The heat due to this, in the device is directly proportional to the change in the electrical resistance due to the combination of increasing resistance and heat production often leads to thermal runaway of the device unless there are mechanisms to control this. High temperatures that cause material degradation leading to failure of the device or its surroundings, for example in electronic packaging the solder joints or the printed circuit board degrade due to heat in integrated circuits, vias that connect the metal layers are the most vulnerable parts in this respect. Another source of failure due to heating can be attributed to Thermal cycling, heating and cooling down during use. These issues lead to fatigue and stress in materials. In particular solder joints are very sensitive for this type of thermal loading and the via edge are prone to electro migration. As demand on performance and reliability increases more and more, combined with an increase in power dissipation and ever growing miniaturization of components, thermal management of electronic devices and circuit components is fast becoming the bête-noire of the electronics industry. It is well known that electronic devices on the micro and nanoscale can generate thousands of watts of heat flux per square centimeter. Today's IC's in the sub nanometer range have several million transistors and the activity of these in different functional units of an integrated circuit create hot spots which although being only a few micrometers in area can be tens of degrees hotter than the rest of the IC. Since thermal design requirements are primarily driven by worst case operating temperature regions, reducing or eliminating these hotspots can help thermal management of not only the integrated circuit but the entire package.

The combination of high power (density) and requirement of relatively low operating temperatures create an important task for the thermal management of electronic components.

Thermal management is usually distinguished in three levels as below:

- Device Level
- Board Level
- System Level

The "device level" is the primary area of interest for package designers. In this level the focus is on the internal conduction heat transport from the device to package boundary. The source of the heat here is the device and since miniaturization makes these devices small, the local heat flux densities are very high which can lead to high temperature gradients. For this reason the aim is to create sufficient heat spreading inside the package and limit the thermal resistance of the heat path from junction to boundary.

The next two levels are the working areas of the system designers. The division in board level and system level has its origin in the content of detail in the analysis. Board level is typically dealing with the interaction between packages, board and the direct environment. System level analysis deals with the overall heat transfer in the system, for instance the distribution of the airflow and differences in air temperatures. System level analysis gives the results of the local thermal conditions for a particular board. Practically it is not feasible to perform analysis of an entire system with detailed thermal models of the boards. As shown in Fig.8.1, the heat flows in an "exposed pad package on a board" are shown. The die is the area of an electronic package where the heat is usually produced and it is transferred through various paths to the surrounding environment. Conduction and internal heat spreading mechanisms are utilized to transfer the heat through die attach layer. Other heat transfer paths are conduction through the body material to the package surface and conduction through bond wires and leads to the board. Heat

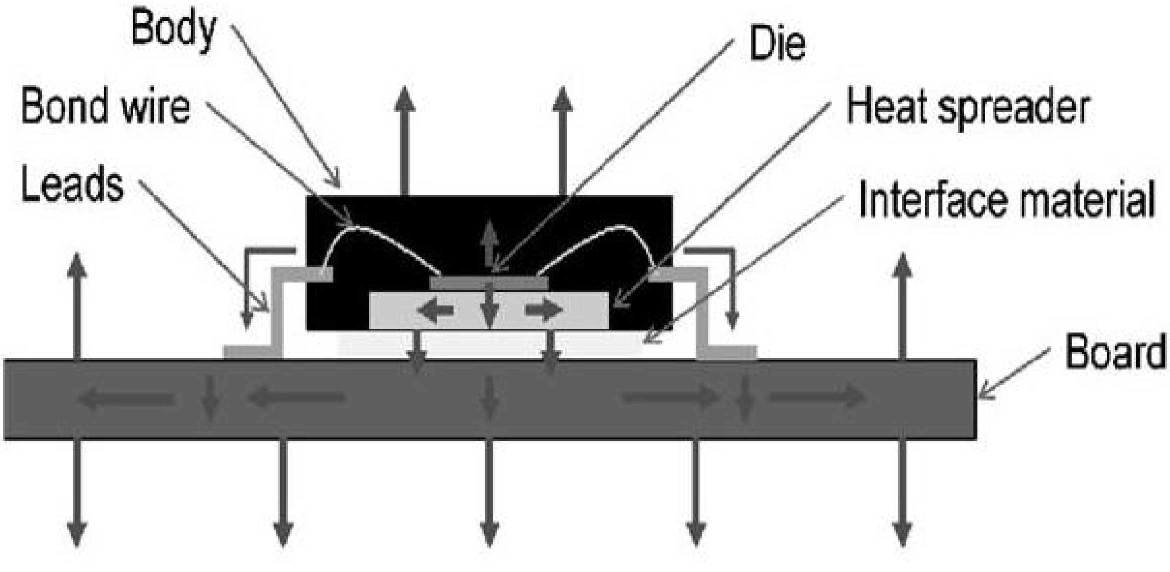


Figure 8.1: Typical heat flows in an IC chip with pad and package exposed [28]

from the heat spreader is transferred to the board across the gap between package and board. This gap can either be filled with air or gap filler material, which has a higher thermal conductivity than air. Inside the board the heat is conducted in two directions: thickness and in-plane. The construction of the board (layer structure and materials) has a strong influence on the heat spreading inside the board. At the board surface the heat is transferred to the environment by means of convection and radiation.

Fig 8.1 shows the schematic of the basic principles in the thermal management of microelectronic devices including their packaging using heat spreading. Heat spreading reduces the heat flux density across interfaces, for instance the gap between package and board, is reduced and thus the temperature rise introduced by the interface is lower. This also implies that the properties of the interface (thermal conductivity and thickness) play an important role. Thermal management in Microelectronics VLSI circuits and packaging is a full fledged area of research and development. A vast number of varied techniques are available that address the issue and provide a myriad of solutions for thermal management each with its own advantages and disadvantages. Depending on the target platform and the application in interest, one or many of such solutions are put in practice. For example, the techniques used to cool an 8000 processor based computer that takes up several rooms in a facility is much different than those used on a notebook computer or an embedded device. Thermal management mechanisms embedded at the integrated circuit level, chip packaging level, package mounted peripherals or board level and system cooling mechanisms are used depending on the power dissipation of the system. The target of this report is to address the current and existing thermal management techniques for microelectronic circuits and packaging.

Material science engineering is the key that provides desirable properties of materials at nanoscale for future improvements in IC thermal management. Nanoscale research at atomic, molecular and supra-molecular length scales of approximately 1-100 nm range, in order to understand and create materials, devices and systems with fundamentally new properties and functions due to their small structure. Miniaturization has been going on in the microelectronic industry for decades and most recently in the field of MEMS. Faster systems with enhanced performance, less space, material, and energy utilization can be realized. In addition, in the nanometer domain unique effects due to quantum phenomena and enhanced surface to volume ratios arise, where existing conventional heat and thermal management solutions are not effective Sheu[30] and Roukes [29]. In this report the various materials that are being used in today's microelectronic industry for IC and package cooling are discussed. Following it, a brief review of nanoscale thermal and thermoelectric transport is given and then active solid-state devices for hot spot removal are treated. Cooling techniques which can be broadly classified as active cooling and passive cooling is discussed. Emerging cooling techniques that have come up due to advancement of technologies are then touched upon. Finally the challenges and limitations of the above thermal management solutions are presented with potential scope for improvement in the long term.

8.1 Materials - Thermal Conductivity

This section presents materials that have been traditionally in use since the earliest advent of IC industry and the latest discoveries of materials that are in use with the special property that they have Ultrahigh Thermal Conductivity.

8.1.1 Traditional Materials

The Microelectronic industry is arguably the most materials intensive Industry. The families of traditional materials included in an integrated circuit and its package include: semiconductors, ceramics, glasses, composites, polymers, and metals. A list of the types of materials used in coming up with a fully functional IC are listed in Table 8.1 Metals are used as conductors and interconnects. Their primary function is electrical but for power devices they also serve thermal solutions. Metals are also used to act as heat sinks for power devices and in RF applications they act as shields. Ceramics and glasses are

Semiconductors	Si, SiGe, GaAs
Metals	Solders for interconnects (Sn-Pb, Sn-Ag, Sn-Ag-Cu, Sn-Au, Sn-Sb) Au wire bonds Cu lead frames (Kovar, CuBe, Alloy 42) Cu traces in substrates W, Mo traces in co-fired ceramics Ag, Au, Pd for thin/thick films on ceramics Ni diffusion barrier metallization Al heat sinks
Ceramics	Al2O3 substrate modified with BaO, SiO2, CuO, etc. LTCC substrates Al2O3 modified with low-temperature glass (e.g., PbO) SiN dielectrics Diamond heat sinks
Polymers	Epoxies (overmold) Filled epoxies (overmold) Silica-filled anhydride resin (underfills) Conductive adhesives (die bonding, interconnects) Laminated epoxy/glass substrates Polyimide dielectric Benzoyclobutene Silicones Photosensitive polymers for photo mask (acrylates, monomers, etc.)
Glasses	SiO2 fibres for optoelectronics Silicate glasses for sealing Borosilicate glass substrates Glass fibres for epoxy/glass substrates (FR-4)

Table 8.1: Materials used for Thermal management in Integrated Circuits and Packaging [31]

used mainly as dielectrics or insulators. Ceramics are used in devices as dielectrics to form capacitors and inductors. In the package, ceramics are used as insulating materials for substrates that provide a structural base that electrically isolates lines and pads. Polymers are used as insulators. In composite form they are used as conductors. As an insulator, polymers function as encapsulants, under fills and substrates. Polymers are also used as insulating adhesives to glue components to a substrate or board to provide mechanical strength. Composite materials are a mix of materials that can be tailored for either mechanical behavior improvements or thermal enhancement as an electrical conductor. Many of the composite materials used in packages are based on a polymer matrix. One of the most important properties of a material in relation to thermal management is its Coefficient of Thermal Expansion (CTE). The CTE is the length of increase of a sample for a given temperature increase. A wide variety of materials with a wide variety of CTEs, such as metals, ceramics, and polymers are joined together. As the temperatures change during processing or in use, the materials expand or contract to various degrees that could result in the formation of extensive and/or nonuniform strains in the integrated circuit or its package assembly. A set of materials typically used in and their CTE's are shown in Table 8.2

8.1.2 Ultrahigh Thermal Conductivity Materials

Thermal management in semiconductor IC industry is universally recognized as a complex issue, but it is clear that materials used are required to have very high thermal conductivities and low CTE's to minimize thermal stresses that affect reliability and performance. Traditional low-CTE materials, which are decades old, have thermal conductivities that are little or no better than that of aluminium (about 200 W/m-K). In addition to thermal requirements, electromagnetic shielding and emission control are also major concerns. Materials are required with both high and low electrical resistivities. In response to these needs, an increasing number of ultrahigh-thermal-conductivity advanced materials have been, and are being developed that offer significant improvements.

Advantaged Include:

- Thermal conductivities ranging up to more than four times that of copper.
- CTEs those are tailorable from -2 to +60 ppm/K.
- Electrical resistivities ranging from very low to very high.
- Extremely high strengths, stiffness and low densities.
- Low-cost, net-shape fabrication processes.

The payoffs are"

- Improved and simplified thermal design.
- Possible elimination of heat pipes and fans.
- Reduced power consumption, thermal stress and warpage.

Material	**CTE (10-6/○ C**
Metals	
304 Stainless Steel	17.8
Ag	19.7
Al	23.5
Alloy 42	4.9
Au	14.2
Cu	16.8
Invar	1.6
Kovar	5.5
Mo	5.1
Ni	13-15
Solder: 63Sn-37Pb	25
Solder: 95Pb-5Sn	28
Ti	10
W	4.5
Ceramics	
AlN	4.3
Alumina (96%)	6.4
Alumina (99.5%)	6.5
BeO	7.8
BN	3.7
Fused silica glass	0.56
Quartz	13
SiC 3	3.8
SiN	3
Semiconductors	
GaAs	5.8
Si	2.7
Organic Materials	
Epoxy resins	50-80
FR-4 (x-y plane)	15.8
FR-4 (z-axis)	80-90
Polycarbonates	50-70
Polyimide glass (x-y plane)	12-14
Polyimide glass (z-axis)	60
Polyimides	40-50
Polyurethanes	180-250
RTV (room temperature vulcanized) polymer	800
Sylgard	300

Table 8.2: Coefficient of Thermal Expansion for materials used in Integrated Circuits and Packaging
[31]

- CTE match allows direct solder attach with hard solders.
- Increased reliability and performance.
- Weight savings up to 90% and size reduction up to 65%.
- Increased manufacturing yield and considerable cost reductions

Over the decades demands from microelectronic and optoelectronic industry has surged material suppliers and integrated circuit and packaging manufacturers to continue to develop, numerous advanced monolithic and composite materials. The last few years have seen revolutionary advances in this domain. To present, these materials are broken down into two categories: those having thermal conductivities of at least 400 W/m-K, which are designated as ultrahigh thermal conductivity, table 8.3, and those having greater than or equal to 300 W/m-K, and less than 400 W/m-K, which are termed as high thermal conductivity, table 8.4.

All the above materials have a very low CTE. It is clear that these advanced materials offer dramatic improvements in properties compared to those of aluminium, copper and copper/tungsten composites. The advantages are particularly striking when density is considered. There is a rule of thumb in the history of technology, which says that when a critical parameter is increased by an order of magnitude, it has a revolutionary effect [32]. In the IC industry with regards to thermal management this is now being seen.

The above advanced packaging materials fall into five main categories:

- Monolithic carbonaceous materials (MCM's)
  Pure carbon is a remarkable material. Carbonaceous materials range from graphite lubricants to diamonds to high-performance carbon fibres. There are a number of monolithic carbonaceous packaging materials of interest. Diamond films and naturally occurring graphite have been used in production applications and as thermal interface materials
- Metal matrix composites (MMCs)
  Silicon-carbide-particle-reinforced aluminium (Al/SiC) is probably the most widely used advanced MMC packaging material. Al/SiC microwave packages and solid and flow through PCB heat sinks are used in numerous applications. The major limitation of Al/SiC materials is that their thermal conductivities are no better than those of aluminium alloys.
- Carbon/carbon composites (CCCs)
  CCCs consist of carbonaceous matrices reinforced with carbon fibres. Although they cannot be considered as high performance materials, they are stronger, stiffer and less brittle than monolithic carbon. Some CCCs have high thermal conductivities. CCCs have been used in a limited number of production thermal management applications.
- Ceramic matrix composites (CMCs)
  IBM's revolutionary disclosure that diamond-particle reinforced silicon carbide (diamond/SiC) heat spreaders were being used in their servers was an historic milestone. Diamond/Sic has a low CTE, low density and a thermal conductivity that is 50

Reinforcement	**Matrix**	**Thermal Conductivity (W/m-K)**	**CTE (ppm/K)**	**Specific Gravity**	**Specific Thermal Conductivity (W/m-K)**
-	CVD Diamond	1100-1800	1-2	3.52	310-510
-	HOPG	1300-1700	-1.0	2.3	740-850
-	Natural Graphite	150-500	-1	-	-
"Thermal Graph"	-	700-800	-0.5	1.8	390-440
Cont. Carbon Fibres	Copper	400-420	0.5-1.6	5.3-8.2	49-79
Cont. Carbon Fibres	Carbon	400	-1.0	1.9	210
Graphite Flake	Aluminium	400-600	4.5-5.0	2.3	174-260
Diamond Particles	Aluminium	550-600	7.0-7.5	3.1	177-194
Diamond and SiC Particles	Aluminium	575	5.5	-	-
Diamond Particles	Copper	600-1200	5.8	5.9	330-670
Diamond Particles	Cobalt	600	3.0	4.12	145
Diamond Particles	Magnesium	500	8	-	-
Diamond Particles	Silver	400-600	5.8	5.8	69-103
Diamond Particles	Silicon	525	4.5	-	-
Diamond Particles	SiC	600	1.8	3.3	182

Table 8.3: Ultra high thermal conductivity and low CTE

[32]

- Polymer matrix composites (PMCs)
  Thermal conductivities of PMCs are not as high as those of other advanced materials, yet they have a number of advantages like ease of fabrication, low density, and relatively low cost. A major advantage of PMCs reinforced with short discontinuous fibres is that they can be formed into complex parts by injection molding. Although thermal conductivities are a lower than those of PMCs using continuous fibres, they are adequate for many applications. They are now being used as integrated circuit pin fin heat sinks and heat spreaders. PMCs reinforced with discontinuous carbon fibres have much higher electrical resistivities than metals, reducing electromagnetic radiation. This has led to their use in video chip heat sinks.

Reinforcement	**Matrix**	**Thermal Conductivity (W/m-K)**	**CTE (ppm/K)**	**Specific Gravity**	**Specific Thermal Conductivity (W/m-K)**
-	Aluminium	218	23	2.7	81
-	Copper	400	17	8.9	45
Copper	Tungsten	157-190	5.7-8.3	15-17	9-13
Natural Graphite	Epoxy	370	-2.4	1.94	190
Cont. Carbon Fibres	Polymer	330	-1	1.8	183
Disc. Carbon Fibres	Copper	300	6.5-9.5	6.8	44
SiC Particles	Copper	320	7-10.9	6.6	48
Cont. Carbon Fibres	SiC	370	2.5	2.2	170

Table 8.4: High Thermal Conductivity and low CTE

[32]

8.2 Thermal Properties at Nanoscale

Research in the thermal properties at nanoscale is quite young as compared to research in nanoscale electronic properties. This is because the technology to measure, model, fabricate devices at micro- and nanoscale for understanding their thermal behavior accurately is quite challenging [33]. Only recently have technological advancements been made that allow measurements of temperature and heat flow at the atomic scale. Scanning Thermal Microscopes (SThM) that contain small thermocouples at their tips have been successfully used to measure temperatures in a spatial resolution of around 100nm [34][7]. Other techniques like Scanning Seebeck voltage measurements are able to measure with accuracy of 2-3 nm spatial resolution by using heated metallic tips in ultra high vacuum [35].

Another recent advancement in thermal measurement is the use of Thermo-reflectance imaging using visible wavelengths. This technique has been used to measure the surface temperature of IC chips with submicron spatial resolution in the sub micron range and in temperature ranges of 1-100mK[36]-[38]. In nanoscale regime, heat transfer mechanisms differ significantly from the mechanisms that govern at the micro and bulk scale. With miniaturization, device dimensions are comparable to the electron mean free path and the Fermi wavelength. At these small scales classical laws are no longer valid. Quantum effects come into play and the wave nature of the electron has to be taken into account to measure, model and predict the heat transfer. To understand the physics that govern this regime and the difference with bulk heat conduction, the concept of phonons is introduced.

8.2.1 Phonons

In physics, a phonon is a quantized mode of vibration occurring in a rigid crystal lattice, such as the atomic lattice of a solid. The study of phonons is an important part of solid state physics, because phonons play a major role in many of the physical properties of solids, including a material's thermal and electrical conductivities. Phonons are a quantum mechanical version of a special type of vibration motion, known as normal modes in classical mechanics, in which each part of a lattice oscillates with the same frequency. Depending on the type of these vibrations, longitudinal or transverse modes result. Acoustic and Optical phonons result when the oscillations between neighboring atoms are in phase or out of phase with each other. Electron transport generates optical phonons in devices.

Optical phonons have low group velocities; hence they decay to acoustic phonons so that heat is transported. When the temperature of one area of the sample is hot, the interaction between neighboring atoms causes the random vibrations to propagate. In order to explain how the phonons are transported, these random vibrations are written in terms of the normal modes. The bulk parameter of thermal conductivity describes heat conduction which is of great importance in thermal management. The parabolic heat conduction equation which assumes instantaneous effect does not provide the correct result when treated in nanoscale regimes (spatial and time). Under conditions where different population of electrons and phonons are not in equilibrium, it is difficult to describe a single temperature. For example, Pop et al [39] has clearly shown that in short length transistors, the strong nonequilibrium between electron and phonon populations and between optical and acoustic phonons leads to submicrometer hot spots and heating in the electron gas.

Ballistic electron and phonon transport over nanometer distances also has a strong impact on heat generation in submicrometer electronic devices. At an interface between two solids, there could be an extra thermal boundary resistance (Kapitza resistance). In practice, there is no theory that can reliably predict heat conduction across boundaries and in nanometers-thick layers [40], [41]. Based on the roughness of the surface and the scattering, phonons are treated as particles and their transport across a boundary can be described using a specular or diffusive reflectance model. In some applications, they are treated as waves and their transport across an interface or a multilayer is described by the acoustic mismatch model. Various experimental results have been explained using these two theories [33], [42], [43].

Thermal conductivity is a bulk effect; hence it is difficult to differentiate between various contributions in order to gain an accurate understanding of heat transport in nanometer structures. In the case of electron transport, charge carriers interact with both electric and magnetic fields. Effects such as Hall voltage and Shubnikov-de Haas oscillations are used to measure electron effective mass, mobility, and number of free carriers. They are also used to differentiate between electrons from different bands in the solid. This has been extremely useful to study electron transport in nanometer-thick layers. Lack of similar complementary measurements in the case of phonon transport, and the added difficulty due to the fact that phonons are "quasi" particles, makes detailed analysis very difficult [45]. Very recently the phonon Hall effect under a magnetic field has been observed. This could provide a useful source of information about phonons and their transport in solids [44].

In case of bulk materials, interface effects are seen only at low temperatures where the mean free path is large. In nanometer scale, confinement comes into the picture and bulk properties are lost. In case of thin films and CNT's which can be considered as confinement in 2D and 1D respectively, phonon scattering affects thermal conductivity considerably.

Fig. 8.2 shows the thermal conductivity of silicon films with varying thickness. At sub-100-nm scales, a reduction of the conductivity by a factor of two to five when compared with bulk data is observed [45].

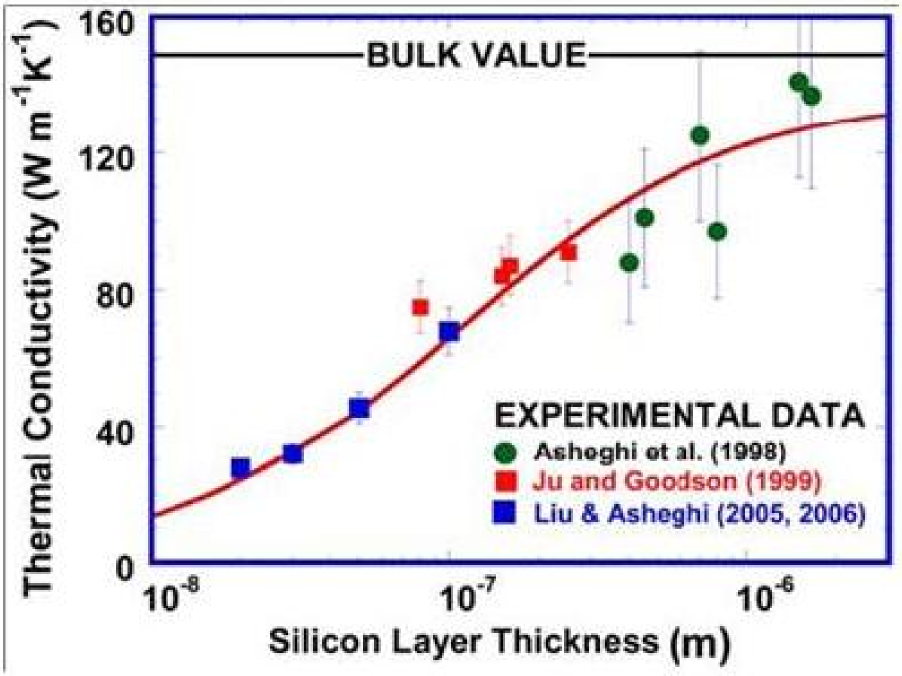


Figure 8.2: Variation of thermal conductivity with film thickness of Silicon thin films [45]

In 1D structures such as CNT's due to confinement, phonons behave much differently than in bulk materials due to the modification of the dispersion relation. In case of silicon nanowires, phonon scattering at the boundaries significantly reduces the thermal conductivity [49], but in case of CNT's which were made of rolled 2D graphite sheets, Mingo and Broido [50] showed that the initial theoretical estimations made by Berber et al [51] and Osman et al [52], were overestimated. At room temperature an upper bound of approximately $4e^9 W/m^2K$ was seen. Also observed was that the phonon transport remained ballistic over quite large distances. Experimental measurements done Yu et al [53] on thermal conductivity of CNT's have reported that the measured thermal conductance of SWCNT with length of 2.76μ is very close to the theoretical value. Experimental measurements on the thermal conductivity of MWCNT's have shown a T^2 dependence on temperature which suggest that MWCNT's behave as 2D solids [54], [55]. Due to their outstanding thermo mechanical properties, CNT's have found wide variety of applications like in Thermal Via's, Thermal Interface Materials, suspended in conductive fluids to improve conductivity etc.

8.3 Thermal Management-Techniques Solutions

Thermal management of ICs and microelectronic packaging is a significant issue because of ever increasing volumetric power densities and the harsh environments where these devices are employed. The increased power densities are the direct result of increasing functionality reducing feature sizes and implementing more processing to be packed into a small area or volume. Modern thermal management systems for Microelectronic devices have to consider die to package, package to board and board to enclosure heat flow. Resistance to this heat flow will lead to device heating with adversely affects performance. The heat resulting from higher power density is a critical issue in semiconductor industry. The International Technology Road-map for Semiconductors (ITRS) predicts that the peak power consumption of high performance electronic systems will jump by 18% (167 W to 198 W) in 2013 and by 51% (91 W to 137 W) in the low and medium performance systems in 2013 **If the planet earth survives 2012**. If cooling methods don't improve, increased die temperatures in ICs and packages will result in performance degradation, lower lifetimes and eventually failure. Uneven thermal distributions across the substrate become more pronounced due to increase in density and failures due to physical stress will increase. Active research is going on to improve chip and package cooling techniques to improve heat transfer, but most of these solutions do not adequately address the problems caused by a non-uniform thermal profile.

IC cooling techniques can be classified into the following 4 categories based on their implementation methodology, Paik et al [56]. The classification is:

8.3.1 Fans

Fan-based cooling is the oldest and the most widespread method used for cooling IC's and packages. This kind of a system used a heat spreader and forced convection. The heat sink usually absorbs heat and a fan coupled to it dissipates the heat. Air is utilized to dissipate the heat and such a system is quite simple.

8.3.2 MEMS

With the advent of advanced micro-fabrication technologies, MEMS-based cooling methods are quite innovative and a new concept. By developing complex structures on an underlying substrate, heat dissipation and conduction is promoted in a very efficient manner. MEMS can be designed to work with either liquid or air as cooling agents. Examples of MEMS based cooling methods include air-impingement methods, jet heat exchangers, and micro membranes. These methods work on the principle of injecting small streams of the coolant through the MEMS device onto the IC substrate's surface to facilitate heat conduction and dissipate heat. Another example of a MEMS-based device is a bulk micro-machined microfin array heat sink. The array consists of warped bimorph micro cantilevers that vibrate when an air jet impinges on them, creating vortices and dissipating heat faster than a plain-wall heat sink.

8.3.3 Refrigeration

This type of cooling technology involves vapor compression, gas compression, or thermoelectric devices. The only requirement of these systems is that the evaporator attached to the IC or the package has to have a contact temperature less than the air-cooling temperature. With this condition, these systems can generate a subzero effective thermal resistance. The result is a manifold increase in heat dissipation from the chip. Although compact vapor and gas compression systems exist, they are limited in their cooling efficiency by their size and hence the current focus is on solid state refrigeration methods which use thermoelectric coolers (TECs) to provide thermal resistance at or below $0\circ$C/W. TECs are based on the Peltier effect, results in a temperature differential when a DC current is applied across two dissimilar materials. In a typical thermoelectric system, the TECs lie between the heat source (the IC die or packaging) and the heat sink. Due to this arrangement, one side of the TEC is hot and the heat sink must dissipate power on the hot side. This results in a higher ambient temperature at the side of the heat sink and at a certain point is unable to dissipate heat which causes the cold side of the TEC element to heat up. Active research is underway to find possible ways to improve the design of heat sinks that are able to dissipate heat quickly.

8.3.4 Fluidics

Fluidics based cooling can be divided into two types: Macro fluidics and Micro fluidics. Macro fluidics-cooling methods work at the macroscale and can be classified into two categories: direct and indirect. In direct cooling the IC chip is completely immersed in an inert dielectric liquid. In indirect cooling there is a two phase flow. Indirect cooling includes thermosyphons and heat pipes.

The idea behind microfluidic cooling technique is that small volumes of liquid are pushed across the surface of the IC to conduct and finally dissipate the heat. Widespread and commercial implementation of this method has not yet been realized due to technological limitations. The first microfluidic cooling device was the microchannel heat sink implemented by Tuckerman and Pease of Stanford University in 1981 [57]. Microfluidic cooling requires supply of a cooling liquid with high pressures. Such a system is inherently unsuitable for compact embedded systems and hence commercialization of such a method has been quite difficult. Over the last couple of decades since their conception, microfluidic cooling has received considerable attention and a wide variety of innovative methods have been implemented. Some interesting implementations that deserve mention are:

- Pumpless loop with microchannel surfaces [58].
- Two-phase microchannel cooling system based on electro-osmotic pumping of liquids [59]
- A MEMS-based microcapillary pumped loop (micro-CPL) has been fabricated on a silicon wafer, integrating an evaporator, a condenser, a reservoir, and liquid lines

[60].

- Piezoelectric droplet generation [61].

There exist other implementations that are only mentioned briefly mainly due to their shortcomings in efficiency and technological difficulty in implementation. Among these are:

- Thermionic refrigerators are a class of refrigerators that are related to thermoelectric refrigerators, and also contain no moving parts. Although an area of active research practical devices have not been implemented yet with this principle. [147] - [64]
- Another class of refrigeration technique is what is known as electron field emission [65], but this concept is even less developed.
- Cooling methods involving the Joule-Thompson refrigerators require the development of an efficient, miniature compressor capable of delivering particularly high pressure ratios, the unavailability of technology to implement such compressors has made cooling by this method not yet feasible.
- Finally, a class of refrigerators that utilize thermo-acoustic phenomenon, which is similar in principle to pulse tube refrigerators, are unattractive option in that the piston is replaced by an acoustic generator, which eliminates the moving parts [66]. However, only relatively small cooling powers have been achieved so far. The primary reason for this being that the acoustic properties required for efficient cooling are insufficient at the sub micron scale.

Fundamentally all of the above cooling methods can be divided into two major classes namely Active cooling methods and Passive cooling methods each having its own advantages, complexities and costs. The next sections will deal with the different variations of active and passive cooling methods in detail.

8.4 Passive Cooling

Thermal conduction, convection and radiation are passive cooling methods. Heat pipes and heat fins and thermosyphons also come under this category but they have much higher efficiency. Thermal interface materials (TIMs) also fall under passive cooling methods. Passive cooling methods are inexpensive and simple but they perform lot worse than their active counterparts. Passive cooling methods included in this report are:

- Thermal Interface Materials
- Heat Pipes
- Heat Sinks
- Vapor Chambers

8.4.1 Thermal Interface Materials

The primary role of Thermal interface materials (TIMs) is to thermally connect various components of the thermal system.

When two solid surfaces are brought in contact with each other the irregularities on the surface limit the actual contact of these two materials. Heat flow across such an interface involving solid-to-solid conduction depends upon the contact area A_c and noncontact area A_{nc} of the interface. This constriction to heat flow due to the non-contact area results in thermal contact resistance R_c at the interface. The solid-solid contact resistance R_{cs} between two nominally flat surfaces 1 and 2 assuming plastic deformation of the surface irregularities is given by [69].

$$R_{cs} = \frac{0.8\sigma}{mk_h}(\frac{H}{P})^{0.95} \tag{8.1}$$

where $\sigma = (\sigma_1^2 + \sigma_2^2)^{0.5}$, σ is the root mean square of roughness $m = (m_1^2 + m_2^2)^{0.5}$, m is mean slope of the irregularity H is the hardness and P is the applied pressure. $k_h = 2k_1k2/(k_1 + k_2)$ is the harmonic mean of the thermal conductivity of the interface. m is the measure of the slope and is given by $\tan(\theta)$ where θ is the slope. Equation 8.1 assumes that at all pressures there is plastic deformation of the surfaces in contact. Fig. 8.3 shows two typical packaging architectures used in microelectronics package cooling. It is seen that the chip is flat when in reality it is far from being flat. Due the coefficient of thermal expansion (CTE) mismatch between the die and the package substrate the chip is usually warped. A schematic of a warped chip is shown in Fig. 8.4

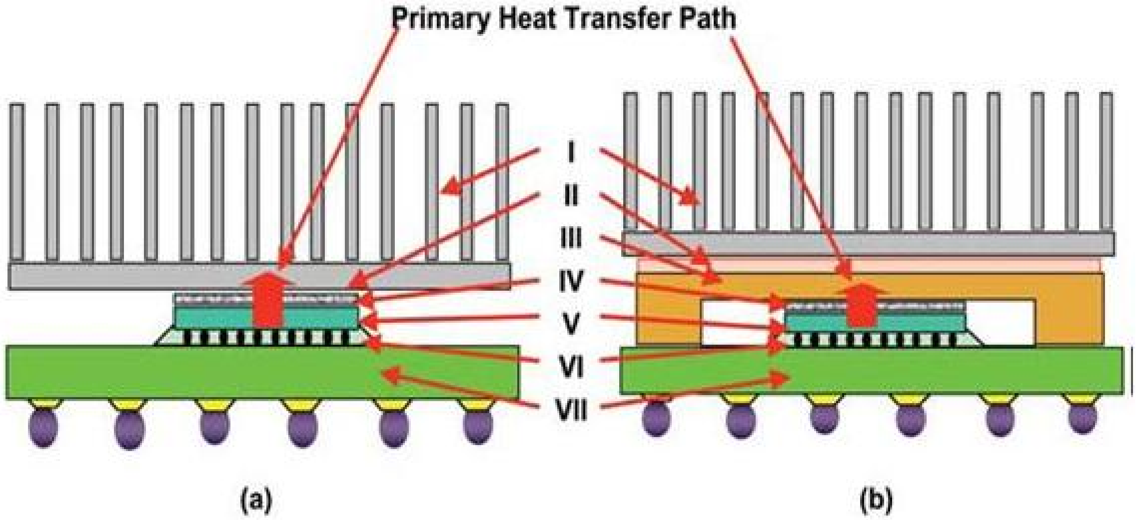


Figure 8.3: Typical packaging applications I-heat sink; II-TIM; III-IHS; IV-TIM; V-die; IV-underfill; and VII-package substrate
[69]

Verma et al [75] measured the warpage and their results showed that warpage has a heavy dependence on temperature. In addition, warpage also depends on many other factors like, geometry of the substrate, CTE mismatch between the substrate and the die, and

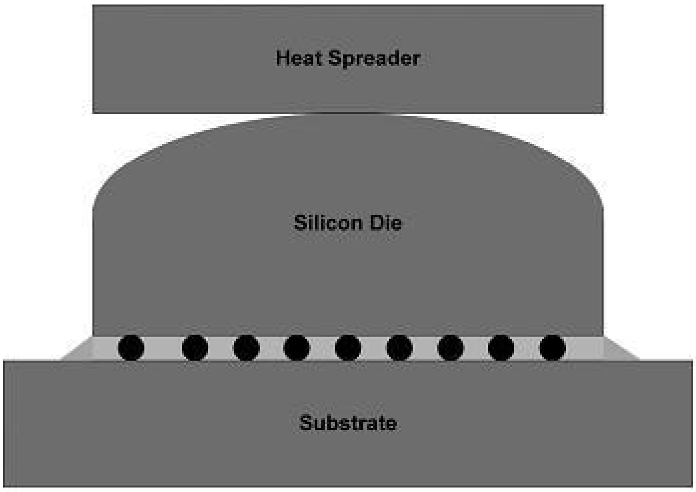


Figure 8.4: Schematic showing a warped die
[69]

the temperature as mentioned above. Due to surface irregularities and the warpage of the die, the contact area will be much less when compared to a flat die. This causes an increase in the resistance, increasing R_c [68] when compared to the value calculated from equation 8.1. Hence to reduce R_c, the gaps caused due to the surface irregularities have to be filled by some material. These materials are highly conductive and are known as Thermal Interface Materials (TIM's). When a TIM is used to fill the gap, equation 8.1 changes as below:

$$R_{cTIM} = \frac{1.53\sigma}{k_{TIM}}(\frac{H}{P})^{0.097} \tag{8.2}$$

[69]

TIM's fill in-between the two materials as shows in Fig 8.5 [69]. Since TIM's have a finite BLT, they do not completely fill the voids because they are not able to wet the surface fully. Due to this inability, the thermal performance of a TIM does not meet the value as obtained from equation 2. Real TIMs looks like that shown in Fig. 8.5(b). This figure shows that real TIMs have finite BLT, and at the interface they do not completely fill the voids because of their inability to completely wet the surface. Therefore, the thermal performance of real TIMs falls short of that predicted by equation 8.2. From Fig. 8.5(b) it can be inferred that the total thermal resistance (RTIM) of real TIMs can be written as:

$$R_{TIM} = \frac{BLT}{k_{TIM}} + R_{c1} + R_{c2} \tag{8.3}$$

[69] where R_c represents the contact resistances of the TIM with the two bounding surfaces. BLT is Bond Line Thickness.

Due to the exponential increase in IC density there is a great interest to reduce the RTIM primarily due to one reason. It is known that the total power (Q) and the heat flux (q) will increase rapidly and the heat flux is nonuniform because the core and the cache are on the same die [70]. The core dissipates more than 90% of the total power which is a

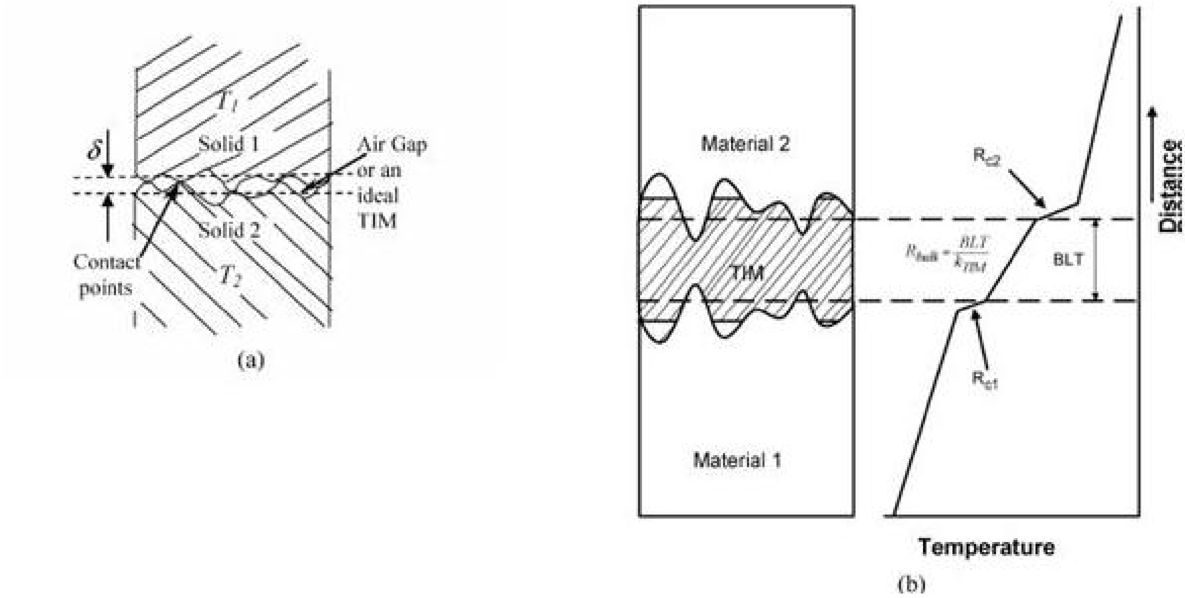


Figure 8.5: (a) Showing the real contact between two solids; (b) represents a real TIM. [69]

very small area in comparison to the complete area of the die. Maintaining an average operating temperature is not the problem with IC cooling methods, but maintaining the temperature of a hot spot in a particular area of the chip below design specifications is certainly a very difficult scenario and very hard to keep up with. When such a scenario is considered, Torresola et al suggest that the total thermal resistance can be written as [71].

$$\psi_{j-a} = DF \times R_{jc} + \psi_{cs} + \psi_{sa} \tag{8.4}$$

Where ψ_{j-a} is the junction to ambient thermal resistance, R_{jc} is the junction to case thermal impedance for a uniformly heated die, cs is the case to sink resistance, and sa is the sink to ambient thermal resistance. In equation 8.4.1, DF accounts for the nonuniformity of the heat flux *(q)* and the die size [71]. DF is measured in cm^2. Due to the nonuniformity of the heat flux, DF will be greater than 1.

Amongst many materials that are used as TIMs, CNTs and CNT based composites have received most attention due to the fact that they have extremely high thermal conductivity. The only limitation of CNTs is that, due to their small diameter, the impact of the thermal resistance between the polymer composite and the CNT will be very large. Nevertheless, Carbon nanotubes (CNTs) are the latest addition to thermal management solutions. CNTs are one dimensional structural materials with an outstanding property of ultrahigh thermal conductivity. A conventional heat dissipation device consists of the heat spreader, and a thin layer of a thermal interface material (TIM) between the heat spreader and the heat sink. The thermal performance of such an implementation is not very good because the thermal contact resistances of TIM with the coupling surfaces are relatively high and the thermal conductivity of conventional TIMs is low. CNTs on the other hand provide extremely high thermal conductivity when dispersed into polymer matrixes such as epoxy resins or other polymer matrix of the TIM. However, the thermal performance of a TIM with randomly dispersed CNTs in epoxy resins or other matrix materials happens to be much below expectation because of the highly anisotropic nature of the thermal conductivity of CNTs. Due to the random dispersion of the CNTs the total thermal contact resistance is high because there are numerous joints between CNTs, between CNT and matrix material. Aligned CNT is the answer. Numerous attempts in trying to come up with an approach to overcome the problems with CNTs have been carried out. Huang et al [76] prepared a prototype of TIM films using in-situ injection molding by embedding aligned CNT arrays in polymer matrices. The incompatibility of the phonon heat transfer modes in polymer and CNTs significantly limits the advantage of heat conduction by CNT. As a result, the thermal conductivity of thermal interface material with aligned CNT arrays in polymer matrix is only $1.21W/m-K$. Another problem is that the cured matrix material is less flexible to fill uneven surfaces of heat source and heat sink. Therefore the contact thermal resistance is much higher. In another attempt Hu et al [73] and Zhang et al [74] were able to form a thermal interface layer with CNT synthesized on the surface of semiconductor Silicon die. The processes are again not compatible for CNT synthesis and device fabrication. If CNTs are grown before device fabrication, the cleanliness of the wafer is decreased and to be able to protect the CNT while device fabrication will make it extremely difficult using normal processes and equipments. On the other hand, if the devices are fabricated before CNT growth, the high temperature required by for growing CNT will damage the device.

To solve these problems and make use of CNT, a novel heat spreader with aligned CNT arrays synthesized on both sides at one fabrication process is demonstrated to provide a low cost, effective thermal management technique.

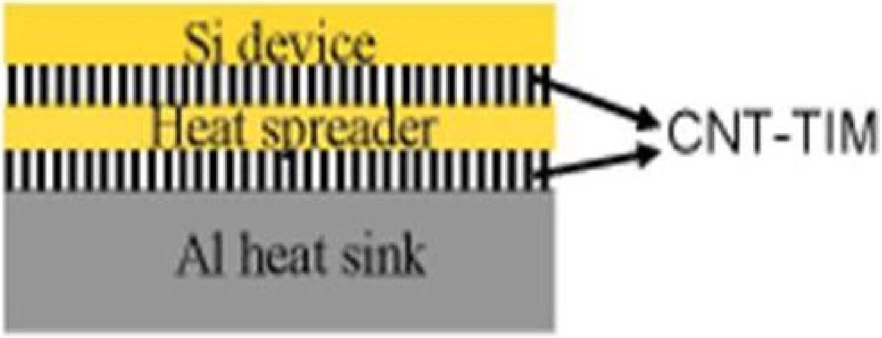


Figure 8.6: Schematic of a sample package using CNT [72]

As shown in Fig. 8.6 [72] the sample consists of a N-type double polished silicon assumed as the silicon device, a silicon substrate with CNT-TIM on both sides which is assumed as

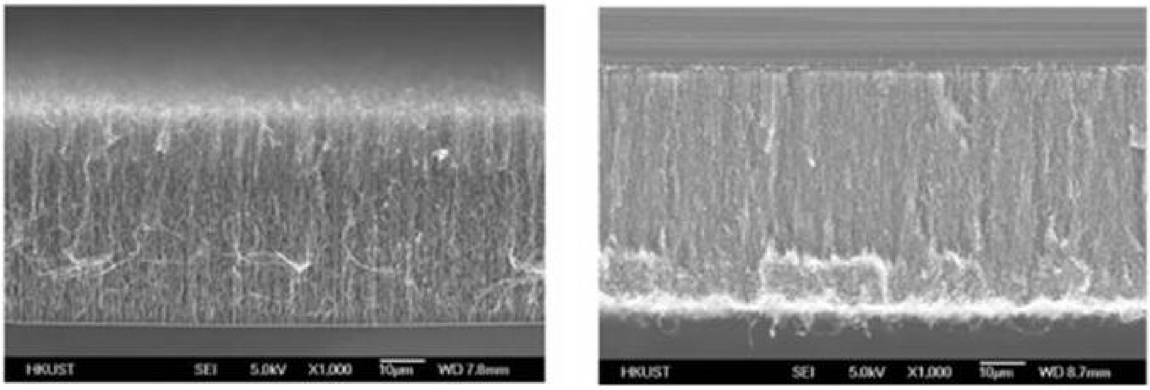


Figure 8.7: SEM images of the CNTs on upward surface (left) and downward surface (Right)

[72]

the heat spreader and TIM on both sides and an aluminium alloy plate which represents the heat sink. With this heat spreader, the silicon device and heat sink can be directly bridged and no additional TIM or individual heat spreader is needed.

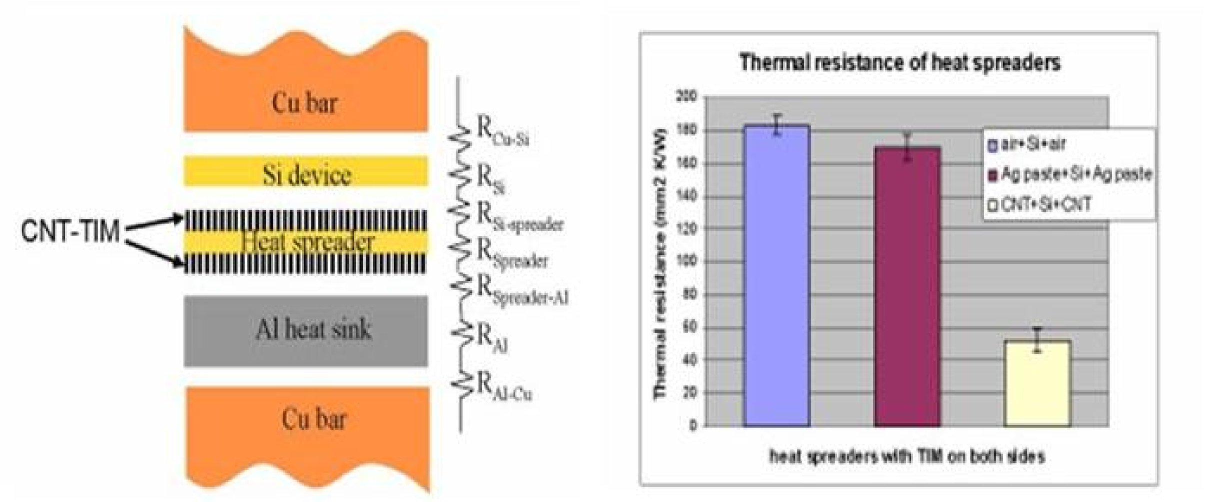


Figure 8.8: Thermal resistance distribution and measurement results

[72]

The thermal resistance of the heat spreader with TIM on both sides is obtained in three steps.First, the whole sample package as shown in Fig. 8.6 is put in the thermal resistance measurement system which gives the complete resistance $R_{package}$ which is an aggregation of a whole series of thermal resistances as shown in Fig. 8.8 left. In the second step the silicon chip is put in the thermal resistance measurement system. This thermal resistance R_{Si} covers R_{cu-si}, R_{si}, R_{Si-Cu} among them, R_{Cu-Si}, is assumed equal to R_{Si-Cu} as the surface roughness conditions of the two Cu block surfaces and the two Si surfaces are the same. Similarly the measured thermal resistance R_{Al} can be obtained. The total thermal resistance of the heat spreader with TIM on both sides is the sum of all these resistances in series. Measurements reveal that he thermal resistance of the novel heat spreader with CNTs on both sides was only about 30% of that of conventional heat spreader with silver epoxy TIM on both sides as shown in Fig 8.8 [72] right. The thermal performance of the novel heat spreader can be further improved by optimizing CNT synthesis parameters.

8.4.2 Heat Pipes

Heat pipes are one of the most effective thermal management techniques. Heat pipe technology has found its acceptance in the microelectronic industry very recently as refinement techniques were unavailable earlier. Heat pipes as the name suggests are pipes which channel the heat of a specific area to another. This is made possible through ducts that carry the heat with minimal change in temperature across the pipe. Fig 8.9 shows the schematic of a heat pipe.

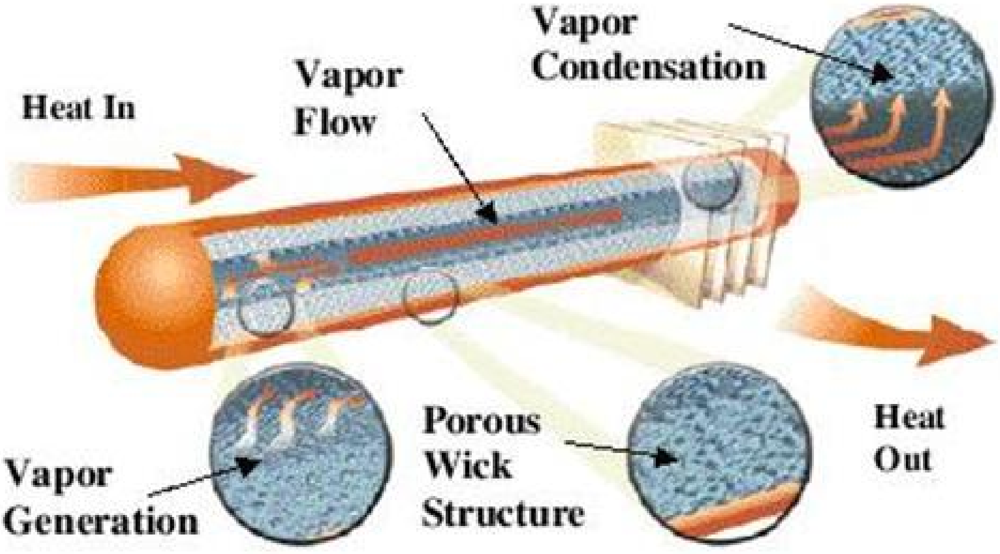


Figure 8.9: Pipe (Courtesy of Thermacore International Inc.)

Proper function of the heat pipe requires that the net capillary pressure difference between the heat source and the heat sink must be greater than the summation of all pressures losses occurring throughout the liquid and vapor flow paths. This is referred to as the capillary limitation and can be expressed mathematically as

$$\Delta P_{cmax} \geq \Delta P_l + \Delta P_y + \Delta P_g \tag{8.5}$$

where ΔP_{cmax} , is the maximum capillary pressure difference generated inside the capillary between the heat source and the heat sink. ΔP_l and ΔP_y are the viscous pressure drops occurring in liquid and vapor phase, respectively and ΔP_g represents hydrostatic pressure drop.

When the maximum capillary pressure is equal to or greater than the summation of these pressure drops as shown by the equation8.5, the capillary structure is capable of returning an adequate amount of working fluid to prevent the dry-out of the evaporator wicking structure. When it is not, the working fluid is not supplied rapidly enough to the evaporator to compensate for the liquid loss through vaporization and the wicking structure dries out. This condition referred to as capillary limitation. It varies according to the wicking structure, working fluid, evaporator heat flux, and operating temperature.

Fundamentally micro-grooved heat pipes operate in the same way as conventional heat pipes. The only difference is the absence of the wicks. The vaporization and condensation processes cause the liquid-vapor interface in the arteries to change continuously along the heat pipe and result in a curvature difference between the hot and the cold end in a micro-grooved heat pipe. A flow of liquid is due to the curvature difference between the hot and the cold ends, allowing the sharp corner regions to be used as liquid arteries and thus no wicking structure is required.

The capillary pressure difference between the vapor generation and vapor condensation causes the flow of the working fluid from the condenser, back to the evaporator, through the corner regions in a in a grooved heat pipe (refer to Fig. 8.10 [77]) and through wicks in a conventional heat pipe. The heat pipe dissipates energy from the heat source by the latent heat of evaporation which is in a nearly isothermal operation. Phase change of the working fluid from liquid to vapor and back to liquid at the evaporator and condenser section makes this device also known as a *two-phase convection device.*

Micro heat pipes have found usage ranging from thermal management systems in spacecrafts to microelectronic circuits. The two phase convection device presents a number of advantages over the conventional single phase systems. Advantages of micro heat pipes can be summarized as follows [77]:

First, due to the high latent heat associated with the closed two phase cycle, heat pipes have heat transfer capacity several orders of magnitude higher than solid conductors

The second advantage is the ability to increase the rate at which the working fluid is vaporized without a significant increase in the operating temperature. This provides an environment in which the evaporation rate automatically adjusts and is able to accommodate a wide range of power inputs while maintaining an almost constant temperature.

The third advantage is that heat pipes work on the principle of capillary action. In comparison with other devices that use a two phase heat transfer, heat pipes require almost no pressure difference whereas for the other case a pressure difference of atleast 5atm is needed.

The fourth advantage is the thermal response time of heat pipes. The thermal response time is is very short since heat pipes utilize a closed two-phase cycle with a relatively small thermal mass. Also another advantage is that it has been shown [78] that this thermal response time does not depend on the length of the heat pipe. Finally, low maintenance

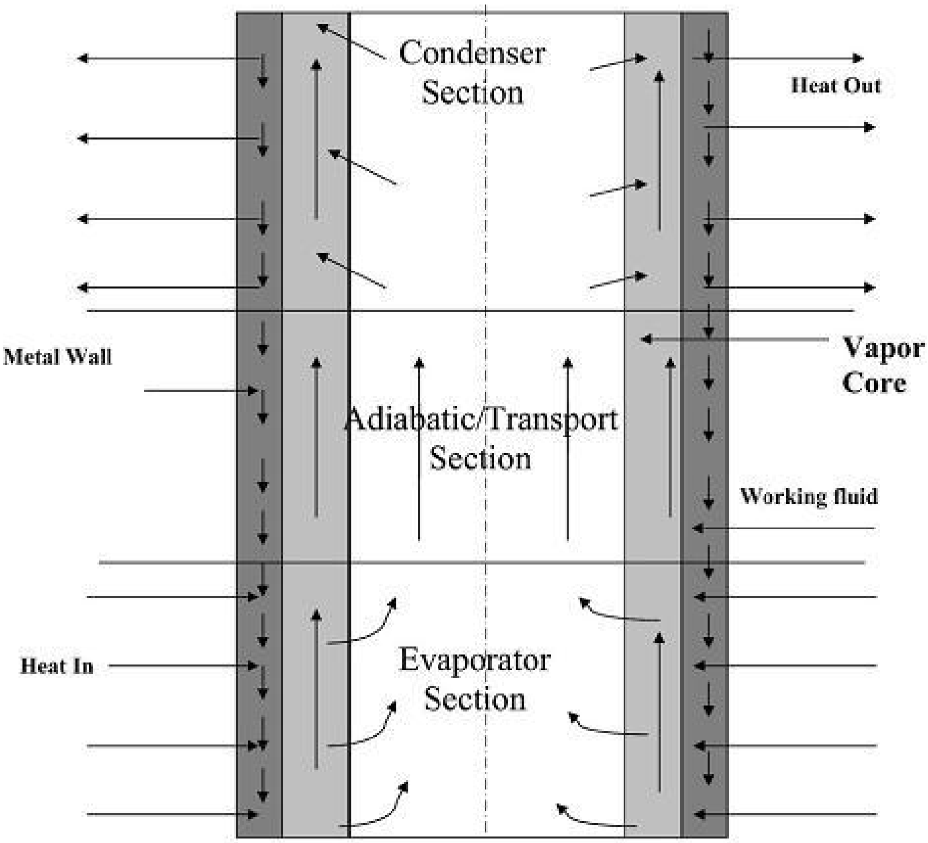


Figure 8.10: Working principle of a heat pipe
[77]

and high reliability and the ability to operate with no external power make them ideal candidates for passive cooling applications.

Cotter [79] was amongst the first to define a micro heat pipe as a device so small that "the mean curvature of the liquid-vapor interface is necessarily comparable in magnitude to the reciprocal of the hydraulic radius of the total flow channel."

Conventional heat pipes can be distinguished from micro heat pipes which have a much lower hydraulic radius to liquid-vapor interface curvature. By definition, a micro heat pipe requires the Bond Number B_o

$$B_o = \frac{\rho g R_h^2}{\sigma} \tag{8.6}$$

where ρ is the density of the working fluid Kgm^{-3} and σ is the surface tension Nm. B_o is of the order of less than or equal to one [79].

When designing a micro heat pipe, two important design parameters are to be taken into account namely, Fill charge and Friction Factor Surface Roughness.

Fill charge defines the amount of working fluid charged in a heat pipe. Affects on the performance of the heat pipe at a given temperature and heat input were clearly demonstrated by Tio et al [80] and Ochterbeck et al [81]. Duncan and Peterson [82] [83] give the mathematical expressions for the fill charge. Suman [84] has presented the deteriorating affects of the overcharge and undercharge on the performance of the heat pipe. If the working fluid is overcharged, the liquid pool thickness in increased due to flooding of the condensing section and resulted in increase of the heat transfer resistance. This flooding in the condenser section caused the area of the condenser to become smaller and a higher rate of heat removal was required. If the working fluid was undercharged, the area available for the flow of the liquid was less which increased the fluid velocity causing frictional loss In addition to this, the available fluid for cooling was less and conditions for dryout were also expected at a high heat input.

From the above two observations, it is clear that there is an optimal charge fill, which is such that the heat pipe is completely filled at the cold end (condenser) and is completely empty at the hot end (evaporator). The exact amount of fluid required for this optimal charge can be calculated by taking the total mass of the coolant and the vapor inside the heat pipe. Suman [84] has shown that the optimal charge depends on the following:

It increases with an increase in heat input, contact angle, and length of heat pipe, and decreases with an increase in inclination, acceleration due to gravity, and surface tension [84]. The second design parameter mentioned above comes into play when the internal diameter or the microgroove of the heat pipe is in the submicron range. Laminar flow of the working fluid is affected by the friction and roughness of the surface and these both together contribute as the channel becomes smaller. With increasing Reynolds number, the Nusselt number and the friction factor of a rough channel surface increase much faster in comparison to a channel which has comparatively smooth surface. More details on the relation between the Reynolds number and the surface roughness can be found in Chen and Cheng [85]. White [86] presented an analytical expression for the friction factor which was a constant and this constant was known as the Poiseuille number and is given by f_{NRe} = constant.

Wu et al [87] carried out several experiments with trapezoidal and triangular cross sections of microchannels and related the laminar friction constant to the cross sectional aspect ratio $\frac{W_b}{W_t}$ which is given as

$$fRe = 11.43 + \exp(2.67\frac{W_b}{W_t}) \tag{8.7}$$

The ratio $\frac{W_b}{W_t} = 0$ represents a triangular channel and $\frac{W_b}{W_t} = 1$ represents a flat trapezoidal microchannel. Suh et al [88] also presented 8.7, but provided a more accurate result for the trapezoidal and triangular channel. The results can be summarized as

$$(fR_e)_v = (-0.94 + 3.8\exp(\frac{\pi h_c}{2h}) + \frac{11.8}{1+\sin\alpha}) + (\frac{W}{L}^2)(52 + 4.6\exp(\frac{\pi h_c}{2h}) + \frac{11.8}{1+\sin\alpha}) \tag{8.8}$$

for the trapezoidal channels and

$$(fR_e)_v = (22.2 + 2.53\exp(\frac{\pi h_c}{2h}) + \frac{4.1}{1+\sin\alpha}) + (\frac{W}{L}^2)(-29.7 + 5.43\exp(\frac{\pi h_c}{2h}) + \frac{42.3}{1+\sin\alpha}) \tag{8.9}$$

for the triangular channels.

Many other factors go into considerations when designing micro heat pipes. When selecting the working fluid, some of the factors to be kept in mind are [48]

- Compatibility
- Thermal Stability
- Wettability
- High conductivity
- Viscosity
- Boiling and freezing points are to be considered.

For the working fluid there is a figure of merit given by Dunn and Reay [89] which is given by

$$M_e = \frac{\rho_l \sigma_l \lambda_l}{\mu_l} \tag{8.10}$$

where λ_l is the latent heat of vaporization of the working liquid JKg^{-1} and
μ_l is the viscosity of the working liquid $Kgm^{-1}s^{-1}$

For the micro channel itself, a few actors that are important are

- Compatibility
- Wettability
- Thermal conductivity
- Strength
- Porosity
- Ease of fabrication

In addition, the material has to be resistant to corrosion; it has to be ductile and malleable to be able to mold it to the appropriate dimensions. Apart from these the environmental conditions of the heat pipe should also be taken into account which can be: operating temperature, power dissipation of the system where the heat pipe is employed, thermal resistances etc etc.

Even after one takes all the best possibilities into account to design a heat pipe, there is a fundamental limit to the amount of heat power a heat pipe can transfer. Heat pipes can be designed to deliver from a few Watts to several kilowatts of heat power. If the heat pipe is operated beyond its capacity for which it is designed, the thermal conductivity is considerably is reduced. Therefore care should be taken to ensure that the heat pipe is operated in the region and within the range for which it is designed. The maximum heat transport capability of the heat pipe is determined by several factors. In all there are eight limitations to the heat pipe to transfer heat. There are all dependent on the temperature of operation and include: viscous, sonic, capillary, pumping, entrainment or flooding, boiling, condenser, vapor continuum and frozen startup. Fig. 8.12, shows the limits graphically.

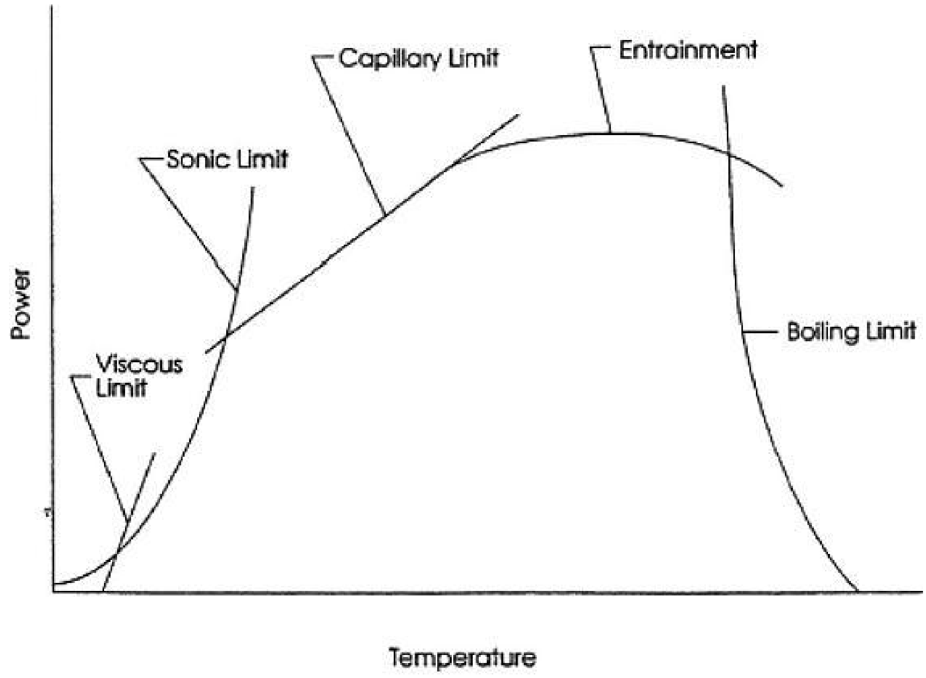


Figure 8.11: Limits of Heat Pipe Operation [48]

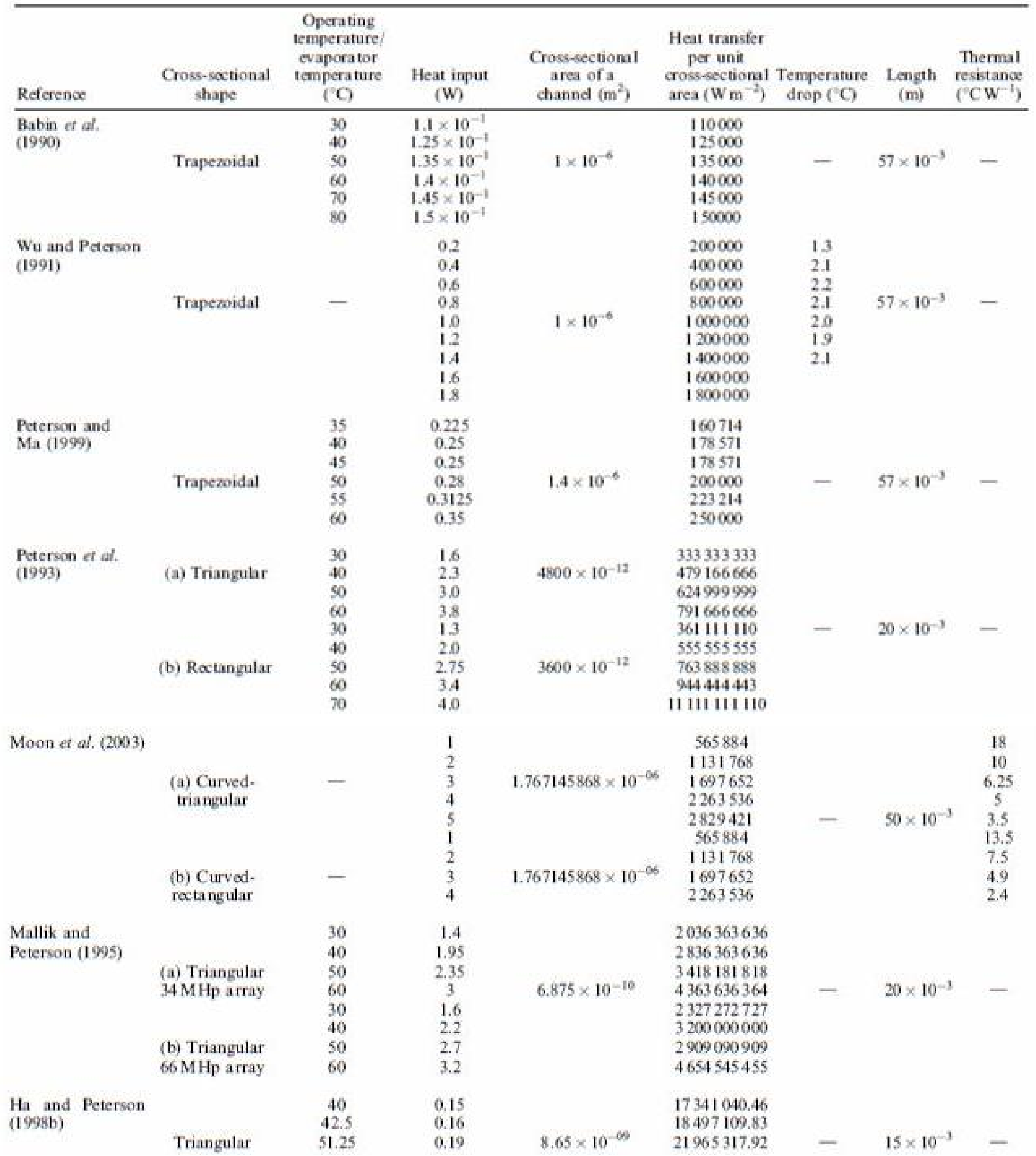
It is beyond the scope to discuss these limits in detail. Refer to [90] and [77] for a complete details on these limitations.

8.4.3 Heat Sinks

Heat sinks are one of the oldest technologies employed to cool electronic systems. Traditionally, heat sinks utilize air cooling with natural convection and radiation or forced convection methods. Although heat sinks are not directly used to cool hotspots or IC's, they are usually mounted on chip packages and provide uniform cooling of the entire package.

Heat sinks that utilize natural convection and radiation are the simplest. They are used in circuit boards and are able to dissipate to 5W of heat power just by natural convection of the atmospheric air. Such heat sinks are commonly utilized in home appliances, desktop computers, televisions, etc.

In an effort to preserve the characteristics of natural convection and at the same time increasing its efficiency, Florio and Harnoy [92] have proposed an innovative and alternative cooling technique that enhances heat transfer utilizing natural convection. In their study whereby appropriately positioning an opening with a transverse vibrating plate,



Reference	Cross-sectional shape	Operating temperature/ evaporator temperature (°C)	Heat input (W)	Cross-sectional area of a channel (m^2)	Heat transfer per unit cross-sectional area ($W m^{-2}$)	Temperature drop (°C)	Length (m)	Thermal resistance (°C W^{-1})
Babin *et al.* (1990)	Trapezoidal	30	1.1×10^{-1}	1×10^{-6}	110 000	—	57×10^{-3}	—
		40	1.25×10^{-1}		125 000			
		50	1.35×10^{-1}		135 000			
		60	1.4×10^{-1}		140 000			
		70	1.45×10^{-1}		145 000			
		80	1.5×10^{-1}		150000			
Wu and Peterson (1991)	Trapezoidal	—	0.2	1×10^{-6}	200 000	1.3	57×10^{-3}	—
			0.4		400 000	2.1		
			0.6		600 000	2.2		
			0.8		800 000	2.1		
			1.0		1 000 000	2.0		
			1.2		1 200 000	1.9		
			1.4		1 400 000	2.1		
			1.6		1 600 000			
			1.8		1 800 000			
Peterson and Ma (1999)	Trapezoidal	35	0.225	1.4×10^{-6}	160 714	—	57×10^{-3}	—
		40	0.25		178 571			
		45	0.25		178 571			
		50	0.28		200 000			
		55	0.3125		223 214			
		60	0.35		250 000			
Peterson *et al.* (1993)	(a) Triangular	30	1.6	4800×10^{-12}	333 333 333	—	20×10^{-3}	—
		40	2.3		479 166 666			
		50	3.0		624 999 999			
		60	3.8		791 666 666			
	(b) Rectangular	30	1.3	3600×10^{-12}	361 111 110			
		40	2.0		555 555 555			
		50	2.75		763 888 888			
		60	3.4		944 444 443			
		70	4.0		11 111 111 110			
Moon *et al.* (2003)	(a) Curved-triangular	—	1	$1.767145868 \times 10^{-06}$	565 884	—	50×10^{-3}	18
			2		1 131 768			10
			3		1 697 652			6.25
			4		2 263 536			5
			5		2 829 421			3.5
	(b) Curved-rectangular	—	1	$1.767145868 \times 10^{-06}$	565 884			13.5
			2		1 131 768			7.5
			3		1 697 652			4.9
			4		2 263 536			2.4
Mallik and Peterson (1995)	(a) Triangular 34 MHp array	30	1.4	6.875×10^{-10}	2 036 363 636	—	20×10^{-3}	—
		40	1.95		2 836 363 636			
		50	2.35		3 418 181 818			
		60	3		4 363 636 364			
	(b) Triangular 66 MHp array	30	1.6		2 327 272 727			
		40	2.2		3 200 000 000			
		50	2.7		2 909 090 909			
		60	3.2		4 654 545 455			
Ha and Peterson (1998b)	Triangular	40	0.15	8.65×10^{-09}	17 341 040.46	—	15×10^{-3}	—
		42.5	0.16		18 497 109.83			
		51.25	0.19		21 965 317.92			

Figure 8.12: Comparison of various heat transfer performance of various micro heat pipes [91]

significant improvements in thermal heat dissipation in comparison with natural convection were observed. They also observed that the geometry of the vibrating plate also influenced the amount of heat dissipation.

Natural convection has its limits which are determined by the airflow and the ambient temperature. When cooling due to natural convection is not adequate, forced convection is employed by using an external source like a fan or a pump to increase the air flow and hence to increase thermal heat dissipation. Heat sinks that utilize forced convection cooling of the hot surface are characterized by extensions on their surfaces known as fins. Many varieties and variations of heat sinks utilizing fins are in commercial use. Some of them are parallel fin, offset fin, louver fin.

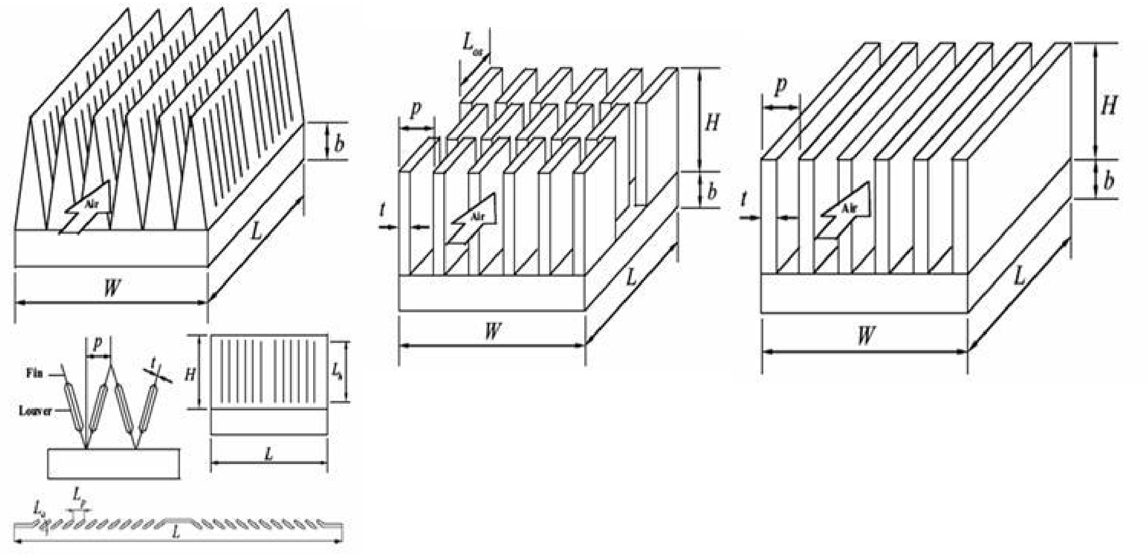
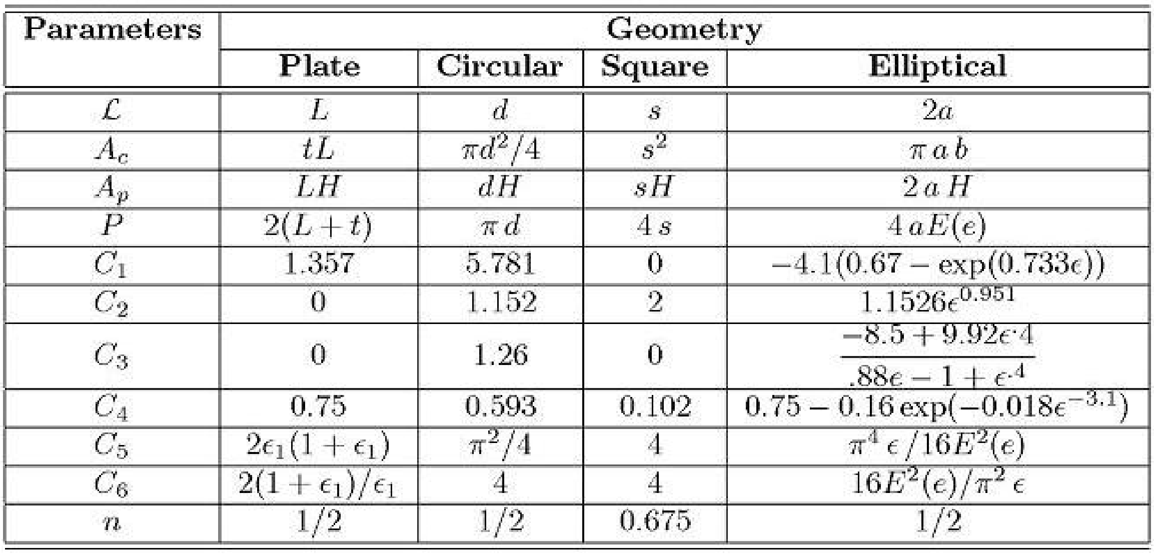


Figure 8.13: Left - Louver Fin heat sink; Middle - Offset fin heat sink; Right - Plate fin heat sink

[93]

Other type of heat sinks that is gaining popularity in free convection are the pin fin heat sinks. Khan et al [94] studied the affect of the fin geometry w.r.t to the thermal performance of the heat fin. Their studies on different fin geometries having the same perimeter were compared from the point of view of the heat transfer drag force and the dimensionless entropy. Their results suggest that there is optimal entropy for every dimension they considered corresponding to the Reynolds number. Their results showed that the square fin performs the worst, when heat transfer is considered and circular geometry performs best when entropy is taken into account and smaller perimeter is considered. Rectangular fin performs best from the point of view of total entropy and a high Reynolds number large perimeter and small aspect ratio. These fins are followed by the Elliptical geometry fins. The parameters they considered are presented below in Fig. 8.14 and 8.15 where:

Nomenclature
A_c = *cross sectional area of the fin [*m^2*]*
A_p = *area for drag force [*m^2*]*
a,b = semi major and minor axes of the elliptical fin [m]
B = duty parameter $\rho\nu^3 kT_a/Q^2$
C_D = *total drag coefficient*
cp = specific heat of the fluid [J/kgK]
D = pin diameter [m]
E(e) = complete elliptic integral of second kind
e = eccentricity in case of elliptical geometry
F_D = *drag force [N]*
H = fin height [m]
k = thermal conductivity [W/mK]



Parameters	Geometry			
	Plate	Circular	Square	Elliptical
$\mathcal{L}$	L	d	s	$2a$
A_c	tL	$\pi d^2/4$	s^2	$\pi\, a\, b$
A_p	LH	dH	sH	$2\, a\, H$
P	$2(L+t)$	$\pi\, d$	$4\, s$	$4\, aE(e)$
C_1	1.357	5.781	0	$-4.1(0.67 - \exp(0.733\epsilon))$
C_2	0	1.152	2	$1.1526\epsilon^{0.951}$
C_3	0	1.26	0	$\frac{-8.5 + 9.92\epsilon^{\cdot 4}}{.88\epsilon - 1 + \epsilon^{\cdot 4}}$
C_4	0.75	0.593	0.102	$0.75 - 0.16\exp(-0.018\epsilon^{-3.1})$
C_5	$2\epsilon_1(1+\epsilon_1)$	$\pi^2/4$	4	$\pi^4\, \epsilon\, /16E^2(e)$
C_6	$2(1+\epsilon_1)/\epsilon_1$	4	4	$16E^2(e)/\pi^2\, \epsilon$
n	1/2	1/2	0.675	1/2

Figure 8.14: Design Parameters for different geometries of Heat Sinks

$$N_s = \frac{1}{Re_{\mathcal{L}}\sqrt{C_5\, Nu_{\mathcal{L}}\, k_{eq}}\tanh(\gamma\sqrt{C_6\, Nu_{\mathcal{L}}\, k_{eq}})} + \frac{1}{2}C_D\; B\,\gamma\, Re_{\mathcal{L}}^2$$

$$\dot{S}_{gen} = \frac{Q^2 R_{th}}{T_\infty^2} + \frac{F_D U_\infty}{T_\infty}$$

$$C_5 = \frac{P\, A_c}{\mathcal{L}^3} \qquad \text{and} \qquad C_6 = \frac{P\, \mathcal{L}}{A_c}$$

$$R_{th} = \frac{1}{k\, A_c\, m\, \tanh(mH)}$$

$$Nu_{\mathcal{L}} = \frac{h\mathcal{L}}{k} \qquad \text{and} \qquad Re_{\mathcal{L}} = \frac{U_\infty\, \mathcal{L}}{\nu}$$

$$F_D = C_D\left(\frac{1}{2}\rho U_\infty^2\right)A_p$$

$$C_D = \frac{C_1}{\sqrt{Re_{\mathcal{L}}}} + C_2 + \frac{C_3}{Re_{\mathcal{L}}}$$

Figure 8.15: Design Equations for different geometries of Heat Sinks

k_{eq} = ratio of thermal conduc tivity of fluid to the fin material[k_f/k]
h = average heat transfer coefficient [W/m^2 K]
L = length of baseplate in flow direction [m]
l = length of plate fin [m]
L = characteristic length of fin [m]
m = fin performance parameter [m^{-1}]
N_s = total dimensionless total entropy generation rate

N_{sf} = *fluid flow irreversibility*
N_{sh} = *heat transfer irreversibility*
N_{uL} = *Nusselt number based on the characteristic length of the fin* hL/k_f
P = *perimeter of the fin [m]*
Pr = *Prandtl number*
Q = *total base heat flow rate [W]*
$R_{eL} = Reynoldsnumberbasedonthecharacteristiclengthofthefin \equiv U_{app}L/\nu$
R_{tot} = *total thermal resistance [K/W]*
$S{\cdot}gen$ = *total entropy generation rate [W/K]*
s = *side of a square fin [m]*
T = *temperature [K]*
t = *thickness [m]*
U_{app} = *approach velocity of the fluid [m/s]*
W = *width of baseplate [m]*
w = *width of plate fin [m]*

Greek Symbols
ϵ = *axis ratio of elliptical fin[b/a]*
ϵ_1 = *ratio of the plate sides[t/L]*
γ = *aspect ratio [H/L]*
ν = *kinematic viscosity of fluid* $[m^2/s]$
ρ = *fluid density* $[kg/m^3]$

Subscripts
$a = ambient$
$b = baseplate$
$f = fluid$
$w = wall$

Pin Fin heat sinks, as the name suggests, are heat sinks that have fins in the shape of long cylindrical pins. Sparrow and Vemuri [95] [96] in their experiments have investigated the heat transfer from arrays of pin fin heat sinks with varying density of the fins. Their experimental observations reveal that there is a significant role of the ratio of the fin diameter to the spacing between the fins. They have observed that there is an optimal value of this ratio as 0,5. In another experiment, Zografos and Sunderland [97] [98] studied square array of pin fins and they came to the optimal value for square array pin fins is 0.33.

Fisher and Torrance [99] have developed a semi analytical determination for optimal geometries for free convection. From their analysis, the "Chimney Effect" was shows to enhance the local heat transfer in a way that the temperature rise does not depend on the height of the heat sink. This effect in turn allows the designer to reduce the volume and the weight of the heat sink without affecting the thermal performance.

In another experiment by Shaukatullah [100], the thermal performance for in-line square pin fins and plate heat sinks for varying fin thick ness, spacing, height and approach angle is studied. Kang and Holahan [101], studied the heat sink performance dependence on the different geometries and formulated an one dimensional thermal resistance model of impingement air cooled plate fin heat sinks. Zeinab S [104] in his experimnents has employed the "Entropy Generation Minimization" method to optimize the thermal performance ans behavior of pin fin heat sinks.The results for optimum design of heat sink for in-line and staggered alignments with circular, square, rhombus, rectangular, and elliptical configurations are investigated and the thermal behavior is compared. The results are shown in table 8.5

Geometries	**Alignment**	**Pin-fin Materials (W/m-K)**	**Opt. Length and Diameter**	**No. of Pin-Fins**	**Entropy Generation Rate (W/K)**
Circular	In-Line	Plastic = 25	21.039, 2.02	10 * 10	$5.8 * 10^{-6}$
		Aluminum=237	18.04, 1.9	10 * 10	$5.1 * 10^{-6}$
		Copper=400	15.2, 1.6	10 * 10	$5.0 * 10^{-6}$
	Staggered	Plastic=25	16.37, 1.9	9 * 9	$6.33 * 10^{-6}$
		Aluminum=237	14.01, 1.6	9 * 9	$5.57 * 10^{-6}$
		Copper=400	10.2, 1.5	9 * 9	$5.2 * 10^{-6}$
Square	In-line	Plastic=25	18.33, 2.2	10 * 10	$7.8 * 10^{-6}$
		Aluminum=237	17.01, 2.01	10 * 10	$6.072 * 10^{-6}$
		Copper=400	15.03, 1.9	10 * 10	$6.01 * 10^{-6}$
	Staggered	Plastic=25	17.003, 2.19	9 * 9	$8.9 * 10^{-6}$
		Aluminum=237	15.12, 2.0	9 * 9	$7.8 * 10^{-6}$
		Copper=400	13.01, 1.99	9 * 9	$7.2 * 10^{-6}$
Rhombus	In-line	Plastic=25	18.24, 2.19	10 * 10	$7.7 * 10^{-6}$
		Aluminum=237	14.4, 2.0	10 * 10	$7.0 * 10^{-6}$
		Copper=400	12.05, 1.99	10 * 10	$6.2 * 10^{-6}$
	Staggered	Plastic=25	17.9, 1.9	9 * 9	$7.99 * 10^{-6}$
		Aluminum=237	13.2, 1.88	9 * 9	$7.6 * 10^{-6}$
		Copper=400	11.24, 1.85	9 * 9	$6.5 * 10^{-6}$
Rectangular	In-line	Plastic=25	16.8, 2.04	10 * 10	$8.0 * 10^{-6}$
		Aluminum=237	15.2, 2.02	10 * 10	$7.7 * 10^{-6}$
		Copper=400	12.01, 2.001	10 * 10	$7.2 * 10^{-6}$
	Staggered	Plastic=25	15.8, 2.01	9 * 9	$8.8 * 10^{-6}$
		Aluminum=237	12.99, 2.0	9 * 9	$7.5 * 10^{-6}$
		Copper=400	11.8, 1.98	9 * 9	$7.3 * 10^{-6}$
Elliptical	In-line	Plastic=25	19.9, 2.0	10 * 10	$6.3 * 10^{-6}$
		Aluminum=237	16.2, 1.9	10 * 10	$6.0 * 10^{-6}$
		Copper=400	14.0, 1.7	10 * 10	$5.5 * 10^{-6}$
	Staggered	Plastic=25	17.9, 1.9	9 * 9	$6.8 * 10^{-6}$
		Aluminum=237	14.6, 1.7	9 * 9	$6.5 * 10^{-6}$
		Copper=400	12.03, 1.67	9 * 9	$5.9 * 10^{-6}$

Table 8.5: Optimum variables for selected pin-fin geometries

[104]

For high density chip packages, more often than not free convection cooling methods mentioned above are not viable options. In such scenarios forced convection cooling is inevitable. Although similar in construction and operation the seminal difference arises from the fact that in forced convection the heat transfer does not occur due the natural forces of buoyancy but from an external source such as a fan or a pump. Detailed literature is present in the works of Bar-Cohen et al [102]. In another approach Brucker et al [103], adopted a porous model that replaces an actual heat sink by the equivalent volume of fluid that covered the fins to study free and forced convection heat transfer. They extended their studies to include different shapes and corresponding flow configurations and studied them under both forced and free convection.

8.4.4 Vapor Chambers

A vapor chamber is schematically shown in Figure 8.16 [105]. It is different from a heat pipe in the sense that the condenser covers the entire top surface of the structure. Liquid flows back to the evaporator which is located on top the heat source or evaporator through capillary action in the wick structure which may be distributed in the form of supporting posts (not shown). For microelectronics applications, the combination of water and sintered copper powder is used. The thermal resistance components are also illustrated in Figure 8.16

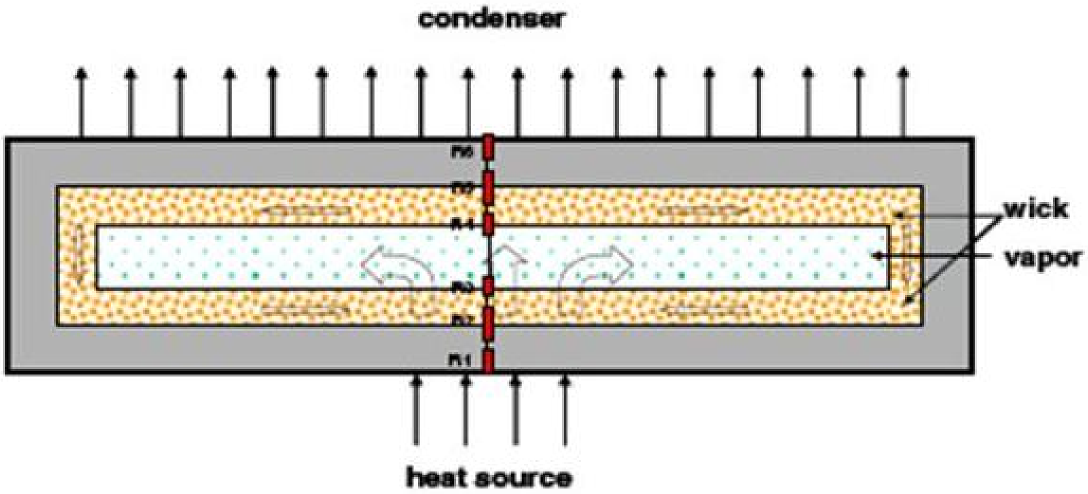


Figure 8.16: Schematic of a Vapor Chamber
[105]

To characterize the thermal performance of heat pipes and vapor chambers, it is necessary to determine the thermal properties of the wick structure and the vapor space. It is beyond the scope of this report to explain the characterization of the wick. Xiojin Wei et al [105] have characterized the wick structure for effective thermal conductivity. They have come up with parallel and serial a assumptions for the wick as idealized cases and have modelled thermal conductivity as

$$K_w = (1-\varepsilon)K_s + \varepsilon K_l \tag{8.11}$$

for the parallel assumption, and

$$K_w = \frac{1}{(1-\varepsilon)/K_s + \varepsilon/K_l} \tag{8.12}$$

for the serial assumption.

Where K_l and K_s are the thermal conductivities for water and copper and ε is the porosity of the wick.

Maxwell et al [107] and Chi [108] give expressions characterizing the wick structure for sintered structure. In addition to equilibrium, the performance of a vapor chamber is governed also by transport parameters. Hence it is interesting t obtain the performance characteristics using the heat transfer rates. Jei Wei [106] et al used the boiling model to come up with the performance of a vapor chamber. It is based on Newton's law of cooling and is given by

$$Q = hA(T_s - T_v) = hA\Delta T \tag{8.13}$$

where Q is is the power dissipated, h is the film coefficient $(Wm^{-2} - K)$. A is the boiling cross sectional area (m^2) T_s is the temperature at the solid wall in contact with the porous wick (K) and T_v is the temperature of vapor (K). The important thing to know here is the value of h. Once h is known the performance of the vapor chamber is easy to calculate.

Yasuhiro Horiuchi et al [109] have implemented a wick structure microchannel vapor chamber as a heat spreading device specifically designed for microprocessor cooling. Their implementation is shown in Fig 8.17

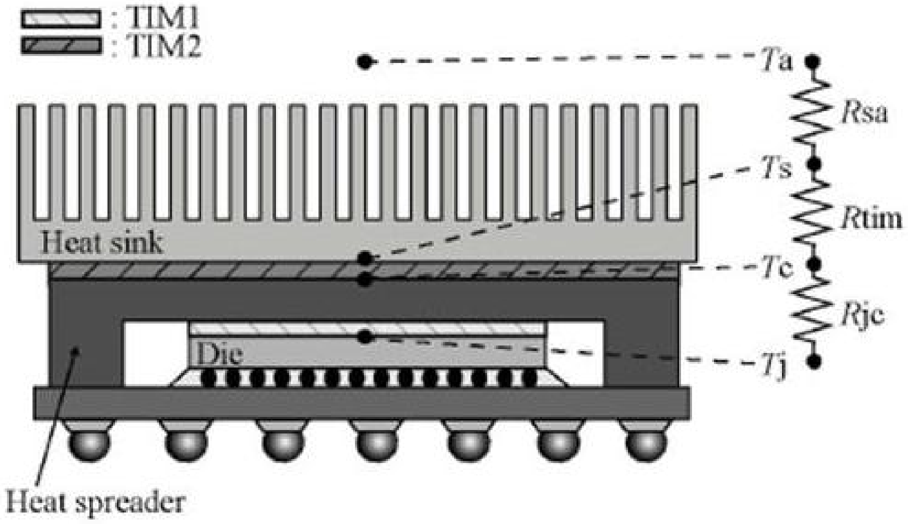


Figure 8.17: Vapor chamber integrated with heat sink and die [80] [109]

In the above implementation, the heat from the die is first spread by the heat spreader

(TIM1) and then by the heat sink. Thermal interface materials are used to reduce the thermal contact resistance. In Fig 8.17 T_a is the ambient, T_s is heat sink, T_c is the heat spreader, T_j is die surface temperature. The contact thermal resistances R_{sa}, R_{tim} and R_{jc} are calculated as below:

$$R_{jc} = \frac{T_j - T_c}{Q} \tag{8.14}$$

$$R_{tim} = \frac{T_c - T_s}{Q} \tag{8.15}$$

$$R_{sa} = \frac{T_s - T_a}{Q} \tag{8.16}$$

Where Q is the heat generated from the die.

$$R_{ca} = R_{tim} + R_{sa} = \frac{T_c - T_a}{Q} \tag{8.17}$$

The total thermal resistance from the surface of the chip die to the ambient surface is given by

$$R_{ja} = R_{jc} + R_{ca} = \frac{T_j - T_a}{Q} \tag{8.18}$$

From the above implementation, Yasuhiro Horiuchi et al were able to conclude the following: The thermal resistance between the heat source and ambient could be improved by 0.1^oC/W by their implementation of a vapor chamber The thermal resistance between the heat source (die) and the heat spreader surface is about 30% better in comparison to vapor chamber implementations with sintered copper. In another study Unnikrishnan et al [110] implemented the vapor chamber in silicon using micromachining technology. Their main idea was to match the CTE of the die to the heat spreader since usual implementation made of copper are limited by the thermal conductivity of copper (401W/mK). In addition the CTE mismatch between silicon dies is about 3ppm/K and between silicon die and copper heat spreader is about 17 ppm/K. Their study clearly showed that using silicon drastically reduced the die stress while maintaining thermal performance comparable to a solid copper block heat spreader. Their observations are listed in Fig 8.18.

Vapor chambers are quite effective in cooling hotspots because their cooling mechanisms. Vapor chambers however lost the race in commercial applications to thermoelectric cooling mainly due to their efficiency and scaling factors in comparison to thermoelectric cooling methods.

8.5 Active Cooling

Active cooling methods require input power. Forced convection, pumped loops, MEMS and refrigerators fall under the category of active coolers. Again referring to Paik et al [56], Active cooling methods can further be classified into adaptive and nonadaptive methods. Adaptive cooling system incorporates a closed-loop feedback mechanism of the

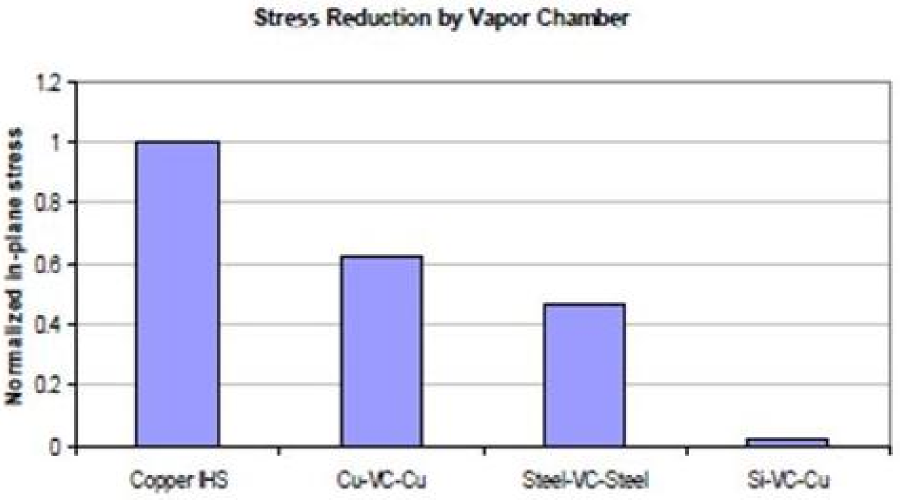


Figure 8.18: Die stress comparison with copper HIS and vapor chamber [81] [110]

chip's temperature. The system is aware of the temperatures at different places on the chip and can dynamically cool these areas or hot spots. As mentioned earlier, thermal non-uniformity is a common factor in modern day IC's due to the integration of many different types of chips on a single die and can reach power densities of 300 W/cm2 or more which is a major cause of detrimental stress on the die. This ability to be able to dynamically cool different areas is an important property. In contrast, although included as an active cooling method, nonadaptive systems do not have the ability to dynamically cool hotspots due to the absence of the temperature feedback mechanism. Such a system can be utilized for uniform cooling of the complete die area or package area. Active cooing methods included are:

- Micropumps
- Microchannels
- Thermoelectric Coolers

8.5.1 Micropumps

Micropumps have emerged as major research area for biological and electronic cooling applications. Use of liquid cooling in electronic applications has proven to be a viable option recently and is also necessary considering the extremely challenging cooling constraints required by the microelectronic industry. Thermal management of electronic devices is of primary interest for manufacturers of portable and reliable electronic components. The need to increase density, functionality, at the same time reducing size is the major contributor for innovations in cooling techniques. One of the available strategies is use of liquid cooling microchannels. The advantage offered by these methods is that the contact and spreading resistances can be reduced by integrating these channels directly on the back side of the die. Another advantage is that the heat generating and heat rejecting components can be separated This makes it possible for the heat exchanger to be placed at any convenient location on the device. The only drawback that limited the implementation of such microchannels was the large pumps that were required. Recent innovative pumping solutions more commonly referred as micropumping solutions have made liquid cooling using microchannels possible. In this section the approach taken by Iverson and Garimella [111] is adopted to classify micropumps into two main categories

1. Mechanical displacement micropumps-defined as those that exert oscillatory or rotational pressure forces on the working fluid through a moving solid-fluid (vibrating diaphragm, peristaltic, rotary pumps), or fluid-fluid boundary (ferrofluid, phase change, gas permeation pumps).
2. Electro- and magneto-kinetic micropumps-defined as those that provide a direct energy transfer to pumping power and generate constant/steady flows due to the continuous addition of energy (electroosmotic, electrohydrodynamic, magnetohydrodynamic, electrowetting, etc.).

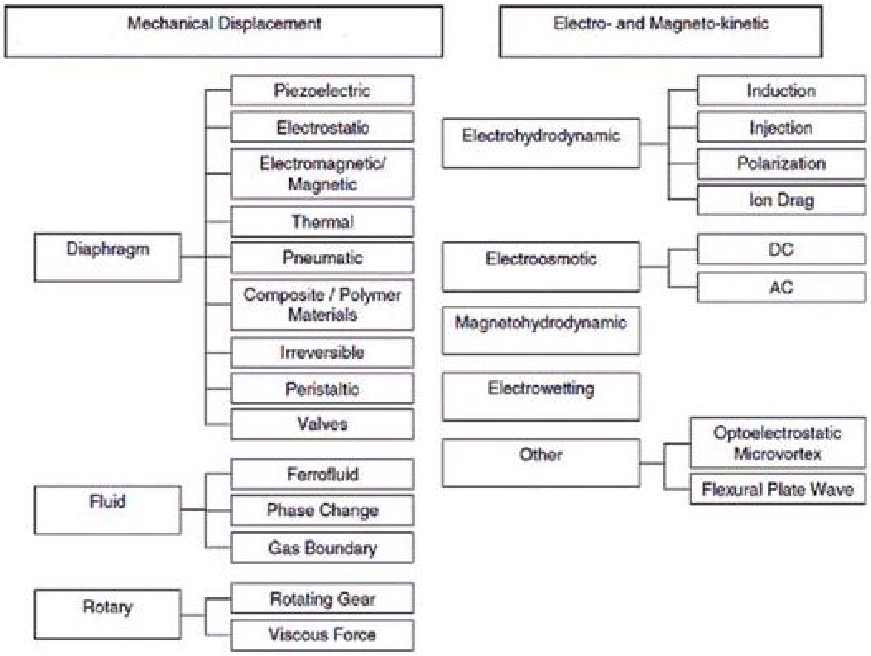


Figure 8.19: Classification of Micropumps
[111]

Presented in this report is only a brief overview of a few variations of the micropumps shown in Fig 8.19. The selection is based on the criteria that a particular type of implementation has been utilized for cooling in the IC or package industry or shows potential to be utilized in this field. For a complete detail on all the above implementations with their applications, advantages and limitations, the reader is referred to reference 82, and further references therein.

Mechanical Displacement techniques

Micropumps that fall under this category use the motion of a gear or a diaphragm or a fluid to generate the required pressure difference to move the working fluid. As seen in Fig 8.19 [111], diaphragm pumps are the most common and contain the most variety of implementations.

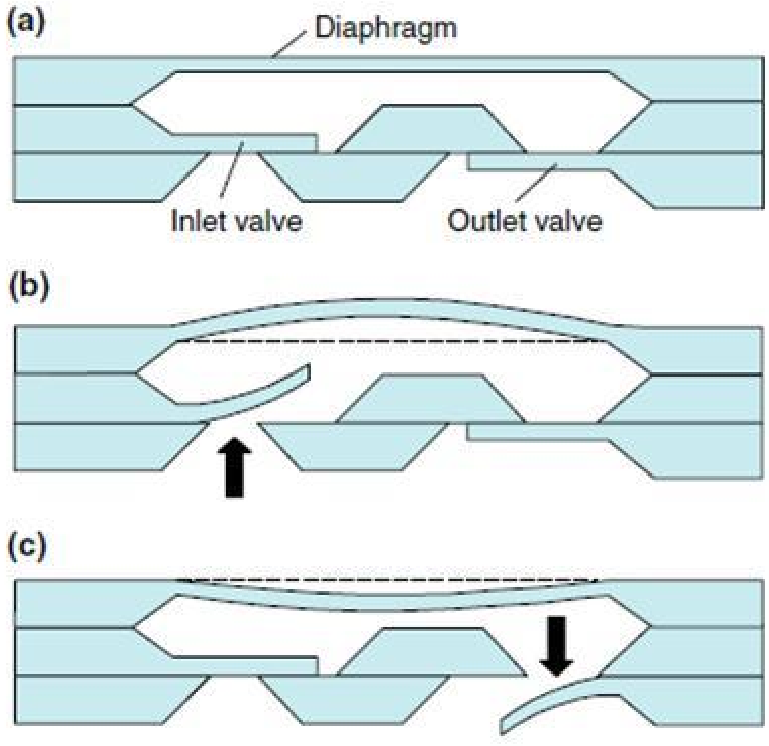


Figure 8.20: Vibrating diaphragm micropump. A: undeflected position b: expansion stroke c: contraction stroke

[111]

Fig 8.20 [111], shows the schematic of a diaphragm displacement pump. Typically such a pump consists of a pumping chamber connected to an inlet and an outlet valve through which the flow takes place. The operating principle is explained as follows:

The diaphragm deflects during the expansion stroke causing the pumping chamber to expand resulting in a decrease in the chamber pressure. When the inlet pressure is higher than the chamber pressure, the inlet valve opens and liquid fills the expanding chamber, Fig 18 b. During the compression stroke, the chamber volume decreases causing the internal pressure to increase and the liquid is discharged through the outlet valve Fig 8.20 c. As seen in Fig 8.19, there are many actuation mechanisms for the diaphragm i.e, Piezoelectric, electrostatic, electromagnetic, pneumatic, and thermopnuematic are among the more common methods.

Piezoelectric

Piezoelectric as the name suggest, are those materials that generate an internal mechanical stress in the presence of an applied electric potential, and vice versa. For diaphragm micropumps, this is a widely used and very popular mechanism. On applying an AC voltage the diaphragm is driven into expansion and compression strokes. The main advantage of using this mechanism is that relatively large displacements of working fluid and forces are possible. Some of the distinguishing characteristics of micropumps utilizing piezoelectric technology are the type of valves used, the material, and the orientation of the shear used for actuation. Active research has been ongoing for quite sometime for developing piezoelectric micropumps.

Nguyen et al [112] have reported a light weight piezoelectric composite actuator as an improvement over traditional piezoelectric actuators. Chen et al [113] have demonstrated novel diaphragm actuation using shear deformation. Such actuation technique can be utilized in droplet injection and diaphragm pumping applications.

Electrostatic

Micropumps employing electrostatic actuation technique use the electrostatic forces that are generated between the electrodes to drive the diaphragm motion. On applying an electrical voltage between an electrode and the diaphragm, it acts as a variable capacitor and the electrostatic forces cause the movement of the diaphragm. This causes a change in the pressure of the pumping chamber and draws the working fluid into it. When the voltage is removed, the membrane reverts back to its original position expelling the fluid inside the chamber. The capacitance between the electrode and the diaphragm can be modelled using the equations below [113]

$$C = \frac{\varepsilon \pi d^2}{4l} \tag{8.19}$$

where d is the diameter or the diaphragm, l is the distance between the electrode and the diaphragm The force acting to pull the plates back to the original position is given by

$$F = \frac{1}{2}\frac{\partial C}{\partial l}V^2 = -\frac{\varepsilon \pi d^2}{8l^2}V^2 \tag{8.20}$$

where V is the applied potential difference.

Magnetic and Electromagnetic

The electromagnetic actuation mechanism consists of a permanent magnet attached to a diaphragm surrounded by a coil. When a current is passed through the coil, the diaphragm deflects due to the Lorentz force. The primary advantage of such an actuation mechanism is that electromagnetic actuation pumps require very low input voltage and they have simple design. These type of micropumps have been discussed in the literature for years. Recently, Chang et al [114] Su et al [115] and Yufeng et al [116] have demonstrated improved diaphragm deflections while reducing the overall size of the micropump by integrating the magnet and the coils directly into the device.

Thermal

Thermal actuation is the expansion or induced stress in a material due to the application of heat. In the context of micropumps using thermally actuated diaphragms, this is usually of the form of thermopneumatic os shape memory alloy (SMA). These actuations are low frequency because they rely on the diffusion of thermal energy. A thermopneumatic actuation occurs when a secondary fluid (separate from the driven fluid) is heated causing it to expand and deflect the pump diaphragm. The expansion stroke occurs as the heater is deactivated allowing the secondary fluid to cool and contract. Phase change materials can also be used in thermal actuation. For example, Paraffin waxes, also used for transient heat absorption from electronics, have been used to actuate a diaphragm pump by exploiting its volume expansion from solid to liquid phase as it melts due to resistive heating. Shape memory actuation uses the shape memory effect. For example, in case of TiNi, a phase transformation between two solid phases is considered:

A high-temperature austenite phase and a low-temperature martensite phase. The martensite is much more ductile than austenite allowing the TiNi to undergo large deformations. Heating above the phase-transformation temperature results in an austenite phase transition in the TiNi during which it assumes its initial shape if it is not constrained. If it is constrained, it exerts a large force in trying to assume its initial shape. Shape Memory Alloys (SMAs) are characterized by large recoverable strain outputs of up to 6-8%. At high frequencies, SMAs do not cool much and their performance suffers. Generally, they operate below 100 Hz.

Peristaltic

As the name suggests, these pumps incorporate the peristaltic motion of actuators in series to generate pumping action. Most peristaltic pumps use three pumping chambers with diaphragms as actuators in series. Figure 8.21 [111], shows the structure and operation of such a pump.

Peristaltic pumps can be considered a subset of the vibrating diaphragm pumps since they utilize many similar transducers (piezoelectric, pneumatic, etc.). When the first diaphragm is actuated, it restricts the flow to the inlet of the pump. The actuation of the second diaphragm pushes the fluid towards the third pumping chamber. Similarly, actuating the third diaphragm in succession pushes the fluid through the outlet of the

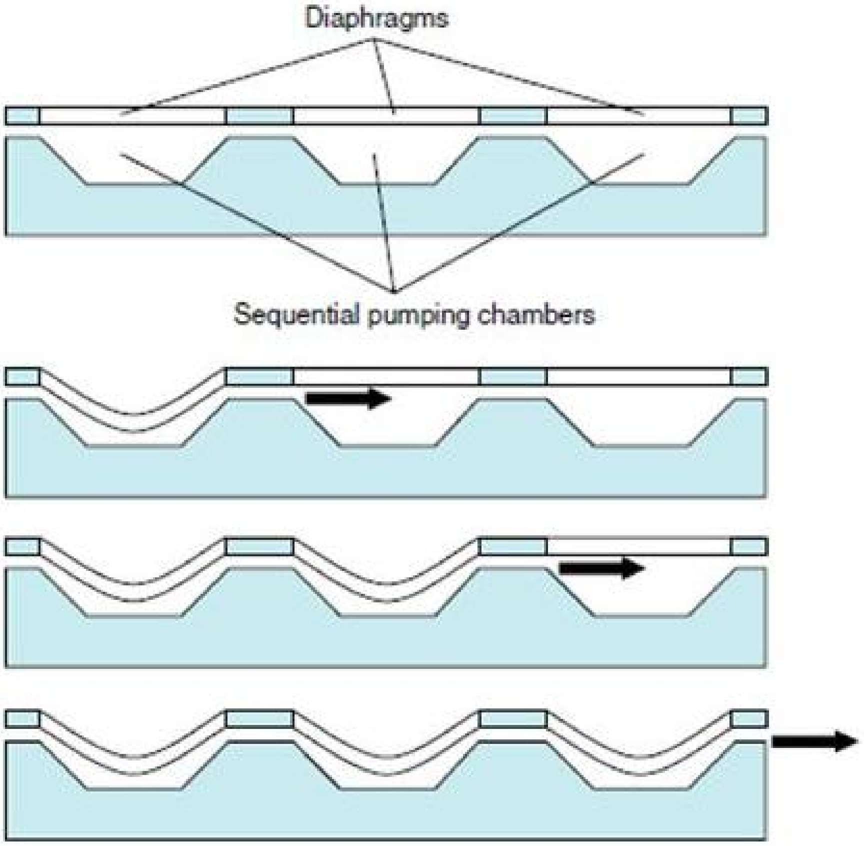


Figure 8.21: Structure and operation of a peristaltic pump
[111]

pump. In essence, the diaphragms act like valves that reduce the flow cross section and provide flow directionality. All three diaphragms are then de-actuated and the sequence is repeated continually for pumping action from left to right. Peristaltic pumps are also bidirectional since reverse pumping actuation is also possible.

Geipel et al. [117] Jang et al. [118] Lin et al. [119], have demonstrated that peristaltic pumps can provide comparatively high-back pressures. Goldschmidtboing et al. [120] developed an analytical model for micro-diaphragm pumps with active valves based on the peristaltic working principle. Their model applies to both fast and slow actuation mechanisms and it can be used for piezoelectric, pneumatic, thermopneumatic and other driving mechanisms. They show that micro-diaphragm pumps suffer from a linear dependence on the flow rate with applied back pressure. In another actuation mechanism for peristaltic pumping, magnetic fluids have been considered. Kim et al. [121] demonstrated a case where the magnetic fluid is attracted and gathered using a permanent magnet into a round-shaped accumulation, which deforms the silicone rubber diaphragm. These lumps are then manipulated by the magnetic field to pump liquid in a peristaltic fashion. With this experiment, a maximum flow rate of 3.8 μ l/ min was achieved.

Another variation of peristaltic pumping is Single-source-actuated peristaltic pumps. The design is similar to the one shown in Fig 8.21, but the pumping chambers are connected serially in a way that the time phased deflection of the diaphragms of the previous chamber generates a peristaltic effect. In this manner these chambers can be connected by using a single source which reduces the failure rate of the components. Huang et al. [122] incorporated a single electromagnetic valve and pneumatic source into a three-chamber design. Yang et al. [123] demonstrated a similar pump which was actuated by a single pneumatic source employing a serpentine pneumatic channel where the intersection of the pneumatic channel and fluid channel constituted areas of membrane deflection. There were seven intersections that provided seven stages of peristaltic actuation. They showed that the flow rate could be increased by increasing the pneumatic pressure, operational frequency or number of membranes (intersections). These s-shaped pneumatic, peristaltic pumps have been used effectively in cell sorting and cytometry applications.

For the above mentioned pumps, the fundamental question pump still remains. When is a pump called micro pump? Does it depend on size of the pump or does it depend on the amount of fluid the pump can handle? Typically mechanical displacement pumps have been fabricated and demonstrated with sizes ranging from 5mm to 1 cm. No mechanical pump has been able to generate a flow rate of less than 1 μ l/min. This is mainly due to its size and the viscous forces that dominate with reduction in size of the operating fluid. The maximum energy density for the above mentioned pumps can be described as below:

For the Electrostatic actuator is given by $\frac{1}{2}\varepsilon E^2$
Where E is the Electric field strength, ε is the dielectric permittivity. The energy density ranges from $0.1\mathrm{Jcm}^{-3}$ to $1\mathrm{Jcm}^{-3}$

For Thermal actuator it is given by $\frac{1}{2}E(\alpha\Delta T)^2$
Where α is coefficient of thermal expansion, ΔT is the temperature difference and E is Young's Modulus. The energy density ranges from $1\mathrm{Jcm}^{-3}$ to $10\mathrm{Jcm}^{-3}$

For the magnetic and electromagnetic actuator it is given by $\frac{1}{2}\frac{B^2}{\mu}$
Where B is the magnetic field strength and μ is magnetic permeability. The energy density ranges from $1\mathrm{Jcm}^{-3}$ to $10\mathrm{Jcm}^{-3}$

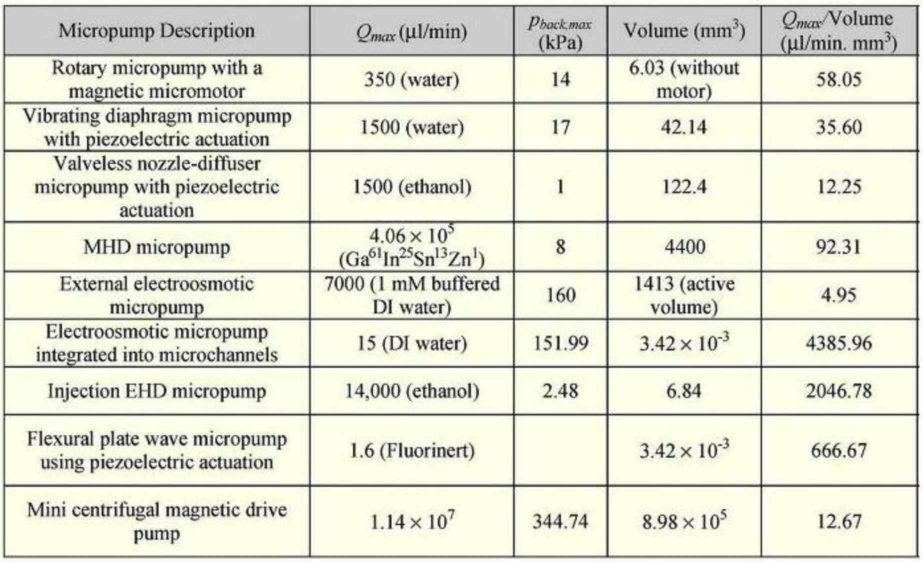
For the piezoelectric actuator it is given by $\frac{1}{2}E_y(d_{33}E)^2$
Where E_y is Young's Modulus, d_{33} is the piezoelectric coefficient E i s the electric field strength. The energy density ranges from $0.1\mathrm{Jcm}^{-3}$ to $1\mathrm{Jcm}^{-3}$

Assuming that for the above mentioned actuation techniques; the energy density does not change by scaling the energy stored decreases by three orders of magnitude [131]. The pumped energy per stroke is given by

$$E_{max} = \frac{p_{max}Q_{max}}{4f} \tag{8.21}$$

Where p_{max} is the max back pressure; Q_{max} is the max flow rate and f is the reciprocating frequency.

The energy efficiency of a pump is defined as the relation between the energy stored in the actuator and the pumped energy. Assuming that the energy efficiency of mechanical micropumps doesn't change compared to its macroscopic counterpart, the energy delivered by a micropump decreases with three orders of miniaturization.



Micropump Description	Q_{max} (µl/min)	$p_{back,max}$ (kPa)	Volume (mm^3)	Q_{max}/Volume (µl/min. mm^3)
Rotary micropump with a magnetic micromotor	350 (water)	14	6.03 (without motor)	58.05
Vibrating diaphragm micropump with piezoelectric actuation	1500 (water)	17	42.14	35.60
Valveless nozzle-diffuser micropump with piezoelectric actuation	1500 (ethanol)	1	122.4	12.25
MHD micropump	4.06×10^5 ($Ga^{61}In^{25}Sn^{13}Zn^1$)	8	4400	92.31
External electroosmotic micropump	7000 (1 mM buffered DI water)	160	1413 (active volume)	4.95
Electroosmotic micropump integrated into microchannels	15 (DI water)	151.99	3.42×10^{-3}	4385.96
Injection EHD micropump	14,000 (ethanol)	2.48	6.84	2046.78
Flexural plate wave micropump using piezoelectric actuation	1.6 (Fluorinert)		3.42×10^{-3}	666.67
Mini centrifugal magnetic drive pump	1.14×10^7	344.74	8.98×10^5	12.67

Figure 8.22: Comparisons of various micropumps
[132]

8.5.2 Electroosmotic Pumps

Electroosmotic (EO) pumping exploits the property of surface charge that develops spontaneously when a liquid comes in contact with a channel wall. Ions with opposite charge in the fluid can be manipulated with an AC or DC electric field resulting in AC and DC electroosmotic pumps.

DC electroosmotic
DC electroosmotic pumps can be constructed using fused silica or glass capillaries. Electric field is provided by the electrodes that are along the channel length. Fig 8.23 [111] shows a schematic of a DC electroosmotic pump.

For the silica-based channels, when an electrolytic solution having a pH ¡ 4 comes in contact with the channel wall, the surface silanol groups in silica spontaneously deprotonate, causing the boundary wall to be negatively charged. This negative surface charge attracts the positive ions and repels the negative ions in the solution resulting in the formation of an electric double layer and a neutral charge density in the center of the channel. Chen et al. [124] have calculated the electric double layer thickness which is given by

$$\lambda_D = \sqrt{\frac{\varepsilon \kappa T}{2q^2 z^2 c}} \tag{8.22}$$

where ε and T are the permittivity and temperature of the fluid, z and c are the valence number and average molar ion concentration, κ is the Boltzmann constant and q is the charge of an electron.

When a DC electric field is applied, the force on the fluid near the capillary wall is much

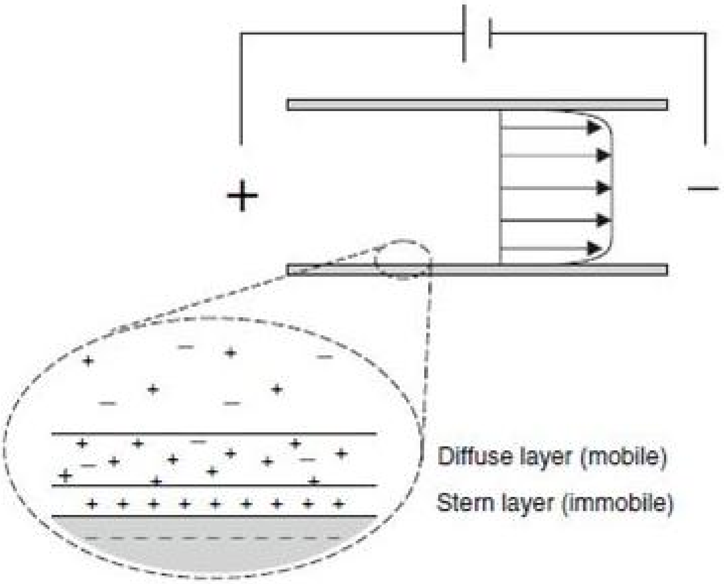


Figure 8.23: DC electroosmotic micropump
[111]

higher due to the high charge density in that area. This charge moves in response to the electric field and the motion of the fluid is propogated due to viscous forces. When the channel hydraulic diameter is much larger than the Debye length, the velocity profile is nearly uniform across the cross-section, and the resulting velocity can be described by the Helmholtz-Smoluchowski as mentioned in Iverson et al. [125]

$$u = -\frac{\varepsilon\zeta}{\mu}\frac{dV}{dx} \tag{8.23}$$

where ζ is the zeta potential at the wall and dV/dx is the voltage gradient of the electric field.

There are two main challenges for electroosmotic pumps.

- First, bubble generation can occur from the large currents in the open channel. Electrolysis and reactions at the electrodes produce ions that can contaminate the sample and generate bubbles, which can block microchannels.
- Second problem with electroosmotic pumps is that an open channel has low stall pressures.

High pressure can be achieved by using very small channels or if the channel is densely packed. In this way some of the disadvantages of electroosmotic pumps are often overcome by controlling particle packing in the channel. Chen et al. [126] developed a monolithic silica matrix EO pump using a sol-gel process and achieved pressures as high as 400 kPa. In another experiment a multi-stage EO pumps was constructed using several porous silica-packed pumping segments in series and was tested with the same driving voltage applied across each of its segments to achieve pressures as high as 10 MPa.

In another experiment a closed-loop two-phase microchannel cooling system using electro-osmotic pumping for the working fluid is presented in Fig. 8.24 [127]. The system consists of an electro-osmotic pump, a microchannel heat exchanger and a heat rejecter. The microchannel heat exchanger is fabricated using plasma etching in a silicon die. The silicon micro machined heat exchanger provides Si-Si compatibility and facilitates multi-chip integration. It also offers a lower thermal resistance. Two-phase forced convection is performed to utilize fluid latent heat and to achieve more uniform temperature distribution on the IC chip. The heat rejecter is a combination of an aluminium fin array heat sink with embedded flow channels and a fan. It is placed at the downstream of the two-phase heat exchanger to condense the two-phase fluid back to liquid state, and to further lower the liquid temperature to a proper temperature for the pump. The cooled liquid is then pumped back into the heat exchanger, thus forming a closed-loop flow system.

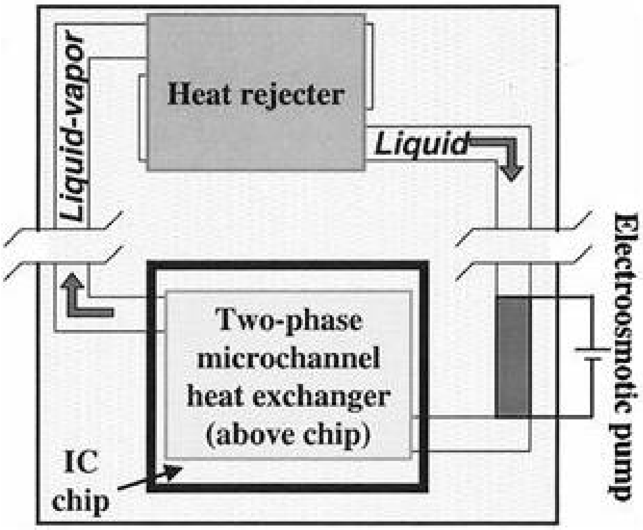


Figure 8.24: Conceptual Schematic of electro-osmotic cooler
[127]

The micro heat exchanger and electro-osmotic pump are extremely compact, and the system can be integrated in a modular fashion with the semiconductor chip (heat source). The entire volume of the whole system is far smaller than that of a heat pipe, a vapor chamber, and a fin-array heat sink capable of removing comparable power from the chip

AC electroosmotic

AC electroosmotic pumps have emerged as a viable alternative when the solution used is conductive or electrolytic. Unlike the deprotonation on the channel surface that happens in DC electroosmotic flow, electrodes in the AC electroosmotic flow are positioned on the channel boundary to provide the charge to establish the electric double layer. The most common geometry used is asymmetric electrodes to induce an electric field and

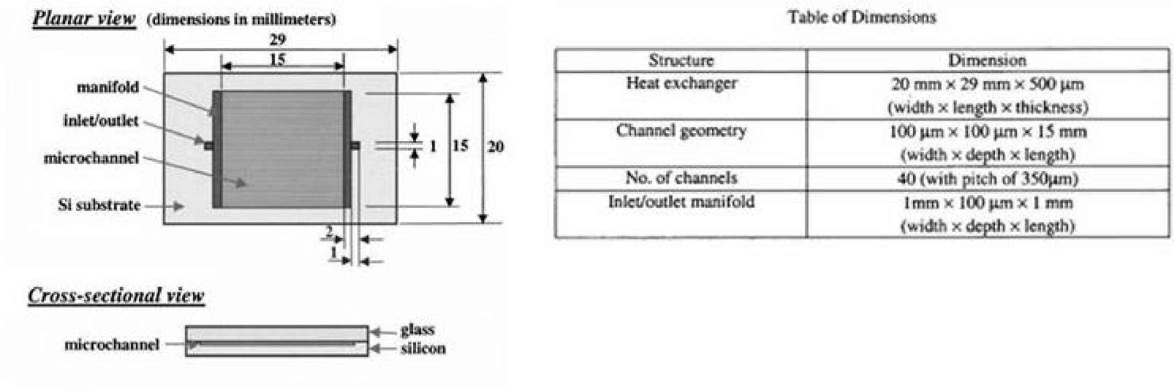


Table of Dimensions

Structure	Dimension
Heat exchanger	20 mm × 29 mm × 500 μm (width × length × thickness)
Channel geometry	100 μm × 100 μm × 15 mm (width × depth × length)
No. of channels	40 (with pitch of 350μm)
Inlet/outlet manifold	1mm × 100 μm × 1 mm (width × depth × length)

Figure 8.25: Schematic of the microchannel heat exchanger
[127]

draw the diffusing layer charges along the surface of the electrodes. Figure 8.26 shows the schematic of an AC electroosmotic micropump [111].

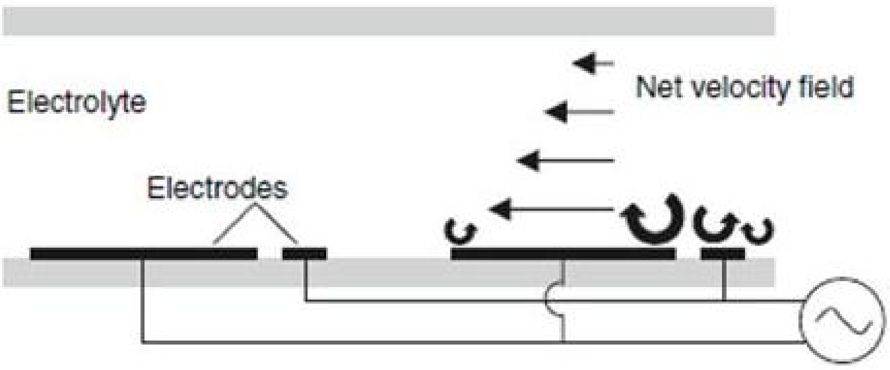


Figure 8.26: AC electroosmotic pump
[111]

AC electroosmotic has the advantage that high velocities of the solution can be achieved for relatively small voltages as described by Debesset et al. [128] and Mpholo et al. [129]. In addition, flow reversal can be achieved by increasing the voltage making these pumps bi-directional as demonstrated by Garcia-Sanchez et al. [130]. Inspite of this research, AC electroosmotic micropumps still remain non commercialized.

8.5.3 Microchannels

As density and the frequency of components in the VLSI industry continues to grow exponentionally, it has also led to the exponential rice of heat generation in IC chips. This has resulted increased heat fluxes that need to be dissipated to maintain the chip operating temperature within specified limits. Conventional cooling methods like air and fan cooling are in no position to satisfy the requirements for high heat flux dissipation requirements. The need for new technologies and techniques for thermal management of these IC chips has become a major and important research area. Many any new innovative techniques have been demonstrated and thoroughly reviewed in literature [133]. Liquid cooling based microchannels are fast becoming the most favourable and effective cooling methods. May other cooling methods such as spray cooling, thermoelectrics, microjets, and thin-film evaporation are being researched but they are either yet to be developed for commercial applications or their implementations suffer from noise, efficiency, or cost issues. Microchannels consist of closed parallel channels with rectangular, trapezoidal, or triangular cross sections having diameters ranging from 100 to 1000 μm. There are two main implementation methods for microchannels: Single-phase flow, in which the heat is transferred from the die via a sensible heat gain by the coolant. In the second method: Two-phase flow, where the latent heat of the coolant is utilized during a phase change (liquid - vapor). In microchannels, the heat transfer in single-phase flow through a heat sink can be calculated using

$$q = N_u * k * \Delta T * \frac{A}{D_k} \tag{8.24}$$

where N_u is the Nusselt number, k is the thermal conductivity of the liquid, ΔT the temperature difference between the die substrate and the operating liquid, A is the area, and D_h is the diameter of the microchannel. The Nusselt number is given by:

$$N_u = \frac{hl}{\kappa} \tag{8.25}$$

Where h is the convective heat transfer coefficient in W/m^2K, L is the characteristic length of the microchannel and κ is the thermal conductivity of the working fluid.

The Nusselt number is a nondimensional temperature gradient at the wall and is a measure of convection heat transfer. The Nusselt number is a constant for fully developed flow in a channel of given shape. The "Fundamentals of Heat and Mass transfer gives the Nusselt number for rectangular channels with a width aspect ratio of 4 with uniform heat flux on all four walls of the channel as $N_u = 5.33$.

Garimella et al [104], describe for a microchannel with water and a driving temperature difference of 30^{o}C between the substrate and water, 433 W of heat can be removed from a chip of dimensions 1 cm × 1 cm using a microchannel with dimensions 100μm-wide by 400μm-deep and separated by 50μm-thick walls. Larger heat transfer rates can be achieved when the coolant is allowed to boil in the channels. The specific heat of water is 4.18 kJ/kg K, while its latent heat is 2400 kJ kg^{-1}. Therefore, for the same mass flow rate, ten times more heat can be removed via boiling (Two phase) than with single-phase convection (allowing a 50^{o}C streamwise temperature rise of liquid water). Conversely, if the heat load is unchanged, only one-tenth the flow rate of coolant would be required for a two-phase microchannel heat sink when compared to the single-phase counterpart.

From the above description it is clear that for single-phase convection, much larger coolant flow rates would be required to obtain the same temperature uniformity when compared with two phase flow. There has been a great deal of research and experimental study on microchannels under both single-phase and two-phase. In Fig. 8.27 [132] high heat transfer rates have been obtained as shown.

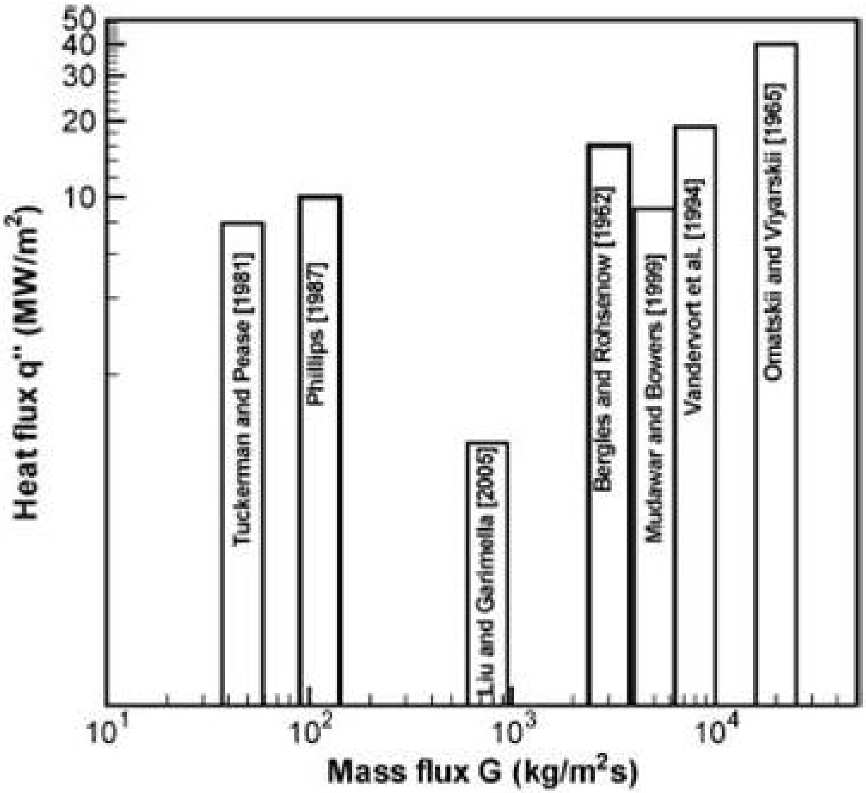


Figure 8.27: : Heat dissipation rates achieved with Microchannels
[132]

Microchannels work on the principle of convection. Heat transfer takes place by the transfer of energy from the substrate die to the coolant. The capacity of a microchannel is proportional to the rate of flow of the coolant. Hence as long as the working fluid or the coolant is able to absorb the heat from the die, the channel is able to dissipate heat. This theoretically implies that there is fundamentally no limit to the heat transfer rate that can be obtained by using the microchannel. The only obstacle is that the convection takes place by a pump and the pumping mechanism determines the heat transfer in an indirect way. This is a major hurdle in case of single phase flow. In case of Two-phase flow, although the flow rate required for the coolant is much lower, the phase transitions between the different boiling regimes in parallel adjacent channels cause instabilities which result in problems and challenges for the pump operation.

Over the last two decades, microchannels utilizing single phase flow have been extensively researched and studies have demonstrated their effectiveness and good modelling is also available. In contrast two phase flow methods have received little attention, despite the fact that they greatly enhance the performance of a microchannel heat sink by providing higher convective heat transfer coefficients, better axial temperature uniformity, and re-

Fluid State	Topic	Isolated Micro-channel	Parallel Micro-Channels
Single-Phase liquid	Laminar: Pressure Drop	Adequate	Adequate
	Laminar: Heat Transfer	Limited	Limited
	Transition	Limited	Limited
	Turbulent: Pressure Drop	Adequate	Adequate
	Turbulent: Heat Transfer	Adequate	Adequate
Two Phase	Incipient Boilng	Limited	Limited
	Flow Patterns	Limited	Limited
	Hydrodynamic instabilities	Very limited	Very limited
	Pressure Drop	limited	Very limited
	Heat Transfer	limited	Very limited
	Critical heat Flux	limited	Very limited

Table 8.6: Present level of understanding of single-phase and two-phase micro-channel

duced coolant flow rate requirements. Table 8.5.3 [134] summarizes the present level of understanding of various aspects of single-phase and two-phase micro-channel heat sinks.

Single phase heat transfer

The phenomenon of heat transfer in microchannels is completely different from the one at macroscale because of the various forces acting in a much different way. The scales and dimensions involved are quite different. Single phase tansport can be characterized in terms of friction factor for pressure drop and Nusselt number for heat transfer as a function of the Reynold's R_e number which is

$$R_e = \frac{u_m D_h}{V} \tag{8.26}$$

a measure of the relative importance of inertial and viscous forces, in which u_m is the mean velocity of the bulk fluid, D_h the hydraulic diameter of the microchannel, and V the kinematic viscosity of the fluid [104].

As mentioned earlier. Pumping requirements for microchannels are high. Garimella et al. [135] define two primary thermal considerations in the design of microchannel heat sinks. These conditions relate the limit on the maximum temperature and the maximum temperature gradient on the chip and are expressed as:

The maximum temperature at any point on the chip should be less than T_{max}, i.e.
$T_d < T_{max}$
The maximum temperature gradient on the chip is limited to $(\frac{dT}{dx})_{max}$ i.e

$(\frac{dT}{dx})_d < (\frac{dT}{dx})_{max}$

They also give the minimum flow rate and the pressure head from the limit on maximum temperature gradient. The limit on maximum spot temperature and also the flow rate versus pressure drop characteristics of a microchannel heat sink of given dimensions is provided and their interdependencies are explained.

Two Phase Transfer

Convective boiling and two-phase flow in small-scale channels has also been widely studied. The complex nature of convective boiling flow has been a major factor for the technology not ebing implemented commercially. In this part two-phase flow and heat transfer research at the microscale for electronics cooling is discussed.

Boiling implies the appearance of vapor in the liquid coolant as the phase change is taking place. Depending on the heat, the intermediate stage of liquid-vapor mixture assumes various morphological structures. These structures in turn affect the flow patterns. Typically, Five different flow patterns:
bubbly, slug, churn, annular, and annular-mist flows can be identified as shown in Fig. 8.28 [132] below. For bubbly, slug, and churn flows, results on the left are for macroscale tubes while those on the right are for smaller tubes with inner diameters.

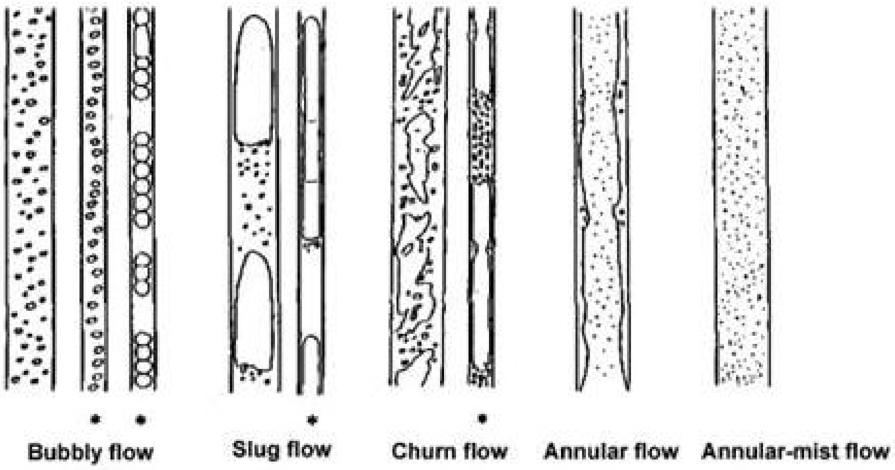


Figure 8.28: Flow Patterns
[132]

A veriety of approaches are present in literature to study the flow patterns. One approach is to study the flow pattern by introducing a mixture of air and water in capillary tubes under adiabatic conditions as demonstrated by Serizawa et al. [136] Another approach is to generate this two-phase flow by heating the coolant in the microchannels. This approach comes close to the actual way heat transfer takes place in microchannels, but research into this method is only recent,

There are two important features of two-phase flow in microchannels as mentioned by

Garimella et al. [132] First, due to the dominance of surface tension over viscous and inertial forces at the microscale, the microchannel flow patterns exhibit significant differences from those in larger channels, such as the absence of stratified flow and the unique structures of bubbly and slug flows. Second, under diabatic conditions (heating the coolant), two-phase flow in microchannels is highly transient and is characterized by the intermittent appearance of different flow patterns. Hence, information from macroscale two-phase flow pattern studies cannot be directly applied to microchannels.

A number of experiments have been carried out to study presssure drop and heat transfer under two-phase flow in microchannels, and a variety of correlations and models exist. These correlations can be divided into two groups: the homogeneous flow model and the separated flow model. The homogeneous flow model assumes that two-phase flow behaves like single-phase flow with averaged fluid properties (weighted means of properties of the liquid and vapor phases). This model is particularly valid for bubbly flow. The separated flow model treats the two phases separately assuming that the velocity of each phase is uniform but not necessarily equal to that of the other phase, and that the two phases are in local thermodynamic equilibrium. This model is more accurate for the slug flow. Weilin Qu [134] gives the mathematical expressions for the homogenous and separated model. The separated model is treated in two different correlations Martinelli-Nelson correlation and Lockhardt-Martinelli correlation. The expressions are presented below. The Martinelli-Nelson correlation and Lockhart-Martinelli correlation share the same principles in their development. However, the former is is based on the assumption of a turbulent liquid-turbulent vapor flow combination because in macrochannels this type of flow is more prevalent, the latter on the other hand is able to account for different combinations of liquid-vapor flow i.e turbulent-turbulent, turbulent-laminar, laminar-turbulent, and laminar-laminar.

8.5.4 Thermoelectric Coolers

The primary motivation for using thermoelectric coolers is the fact that they allow ICs to be cooled below ambient temperatures. The present levels of power dissipation and power density require innovative cooling solutions and that is where thermoelectric coolers find their place. Thermoelectric coolers can be used to manage cooling of the entire chip area similar to a hot sink or heat spreader and also be utilized to efficiently only hot spots on the chip.

The thickness of the thermoelectric material used determines the amount of cooling possible. This thickness is a limitation or a determining factor to decide the usage of a thermoelectric cooling solution. Since thermoelectric cooling can be utilized for cull chip cooling as well as hot spot cooling the efficiency or the Coefficient of Performance (COP) is an important parameter that should not be overlooked when the thermoelectric material is being used for full chip cooling. In the other case this is not critical as the thermoelectric device will handle only a very small area for cooling.

Sharp et al [109] give the mathematical equations that are relevant in determining the performance and efficiency of thermoelectric coolers. The amount of heat that a thermoelectric cooler is able to dissipate is determined by the contributions of the following three factors

Homogenous Equilibrium Model (HEM)	$\Delta P_{tp,f} = \frac{2f_{tp}G^2L_{tp}v_f}{d_h}\ [1 + \frac{x_{e,out}}{2}(\frac{v_{fg}}{v_f})]$ $f_{tp} = 0.003$ $\Delta P_{tp,a} = G^2 v_f x_{e,out}$
Martinelli-Nelson (M-N) Correlation	$\Delta P_{tp,f} = \frac{2f_{tp}G^2L_{tp}v_f}{d_h}\ r_1(x_{e,out}, P)$ $\Delta P_{tp,a} = G^2 v_f x_{e,out}, P$
Lockhardt-Martineli(L-M) Correlation	$\Delta P_{tp,f} = \frac{L_{tp}}{x_{e,out}} \int_0^{x_{e,out}} \frac{2f_fG^2(1-x_e)^2v_f}{d_h}\ \phi_f^2 dx_e$ $f_f = \frac{24(1-1.335\beta+1.947\beta^2-1.701\beta^3+0.956\beta^4-0.254\beta^5)}{\frac{G(1-x_e)d_h}{\mu_f}}$ $\phi_f^2 = 1 + \frac{C}{X_{vv}} + \frac{1}{X_{vv}^2}$ $C = 5$ $X_{vv} = (\frac{\mu_f}{\mu_g})^{0.5}(\frac{1-x_e}{x_e})^{0.5}(\frac{v_f}{v_g})^{0.5}$ $\Delta P_{tp,a} = G^2 v_f[\frac{x_{e,out}^2}{\alpha_{out}}(\frac{v_g}{v_f}) + \frac{(1-x_{e,out})^2}{1-\alpha_{out}} - 1]$ $\alpha_{out} = 1 - \frac{1}{\sqrt{1+\frac{20}{X_{vv,out}}+\frac{20}{X_{vv,out}^2}}}$

Table 8.7: : Correlation models for pressure drop

[134]

1. the Peltier heat pumping, which is opposed by
2. conduction across the thermoelectric elements and
3. resistive heating within the elements.

$$Q_c = N[\overbrace{IT_c(\alpha_p - \alpha_n)} - \overbrace{(T_h - T_c)K} - \frac{1}{2}I^2R] \tag{8.27}$$

where $R = \rho_p(\frac{L}{A}) + \rho_n(\frac{L}{A})$
and $K = \lambda_p(\frac{A}{L}) + \lambda_n(\frac{A}{L})$

The geometry of the thermoelectric material is characterized by the number N of the p or n couples, the length L and the area A of the thermoelectric material. The voltage V that is generated when a current I is applied across a TEC is the sum of the IR part, and the product of the difference between the Seebeck coefficients (α) of the n and p materials and the difference of temperature between the hot and cold junctions as shown below

$V = N[(\alpha_p - \alpha_n)(T_h - T_c) + IR]$ The total input power then is just the total voltage V multiplied by the input current I as shown below

$W = NI[(\alpha_p - \alpha_n)(T_h - T_c) + IR]$

The Coefficient of performance or COP is then determined as the ratio of the heat pumping capacity to the input power.

$$COP = \frac{Q_c}{W} \tag{8.28}$$

$$COP = \frac{[IT_c(\alpha_p - \alpha_n) - (T_h - T_c)K - \frac{1}{2}I^2R]}{I[(\alpha_p - \alpha_n)(T_h - T_c) + IR]} \tag{8.29}$$

Thus the overall performance of the thermoelectric element which is known as the figure of merit for the element and denoted by Z is given by

$$Z = \frac{\alpha^2}{\rho\lambda} \tag{8.30}$$

For a given ratio of L and A, there is an optimum current depending on the temperatures of the hot and cold sides. It is this current that gives the maximum COP as given below

$$I_{opt} = \frac{[(\alpha_p - \alpha_n)(T_h - T_c)]}{R\sqrt{ZT_{avg} + 1} - 1} \tag{8.31}$$

As seen, the optimum current depends upon the properties of the TEC material. Using the expression for the COP, and I_{opt}, one can obtain the optimum COP as below

$$COP_{opt} = \frac{T_c[\sqrt{ZT_{avg} + 1} - \frac{T_h}{T_c}]}{(T_h - T_c)[1 + \sqrt{ZT_{avg} + 1}]} \tag{8.32}$$

It is extremely important that this optimum COP is known and kept in mind when designing for a thermoelectric cooler. As it is clear, a suboptimum COP will not be able to give the performance expected from a thermoelectric cooler.

It is extremely important to use TECs that operate near this optimum COP. Selection and use of a suboptimum TEC will result in performance significantly poorer than that of an optimum TEC. Figure 8.29. shows the dependence of the COP with respect to the temperature and the figure of merit.

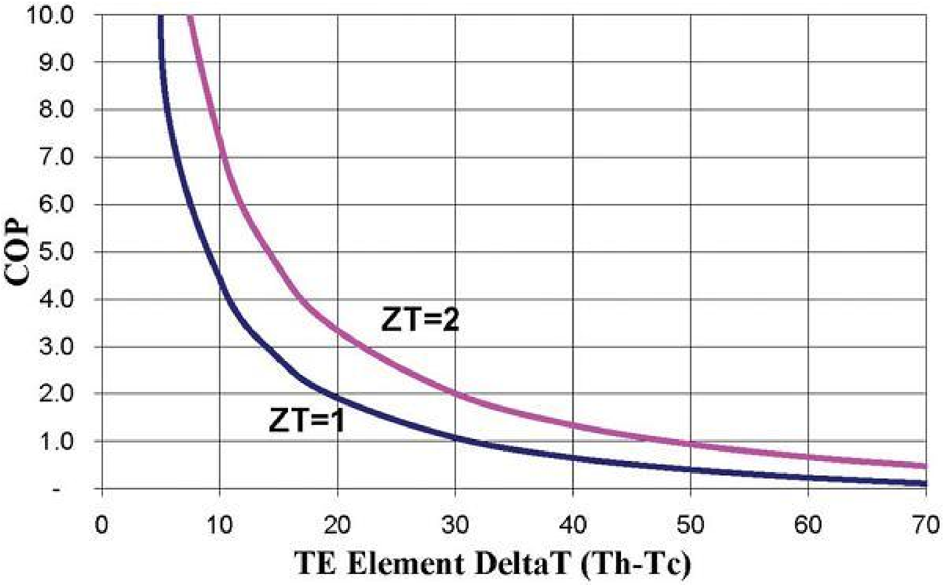


Figure 8.29: Optimum COP Vs. temperature difference between hot and cold sides of the thermoelectric element

[45]

Nanoscale Thermoelectrics

Dresselhaus et al [140] showed that electrons in low-dimensional semiconductors such as quantum wells and wires have an improved thermoelectric power factor and Z greater than 2 -3 can be achieved as shown in Fig 8.30 below

This increase in the figure of merit is due to the fact that due to the electron confinement, its motion perpendicular to the potential barrier is quantized. This quantum confinement eliminates some energy states the electrons occupy as they do not obey the boundary conditions. The main advantage of this confinement is that it changes the energy of the band edge of the material.thereby creating sharp features in the DOS. These sharp features can be utilized to increase the asymmetry between the hot and cold regionsand obtain a large number of carriers.

Hicks et al [141] have rigorously demonstrated the this concept of low dimensional thermoelectrics but so far no considerable increase utilizing the confinent principle for the figure of merit has been shown. Shakouri [45] mentions two important reasons and are quoted: First, we live in a 3-D world and any low dimensional quantum well structure should be imbedded in barriers. These barriers are electrically inactive but they add

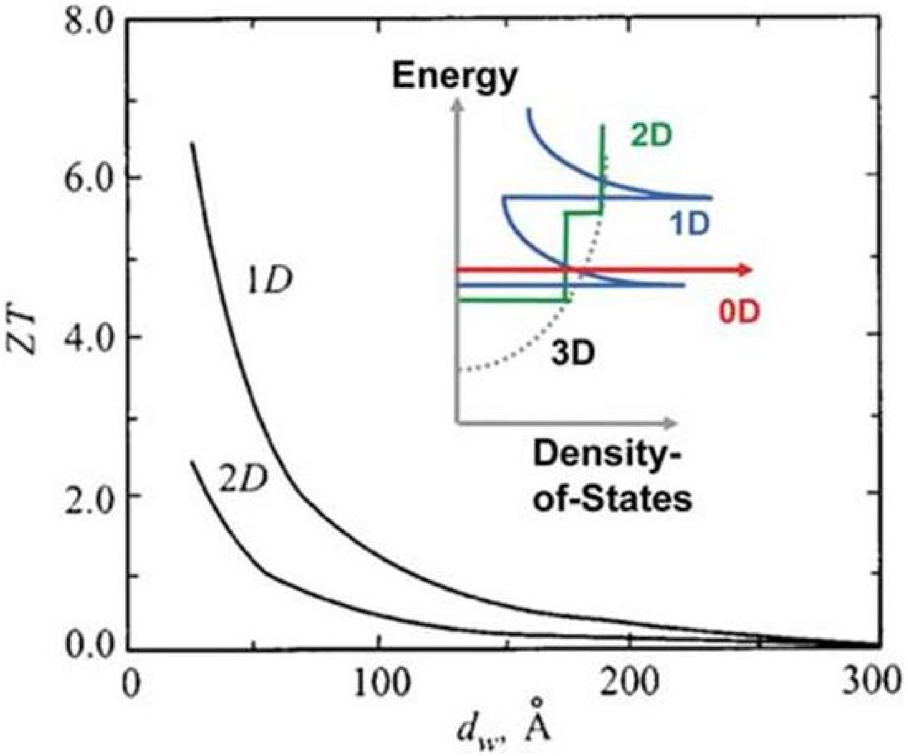


Figure 8.30: Thermoelectric figure of merit Z and DOS for 1-D 2-D 3-D and 0-D structures [45]

to thermal heat loss between the hot and the cold junctions. This can reduce the performance significantly [50]. One cannot make the barrier too thin, since the tunneling between adjacent quantum wells will broaden energy levels and reduce the improvement due to the DOS.

The second reason is that the sharp features in the DOS of 2-D nanostructures disappear quickly as soon as there is size nonuniformity in the material. A natural extension of quantum wells and superlattices is to quantum wires [143], [144], [145]. Theoretical studies predict a large enhancement of the figure of merit Z inside quantum wires due to additional electron confinement as seen in Fig 8.30 [45] AS yet, there are no significant claims of enhancement of thermoelectric properties due to quantum wires mainly because making a contact to a nanowire and doing thermoelectric measurements is not easy with all the technological advancements yet.

Subramanian et al [146] have suggested quantum dots (0-D) as an extension to the low dimensional thermoelectrics. The theory developed by Dresselhaus [140], limits itself to 1-D and 2-D structures and is not extended to 0-D structures. The power factor enhancement in1-D and 2-D confinement happens perpendicular to the direction of confinement. In case of 0-D, electrons have to move between the dots in order to transfer heat from one location to another. If the electronic bands in the dots are very narrow, then electrons are highly confined and it is not easy to take them out of the dots. On the other hand, it is easy to take electrons out of shallow quantumdots but at the same time the density of electronic states in the dot will have broad features all the details. Readers are advised to refer to reference [45] of this report and the corresponding references. There is abundant literature available on this topic and it is beyond the scope of this report to go into all the details.

8.5.5 Conclusions

In this report, an overview of thermal materials, thermal metrics and thermal management method for cooling IC chips and packages is presented which aids to understand the thermal management techniques explained. Active and passive cooling forms the two major branches of thermal cooling in Microelectronic devices and a few handpicked innovative and interesting examples have been discussed.

Table 8.31 [143] sums up the thermal management techniques we have taken a look and what is in store in the near future.

In addition to the ones described and to the ones mentioned in table 8, a few other types of refrigeration systems are also present, but they find only a brief mention for reasons cited below:

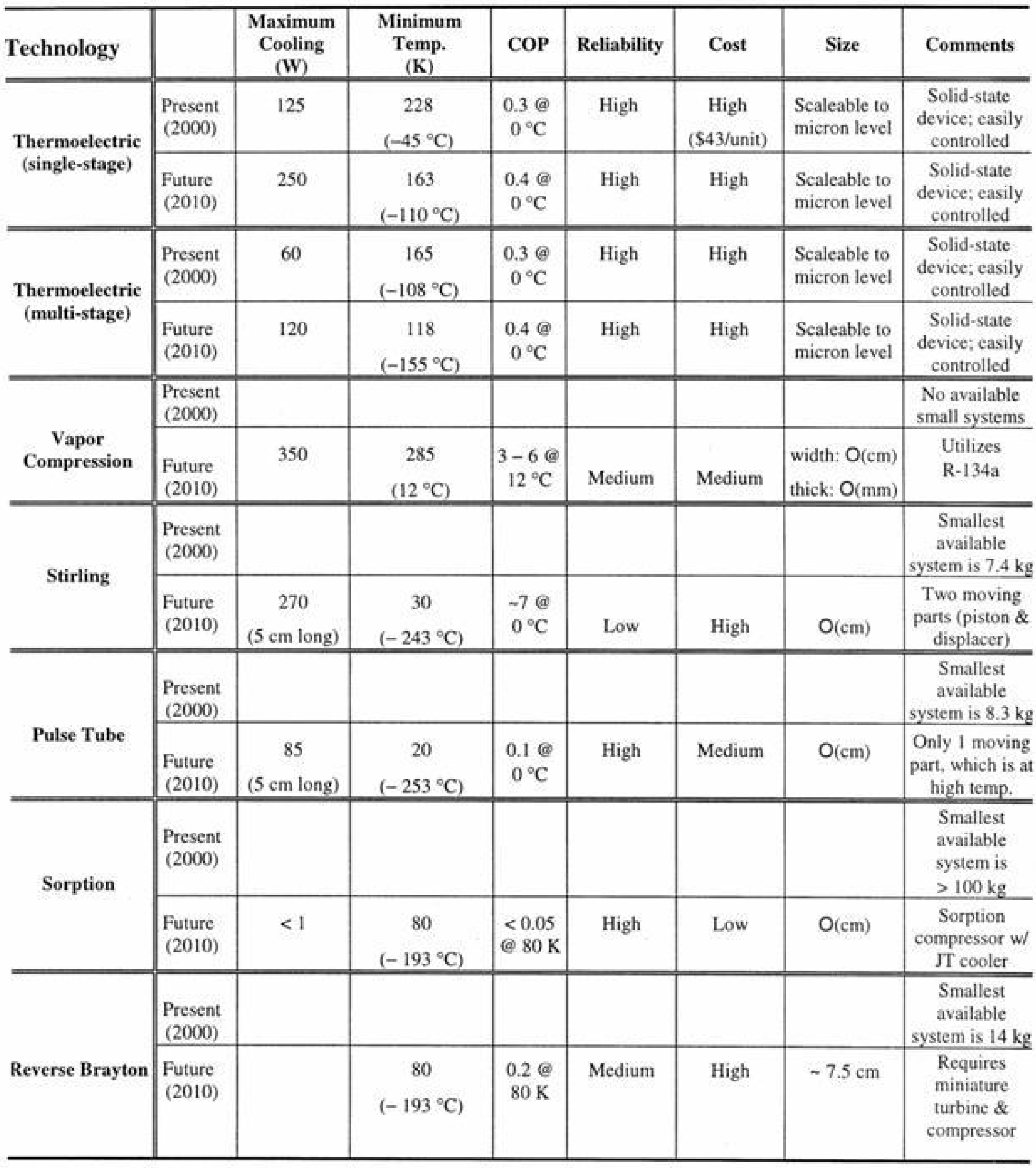
Thermionic refrigerators are a class of refrigerators that are related to thermoelectric refrigerators, and also contain no moving parts. Although an area of active research practical devices have not been implemented yet with this principle. [147] - [152]

Another class of refrigeration technique is what is known as electron field emission [153], but this concept is even less developed.

Cooling methods involving the Joule-Thompson refrigerators require the development of an efficient, miniature compressor capable of delivering particularly high pressure ratios, the unavailability of technology to implement such compressors has made cooling by this method not yet feasible.

Finally, a class of refrigerators that utilize thermo-acoustic phenomenon, which is similar in principle to pulse tube refrigerators, are unattractive option in that the piston is replaced by an acoustic generator, which eliminates the moving parts [154]. However, only relatively small cooling powers have been achieved so far. The primary reason for this being that the acoustic properties required for efficient cooling are being achieved at the sub micron scale.

The constraint on a cooling technique for heat removal is a major factor that limits its performance. The thermodynamic limit of performance for a thermal management system utilizing air cooling to reject thermal energy varies inversely as its mass flow-heat capacity product. The minimum resistance offered by a cooling solution to heat flow is refined by taking into account the effects of thermal interface materials, substrate materials, and the impact the non uniform device layer heating has on it. In this last part of the report the minimum efficiency that an active cooling solution must provide to offer a thermal advantage over passive cooling is described. This depends on the thermal resistances involved in drawing heat into the active cooler and expelling it out to the ambient environment. Engineers and designers with knowledge of these fundamental limits of thermal solutions are in a much better position to set roadmaps, define architecture



Technology		Maximum Cooling (W)	Minimum Temp. (K)	COP	Reliability	Cost	Size	Comments
Thermoelectric (single-stage)	Present (2000)	125	228 (−45 °C)	0.3 @ 0 °C	High	High ($43/unit)	Scaleable to micron level	Solid-state device; easily controlled
	Future (2010)	250	163 (−110 °C)	0.4 @ 0 °C	High	High	Scaleable to micron level	Solid-state device; easily controlled
Thermoelectric (multi-stage)	Present (2000)	60	165 (−108 °C)	0.3 @ 0 °C	High	High	Scaleable to micron level	Solid-state device; easily controlled
	Future (2010)	120	118 (−155 °C)	0.4 @ 0 °C	High	High	Scaleable to micron level	Solid-state device; easily controlled
Vapor Compression	Present (2000)							No available small systems
	Future (2010)	350	285 (12 °C)	3 – 6 @ 12 °C	Medium	Medium	width: O(cm) thick: O(mm)	Utilizes R-134a
Stirling	Present (2000)							Smallest available system is 7.4 kg
	Future (2010)	270 (5 cm long)	30 (– 243 °C)	~7 @ 0 °C	Low	High	O(cm)	Two moving parts (piston & displacer)
Pulse Tube	Present (2000)							Smallest available system is 8.3 kg
	Future (2010)	85 (5 cm long)	20 (– 253 °C)	0.1 @ 0 °C	High	Medium	O(cm)	Only 1 moving part, which is at high temp.
Sorption	Present (2000)							Smallest available system is > 100 kg
	Future (2010)	< 1	80 (– 193 °C)	< 0.05 @ 80 K	High	Low	O(cm)	Sorption compressor w/ JT cooler
Reverse Brayton	Present (2000)							Smallest available system is 14 kg
	Future (2010)		80 (– 193 °C)	0.2 @ 80 K	Medium	High	~ 7.5 cm	Requires miniature turbine & compressor

Figure 8.31: Comparison of Active Cooling methods
[45]

specifications, and evaluate the validity of thermal system performance claims. These limits to be known are

- **Thermal Resistance**
  The performance of a thermal solution is characterized by thermal resistance which is the ratio of the temperatures between the two locations of interest where energy transfer takes place. This ratio is given by

$$R_{ja} = \frac{(T_j - T_a)}{Q} \tag{8.33}$$

Where T_a is the temperature of the ambient environment of the device and T_j may be the maximum or average temperature of the device layer of an integrated circuit, or any other point of interest that can be estimated experimentally, analytically, or numerically for design purposes

- **Air side limit**
  As we seek to minimize thermal resistance an examination of its minimum provides a useful guide in developing road maps setting design goals. A knowledge of the limits of thermal resistance provides the designers with estimations to quantify the limits to which a current thermal solution technology may be extended, and assist in decisions to implement alternative cooling techniques (such as active cooling).

  - **Simple single fluid thermal system**
    In this case a simple heat sink cooled by a stream of air can be considered as a solid-fluid heat exchanger. Traditionally, heat exchanger design methodology defined the effectiveness of an exchanger as the ratio of the heat removed to the maximum theoretical heat that may be removed. For a solid-fluid heat exchanger, this effectiveness may be expressed in terms of the temperatures of the incoming air stream, exit stream, and heat source i.e

    $$\epsilon = \frac{T_h - T_a}{T_j - T_a} \tag{8.34}$$

    If a thermal solution is highly effective at heating the air stream to the junction temperature T_j, effectiveness approaches unity. On the other hand, if it poorly heats the air stream, the effectiveness goes to zero.The actual rate that energy is gained by the fluid stream, Q , must be the same as the rate that energy passes from the device to be cooled, as required by the first law of thermodynamics [158]

    $$Q = mC_p(T_h - T_c) \tag{8.35}$$

    In equation 8.35, m is the mass flow rate and C_p is the heat capacity at a constant pressure of the fluid.

    From equations 8.34 and 8.35, the thermal resistance is given by

    $$R = \frac{T_j - T_a}{Q} = \frac{1}{mC_p\epsilon} \tag{8.36}$$

    The thermodynamic minimum of thermal resistance is then found when $\epsilon = 1$

  - **Two fluid thermal system**
    Thermal solutions may employ two separate fluid streams. A thermal reservoir may act as an ambient air surrounding and a pump may act as a fan providing a stream of air to a fluid-fluid heat exchanger. Another pump may drive a single or two phase flow of fluids such as water, liquid metal, etc, hence transferring heat between the fluid-fluid heat exchanger and the fluid-solid heat exchanger. The performance of each heat exchanger may be described by thermal effectiveness, and is a function of the heat capacity rate of the

respective fluid streams. Thermal effectiveness can be defined as the ratio of the actual temperature difference achieved to the maximum achievable. i.e

$$\epsilon_{ff} = \frac{(mC_p)_2(T_{fh} - T_{fc})}{(mC_p)_1(T_{fh} - T_a)} \tag{8.37}$$

and

$$\epsilon_{sf} = \frac{T_{fh} - T_{fc}}{T_j - T_{fc}} \tag{8.38}$$

Hence from equation 8.33, 8.37 , 8.38 we can define the thermal effectiveness of the exchanger as

$$R = \frac{T_j - T_a}{Q} = \frac{1}{m_1 C_{p1}\epsilon_{ff}} + \frac{1}{m_2 C_{p2}\epsilon_{sf}} - \frac{1}{m_2 C_{p2}} \tag{8.39}$$

When $\epsilon_{ff} = \epsilon_{sf} = 1$, the minimum thermal resistance is obtained which is

$$\Theta = \frac{1}{mC_p} = \Theta_{air} \tag{8.40}$$

which is the fundamental limitation of a thermal resistance for a two fluid thermal system.

- **Source side limit**
  The path for heat flow, for example from the device layer of an IC to the thermal heat sink will inadvertantly include many elements that impede the flow of heat. We shall briefly take a examine these and see their limitations

  - **Thermal interface material resistance**
    As the name suggests, this is the thermal impedance of the TIM. The minimum value of which is given by

    $$\Theta_{tim} = \frac{r_{tim}}{A} \tag{8.41}$$

  - **Semiconductor die resistance**
    The die on which the circuit is designed and fabricated itself has some thermal resistance. This minimum resistance is given by

    $$\Theta_{die} = \frac{t}{A\lambda} \tag{8.42}$$

    The fundamental limit for the total resistance including the air, the TIM and the die can be defined as

    $$\Theta = \Theta_{air} + \Theta_{tim} + \Theta_{die} = \frac{1}{mC_p} + \frac{r_{tim}}{A} + \frac{t}{A\lambda} \tag{8.43}$$

  - **Hot spot limitations on a minimum thermal resistance**
    Non-uniform heat distribution and/or creation causes hotspots. The actual power distribution on a microelectronic chip can be modeled as two components, 1. A hot spot of uniform power density and radius a and 2. a uniform

background power density. With this, the total rise in temperature is the superposition of the individual temperature rises due to each. Since we consider the source to be circular, the temperature rise may be approximated by considering the die to extend infinitely laterally. The maximum temperature rise due to 3-D heat flow from a circular heat source of radius a into material of thickness of infinite lateral extent is

$$\Delta T_{spot} = \frac{q_{spot} t}{\lambda}(\frac{a}{t} + 2 - \sqrt{(\frac{a}{t})^2 + 4}) \tag{8.44}$$

We know the die resistance is given by eq **??**. If we take into account, the die now has a hot spot, the resistance of such a uniformly heated die $\gamma = R/R_{1D}$ can be expressed in terms of the ratio of the heat flux within the area of the spot q_{spot} to the heat flux on the total die q^{-}

$$\gamma = \phi(\frac{a}{t} + 2 - \sqrt{\frac{a^2}{t} + 4} - \frac{\pi a^2}{A}) + 1 \tag{8.45}$$

$$\phi = \frac{q_{spot}}{q^{-}} \tag{8.46}$$

With the above, the fundamental limit defining the thermal resistance of the substrate with the hotspots can be given by

$$\Theta_{limit} = \Theta_{air} + \Theta_{tim} + \gamma \frac{t}{A\lambda} \tag{8.47}$$

- **Impact on system design**

  - **Air side vs. Source side**
    Using the relations provided above, one can determine the fundamental limit for a complete system. The ratio of the fundamental limit to air side to the total fundamental limit can be expressed as

    $$\chi = \frac{\Theta_{air}}{\Theta_{limit}} = \frac{\Theta_{air}}{\Theta_{air} + \Theta_{tim} + \gamma \frac{t}{A\lambda}} \tag{8.48}$$

    A system with $\chi > 1/2$ is air side limited. As χ approaches to 1, the system is strongly limited by the thermodynamic limit of the air flow available to the system. In such a scenario, the total thermal resistance of the system cannot be reduced by any heat spreading enhancements.
    As χ approaches to 0, it is beneficial to include head spreading enhancements to take advantage of the thermal mitigation at the die level since the system now is limited by the source side.

  - **Maximum power implication**
    Power dissipation also impacts the limits of thermal resistance. From 8.33, restating it, we have

    $$Q_{max} = \frac{T_j - T_a}{\Theta_{limit}} \tag{8.49}$$

The temperature budget given by $T_j - T_a$ and the fundamental limit to thermal resistance Q_{limit}, allow the designer or system architect to evaluate the theoretical maximum power handling capacity of a given thermal system design.

- **Active Cooling**

  In cases where passive thermal solutions are unable to provide the performance required by an application, active cooling solution become an attractive solution. The limitations of passive cooling techniques can be included in the analysis of active cooling systems to enable one to evaluate the costs and benefit of moving to active cooling.

  The active cooler requires input power, W, and draws in energy at a rate of Q_c. Thermal resistances on the source side may include hot spot spreading, die, thermal interface material resistance or thermal impedances of other origin. The operation of active cooling systems may be characterized by the coefficient of performance, η, and an efficiency, μ, normalized to the maximum possible coefficient of performance, η_c, or Carnot Limit [158] as

$$\eta_c = \frac{T_c}{T_h - T_c} \tag{8.50}$$

$$\eta = \frac{Q_c}{Q_h - Q_c} = \mu.\eta_c = \mu(\frac{T_c}{T_h - T_c}) \tag{8.51}$$

  Heat flow is further impeded by the resistances R_s and R_a which are the thermal resistances of the source and the ambient environment. The thermal resistance of a system without active cooling can be expressed as

$$R = R_a + R_s \tag{8.52}$$

  With an active cooler now incorporated, the thermal resistance R_{active} is derived from eq 8.51 and from the energy conservation, $Q_c = (T_j - T_c)/R_c$ and $Q_c + W = (T_h - T_a)/R_a$, is given by

$$R_{active} = (T_j - T_a)/Q_c \tag{8.53}$$

  With this the minimum efficiency of a system utilizing an active cooler can be expressed as

$$\mu_{min} = \frac{T_j - T_a}{(\frac{R_s}{R_a})T_a + T_j} \tag{8.54}$$

  With the help of the above equation one is able to determine the type of active thermal solution that offers a advantage over passive cooling, and along with other factors like price, manufacturablility, market data, etc. allows one to evaluate the merits of moving from passive to active cooling.

  In case of thermoelectric cooling, the performance is measured by a dimensionless figure-of-merit Z which is given by

$$Z = \frac{S^2\sigma}{\lambda} \tag{8.55}$$

  where S is the materials Seebeck coefficient, σ is the electrical conductivity and λ is the thermal conductivity. The efficiency, normalized to the Carnot coefficient of performance, η_c, can be expressed as a function of $Z$and the temperatures of the hot and cold sides of the cooler as

$$\mu = \frac{\eta}{\eta_c} = \frac{\sqrt{1+(ZT)} - \frac{T_a}{T_c}}{\sqrt{1+(ZT)}+1} \tag{8.56}$$

where T denotes the mean temperature of the hot and cold regions.

Chapter 9

Conclusions

The objective of this thesis was address the the refrigeration of solid state devices and with solid state devices. In this regard we have presented a study of the state of the art cooling methods in the Microelectronic industry as well as refrigeration at cryogenic temperatures.
We also investigate the energy relaxation mechanisms of quasiparticles in a superconducting mesoscopic wire which is driven out of equilibrium in this regard we conduct an experiment in which we study the energy relaxation mechanisms of the nonequilibrium quasiparticles which heps us to understand the physical processes involved that cause heating of the superconducting leads ultimately decreasing their efficiency.

From the literature presented, it is clear that better cooling is possible that that is achieved currently. The research in cooling electronic devices is mainly focussed on cooling the complete device, more so its concentrated towards package cooling. This is mainly due to fact that implementing a cooling method directly on the device is a costly and time consuming affair which will require the industry to make major changes in their manufacturing process. Secondly implementing cooling solutions on chip such as hot spot cooling with Peltier coolers requires designers to implement the peltier elements themselves since they are the ones with the knowledge of the design layout and power dissipation. Designers rarely would want to do more than design. Even if these were implemented in the design, major changes would have to be incorporated while fabrication since the cooling elements would require one layer of the routing for themselves. In an industry where routing over multiple metal stacks is highly expensive, this is not an option. Finally among all the varieties seen, there is no solution which would fit all the requirements. Each of the variants of the cooling methods present will fit in an environment based on the size and the required power. More research is required in this area to come up with any ground breaking inventions which fit with the semiconductor industry's manufacturing process, are cheap, reliable, and efficient.

In the area of cryogenic cooling too, efficiency and ease of manufacturing are the main bottlenecks. The working principles of these devices is entirely different from the main stream electronic devices (although some similarities can be drawn with Peltier coolers) and hence their applications are restricted to low temperature implementations.

In the experimental study where we try to address the issue encountered by the cryogenic coolers we observe that inducing non-equilibrium does affect the distribution function of the superconductor which affects the energy gap analogous to heating. This is mainly due to the quasiparticle current since the Josephson current does not have any affect. We conclude that the nonequilibrium affects are observable for quite a large distances. This implies that the diffusion happens over long distances which helps to avoid hotspots. It is clear that efficient extraction of this nonequilibrium is important to maintain the efficiency of the cooler. Placing quasiparticle traps can be a solution.

From the analysis of the charge imbalance we can conclude that, charge imbalance is seen only at high injection currents. This does not have much affect on the efficiency or on the operation of the cooler since the device is operated at energies close to the gap. It is clear that charge imbalance does not manifest when the injection is close to the energy gap.

Bibliography

[1] C. Harmans, *Mesoscopic Physics: An Introduction* .

[2] P.G. DeGennes, *Superconductivity of Metals and Alloys.* Advanced Book Classics.

[3] J. Bardeen, L.N. Cooper and J.R. Schriefer P.G. DeGennes,
Phys.Rev. **106**, 162(1957), **108**, 1175(1957)

[4] C. Kittel, *Introduction to Solid State Physics.* Wiley Inc, New York.

[5] Michael Tinkham, *Introduction to Superconductivity.* McGraw Hill Inc, New York.

[6] T.G.A Verhagen *MSc Thesis: Section 4-3-2.* T.U Delft.

[7] Pieter de Groot *Friendly communications in the clean room*

[8] David Berman *PhD Thesis, Chapter 2.* MIT.

[9] N.W Ashcroft and N.D Mermin *Solid State Physics* Saunders College Publishing.

[10] S. Wind, M.J. Rookes et al *One dimensional electron electron scattering with small energy transfers* Phys. Rev. **57**, 633(1986).

[11] E. Burstein, S. Lundquist *Tunneling Phenomena in Solids* Lectures presented at 1967 NATO Advanced Study Institute, Chapter 20.

[12] G.R Boogard, A.H Verbruggen, W.Belzig, T.M Klapwijk *Resistance of superconducting nanowires connedted to naomal metal leads* Phys. Rev. B. 69:220503(2004)

[13] F. Giazotto, et. al. *Opportunities for mesoscopics in thermometry and refrigeration: Physics and applications* Phys.Rev.B. 78:217(2006)

[14] T.M. Klapwijk J.E. Mooij *Microwave ehnanced superconductivity in aluminium films* Physica B81, 132 (1976)

[15] B.I. Ivlev, G.M Eliashberg *Nonequilibrium excitations in supercondutors in high frequency fields* Jour. of Low temp. physics 10:449 (1973)

[16] T. Kommers, J. Clarke *Measurement of Microwave-enhanced energy0gap in superconduting aluminium by tunneling* Phy. Rev. B 38:19:1019 (1977)

[17] T.C Tredwell, E.H Jacobsen *Phonon induced enhancement of superconducting gap* Phy. Rev. Let. 35:4:244 (1975)

[18] C.S. Owen, D.J. Scalapino *Superconducting state under the influence of external dynamic pair breaching* Phy. Rev. Let. 28:24:1559 (1972)

[19] W.H. Parker *Modified heating theory for nonequilibrium superconductors* Phy. Rev. B. 12:9:3667 (1975)

[20] A.Rothwarf, B.N. Taylor *Measurements of recombination lifetimes in superconductors* Phy. Rev. Let. 19:1:27 (1967)

[21] I.Iguchi *Study of the nonequilibrium state of supercondutors by large quasiparticle injection from an external source* Phy. Rev. B. 16:5:1954 (1977)

[22] M.Nahum, T.M Eiles, J.M Martinis *Electronic microrefrigerator based on a normal-insulator-superconductor tunnel junction* Appl. Phys. Lett 65:24:3123 (1994)

[23] M.M Levio, J.P Pekola *Efficient Peltier refrigeration by a pair of normal metal/insulator/ superconductor junctions* App. Phys. Lett 68:14 (1996)

[24] A.M Clark, N.A Miller *Cooling of bulk material by electron-tunneling refrigerators* App. Phys. Lett. 86:175308 (2005)

[25] R. Leoni, et al. *Electron cooling by arrays of submicron tunnel junctions* Journal of Applied Physics 85:7:3877 (1999)

[26] R.H Parmenter *Enhancement of superconductivity by extraction of norlam carriers* Phy. Rev. Lett 7:7:274 (1961)

[27] R.G. Melton, J.L Kaplan *Superconduting tunnel junction refrigerator* Phy. Rev. B 21:5:1857 (1980)

[28] H. Eggink and J.Janssen, Chapter 3, *Thermal Management, Mechanics of Microelectronics.* Springer Link

[29] M.L Roukes, S.Fritz *Understanding Nanotechnology* Scientific American, Dec 2002

[30] B.Sheu, et al. *Special Issue on Nanoelectronics and Nanoscale Processing, Proc.* IEEE, 91:11:1745-1988 (2003)

[31] Darrel Frear. *Packaging Materials*, Springer Handbook of Electronic and Photonic Materials, Part E, 55.

[32] Carl Zweben, *Ultrahigh Thermal Conductivity Packaging Materials*

[33] D.G. Cahill et al. *Nanoscale Thermal Transport* J. Appl. Phys. 93:2:793-818 (2003)

[34] L.Shi et al. *Design and batch fabrications of probes for sub-100nm scanning thermal microscopy* J. Microelectromech. Systems 10:3:370-378 Sep. 2001

[35] H. K. Lyeo et al. *Profiling thermoelectric power of semiconductor junctions with nanometer resolution* Science, 303:816-818 2004

[36] Y. S. Ju and K. E. Goodson, *Short-time-scale thermal mapping of microdevices using a scanning thermoreflectance technique* J. Heat Transf., Trans. ASME, 120:2:306313, May 1998.

[37] J. Christofferson and A. Shakouri, *Thermoreflectance based thermal microscope* Rev. Sci. Instrum., 76:024903-1024903-6, Jan. 2005.

[38] —— *Thermal measurements of active semiconductor micro-structures acquired through the substrate using near IR thermoreflectance* Microelectron. J.–Circuits Syst., 35:10:791796, Oct. 2004.

[39] E. Pop, S. Sinha, and K. E. Goodson, *BHeat generation and transport in nanometer-scale transistors* Proc. IEEE, 94:8:15871601, Aug. 2006.

[40] E. T. Swartz and R. O. Pohl, *Thermal boundary resistance* Rev. Mod. Phys., 61:605668, 1989.

[41] R. S. Prasher and P. E. Phelan, *A scattering-mediated acoustic mismatch model for the prediction of thermal boundary resistance* J. Heat Transf., Trans. ASME, vol. 123:1:105112, Feb. 2001.

[42] G. Chen, *Phonon transport in low-dimensional structures* Semicond. Semimet., 71:203259, 2001.

[43] A. R. Abramson, C.-L. Tien, and A. Majumdar, *Interface and strain effects on the thermal conductivity of heterostructures: A molecular dynamics study* J. Heat Transf., Trans. ASME, 124:5:963970, Oct. 2002.

[44] C. Strohm, G. L. J. A. Rikken, and P. Wyder, *Phenomenological evidence for the phonon hall effect* Phys. Rev. Lett., 155:901155:904, 2005.

[45] A. Shakouri *Nanoscale Thermal Transport and Microrefrigerators on a Chip* Proc. IEEE 94:8:1613:1638 Aug. 2006

[46] W. Liu and M. Asheghi, *Phonon-boundary scattering in ultrathin single-crystal silicon layers* Appl. Phys. Lett., 84:3819, 2004.

[47] Y. S. Ju, *Phonon heat transport in silicon nanostructures* Appl. Phys. Lett., 87:153106, 2005.

[48] W. Liu and M. Asheghi, *Modeling and data for thermal conductivity for single crystal SOI layers at high temperature* IEEE Trans. Electron Devices, 2006, to be published

[49] D. Li, Y. Wu, P. Kim, L. Shi, P. Yang, and A. Majumdar, *Thermal conductivity of individual silicon nanowires* Appl. Phys. Lett., 83:29342936, 2003.

[50] Mingo and D. A. Broido, *Carbon nanotube ballistic thermal conductance and its limits* Phys. Rev. Lett., 95:096105, 2005.

[51] S. Berber, Y.-K. Kwon, and D. Tomanek, Phys. Rev. Lett., 84:4613, 2000.

[52] M. A. Osman and D. Srivastava, *Temperature dependence of the thermal conductivity of single-wall carbon nanotubes* Nanotechnology, 12:1:2124, Mar. 2001.

[53] C. Yu, L. Shi, Z. Yao, D. Li, and A. Majumdar, *Thermal conductance and thermopower of an individual single-wall carbon nanotube* Nano. Lett., 5:18421846, 2005.

[54] P. Kim, L. Shi, A. Majumdar, and P. L. McEuen, *Thermal transport measurements of individual multiwalled nanotubes* Phys. Rev. Lett., 87:215 502, 2001.

[55] M. Fujii, X. Zhang, H. Xie, H. Ago, K. Takahashi, T. Ikuta, H. Abe, and T. Shimizu, *Measuring the thermal conductivity of a single carbon nanotube* Phys. Rev. Lett., 95:065502, 2005

[56] P.Y.Paik et al. *A Digital-Microfluidic Approach to Chip Cooling* IEEE Design and Test Computers 372-381, Aug 2008

[57] D.B. Tuckerman and R.F.W. Pease, *High-Performance Heat Sinking for VLSI* IEEE Electron Device Letters, 2:5:126-129, May 1981.

[58] S. Mukherjee and I. Mudawar *Smart Pumpless Loop for Micro-Channel Electronic Cooling Using Flat and Enhanced Surfaces* IEEE Trans. Components and Packaging Technologies, 26:1:99-109 Mar. 2003.

[59] L. Jiang et al., *Closed-Loop Electroosmotic Microchannel Cooling System for VLSI Circuits* IEEE Trans. Components and Packaging Technologies, 25:3:347-355, Sept. 2002.

[60] K. Pettigrew et al., *Performance of a MEMS Based Micro Capillary Pumped Loop for Chip-Level Temperature Control*Proc. 14th IEEE Intl Conf. MicroElectro Mechanical Systems (MEMS 01), IEEE Press, 427-430 2001.

[61] S.N. Heffington, W.Z. Black, and A. Glezer *Vibration-Induced Droplet Atomization Heat Transfer Cell for High-Heat Flux Applications* Proc. 8th Intersociety Conf. Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm 02), IEEE Press, 408-412 2002.

[62] A. Shakouri and J. E. Bowers, *Heterostructure Integrated Thermionic Refrigeration* 16th International Conference on Thermoelectrics, 1997.

[63] G. D. Mahan *Multilayer Thermionic Refrigeration* 18th International Conference on Thermoelectrics, 1999.

[64] A. Shakouri, C. LaBounty, P. Abraham, J. Piprek, and J. E. Bowers, *InP Based Thermionic Coolers* 11th International Conference on Indium Phosphide and Related Materials,Davos, Switzerland, May 16-20, 1999.

[65] T. S. Fisher and D. G. Walker, *Direct Refrigeration by Electron Field Emission from Diamond Microtips* ASME International Mechanical Engineering Congress & Exposition, Orlando, Florida, November 5-10, 2000.

[66] M. E. H. Tijani, J. Zeegers, and A. T. A. M. de Waele, *Design, Development, and Operation of a Thermoacoustic Refrigerator Cooling to Below 60 C* Cryocoolers 11, R. Ross Jr, Ed. New York: KluwerAcademic/Plenum Publishers, 2001.

[67] M. M. Yovanovich, E. E. Marotta, *Thermal spreading and contact resistances* Heat Transfer Handbook A. Bejan and A. D. Kraus Eds. Hoboken, New Jersey: Wiley, 2003, 261395.

[68] C. V. Madhusudana, *Thermal Contact Conductance.* New York: Springer-Veralag, 1996.

[69] R. Prasher *Thermal Interface Materials: Historical Perspective, Status, and Future Directions* Proceedings of the IEEE. 94:8:1571-1586, August 2006

[70] R. Mahajan, C.-P. Chiu, and G. Chrysler, *Cooling a chip: A packaging perspective* Proc. IEEE, 94:7, Jul. 2006.

[71] J. Torresola et al. *Density factor approach to representing die power map on thermal management* IEEE Trans. Adv. Packag., to be published

[72] Kai Zhang, Matthew M. F. Yuen *Heat Spreader with Aligned CNTs Designed for Thermal Management of HB-LED Packaging and Microelectronic Packaging* IEEE Electronic Packaging Technology, 2006. ICEPT '06. 7th International Conference, Aug 2006

[73] J. Xu, T.S. Fisher, *Enhanced thermal contact conductance using carbon nanotube arrays* Proc 9th Intersociety Conference on Thermal and thermomechanical Phenomena in Electronic Systems, Vol.2, June. 2004.

[74] K. Zhang, et al. *Thermal Interface Material with Aligned CNT and Its Application in HBLED Packaging* Proc 56th Electronic Components and Technology Conf, San Diego, CA, May. 2006.

[75] K. Verma, D. Columbus, B. Han, and B. Chandran, *Real time warpage measurement of electronic components with variable sensitivity* presented at the 1998 ECTC, Seattle, WA.

[76] H. Huang, C.H. Liu, Y. Wu, S.S. Fan, *Aligned Carbon Nanotube Composite Films for Thermal Management* Advanced Materials, Vol.17 2005

[77] Balram Suman *Modeling, Experiment, and Fabrication of Micro-Grooved Heat Pipes: An Update* App. Mech. Rev 60:107-119 2007

[78] A. Nouri-Borujerdi and M. Layeghi *A Review of Concentric Annular Heat Pipes* Heat Transfer Eng., 26:4558, 2005

[79] T.P. Cotter *Principles and Prospects of Micro Heat Pipes* Proceedings of the 5th Int. Heat Pipe Conference, Tsukuba, Japan, pp. 328332, 1984.

[80] K.K. Tio et al. *Thermal Analysis of Micro Heat Pipes Using a Porous Medium Model* Heat Mass Transfer, 36:2128 2000

[81] J.M. Ochterbeck*Heat Pipes in Heat Transfer Handbook* A. Bejan and A. D. Kraus, eds., Wiley, New York.

[82] A.B. Duncan *Charge Optimization for Triangular Shaped Etched Micro Heat Pipe* HTD Am. Soc. Mech. Eng. 278:110 1994

[83] A.B. Duncan,G.P. Peterson*Charge Optimization for Triangular Shaped Etched Micro Heat Pipe* J. Thermophys. Heat Transfer, 9:2:365367. 1995

[84] B. Suman *On the Fill Charge and the Sensitivity Analysis of a V-Shaped Micro Heat Pipe* AIChE J., 52:9:30413054 2006

[85] Y.P. Chen, P. Cheng *Fractal Characterization of Wall Roughness on Pressure Drop in Microchannels* Int. J. Heat Mass Transfer, 30:1:111 2003

[86] F.M. White*Viscous Fluid Flow* 2nd ed., McGrawHill, New York. 1991

[87] H.Y. Wu P. Cheng*Friction Factors in Smooth Trapezoidal Silicon Microchannels With Different Aspect Ratios* Int. J. Heat Mass Transfer, 46:25192525. 2003

[88] J.S. Suh et al. *Friction in Micro-channel Flows of a Liquid and Vapor in Trapezoidal and Sinusoidal Grooves* Int. J. Heat Mass Transfer, 44:31033109. 2001

[89] P.D. Dunn D.A Reay*Heat Pipes* Pergamon, Oxford, UK 1982

[90] A. Faghri *Heat Pipe Science and Technology* Taylor Francis, Washington, D.C. 1995

[91] C.B Sobham et al *A review and comparative study of the investigations on micro heat pipes* Int. Jour. of Energy Research, Wiley Inter, 31:664-668 2007

[92] L.A Florio, A. Harnoy *Combination Techique for Improving Natural Convection Cooling in Electronics* Int. Journal of Thermal Sciences, 46:11:76-92 2007

[93] Hun Sik Han et al *Heat Sink Design for a Thermoelectric Cooling System* IEEE Thermal and Thermomechanical Phenomena in Electronic Systems, 2008. 11th Intersociety Conference. May 2008

[94] W.A. Khan *The Role of Fin Geometry in Heat sink Performance* Jour, of Elec. Pack., Tran. of the ASME 128:324-330 2006

[95] E. M Sparrow, S.B Vemuri, *Natural Convection Radiation Heat Transfer from Highly Populated Pin Fin Arrays* J. of Heat Transfer, 107:1:190-197 1985

[96] E. M Sparrow, S.B Vemuri, *Orientation Effects on Natural Convection/ Radiation in PinFin Arrays* J. of Heat Transfer, 29:3:359-368 1986

[97] A.I Zografos, J.E Sunderland,*Natural Convection from Pin Fin Arrays* Exp. Ther mal fluid Sci., 3:440-449 1990

[98] A.I Zografos, J.E Sunderland,*Numerical Simulation of Natural Convection from Pin Fins Ar rays* ASME, 157:55-66 1990.

[99] T.S Fisher, K. E Torrance *Free Convection Limits for Pin Fin Cooling* Proceedings, 32nd National Heat Transfer Conference, Balti more, Md., USA, 5:129-138 1997

[100] H. Shaukatullah et al. *Design and Optimization of Pin-Fin Heat Sinks for Low Velocity Applications* Proceedings, 12th Annual IEEE SEMI-THERM Symposium, Austin, Tex., USA, 1996, 151-163

[101] S.S Kang M.F Holahan *Impingement Heat Sinks for Air Cooled High Power Electronic Modules* ASME 303, National Heat Transfer Conf. 1:139-146 1999

[102] A. Bar-Cohen et al.*Characterization of Chippackages Justification, Limitations, and Future* IEEE Transactions on Components Hybrids and Manufacturing Technology, 12:4:724-731 1989

[103] W. Krueger A. Bar-Cohen*Thermal Characterization of PLCC-Expanded Rjc Methodology* IEEE Transactions on Components Hybrids and Manufacturing Technology, 15:5:691-698 1992

[104] S. Zeinab et al. *Optimization and Thermal Performance Assessment of Pin-fin Heat Sinks* Jour. of App. Sci. Res. 3:3:227-235 2007

[105] Xiaojin Wei, Kamal Sikka *Modelling of vapour chamber as heat spreading devices* IEEE Thermal and Thermomechanical Phenomena in Electronics Systems, 2006. The Tenth Intersociety Conference June 2006

[106] Jie Wei et al *Measurement of Vapor Chamber Performance* 19th IEEE Therm Symposium 191-194 2003

[107] J.C. Maxwell *A Treatise on Electricity and Magnetism* 1 3rd edition, reprinted by Dover, New York, 1954

[108] S. Chi *Heat Pipe Theory and Practice* Hemisphere Publishing Corporation, 1976

[109] Yashuhiro et al *Micro Channel Vapor Chamber for high heat Spreading* IEEE Electronics Packaging Technology Conference, 2008. EPTC 2008. Dec 2008

[110] V. Unnikrishnan et al *Silicon/Water Vapor Chamber as Heat Spreaders for Microelectronic Packages* 21st IEEE SEMI-Therm Symposium 2005

[111] B.D. Iversion et al *Recent advances in microscale pumping technologies: a review and evaluation* Microfluid Nanofluid, Springer 5:145-174 2008

[112] T-T. Nguyen et al. *A novel lightweight piezo-composite actuator micropump* Proceedings of SPIE the international society for optical engineering, smart structures and materials 1672:617212 2006

[113] S.C Chenet al *Analysis and experiment of a novel actuating design with a shear mode PZT actuator for microfluidic application* Sensors Actuat APhys 135:1:19 2007

[114] H.T. Chang *Design and modeling of electromagnetic actuator in mems-based valveless impedance pump* Microsyst Technol 13:1112:16151622 2007

[115] Y. Su et al *Analysis and fabrication process of an electromagnetically actuated valveless micropump with two parallel flexible diaphragms* Proc Inst Mech Eng C J Mech Eng Sci 219:9:10071014 2005

[116] S. Yufeng S et al *Electro-magnetically actuated valveless micropump with two flexible diaphragms* Int J Adv Manufact Technol 30:34:215220 2006

[117] A. Geipel et al *A novel two-stage back pressure independent micropump: modeling and characterization* J Micromech Microeng 17:5:949959 2007

[118] L.S Jang et al *A stand-alone peristaltic micropump based on piezoelectric actuation* Biomed Microdevices 9:2:185194 2007

[119] Q. Lin et al *Dynamic simulation of a peristaltic micropump considering coupled fluid flow and structural motion* J Micromech Microeng 17:2:220228 2007

[120] F. Goldschmidtboing et al *A generic analytical model for micro-diaphragm pumps with active valves* J. Micromech Microeng 15:4:673683 2005

[121] E.G Kim et al *A study on the development of a continuous peristaltic micropump using magnetic fluids* Sensors Actuat A Phys 128:1:4351 2006

[122] C.W Huang et al *Pneumatic micropumps with serially connected actuation chambers* J Micromech Microeng 16:11::22652272 2006

[123] S.Y Yanget al *A cell counting/sorting system incorporated with a microfabricated flow cytometer chip* Meas Sci Technol 17:7:20012009 2006

[124] Z. Chen et al *An electro-osmotic micropump based on monolithic silica for micro-flow analyses and electro-sprays* Anal Bioanalyt Chem 382:3:817824 2005

[125] B.D. Iverson et al *Thermally developing electroosmotic convection in rectangular microchannels with vanishing Debye-layer thickness* J Thermophys Heat Transfer 18:4:486493 2004

[126] L. Chen et al *Fabrication and characterization of a multi-stage electroosmotic pump for liquid delivery. Sensors* Actuat B Chem 104:1:117123 2005

[127] Linan Jiang et al. *Closed-Loop Electro-osmotic Microchannel Cooling System for VLSI Circuits* IEEE Transcations on Components and Packaging Technology, 25:3 2002

[128] S. Debesset et al *An AC electroosmotic micropump for circular chromatographic applications* Lab On A Chip 4:396400 2004

[129] M. Mpholo et al *Low voltage plug flow pumping using anisotropic electrode arrays* Sensors Actuat B Chem 92:3:262268 2003

[130] Garcia-Sanchez et al *Experiments on AC electrokinetic pumping of liquids using arrays of microelectrodes* IEEE Trans Dielectrics Electrical Insulation 13:3::670677 2006

[131] Nam-Trung et al *MEMS-Micropumps: A Review* Transactions ASME 124:384-392 June 2002

[132] S.V. Garimella et al *On-Chip Thermal Management With Microchannel Heat Sinks and Integrated Micropumps* Proc. IEEE 91:8:1534-1548 2006

[133] S.V. Garimella *Advances in mesoscale thermal management technologies for microelectronics* Microelectronics Journal 37:1165-1185 2006

[134] Q.Weilin *Transport Phenomena in Two-Phase Micro-Channel Heat Sinks* Journal of Electronic Packaging, ASME, 126:213-224 2004

[135] S.V Garimella, Vishal Singhal *Single-Phase Flow and Heat Transport and Pumping Considerations in Microchannel Heat Sinks* http://docs.lib.purdue.edu/coolingpubs/58; Heat Transfer Engineering 25:1:15-25 2004

[136] A. Serizawa, et al. *Two-phase flow in microchannels* Exp. Thermal Fluid Sci., 26:703714 2002.

[137] "http://en.wikipedia.org/wiki/Heike-Kamerlingh-Onnes"

[138] J. Sharp et al *Overview of Solid-State Thermoelectric Refrigerators and Possible Applications to On-Chip Thermal Management* IEEE Proc.94:8:1602-1612 2006

[139] H. J. Goldsmid *Electronic Refrigeration* London, U.K Pion Ltd., 1985.

[140] L. D. Hicks and M. S. Dresselhaus *Effect of quantum-well structures on the thermoelectric figure of merit* Phys. Rev. B, Condens. Matter, 47:19:12 72712 731, 1993.

[141] L. D. Hicks et al. *Experimental study of the effect of quantum-well structures on the thermoelectric figure of merit* Phys. Rev. B, Condens. Matter, 53:16:R10493R104936, 1996.

[142] D. A. Broido and T. L. Reinecke *Theory of thermoelectric power factor in quantum well and quantum wire superlattices* Phys. Rev. B, Condens. Matter, 64:045324 2001.

[143] M. S. Dresselhauset al *Advances in 1-D and 2-D thermoelectric materials* in 18th Int. Conf. Thermoelectrics Proc., 9299. 1999

[144] J. Zou and A. Balandin, *Phonon heat conduction in a semiconductor nanowire* J. Appl. Phys., 89:5:29322938 2001.

[145] Y.-M. Lin and M. S. Dresselhaus *Thermoelectric properties of superlattice nanowires* Phys. Rev. B, Condens. Matter, 68:075304, 2003.

[146] Patrick E. Pheln et al *Current and Future Miniature Refrigeration Cooling Technologies for High Power Microelectronics* IEEE Transactions on Components and Packaging Technology, 25:3 2002

[147] A. Shakouri and J. E. Bowers, *Heterostructure Integrated Thermionic Refrigeration* in 16th International Conference on Thermoelectrics, 1997.

[148] G. D. Mahan, *Multilayer Thermionic Refrigeration* in 18th International Conference on Thermoelectrics, 1999.

[149] A. Shakouri, C. LaBountyet al *InPBased Thermionic Coolers* in 11th International Conference on Indium Phosphide and Related Materials,Davos, Switzerland, May 16-20, 1999.

[150] A. Shakouri et al *BThermoelectric effects in submicron heterostructure barriers* Microscale Thermophys. Eng.2:3742, 1998.

[151] X. F. Fan G. H. Zeng et al *SiGeC/Si superlattice microcoolers* Appl.Phys. Lett., 78:15801582, 2001.

[152] C. LaBounty, A. Shakouri, et al. *Monolithic integration of thin-film coolers with optoelectronic devices* Opt. Eng., 39:28472852, 2000

[153] M. E. H. Tijani, J. Zeegers, and A. T. A. M. de Waele, *Design, Development, and Operation of a Thermoacoustic Refrigerator Cooling to Below 60 C* in Cryocoolers 11, R. Ross Jr, Ed. New York: KluwerAcademic/Plenum Publishers, 2001.

[154] Andrew Miner and Uttam Ghoshal *Limits of Heat Removal in Microelectronic Systems* IEEE Transactions on Components and Packaging Technology, 2:4, Sept 2006

[155] H.L. Edwards et al *Cryogenic cooling using tunneling structures with sharp energy features* Phy. Rev. B. 52:8:5714-5736 1995

[156] J. Clarke and J.L. Patterson *Measurements of Relaxation of Quasiparticle Branch Imbalance in Superconductors* Jour. Low Temp. Physics 15:5/6:491-522 1974

[157] M. Tinkham *Tunneling Generation, Relaxation and Tunelling Detection of Hole-Electron Imbalance in Superconductors* Phy. Rev. B. 6:5:1747-1756 1972

[158] M. J. Moran and H. N. Shapiro *Fundamentals of Engineering Thermodynamics* 5th ed. New York: Wiley, 2003

[159] J.P. Pekola et al *Limitations in Cooling Electrons using Normal-Metal-Superconductor Tunnel Junctions* Phy. Rev. Lett. 92:5:0568041-4 2004

[160] J. Jochum et al *Modelling the power flow in normal conductor-insulator-superconductor junctions* Jour, App. Phy. 83:6:3217-3224 1998

[161] J.P. Pekola *Trapping of quasiparticles of a non-equilibrium superdonductor* App. Phy. Lett. 76:19:2872-2874 2000

[162] Donald Neamen *Introduction to Semiconductor Device Physics* McGraw Hill Engineering Science 2005

[163] H. Pothier et al *Energy distribution of electrons in an out or equilibrium metallic wire* Z. Phys. B 104:178-182 1997

[164] H. Pothier et al *Energy distribution function of Quasiparticles in Mesoscopic Wires* Phy. Rev. Lett. 79:18:3490-3493 1997

[165] C.M.Wilson and D.E. Prober *Quasiparticle number fluctuations in superconductors* Phy.Rev.B 69:094524-1:11 2004

List of Figures

1.1 Power dissipation in Intel's microprocessor family over the years 3
1.2 Summary of constant field device scaling 4

2.1 Schematic of a typical setup 8
2.2 (a) Hypothetical cryogenic refrigerator (b) Operating region of the refrigerator 8
2.3 Energy Band Diagram of a NIS Tunnel Junction. 9
2.4 Energy Band Diagram of a SINIS Structure with different tunneling processes. 11
2.5 Schematic showing the SIS junction with the large junction used to inject quasiparticles and the probes at various distances used to measure non-equilibrium 13

3.1 Test Structure schematic. Green → connections to pads. Red → varying widths of probes and the wire over which the tunnel barrier is to be made. Yellow → bridge for the undercut. 15
3.2 SEM image of the test structure developed using the PMMA-MMA bilayer of photoresist showing not enough undercut (no overlap of the probes and wires) 16
3.3 Proximity effect realized by e-beam strengths of 20KeV and 100KeV . . . 17
3.4 SEM image of the test structure developed using PGMI-PMMA recipe . 18
3.5 Flowchart for EBPG flow. 19
3.6 Principle of shadow evaporation. 21
3.7 SEM image of the final device. 22
3.8 SEM image of the final device 23

4.1 Schematic to calculate the energy gap using the Rowell Criteria. 26
4.2 Final design showing the complete structure. 28
4.3 Final design showing details of the probes with measurements. 29

5.1 Two point and Four point resistance measurement techniques 31
5.2 Schematic of the voltage and current sources and meters used in the measurements. 32
5.3 Schematic of a dilution fridge (Left) Photograph of the dilution fridge used for measurement. 34
5.4 Schematic of the electrical wirings in a dilution fridge 35

5.5 Schematic of the device structure showing the injector junction as well as the 3 detector junctions with the respective sources and measurement connections. 35
5.6 IV characteristics of the injector junction. Green curve is after taking into account the internal resistance of the current source, Red is the original measured. 36
5.7 IV characteristics of the detector junction. Green curve is after taking into account the internal resistance of the voltage source, Red is the original measured. 37
5.8 IV characteristics of the detector junctionshowing the non local behavior for different injector currents. 38

6.1 Fermi-Dirac distribution showing distribution for T = 0, T> 0 and T >> 0 Quasi-Equilibrium and non-equilibrium 41
6.2 Experimental layout: diffusive wire of length L connected to the two reservoirs biased at potentials 0 and U 42
6.3 Distribution function $f(x, E)$ as a function of position x and energy E for non-interacting quasiparticles 43
6.4 Distribution function $f(x, E)$ as a function of position x and energy E for strong quasiparticle scattering but negligible phonon scattering 44
6.5 Distribution function $f(x, E)$ as a function of position x and energy E for strong phonon scattering 44
6.6 Temperature dependence of electrical resistivity 45
6.7 Weak external magnetic field expelled from the superconductor 45
6.8 Variation of the gap with temperature 46
6.9 (a) Dispersion curves of excitation energies in Normal and Superconducting states. (b) Energy-Mode excitation with $T > T^*$. (c) Charge-Mode excitation showing charge imbalance with $Q > T^*$ and shift of μ_n and μ_p w.r.t equilibrium Fermi E_f 47
6.10 Semiconductor model description of tunneling. (a)*N*-*S* tunneling at $T = 0$ with bias eV slightly exceeding the energy gap Δ. (b) *S*-*S* tunneling at $T = 0$ with bias voltage $eV < \Delta_1 + \Delta_2$. Arrows depict tunneling. 49
6.11 Tunneling Characteristics of Superconductor-Superconductor tunneling. Solid Lines are for T = 0 and dotted lines are for $T > 0$. 50

7.1 Schematic of the device structure showing the injector junction as well as the 3 detector junctions with the respective sources and measurement connections. 52
7.2 IV Characteristics of injector junction showing specific bias points where detectors are probed 53
7.3 Plot showing the quasiparticle density as a function of temperature for Aluminium 54
7.4 IV Characteristics of detector junction 1 for different injector currents shown in Fig 7.2 55
7.5 IV Characteristics of detector junction 2 for different injector currents shown in Fig 7.2 56

7.6 IV Characteristics of detector junction 3 for different injector currents shown in Fig 7.2 57
7.7 Plot showing the variation of the gap voltage at the three detectors for different injector currents 58
7.8 Plot showing the variation of the gap voltage as a function of the bias points along the injector IV characteristics for the three junctions 58
7.9 Plot showing the variation of the excess current with respect to the injector current as a function of the voltage across the injector 59
7.10 Plot showing the excess current at the three detectors for an applied bias at the detector 60
7.11 Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 1 for injection current of 1μA 61
7.12 Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 1 for injection voltage of 0.85μA 62
7.13 Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 2 for injection current of 1μA 62
7.14 Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 2 for injection voltage of 0.85μA 63
7.15 Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 3 for injection current of 1μA 63
7.16 Plot showing the current at the detector for the 4 cases enumerated in section 7.3 for detector junction 3 for injection voltage of 0.85μA 64
7.17 Plot showing the excess current as a function of injector current for different voltages at detector junction 1 65
7.18 Plot showing the symmetric part of the excess current for junction 1 which is independent of the injector current 65
7.19 Plot showing the asymmetric part of the excess current for junction 1which is independent within the Δ region but starts becoming heavily dependent upon injector current 66
7.20 Plot showing the excess current as a function of injector current for different voltages at detector junction 2 66
7.21 Plot showing the symmetric part of the excess current for junction 2 which is independent of the injector current 67
7.22 Plot showing the asymmetric part of the excess current for junction 2. Asymmetric part is dependent on injector current 67
7.23 Plot showing the excess current as a function of injector current for different voltages at detector junction 3 68
7.24 Plot showing the symmetric part of the excess current for junction 3 which is independent of the injector current 68
7.25 Plot showing the asymmetric part of the excess current for junction 3. Asymmetric part is dependent on injector current 69
7.26 Plot showing the variation of the asymmetric part of the excess current for the three detector junctions 69

8.1 Typical heat flows in an IC chip with pad and package exposed 73
8.2 Variation of thermal conductivity with film thickness of Silicon thin films 82

8.3 Typical packaging applications I-heat sink; II-TIM; III-IHS; IV-TIM; V-die; IV-underfill; and VII-package substrate 86
8.4 Schematic showing a warped die 87
8.5 (a) Showing the real contact between two solids; (b) represents a real TIM. 88
8.6 Schematic of a sample package using CNT 89
8.7 SEM images of the CNTs on upward surface (left) and downward surface (Right) 90
8.8 Thermal resistance distribution and measurement results 90
8.9 Pipe (Courtesy of Thermacore International Inc.) 91
8.10 Working principle of a heat pipe 93
8.11 Limits of Heat Pipe Operation [48] 96
8.12 Comparison of various heat transfer performance of various micro heat pipes 97
8.13 Left - Louver Fin heat sink; Middle - Offset fin heat sink; Right - Plate fin heat sink 98
8.14 Design Parameters for different geometries of Heat Sinks 99
8.15 Design Equations for different geometries of Heat Sinks 99
8.16 Schematic of a Vapor Chamber 102
8.17 Vapor chamber integrated with heat sink and die [80] 103
8.18 Die stress comparison with copper HIS and vapor chamber [81] 105
8.19 Classification of Micropumps 106
8.20 Vibrating diaphragm micropump. A: undeflected position b: expansion stroke c: contraction stroke 107
8.21 Structure and operation of a peristaltic pump 110
8.22 Comparisons of various micropumps 112
8.23 DC electroosmotic micropump 113
8.24 Conceptual Schematic of electro-osmotic cooler 114
8.25 Schematic of the microchannel heat exchanger 115
8.26 AC electroosmotic pump 115
8.27 : Heat dissipation rates achieved with Microchannels 117
8.28 Flow Patterns 119
8.29 Optimum COP Vs. temperature difference between hot and cold sides of the thermoelectric element 123
8.30 Thermoelectric figure of merit Z and DOS for 1-D 2-D 3-D and 0-D structures124
8.31 Comparison of Active Cooling methods 126

List of Tables

6.1 Comparison of different length scales in transport regimes. 40

8.1 Materials used for Thermal management in Integrated Circuits and Packaging 75
8.2 Coefficient of Thermal Expansion for materials used in Integrated Circuits and Packaging 77
8.3 Ultra high thermal conductivity and low CTE 79
8.4 High Thermal Conductivity and low CTE 80
8.5 Optimum variables for selected pin-fin geometries 101
8.6 Present level of understanding of single-phase and two-phase micro-channel 118
8.7 : Correlation models for pressure drop 121