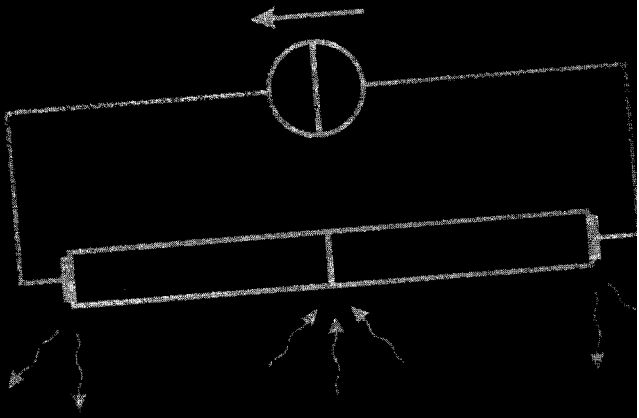
Lateral on-chip integrated

[illegible]

based on polycrystalline silicon germa[illegible]



[illegible] Wang[illegible]

Propositions belonging to the dissertation

Lateral on-chip integrated Peltier elements
based on polycrystalline silicon germanium

of **Davey Wijngaards**

Delft, 8 September 2003

---

1. The performance of the lateral on-chip integrated thermoelectric cooler is determined primarily by the parasitic thermal conductance. (Chapters 2 and 5)
2. The better the thermoelectric performance of the materials applied, the more sensitive the thermoelectric structure becomes to parasitic effects. (Chapter 2)
3. Regarding the cooling of on-chip hot spots, lateral on-chip Peltier devices are inferior to simple on-chip integrated passive heat spreaders. (Chapter 5)
4. Polycrystalline silicon germanium has the potential to improve the detectivity of polySi-based infrared detectors by a factor of 3 to 6. (Chapter 5)
5. The temperature range of a micro-thermostat is maximised by using a Peltier device only, for cooling, heating and temperature sensing. (Chapter 5)
6. Science should not just be about understanding difficult things, but also about making difficult things easy to understand.
7. Research at universities becomes increasingly driven by political arbitrariness.
8. Fuzzy logic is applied the most by politicians and civil servants.
9. Woe to you, o land, whose king is a child—Ecclesiastes 10:16 (NIV). Known for so long, it still hurts us on many levels of society.
10. In MEMS, carelessness leads to much stress, both mechanically and psychologically.
11. Ease keeps the economy running.

These propositions are considered defendable and as such have been approved by the promotors prof.dr.ir. J.H. Huijsing and prof.dr.ir. G.C.M. Meijer

Stellingen behorende bij het proefschrift

Lateral on-chip integrated Peltier elements
based on polycrystalline silicon germanium

van **Davey Wijngaards**

Delft, 8 september 2003

---

1. De prestaties van een lateraal geïntegreerde thermoelektrische koeler wordt hoofdzakelijk bepaald door de parasitaire thermische geleiding. (Hoofdstukken 2 en 5)
2. Hoe beter de thermoelektrische prestaties van de gebruikte materialen, deste gevoeliger wordt de thermoelektrische structuur voor parasitaire effecten. (Hoofdstukken 2)
3. Met betrekking tot het koelen van hete gebieden binnen een chip zijn lateraal geïntegreerde Peltier elementen inferieur aan eenvoudige geïntegreerde passieve warmtespreiders. (Hoofdstuk 5)
4. Polykristallijn silicium germanium heeft de potentie om de gevoeligheid van polySi-gebaseerde infrarood detectoren met een factor 3 tot 6 te verhogen. (Hoofdstuk 5)
5. Het temperatuurbereik van een micro-thermostaat wordt maximaal indien enkel een Peltier element wordt gebruikt, voor zowel koeling, verwarming als temperatuurmeting. (Hoofstuk 5)
6. Wetenschap moet niet alleen gaan om het begrijpen van moeilijke zaken, maar ook om het begrijpelijk maken ervan.
7. Onderzoek op universiteiten wordt steeds meer gedreven door politieke willekeur.
8. Vage logica wordt nog het vaakst toegepast door politici en ambtenaren.
9. Wee u, o land, welks koning kind is—Prediker 10:16 (NBG). Al zo lang bekend, raakt het ons nog steeds in vele lagen van de maatschappij.
10. In MEMS leidt onzorgvuldigheid tot veel stress, zowel mechanisch als psychisch.
11. Gemak houdt de economie in beweging.

Deze stellingen worden verdedigbaar geacht en zijn als zodanig goedgekeurd door de promotoren prof.dr.ir. J.H. Huijsing and prof.dr.ir. G.C.M. Meijer

LATERAL ON-CHIP INTEGRATED PELTIER ELEMENTS

BASED ON POLYCRYSTALLINE SILICON-GERMANIUM

DAVEY WIJNGAARDS

LATERAL ON-CHIP INTEGRATED PELTIER ELEMENTS

BASED ON POLYCRYSTALLINE SILICON-GERMANIUM

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op 8 september 2003 om 13:00

door

David Daniel Leonard WIJNGAARDS

elektrotechnisch ingenieur
geboren te Bergen op Zoom

Dit proefschift is goedgekeurd door de promotoren:
Prof.dr.ir. J.H. Huijsing
Prof.dr.ir. G.C.M. Meijer

Toegevoegd promotor:
Dr.ir. R.F. Wolffenbuttel

Samenstelling promotiecommissie:

Rector Magnificus	Technische Universiteit Delft, voorzitter
Prof.dr.ir. J.H. Huijsing	Technische Universiteit Delft, promotor
Prof.dr.ir. G.C.M. Meijer	Technische Universiteit Delft, promotor
Dr.ir. R.F. Wolffenbuttel	Technische Universiteit Delft, toegevoegd promotor
Prof. D.M. Rowe	Cardiff University
Prof.dr.rer.nat. G. Wachutka	Technische Universität München
Prof.dr. P.P.L. Regtien	Universiteit Twente
Prof.dr. P.M. Sarro	Technische Universiteit Delft

ISBN: 90-9017216-5

To Sabine,
For letting me be who I am

TABLE OF CONTENTS

Chapter 2:
Theoretical device performance 25

Chapter 3:
Material characterisation 63

LIST OF CONSTANTS AND SYMBOLS

Physical constants

Symbol	*Name*	*Value*	*Unit*
e	Electron charge	$1.6022 \cdot 10^{-19}$	[C]
h	Planck's constant	$6.62617 \cdot 10^{-34}$	[Js]
$\hbar$	Modified Planck's constant	$h/2\pi$	[Js]
k_B	Boltzmann's constant	$1.380662 \cdot 10^{-23}$	[JK^{-1}]
π	Pi	3.141592654	[-]
σ_B	Stefan-Boltzmann constant	$5.66961 \cdot 10^{-8}$	[$Wm^{-2}K^{-4}$]

Variables (Roman)

Variable	*Name*	*Unit*
A	Area	[m^2]
c_v	Specific heat for constant volume	[$Jg^{-1}K^{-1}$]
C	Electrical capacitance	[F]
C_{th}	Thermal capacitance	[JK^{-1}]
C_s	Surface dopant concentration	[m^{-3}]
D_0	Diffusion coefficient	[m^2s^{-1}]
D	Diffusivity	[m^2s^{-1}]
D	Depth	[m]
D^*	Detectivity	[$mHz^{1/2}W^{-1}$]
E_a	Activation energy	[J]
E_g	Bandgap energy	[J]
E_c	Conduction band energy	[J]
E_v	Valence band energy	[J]
f	Frequency	[Hz]
h	Heat transfer coefficient	[$Wm^{-1}K^{-1}$]
H	Height, thickness	[m]
I_{el}	Electrical current	[A]
I_{th}	Thermal flow	[W]

J	Electrical current density	[Am^{-2}]
K	Thermal conductance	[WK^{-1}]
K'	Modified thermal conductance (= $K + K_m$)	[WK^{-1}]
K_m	Thermal conductance of the membrane	[WK^{-1}]
L	Length	[m]
L	Lorentz number	[V^2K^{-2}]
$\mathcal{L}$	Lorentz factor	[-]
m^*	Effective mass	[g]
N	Doping concentration	[m^{-3}]
n	Electron/implantation concentration	[m^{-3}]
q	Power, thermal energy generation/removal rate	[W]
q'	Volumetric heat generation rate	[Wm^{-3}]
q''	Heat flux	[Wm^{-2}]
q_c	Heat removal rate at cold junction	[W]
q_e	Heat generation rate from external source	[W]
q_h	Heat release rate at hot junction	[W]
q_J	Joule heating	[W]
q_P	Thermal energy transfer rate by Peltier effect	[W]
q_{conv}	Convective heat transportation rate	[W]
q_{rad}	Radiant heat exchange rate	[W]
r	Radius	[m]
R	Electrical resistance	[Ω]
R'	Modified electrical resistance (= $R + R_c$)	[Ω]
R_c	Electrical contact resistance	[Ω]
R_{sh}	Electrical sheet resistance (= ρ/H)	[Ω/square]
R_p	Projected range	[m]
ΔR_p	Projected straggle	[m]
$\Delta R_\perp$	Lateral straggle	[m]
s	Spacing	[m]
s	Scattering parameter	[-]
S	Dopant dose	[m^{-2}]
S	Sensitivity	[VKW^{-1}]
t	Time	[s]
T	Absolute temperature	[K]
T_a	Ambient temperature	[K]
T_c	Cold junction temperature	[K]
T_h	Hot junction temperature	[K]
T_∞	Temperature far away	[K]
ΔT	Temperature difference	[K]
v	Speed of sound in a material	[ms^{-1}]
V	Volume	[m^3]
V	Electrical potential	[V]

ΔV	Electrical potential difference	[V]
W	Width	[m]
ΔW	Spread in width	[m]
z_{th}	Thermal impedance	[KW^{-1}]
z	Figure-of-merit of a material	[K^{-1}]
Z	Composite figure-of-merit	[K^{-1}]
$Z(K_m)$	Figure-of-merit compensated for K_m	[K^{-1}]
$Z(R_c)$	Figure-of-merit compensated for R_c	[K^{-1}]

Variables (Greek)

Variable	*Name*	*Unit*
α *(alpha)*	Seebeck coefficient	[VK^{-1}]
β *(beta)*	Thomson coefficient	[VK^{-1}]
β	Temperature coefficient of resistance	[K^{-1}]
χ *(chi)*	Electron affinity	[V]
δ *(delta)*	Grain boundary width	[m]
δ_{th}	(Thermal) diffusivity	[m^2s^{-1}]
ε *(epsilon)*	Emissivity	[-]
ε	Permittivity	[$C^2N^{-1}m^{-2}$]
η *(eta)*	Efficiency	[-]
η	Reduced fermi potential	[-]
λ *(lambda)*	Thermal conductivity	[$Wm^{-1}K^{-1}$]
μ *(mu)*	Mobility	[$m^2V^{-1}s^{-1}$]
ϕ_B *(phi)*	Barrier height	[V]
φ *(phi)*	Angle	[°]
φ	Phonon drag contribution	[VK^{-1}]
π *(pi)*	Peltier coefficient	[V]
θ *(theta)*	Temperature difference $T(x)$-T_a	[K]
ρ *(rho)*	Electrical resistivity	[Ωm]
ρ_c	Electrical contact resistivity	[Ωm^2]
σ *(sigma)*	Electrical conductivity	[$\Omega^{-1}m^{-1}$]
σ	Standard deviation	[-]
τ	Time constant, relaxation time	[s]
ψ *(psi)*	Power factor	[$Wm^{-1}K^{-2}$]
ω *(omega)*	$2 \cdot \pi \cdot f$	[rad s^{-1}]
ζ_n *(zeta)*	E_c-E_F	[J]
ζ_p	E_F-E_v	[J]

ACKNOWLEDGEMENTS

In some cases you do not realise how difficult the task at hand really is until you act-actually start it. Writing of these acknowledgements is just such a task. It is easy to give credits to the people who have contributed directly to the work in this thesis. Yet most people have contributed indirectly, either simply by providing some desperately needed relaxation or by taking some work off my shoulders. My thanks go to each of you! Nonetheless, some people deserve a special mention.

First off, I would like to thank the persons who sparked my enthusiasm for the fields of electronic instrumentation and MEMS. Without the encouragement by Reinoud Wolffenbuttel, the technological insight of Marian Bartek and the eloquence and wit of Paddy French, it is quite likely I would not have embarked upon the mental journey that led to this thesis. Moreover, the help of Ger de Graaf with the electronics is greatly appreciated, as it saved me a lot of time (and swearing.)

Without the assistance provided by the people at the DIMES Technology Centre it would have been impossible to obtain the required level of technology for the fabrication of the devices discussed in this thesis. A few persons in particular deserve a word of gratitude: Prof. Lina Sarro, for her technological expertise and experience, and her willingness to share these with with me. Wim van der Vlist and Ruud Klerks, for their skills, patience and ingenuity required to dice even the most delicate of wafers. Mario Laros and Charles de Boer, for their skills and improvisational talents required to process the Peltier devices. And finally, Cassan Visser for his experience in deposition of polycrystalline silicon germanium.

Special regards go to my former roommates at the university. A richer mix of nationalities could hardly exist: Seong (South-Korea) and Heng (China) taught me characteristics that the Dutch so frequently lack, i.e., compassion and respect. I hope these characterisics have rubbed off on me. Edmond (Romania) and Lukasz (Poland) showed the clarity of the Eastern-European mind. And finally, José (Portugal) who in his own way showed what passion is, 'you understand?' Furthermore, I greatly appreciate the help I received from Seong, Edmond and Lukasz to further my research.

Additionally there is that group of colleagues, of which one has special memories: Hans Goosen, Edward van Kampen, Péter Turmezei, Luis Rocha, Robert Kazinczi, Colin Ashruf, Vladimir Kutchoukov, Dafina Tanase, Frederique Laugère, Bonny Gray, Jeroen Bastemeijer, Maureen Meekel, Piet Trimp, Harry Kerkvliet, Kofi Makinwa and Michiel Pertijs provided an atmosphere that was not only pleasant to work in but allowed one to discuss and socialise as well. Prof. Simon Middelhoek and Prof. Han Huijsing showed how tenacity can keep a group of people together and finally, Prof. Michael Vellekoop and Richard den Dulk reflected my actions, allowing me to grow mentally.

Next to the people at the Delft University of Technology, I am indebted to the people at Océ for leaving me the necessary room to finish this work. In particular Aart Polderman, Lars Idema and Ton Albers deserve to be thanked for their understanding.

Then there is that other group of people that helped me succeed. Not in my work, by contributing to it, but in life, by simply taking my mind off my work. And by encouraging me to grow through their help, discussions and scrutiny. These people are my friends and family.

Léon and Ivanka, Cindy and Rogier, Roy and Evelyne, Bas and Suzan, Richard and Jacqueline, Willeke and Johan, Ad and Wendie, Jeroen and Saskia, Richard and Karin, Evelyn and Zion and, how could I possibly forget, Willem and Heleen, thank you for providing the love and warmth required to help me through what often felt like an uphill struggle.

One of my unfathomablenesses is that I always keep the best and/or most difficult for last. Often to such an extent that the people around me start getting nervous; My family in particular. Still, I hope that over the years I have proven that this extreme case of just-in-time-management is what works for me. Therefore I would like to thank my parents Wil and Adriënne, my brother Micha, and Ad and Tonny for their patience and trust, even though this was not always easy to maintain.

For no one these last few years were as hard as for Sabine. She could do nothing but accept my peculiar ways, the late working hours and limited spare time that came with those ways. Without her dedication this thesis would never have been completed. Nor would I be able to feel complete. And that is her greatest achievement yet: She gives me the chance to be me. For this I dedicate this thesis to her.

Chapter 1
INTRODUCTION

Well, if you start with a 10-Watt device
and go up 40-fold, the darn thing smokes!

Gordon. E. Moore

In this introductory chapter the basic functionality of the subject of this thesis—the lateral on-chip integrated Peltier device—is explained. The development of such a device, or any on-chip integrated transducer for that matter, is strongly cross-bordering work. Succesful implementation requires the fields of (quantum) physics, electronics, materials science and IC fabrication technology to be directly linked. To introduce the reader to this intersection of technologies, the most fundamental aspects of my work are outlined below. Three aspects are addressed:

1. What a Peltier element is (and, thus, what *thermoelectricity* is)
2. What is implied by *on-chip integration*
3. What applications a Peltier device can be used for (i.e., not just for *cooling*)

Each of the subsequent sections in this chapter discusses an aspect in more detail. Section 1.1 discusses the basics of thermoelectricity and its merits. Sections 1.2 and 1.3 provide a high-level discussion on the issues associated with downscaling and on-chip integration. Finally, Section 1.4 discusses why a Peltier device is particularly suitable for localised cooling of thermally isolated volumes of a chip rather than an entire chip.

Before proceeding with the first section, I would like to make a remarks about the nomenclature used in this thesis. Even though the term 'Peltier device' is the more generic designation, the term 'thermoelectric cooler' (or 'TEC' for short) is applied frequently throughout this thesis. Both terms refer to the same device. Even though a Peltier device is capable of more than just cooling, the term 'TEC' emphasises the primary intent of the device, and—of a more practical reason—clearly stands out from the text and is shorter to write.

1.1 THERMOELECTRICITY

Simply speaking, thermoelectricity encompasses three physical effects in which thermal energy is converted directly into electrical energy and *vice versa*. To provide an idea of the history of thermoelectricity, Figure 1.1 lists some of the milestones. Even though the existence of thermoelectric effects is known for over 180 years, the use of these effects remains limited right to this date. The effect first discovered, the Seebeck effect, finds the most widespread use of the three effects, primarily in temperature sensors and thermal power generators. The second, the Peltier effect is already encountered much less frequently, primarily in coolers and thermostats. The third effect, the Thomson effect, generally is so small that no suitable application for this effect exists.

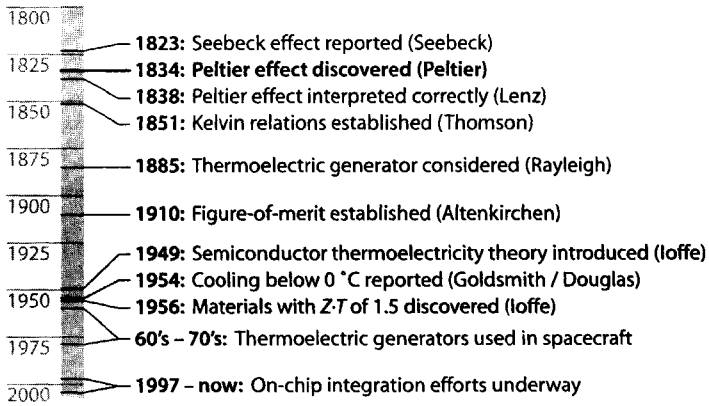


Figure 1.1 Historical timeline of thermoelectric milestones.

The discovery of the Seebeck effects dates back to 1823. Even though Seebeck managed to devise a good measure for the magnitude of the effect, he failed to properly explain the nature of it, as he thought it to be magnetic. Nevertheless, his measure, the product of the Seebeck coefficient α [VK^{-1}] named after himself and the electrical conductivity σ [$\Omega^{-1}m^{-1}$], is used right to this date and is known as the power factor ψ.

Then, in 1834 the French watchmaker Jean Charles Athanase Peltier (1785–1845) discovered the second thermoelectric effect. Similar to Seebeck, he too failed to properly understand the effect he discovered [1.1]. Some four years later Becquerel and Lenz confirmed his experiments and managed to correctly interpret the effect. It is observed at the region where two different electrically conducting materials are in direct contact. If an electrical current runs through such a junction, thermal energy is absorbed or released at the junction, due to physical dissimilarities between the two materials. By proper placement of arrays of such junctions—physically splitting

up those where heat is absorbed and those where heat is released—thermal energy can be transported from one set of junctions to another. Thus, in simple terms a Peltier device is an electrical (non-mechanical) thermal energy pump.

The third effect, the Thomson effect, was discovered in 1851 by W. Thomson (who later became known as Lord Kelvin) when he realised that there is a direct relationship between the Seebeck coefficient α and the Peltier coefficient π, i.e., $\pi = \alpha T$. In this relationship T the absolute temperature in Kelvin. He predicted—and experimentally verified—the existence of a third effect that is needed to balance the energies in the aforementioned two. The Thomson effect is observed in electrical conductors across which a thermal gradient exists. When a current is ran through such conductors, heat is absorbed or released within the conductor itself, again depending on the material properties. Stated at the beginning of this section, the Thomson effect is not applied in itself but should nevertheless be taken into consideration with respect to the other two effects, in order to complete the energy balance. Still, the Thomson effect will be addressed only marginally in this thesis.

Based on first two effect, four modes of operation can be distinguished, each of which is shown in Figure 1.2. In every mode, two strips of physically dissimilar materials A and B combine to form the key component of the thermoelectric conversion, the thermocouple. (Each individual strip is referred to as a thermoelement or thermoelectric leg.)

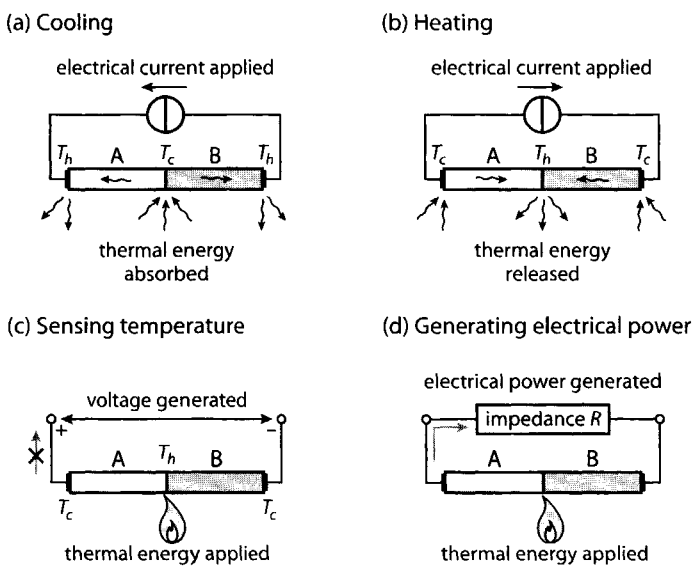


Figure 1.2 The four fundamental means of applying thermoelectric effects. The electrical-to-thermal Peltier effect can be used to construct (a) a cooler and (b) a heater. The thermal-to-electric Seebeck effect can be used to construct (c) a temperature sensor and (d) a generator.

Figure 1.2a shows the most fundamental mode of operation of a Peltier device. By applying an electrical current from conductor A to conductor B, thermal energy is absorbed at the interface between the two materials when conductor B has a more positive Seebeck coefficient than conductor A. (Thus, when $\alpha_A < \alpha_B$ or, equivalently, $\pi_{AB} = \alpha_{AB}T = (\alpha_A - \alpha_B)T < 0$.)

Figure 1.2b shows a second mode of operation of a Peltier device, even though it is used less frequently. Simply by reversing the direction of the electrical current, heat is released instead of absorbed at the interface between A and B. This reveals a very pleasant characteristic: (all three) thermoelectric effects are reversible. Keep in mind though, that even though the Peltier effect itself is perfectly reversible, a device employing this effect will not be: Simply switching the direction of a current to go from heating mode (Figure 1.2b) to cooling mode (Figure 1.2a) does not automatically mean that cooling is obtained. This is because the electrical current running through the thermoelements causes Joule heating, which is irreversible. In heating mode this Joule heat works constructively (adding to the thermal energy released by the Peltier effect), but in cooling mode it works destructively (cancelling out the thermal energy absorbed by the Peltier effect.) This requires these two powers to be carefully balanced during the design of the thermoelements, in particular as the Joule heat is a square function of the electrical current, while the Peltier effect is a linear function of the current. Chapter 2 discusses this balance in detail.

Note that the Peltier effect always converts electrical energy to thermal energy (hence it should be called an electrothermal rather than a thermoelectric effect.) The Seebeck effect does just the opposite. In Figure 1.2c the conventional mode of operation for the Seebeck effect is shown. The interface between conductors A and B is heated, so that this interface is at a higher temperature than the open ends of the conductors (thus, $T_h > T_c$). As a result of the Seebeck effect, a voltage is generated between the open ends. Under the same condition as before where $\alpha_A < \alpha_B$, the open end of A will be at a positive voltage compared to that of B. Note that this Seebeck potential or Seebeck voltage is only accurately measurable if there is absolutely no current flowing through the device, i.e., when it is an open-circuit potential. Luckily this condition is easily met by most voltage read-out circuits. The Seebeck voltage that is measured by such a circuit is $V_{AB} = \alpha_{AB}(T_h - T_c)$ [V]. Thus, the Seebeck voltage is linearly dependent on the temperature difference, rather than an absolute temperature. In the situation sketched in Figure 1.2d, the same Seebeck potential is generated, but this time an electrical device with resistance R is placed between the open ends. As a result, an electrical current I_{el} [A] will flow, equivalent to V_{AB}/R, thereby generating an electrical power, q_J [W] equivalent to $(V_{AB})^2/R$.

So, with the use of Figure 1.2 the first question, what thermoelectricity is, has now been answered. Simply speaking thermoelectricity is the *direct* and *reversible* conversion of energy between the thermal and electrical domain. From the key element in this conversion—the thermoelement—three fundamentally different devices can be constructed:

- The Peltier device, where an electrical current can transport thermal energy from one junction to another ($I_{el} \rightarrow q_{th}$).
- The thermocouple as such, which—even though its name is generic—is used to sense a temperature difference ($\Delta T \rightarrow \Delta V$).
- The thermoelectric generator (TEG), which is used to convert a temperature difference into electrical power ($\Delta T \rightarrow q_{el}$).

One thing has to be kept in mind, though. In theory, a single thermocouple is capable of operating in any of the four modes. In practise, however, a thermocouple will be optimised for only one of them. For example, a Peltier device needs to handle large current densities, while for temperature sensing (which is currentless) the number of elements per unit surface are is dominant, as the resistance of the thermopiles sets the thermal noise level. Finally, the thermoelectric generator needs to transfer as much power from the thermal to the electrical domain, so the impedance of the TEG needs to be matched to that of the electrical circuit it is driving.

Having stated that, in Chapter 5 it will be shown, that the best micro-thermostat concept is based on a Peltier device, that is used for cooling, heating and sensing alike. Even though this device is optimised for cooling and not for the other two functions, it still provides the best performance in terms of thermal operating range.

1.2 CONSEQUENCES OF DOWNSCALING

In essence, on-chip integration deals with two issues: 1) How to shrink conventional discrete devices to the dimensions of a chip—at most 10 mm wide × 10 mm long × 0.5 mm high. 2) How to integrate / merge different fabrication technologies, namely integrated circuit (IC) technology and micro-electro-mechanical systems (MEMS) technology. Merely shrinking the discrete device to the size of a chip generally leads to loss of performance and structures that are incompatible with IC technology. Instead, a compromise between performance and compatible has to be found.

In terms of the Peltier device in this thesis, on-chip integration encompasses scaling of the conventional discrete thermoelement, typically 1 mm × 3 mm × 1 mm in size, to thin film structures with dimensions in the order of 50 μm × 200 μm × 1 μm. So, while the length of such an element has to be scaled down by a factor of 15 'only', the thickness has to be scaled down a factor of 1000 already! On top of that, this has to be done whilst minimising the parasitics that occur due to downscaling. A detailed analysis of these parasitics and their influence on device performance can be found in Chapter 2. Moreover, if a micro-thermostat is to be fitted with a Peltier device in order to stabilise particular regions of an electronic circuit, the device has to be compatible with the electronics already residing on the chip. The following subsections are intended to provide the reader with a overview of what the considerations and consequences are of downscaling and merging of two very different technologies. The microelectronics industry serves as an example.

1.2.1 Shrinking dimensions

When looking at the microprocessor industry, increasing the component density is the undisputed main driver: The higher the density, the higher the level of functionality per square millimetre. Moreover, smaller transistors are faster, allowing higher clock speeds, increasing the number of routines per second, which in turn increases the functionality.

Increasing the component density has another advantage besides increasing the functionality (where new chips do more with the same space than old chips.) It also facilitates downsizing the current generation of devices (so new chips do the same as old chips in a smaller space.) This allows the manufacturer to collect more devices off the same wafer area, thereby reducing the cost per device. At the same time as the feature size is reduced, the industry can switch to larger wafers, which should allow another 30%–40% cost decrease [1.2].

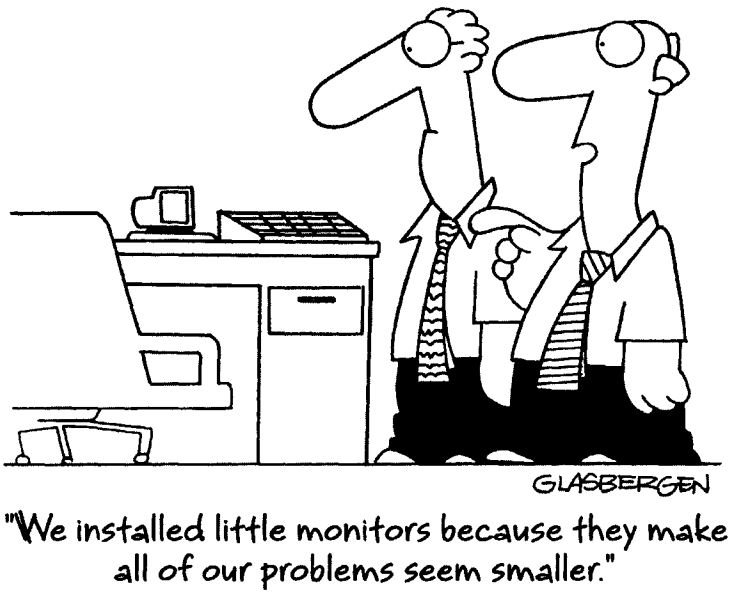
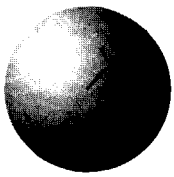
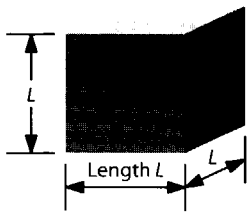


Figure 1.3 Undoubtedly the most specious argument of all for downscaling. (Source: "Technology Bytes" by Randy Glasbergen [1.3].)

It certainly appears true that downscaling provides the end-user (as well as the short-sighted sales manager) with more functionality for less, which quickly leads to the belief that smaller is simply always better. Strangely enough, when stretching this statement to the situation depicted in Figure 1.3, everyone agrees that downscaling to reduce existing problems is ridiculous. In particular from a technological perspective just about the opposite is true. Indeed downscaling can improve the level of functionality. But at a significant cost only: Each new and smaller generation of electronics



Sphere:
Radius r
Surface area $A = 4 \cdot \pi \cdot r^2$
Volume $V = 4/3 \cdot \pi \cdot r^3$
$A/V = 1/(3r)$



Cube:
Side length L
Surface area $A = 6 \cdot L^2$
Volume $V = L^3$
$A/V = 6/L$

Figure 1.4 Two basic shapes, with their respective side length (radius), area and volume. Irrespective of the shape, the area:volume ratio is inversely proportional to the size.

has an increased number of technological problems that need solving. The easiest way to explain why these problems occur is by means of Figure 1.4. Take the sphere with radius r. The surface area is equal to $4 \cdot \pi \cdot r^2$ [m^2] and the volume is equal to $4/3 \cdot \pi \cdot r^3$ [m^3]. Hence, the area:volume ratio scales with $1/(3r)$, inversely proportional with the radius. This inversely proportional ratio is found for any three-dimensional structure.

In practise the above means that when going from millimetre-scale to micrometre-scale devices the area:volume ratio is increased by a factor 1000. This is precisely what enables the extremely high aspect ratios encountered in MEMS technology, like the 50 μm × 200 μm × 1 μm thermoelements presented in this thesis. Due to the decreased area:volume ratio, the forces that act upon the volume (i.e., the bulk forces like gravity) are significantly reduced compared to the forces acting on an area (i.e., the surface forces like electrostatic attraction.)

Still, even if a very thin film is technologically manufacturable, yield is a concern. Many thin films natively exhibit very large levels of mechanical stress (most noticeable if the film consists of several stacked layers.) These stresses can lead to heavy buckling, bending or even fracturing. Stress compensation can often be obtained by careful layer stacking, while stress relief can often be obtained by proper thermal treatment of the materials. These considerations are addressed in Chapter 4.

The rate at which capital is invested in the IC industry is a clear indicator that an end to transistor downscaling is not expected shortly. Nevertheless, more and more indications are found that smaller might not be the way to go for very much longer. For a considerable part this is caused by market maturation, where the customer does not necessarily need faster equipment, but rather more versatile and tailor-made equipment. This issue is discussed in some detail in Section 1.3.1. In my opinion it will be technological rather than economic difficulties that will slow down the rate of progress observed in IC technology. As a matter of fact—and of significant relevance to this thesis—I believe it is the increasing power density that is the major threat to progress in the IC industry. Only months after writing this chapter, my view received a serious boost from an article published in The Economist Technology Quarterly [1.4], that all but summarises the arguments put forward in this section.

1.2.2 Surmountable technological barriers

Consider the recent advances made in lithography. In the 1997 International Technology Roadmap for Semiconductors (ITRS 1997), the 0.13-micron[1] barrier was not expected to be reached until 2003 [1.5]. Already in 2001 this prediction was superseded, when 0.13-micron technology was used routinely for the first time: In just five years, the Roadmap was pulled in by two years already. Now, following the predictions of the ITRS 2001 [1.6], lithography will not be a major concern until 2007. (Which does not mean that significant technological and financial efforts from industry are no longer required to stay on the projected path.)

Interconnect concerns

The downscaling issues encountered on a physical level are of a much more acute and complex nature. With each new technological generation, wires become narrower and interconnect becomes denser. As a result both the electrical resistance and capacitance of the interconnect increase. One of the consequences is enlargement of signal delays. This is most noticeable in the global chip wiring, where it may soon take several clock cycles for a signal to travel across a chip [1.7]. An even more severe consequence of the increasing component and wiring density is the rapidly rising level of power dissipation. This issue will be discussed shortly.

Physical imperfections of transistors

Due to the non-idealities during processing, there are small fluctuations between devices. As the transistor shrinks, these fluctuations become more noticeable and will increase the spread of the transistor parameters. As component matching is often critical for device performance this spread leads to considerable problems.

With a further decrease in size the behaviour of the transistor is changing on a fundamental level. Beyond the 0.03 μm barrier all kinds of quantum physical effects become prominent. Electrons start tunnelling, and due to the extremely short paths electrons may not be considered to travel at a constant velocity any more, but keep accelerating towards the anode [1.8]. When electrical field strengths are very high this could lead to avalanche breakdown. In order to keep the electrical field strength within reason, the supply voltage needs to be scaled down with the process. Unfortunately, the associated reduction in threshold voltage increases the leakage current, which (like the increasing component density) has negative impact on power dissipation.

1.2.3 Insurmountable technological barriers

For all the issues mentioned in Section 1.2.2, technological solutions either are already in the making or have at least been identified. Unfortunately, for many issues a concrete solution is not yet known, but still need to be dealt with within a few years

1. The measure '0.x micron' refers to the so-called minimum metal half pitch—a feature size in IC technology—used to capture the rate of downscaling in DRAM and microprocessor technology.

from now. One is the electrical current density, notorious for electromigration, which causes metal connections to be gradually displaced, leading to open circuits. The ITRS shows that already in 2006 the maximum current density J_{max} increases to such a level that a manufacturable solution is no longer known. By that time J_{max} is expected to reach $1.9 \cdot 10^6$ A/cm^2, practically twice the present value.

Still, the most urgent matter is the increasing power density, due to increased component and wiring densities, as well as increasingly lossy behaviour of electronics. This issue is enunciated below.

1.2.4 Increasing power density

Compared to the technological barriers mentioned in Sections 1.2.2 and 1.2.3, the most often reported (and by now the most feared) technological barrier is power density. The reason why power density poses such a threat to the microprocessor industry is that power is directly linked to performance. The total power consumed by a microprocessor, q_{tot}, can be expressed as:

$$q_{tot} = C \cdot V_{DD}^2 \cdot f \tag{1.1}$$

Here, C is the effective load capacitance of the microprocessor, V_{DD} the supply voltage and f the clock frequency. Theoretically, the ideal downscaling laws are in favour of downscaling: With the ideal shrink (when shrinking the existing process by 30% using constant electric field scaling[2]) the performance should increase by 50%, the frequency should go up by 43%, and the die size should reduce by 50%, while using 50% less power. With the ideal shrink (to a factor of 0.7), Equation (1.1) yields $C \cdot V_{DD}^2 \cdot f = 0.7 \cdot 0.7^2 \cdot 0.7^{-1} \approx 0.5$ (which is the 50% power reduction just mentioned.) Besides the ideal shrink, there is the ideal new generation, which employs twice the amount of transistors to obtain three times the performance, while the die size and power consumption remain unaffected. With both the ideal shrink and the ideal new generation, the theoretical gain in the number of instructions per second, per unit of power should increase three times without a change in power density [W/cm^2]. In reality the increasingly non-ideal behaviour of the electronics mentioned in Sections 1.2.2 and 1.2.3 cause a growing disparity between the idealised models and practise.

The required downscaling of the supply voltage and the threshold voltage significantly increase the leakage power due to the reverse-bias and sub-threshold leakage currents. As the gate oxide thickness scales down with feature size, also the gate-substrate leakage is increased. These effects cause the overall power density to increase. In turn this increases the device temperature, which further increases the junction leakage.

---

2. Constant electric field scaling means scaling down the supply voltage linearly with the feature size, to maintain the same electric field strength across process generations. With constant voltage scaling, the supply voltage is kept constant. The first is preferred as it provides a faster transistor for the amount of energy consumed [1.9].

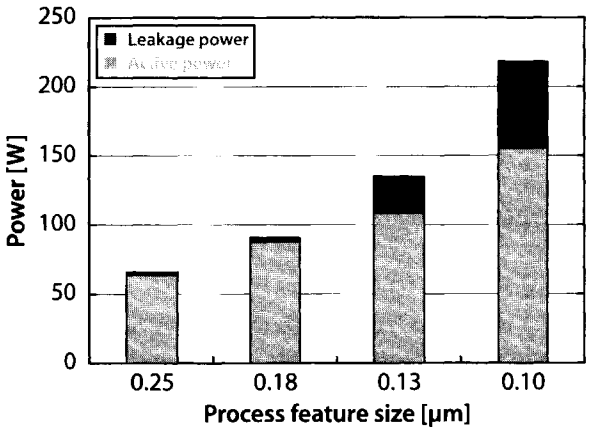


Figure 1.5 Trend graph revealing the increase in total power consumption and the increase in in leakage power consumption. (Source: [1.11].)

From Figure 1.5 it is clearly illustrated that leakage power is no longer a small fraction of the total power consumption of a microprocessor. In [1.10] the total leakage current is estimated to increase as much as 7.5 times per generation. Together with the increasing component density, the leakage power is expected to quintuple per generation. Thus, a few generations from now the leakage power will outweigh the active switching power. The figure also reveals that the total power consumption of a microprocessor increases exponentially over time. This is quite the opposite of what the ideal scaling laws predict.

Until now the existing packaging and (external) cooling solutions are sufficient to cope with increasing power consumption. However, in a few years from now the power density will become a serious problem, as existing cooling solutions become too expensive [1.12] or simply cease to exist. According to [1.13], microprocessors that continue to use the existing circuit technology and architectures would exceed package power limits by over a factor 25 in 2015. Figure 1.6 shows the general trend in the power density increase, roughly increasing by 80% each generation. At this rate of increase, the power density will soon cause the on-chip temperatures to reach levels that push beyond the range of what even the (most expensive) state-of-the-art heat sinking methods can handle.

1.3 ON-CHIP INTEGRATION

Even though the beginning of this section will seem to have very little to do with on-chip integration, the trend in microelectronics is addressed to explain the increased need for diversified products and the technological means to deal with this need.

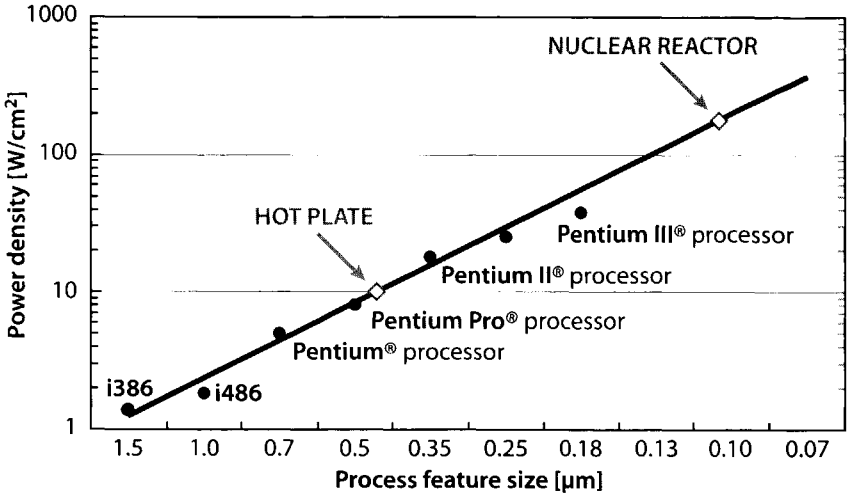


Figure 1.6 Extrapolation of the exponential increase in power density reveals the power density increases roughly 80% per generation. At this rate it is only a matter of years until the power densities normally found in nuclear reactors are reached. (Source: [1.7].)

1.3.1 User-driven technology crossing

Already in 1997, Hutcheson *et al.* made a serious remark regarding pace of developments in the microelectronics industry [1.14]: "... *slowing the rate of progress in semiconductors could have unexpected advantages, such as giving computer architectures and software time to begin assimilating the great leaps in chip performance.*" This remark reveals a serious problem, which remains underexposed: The microelectronics industry behaves as if it has to keep up with Moore's Law[3]. It becomes questionable whether the tremendous downscaling efforts required to achieve this are desirable for very much longer. This view received a big stimulus in April 2002 as, according to Bass and Christensen, there is a serious indication that the focus of the market tiers is shifting away from the high-end (Moore's Law dominated) microprocessor industry [1.15]: "...*The next stage of an industry's development begins when the performance of its products has overshot the needs of customers in the less-demanding tiers of the market. These customers won't pay premium prices for more performance ... but will pay extra for a product that is extraordinarily reliable, or one that has been customized to meet their specific needs—especially if they can get that ultra-reliable or customized product quickly and conveniently...*"

3. In general, Moore's Law states that the performance of microelectronics increases exponentially. In practise this means, for example, that the number of transistor doubles every 18 months.

A historic scenario—the aircraft industry

The demand for higher levels of customisation and diversity is largely equivalent to the crossroads encountered previously in other major industries like the aircraft industry. That industry matured in only 70 years, going from the Wright brothers in 1903 till the introduction of the biggest and fastest passenger jets to date. Boeing's 747 (1970) currently carries over 500 passengers. BAC/Aerospatiale's Concorde (1976) and travels at 2200 km/h, equivalent to Mach 2.

During the 25 years that followed, there was no need for even larger and faster jets. In terms of Bass and Christensen, the performance-dominated phase was over. Nevertheless, various new aircraft were developed in that period, strongly diversifying the range of jets with features like increased fuel efficiency, quieter operation or customer-based tailoring like private and corporate jets. The aircraft industry also reveals a market will never become truly saturated: In 2000 Airbus announced its A380, intended to carry over 555 passengers. A replacement for the Concorde is spotted on the drawing boards at NASA, designated the High-Speed Civil Transport (HSCT). Even though not much faster than the Concorde (Mach 2.4), it can carry three times as many passengers twice the distance, at a much lower cost per flight.

The current-day scenario—the microprocessor industry

After only 55 years, he microelectronics industry is already showing signs of maturation [1.14]: "*Price and performance, fuelled by the industry's collective preoccupation with Moore's Law, are still the metrics valued in essentially all tiers of the market today. Even so ... the initial, performance-dominated phase is giving way to a new era in which other factors, such as customization, matter more.*" The demand for raw computational power is gradually making room for a serious increase in the need for higher robustness and reliability, a shorter time-to-market and an increased level of customisation.

1.3.2 The System-on-a-Chip (SoC) and System-in-a-Package (SiP)

As the microelectronics industry struggles to keep up with Moore's law, diversification/customisation is likely to become the next primary source for profit. So, instead of the industry debating whether it *can* keep up with Moore's law—which can't be answered indefinitely anyway [1.16]—it should be asking whether it *should* keep up.

At current, the microprocessor industry is insufficiently capable of answering the demand for a more diversified range of products. To deal with this, the concept referred to as the System-on-a-Chip (SoC) is under development, which should enable a mix of different technologies to be accommodated in a single chip. For the concept to become successful it must be versatile, which makes it essential that the SoC device modules become available off the shelf. This necessitates the intellectual property (IP) of the microelectronics developers and design houses to be reusable and to be independent of the other technologies used, i.e., these modules may not interfere with each others processing and physical behaviour: they must be fabrication compatible.

The main technologies covered by the SoC concept are the digital, analogue, high-voltage and radio frequency (RF) technologies [1.17]. This attempt to merge very different technologies, creates the exact same problem encountered when joining a Peltier device with electronics: different technologies are far from compatible *per se*. In the short term, it is expected that only analogue and digital technologies will benefit from the SoC concept. Only these two technologies are related closely enough to make most of the required mutual trade-offs bridgeable. The Peltier device presented in this thesis also suffers from such trade-offs, as fabrication compatibility forces the use of polycrystalline silicon germanium rather than bismuth-ides, or other compounds that are favourable from a perspective of performance.

Consequently, not all minds are favourably disposed towards the SoC concept (including me). Very likely the mutual technological trade-offs are far too large for the concept to become successful, as these trade-offs compromise the performance of the individual technologies [1.18]. To circumvent the necessary trade-off, the concept of a System-in-a-Package (SiP) was introduced, where each technology is processed individually, both in time and space. The dies can then be placed in a single package, still saving a considerable amount of space.

1.3.3 Co-integration of MEMS with microelectronic circuits

The MEMS market is substantially smaller than the microelectronics market. Just the microprocessor market totals about € 45 billion a year. In contrast, the total MEMS industry represented a € 2 billion – € 5 billion market in 2000, expected to grow to € 9 billion – € 16 billion by 2004 [1.19]. The MEMS market is highly segmented, both in terms of products as well as technologies. The result is that per MEMS device/technology there is less money to be invested. Unfortunately, such a smaller and more diversified market tends to prevent a unified approach towards the development of the reusable IP, required for co-integration to become commonplace.

Similar to the need for diversification in microelectronic circuits, there is a need for diversification of transducer materials and technologies. Unlike with microelectronics, this trend is not so much driven by consumer needs, but more by the limitations of the standard materials. Silicon for example has excellent optical and mechanical properties, but is far from the best transducer materials in other domains. The main considerations regarding the integration of non-standard materials with standard electronics are discussed in the remainder of this section.

1.3.3.1 Hybrid integration versus co-integration

Consider the differences between hybrid integration and co-integration. In the case of the hybrid system in Figure 1.7a—as with the SiP—all transducer and circuit dies are fabricated separately and mounted together on a passive carrier (i.e., a carrier not containing any electronics.) Electrical connections between the dies are established by means of wire bonds or flip-chip bonding, for example. In the semi-hybrid configuration (Figure 1.7b) the electronic circuits are integrated as much as possible in a

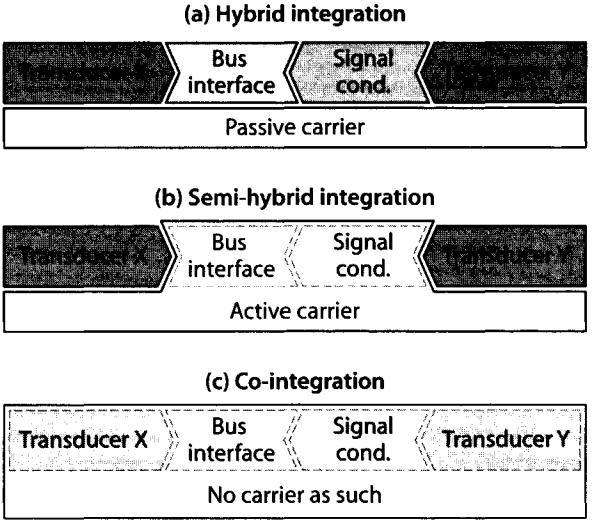


Figure 1.7 A schematic view of the difference between (a) hybrid integration (b) semi-hybrid integration and (c) co-integration.

single die, that replaces the passive carrier. Thus, the die holding the electronics acts as active carrier. Finally, in the case of co-integration, (Figure 1.7c) all transducers and electronic circuits are merged to form a single chip.

As a result of the market segmentation, the mainstream MEMS industry only focuses on the obvious, financially attractive transducer applications. After pressure sensors and accelerometers, now the industry focus is mainly on gyroscopes, biomedical transducers and optical switching technology. All of these are essentially high-volume markets, with expected annual sales volumes in the millions of transducer units. Even within such a relatively large industry focus, co-integration is hardly ever considered [1.20]. Take the early on-chip accelerometers. Only Analog Devices co-integrate their accelerometers with read-out circuits. It should be noted that substantial US governmental funding (through DARPA) has been invested in this development. Thus, the commercial drive is questionable—even here. All other accelerometers commercially available are mounted in a single package as a hybrid [1.21] or semi-hybrid device [1.22]. Each of the three integration approaches (hybrid, semi-hybrid and co-integration) are schematically shown in Figure 1.7.

Apparently there is a reason to go for a hybrid solution rather than a co-integrated device. The reason is that the first approach has both practical and economic advantages. First and foremost, as with the SiP compatibility of transducer fabrication technologies and electronics fabrication technologies is not an issue, as the fabrication is physically separated. Second, as a result of having individual dies, the yield is intrinsically higher [1.22]: Co-integration requires more complex and more elaborate

processing, which increases the probability of a fatal fabrication complication. Moreover, the separate fabrication of transducers and circuits allows each component to be tested individually. This way, the 'known good dies' can be selected prior to bonding, increasing yield. The physical separation of the dies also maximises the mix-and-match capabilities of the individual components and allows easy implementation of changes in one of the components without influencing the others.

Whilst hybrid and semi-hybrid techniques strive for a *single-package solution*, only the co-integration technique strives for a *single-chip solution*. Given the advantages of hybrid and semi-hybrid integration, such a single-chip solution doesn't seem favourable. However, co-integration has the advantage of having minimised component and bond pad counts as well as having the smallest possible device volume. The latter issue must not be underrated as the need for extreme miniaturisation is increasing, in particular in mobile telephony as well as in medical applications.

The cost of a microsystem—whichever integration approach is chosen—remains ambiguous, as the total cost of a co-integrated device is likely to be dominated by testing and (non-standard) packaging rather than processing. The addition of on-chip self-test and self-calibration facilities during co-integration provides a means to simplify testing during and after fabrication, which is a potent time and cost saver.

Even though the economic consequences remain debatable, the work reported in this thesis is directed towards a co-integrated approach, as this provides the largest possible flexibility, both in the short term and the long term. After all, a transducer designed for co-integration can eventually be integrated using a hybrid technique. *Vice versa* a transducer designed for hybrid integration can almost never be co-integrated. The Peltier devices are optimised and qualified for co-integration, but are initially packaged using a hybrid configuration to facilitate rapid prototyping.

Clearly, co-integration is a lot more challenging than hybrid and semi-hybrid integration. In fact, considering aspects like flexibility, development time and optimised transducer/electronics performance, I believe that (semi-)hybrid integration will triumph over co-integration; In particular when the advances in packaging and bonding are taken into consideration. Nevertheless, on-chip co-integration can be achieved successfully if a number of constraints are fulfilled. These are treated in the the following two subsections.

1.3.3.2 Fabrication compatibility

The electronics industry is—justifiably—very conservative with little or no room for experimental processing. As a result, non-standard transducer materials and non-standard fabrication processes can significantly compromise co-integration. To prevent this fabrication compatibility has to be assured, which encompasses:

- *Material compatibility.* The properties of a non-standard material or a combination of non-standard materials (in relation to the conventional IC fabrication process) are not allowed to interfere the performance of the other components in the system, and *vice versa*.

- *Process compatibility.* Application of a non-standard fabrication process, in particular when micromachining is involved, is not allowed to influence the fabrication of the electronics, and *vice versa.*
- *Equipment contamination restrictions.* Cleanroom re-entrance is generally restricted (and often even prohibited) once non-standard materials or non-standard processing have been applied.

In the work reported in this thesis, only the thermoelements are made of non-standard materials. Therefore, material compatibility needs to be ensured for the thermoelectric material only. Still, the observation that the highest-performance materials are (nearly always) the least compatible means a trade-off in performance is required. Material performance is addressed in detail in Chapter 3.

Process compatibility needs to be considered throughout the entire fabrication stage of the thermoelectric cooler. This not only involves scrutinising the growth, doping, and thermal treatment of the thermoelectric material, but particularly how and when to apply specific etching steps. The latter are required to properly shape the thermoelectric coolers as well as the devices in which these are applied.

Finally, cleanroom restriction issues are best dealt with during drafting of the process flow, by properly arranging the order of all fabrication steps required in the co-integration process. The most structural way to achieve this is to separate the fabrication steps of the circuit and transducer as much as possible. This avoids frequent transitions between the two processes. Ideally, the transducer fabrication process is even designed as an entirely separate module from the standard IC process. The advantage of such fabrication modularity is addressed below.

1.3.3.3 Fabrication modularity

To maximise performance, a fully customised fabrication process is unavoidable. This fully intertwines the fabrication of the circuits and transducers. Because of the delicate balance between all fabrication steps, a customised process becomes very inflexible. Virtually no changes can be made without a significant influence on the overall process. Moreover, as only a small portion of the standard IC fabrication line may still be used (due to cleanroom restriction issues) a huge investment in dedicated fabrication equipment is required. So, unless high-volume production is anticipated, a fully customised process is very impractical.

The drawbacks of a dedicated co-integration process can be circumvented by consecutive rather than intertwined processing of electronics and transducer [1.24]. Figure 1.8 indicates the three common positions of the transducer fabrication module, in relation to the standard IC fabrication process. The philosophy is to leave the standard IC fabrication process intact, which allows the existing microelectronics infrastructure to be used with little restrictions. As a result, intermediate processing should be avoided if possible, as this interrupts the standard IC fabrication process. The only common exception is fabrication of a transducer directly in advance of the final metallisation. Front-end transducer processing is also restricted in general. In

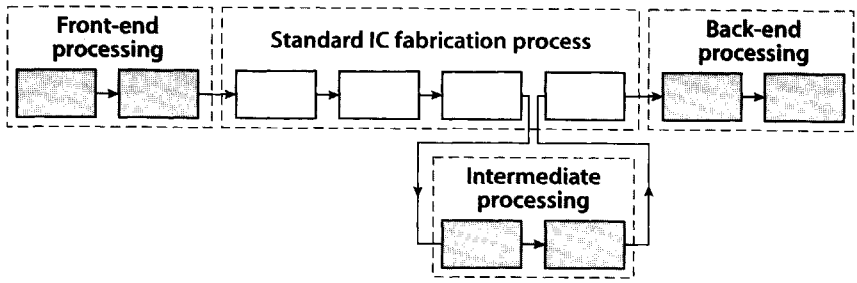


Figure 1.8 The three ways in which a MEMS fabrication process can be co-integrated with a standard IC fabrication process, when using a modular fabrication approach.

particular when micromachining is involved, the wafer surface and wafer thickness are modified very likely, so that the fabrication compatibility constraints are violated. As a result the wafer becomes unsuitable for the consecutive standard IC processing. Therefore, of the available transducer fabrication modules, back-end processing is preferred, as this saves the 'messy' and most intrusive processing until after the standard IC fabrication process.

Keeping both the fabrication compatibility as well as the fabrication modularity in mind, the fabrication of the TEC can be implemented using a fabrication compatible back-end process. This fabrication process is discussed in detail in Chapter 4.

1.4 METHODS OF APPLYING A THERMOELECTRIC COOLER

The final question that needs to be answered is what a thermoelectric cooler, that has been co-integrated on-chip successfully, can be used for. Besides the obvious application, cooling, two more application ranges are distinguished, i.e., thermal stabilisation and heat spreading. Each one is introduced, respectively in Section 1.4.1, 1.4.2 and 1.4.3.

1.4.1 Cooling

Of the ways in which a thermoelectric cooler can be applied, cooling is the obvious one. What is less obvious to most people is what you can cool with an on-chip integrated TEC. Unlike its conventional discrete counterpart, it's not capable of cooling an entire chip (which is explained in a moment.) Also, living in The Netherlands—which has more brands of beer than people working on thermoelectrics—a frequently encountered question is whether such an on-chip integrated TEC can cool the inquirer's precious liquid. Answering: "Only a drop at a time," doesn't capture their attention. So, there aren't many 'useful' applications, right? Not quite. It is important to realise how the behaviour of the thermoelectric cooler changes with downscaling. This requires the performance of the on-chip integrated thermoelectric cooler to be understood. To explain all the detail requires a chapter in itself (Chapter 2, in fact).

However, from a few simple assumptions, an easy to understand estimate can be made. Three factors are of importance:

1. The lowest temperature (T_c) below ambient temperature (T_h) that can be obtained. Throughout this thesis this will be expressed as the maximum temperature difference $\Delta T_{max} = T_h - T_c$. (See Figure 1.2a on page 3.)
2. The amount of power q_c that can be removed by the thermoelectric cooler.
3. The operating efficiency η of the thermoelectric cooler.

Without going into detail on why (which is done in Chapter 2) the maximum temperature difference ΔT_{max} can be approximated by:

$$\Delta T_{max} = \frac{1}{2}\frac{\alpha^2}{\rho\lambda}T_c^2 \tag{1.2}$$

This equation reveals that ΔT_{max} only depends on the cold-side temperature (T_c) as well as on material properties; These are the Seebeck coefficient α, the electrical resistivity ρ and the thermal conductivity λ. In other words, the maximum temperature difference does not depend on the size of the device: In a first-order approximation it is reasonable to assume that downscaling does not influence the maximum temperature difference.

With high-performance thermoelectric materials like bismuth telluride (Bi_2Te_3), cooling to 60 °C below ambient temperature is possible[4]. Based on the above reasoning, this should also be obtainable with an integrated TEC. However, due to the fabrication compatibility constraints, a medium-performance material must be used, like polycrystalline silicon-germanium (poly$Si_{0.7}Ge_{0.3}$). Based on Equation (1.2), polySiGe should still provide a ΔT_{max} of 15 °C.

Next, the amount of power that can be removed by the TEC is expressed as:

$$q_c = \frac{A}{L}\underbrace{\left(\frac{(\alpha_p - \alpha_n)^2 T_c^2}{2\rho} - \lambda\Delta T_{max}\right)}_{\text{size-independent}} \tag{1.3}$$

Besides a large size-independent portion, this equation reveals that q_c is also dependent on geometry, through the ratio of the cross-sectional area A and length L of the TEC. Taking a typical size for the conventional thermoelement, 1 mm (W)ide × 3 mm (L)ong × 1 mm (H)igh, the ratio $A/L = W{\cdot}H/L = 1{\cdot}10^{-3}{\cdot}1{\cdot}10^{-3}/3{\cdot}10^{-3} = 3{\cdot}10^{-4}$ m. Similarly, a typical on-chip integrated thermoelement is in the order of 50 µm × 200 µm × 1 µm, so $A/L = 50{\cdot}10^{-12}/200{\cdot}10^{-6} = 2.5{\cdot}10^{-7}$ m. So, due to shrinking to chip-scale dimensions, in particular due to the reduction of the height of the TEC, the level of power that can be handled by a TEC drops by a factor of 10^3. Typically, as the

4. Under zero-load conditions, e.g., when there is no power generated within the region from which thermal energy is removed by the thermoelectric cooler.

conventional TEC removes a few Watts per thermoelement, an integrated thermoelement will be capable of removing in the order of milliWatts of thermal energy only. This explains why an on-chip integrated thermoelectric cooler is not capable of cooling an entire chip.

The estimates with respect to ΔT_{max} as well as q_c were made under one very important assumption: the thermoelement is the only physical object in between the cooled area and the ambient. If this assumption is not met, both ΔT_{max} and q_c rapidly approach to zero. As such, it is essential that the on-chip integrated thermoelements are thermally isolated; Placing them on a substrate with a large thermal conductivity like Si is simply not done, as the thermal volume of the substrate practically nullifies the effect of the thermoelement. Typically thermal insulation this results in thin-film membranes, beams and cantilevers, with a thickness of 1–5 μm.

With the above estimates for the temperature and power range (i.e., ΔT_{max} around 15 °C and a q_c in the order of a few milliWatts) typical applications that can be targeted for cooling are small thermally isolated circuits containing low power electronics or passive devices like integrated dew-point sensors.

1.4.2 Thermal stabilisation

In most cases of practical importance a thermoelectric cooler will be part of a larger thermal control system. The most frequently encountered intent of such a system is to accurately stabilise the temperature of a component at a user-specified value. Figure 1.9 helps illustrate this.

First consider the system that has no active temperature control. This corresponds to the open-loop system indicated by the light grey blocks. The problem with such a system is that any changes in heat generation will cause a nearly uncontrollable temperature change in the system. As in particular the heat generation in the ambient of the system is difficult to control, the physical coupling with external heat sources must be eliminated whenever possible, i.e., the thermal resistance between the ambient and the system must be maximised. Generally, the thermal resistance K^{-1} of a thermally conducting object can be expressed as

$$K^{-1} = \lambda^{-1} \frac{L}{A}. \tag{1.4}$$

So, in order to increase K^{-1}, the thermal conductivity λ and the cross-sectional area A must either be decreased, or the thermal path length L must be increased. In practise, these three options are applied simultaneously. Unfortunately, even though a high thermal resistance is able to keep out the externally generated heat, it also traps the internally generated heat. Therefore, the temperature of a thermally isolated structure is highly sensitive to fluctuations in internal heat generation.

To improve the accuracy and swiftness of thermal stabilisation, feedback is required, indicated by the path that includes the medium grey block in Figure 1.9. According to the difference between the system temperature measured and the

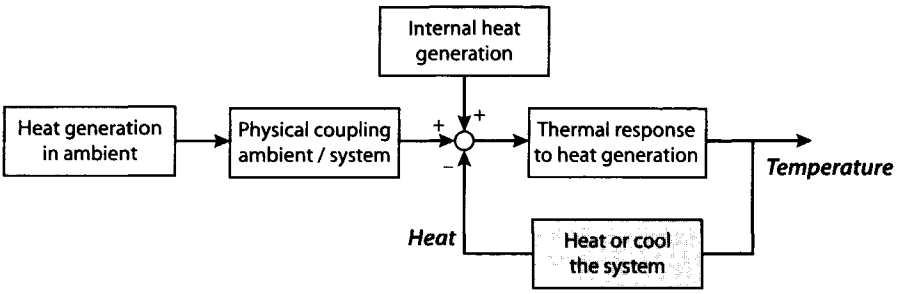


Figure 1.9 Schematic diagram of a thermal control system with feedback.

temperature required, the amount of heat within the system is adjusted in such a way that the system temperature becomes equal to the required temperature.

On a true micro-scale level, only one kind of closed-loop thermal control system is encountered at current, the micro-hotplate [1.25][1.26]. In such a device, the feedback consists of a temperature sensor and a heater only. The drawback of the micro-hotplate is that, as the name already suggests, such a device can only operate at elevated temperatures. The reason for this is that the feedback is only capable of adding thermal energy to the system. As a result, when the temperature in the hotplate is above the required temperature, thermal energy can only be removed from the system passively, through K^{-1}, which was designed to be very large. The heat release from the hotplate to ambient, q_{sa}, can be expressed as

$$q_{sa} = K^{-1}(T_s - T_a). \tag{1.5}$$

For the system to respond significantly fast, as much heat as possible must be released (i.e., q_{sa} must be considerable.) This can only be achieved if $T_s \gg T_{a,\max}$, so a typical micro-hotplate can only operate reliably at temperature of 90 °C and above.

When understanding the above, the integration of a thermoelectric cooler with a micro-hotplate should be obvious. It directly transport heat from the system to ambient, 'bypassing' the high thermal resistance K^{-1}. This not only makes active heat removal (by means of the TEC) much faster than passive heat removal (through K^{-1}), it also allows thermal stabilisation at a much lower temperature.

As a result of the downward extension of the thermal operating range, when using on on-chip integrated TEC, it is more suitable to talk about a micro-thermostat than a micro-hotplate. Figure 1.10 compares the operating range of the typical micro-hotplate to that of a micro-thermostat, for different TEC materials. In particular, the micro-thermostat opens up the way to thermal control/stabilisation in the 0 °C to 75 °C range, in which many biomedical applications are found, e.g., cell stimulus, DNA sequencing, etc.

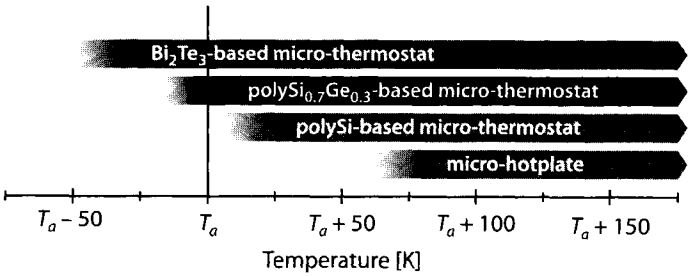


Figure 1.10 Operating ranges of micro-thermostats with different thermoelectric materials, in comparison to the operating range of a micro-hotplate. (T_a refers to the ambient temperature.)

1.4.3 Heat spreading

Even though an on-chip integrated thermoelectric cooler is only capable removing milliWatts of power from a small thermally isolated volume, it can remove substantially more if the volume is above ambient temperature for two reasons. First, the heat removal rate by means of the Peltier effect is $q_P = \alpha \cdot T \cdot I_{el}$. Thus, as T increases, so does the amount of heat removed. Second, the heat transported by thermal conductivity is always from the body with a higher temperature to the body with a lower temperature. As the object to be cooled is above ambient temperature, passive cooling cooperates with Peltier cooling and may even outweigh it.

Such a situation occurs exactly in high-frequency gallium arsenide (GaAs) power circuits, used for telecommunication applications. The major drawback of GaAs is that its thermal conductivity decreases with temperature. In other words, a hot spot becomes more thermally isolated as the temperature increases. If no measures are taken, this can lead to thermal runaway, resulting in certain destruction of the circuit. For this reason, many telecommunications chips are cooled as a whole, which is not very economic. In contrast, as the hot spot is very localised, it helps to spread the heat of the hot spot over the rest of the chip. An on-chip integrated thermoelectric cooler is well suited for this purpose, as it is located along the surface of a chip. This may provide a more economic way of keeping the operating temperature of a chip within safe boundaries. The idea is illustrated in Figure 1.11. By spreading the heat over a large area, the ambient temperature will rise slightly, while the hot spot temperature is decreases significantly.

Moreover, in many occasions, the power transients in the high-frequency circuits are very large and occur very sudden. Reference [1.12] even mentions that response times in the order of 100's of microseconds are required for a typical Pentium® microprocessor, in order to prevent overheating. An external cooling solution just isn't fast enough to provide such response times. But an on-chip integrated TEC might just be, as it can be integrated only microns away from the hot spot.

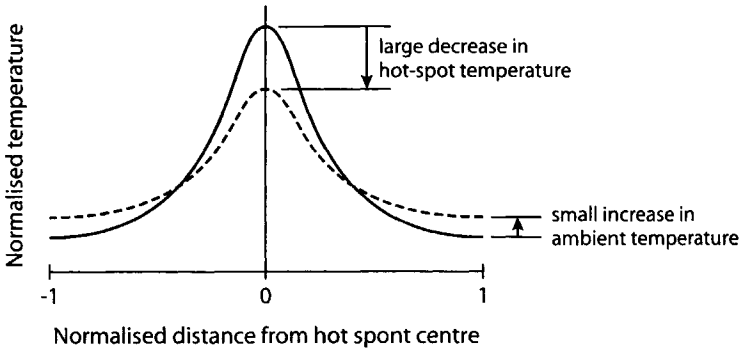


Figure 1.11 Sketch to illustrate the principle of heat spreading / hot-spot avoidance.

1.5 MOTIVATION AND OBJECTIVES

The field of thermoelectrics is slowly gaining interest from science and industry. In particular the recent advances in high-performance thermoelectric materials have spurred this interest. On the downside, most of the work performed remains limited to either materials research or optimising and improving the conventional thermoelectric devices. Research on the on-chip integration as well as the development of chip-scale thermoelectric applications is lagging behind.

The objective of my work has been to fill this void: This thesis proves it is possible to successfully develop and apply on-chip integrated thermoelectric coolers and to correctly predict their performance. The philosophy behind this work is not to produce a physically exhaustive piece of research, but to provide a straightforward and easy to read guide to developing on-chip integrated TECs.

In this introductory chapter I have tried to point out that my work requires far different problems to be addressed than those commonly considered by the majority of the thermoelectrics community. In particular the practical considerations of a TEC have been considered, with respect to performance (Chapters 2 and 3), fabrication (Chapter 4) and potential application (Chapter 5).

1.6 ORGANISATION OF THIS THESIS

This first chapter introduces the design and application aspects of the on-chip integrated thermoelectric cooler, as these are treated in the subsequent chapters. First the basic thermoelectric effects and the theoretical device modelling are treated. Next, the important fabrication aspects are discussed. Finally the three major application ranges for an on-chip integrated TEC are outlined, i.e., cooling, thermal stabilisation and heat spreading.

Chapter 2 discusses the theoretical performance of the on-chip integrated thermoelectric cooler. First of all this includes the non-idealities that become prominent

when downscaling (contact resistance, parasitic thermal conduction, radiant heat exchange.) Also, the influence of the substrate on device performance in analysed, as this substrate has to act as the integrated TEC's heat sink. Finally, the thermal-electrical interface is considered, in order to incorporate the thermoelectric feedback by the Seebeck effect, which can modify the magnitude of electrical current required to drive the thermoelectric cooler.

Chapter 3 presents the full thermophysical characterisation of the thermoelectric material used, polycrystalline silicon germanium ($polySi_{0.7}Ge_{0.3}$). This includes the resistivity, Seebeck coefficient, thermal conductivity and contact resistance. Where possible, the temperature coefficients of the material properties have been determined.

Chapter 4 goes into more detail on the fabrication issues introduces in this chapter, i.e., as thermoelectric materials are relatively difficult to use with electronics, co-integration and fabrication compatibility have to be investigated. The result of our design efforts has lead to a number of fabrication compatible processing modules, primarily for the construction of the thermoelectric cooler and related applications, but also for the deposition of non-standard metallic thin-film resistors, for example.

Chapter 5 finally goes into detail on the applications mentioned in Section 1.4, cooling, thermal stabilisation and heat spreading. Even though the latter application has been limited to a feasibility study, the first two applications have been experimentally verified to work.

In *Chapter 6* the conclusions that can be drawn from this work are presented. After the conclusions, two appendices have been added. *Appendix A* goes into more detail on the theory of thermoelectricity. *Appendix B* provides an overview of thermoelectric materials, together with some indications where thermoelectric materials research is headed next.

1.7 REFERENCES

[1.1] M. Peltier, "Nouvelles expériences sur la caloricité des courans électriques," *Annales de Chimie et de Physique*, LVI (56), pp. 371-386, 1834.

[1.2] K.E. Mautz, T. Morgenstern and R. Schuster, "Evaluation techniques for 300-mm equipment," *Proc. of the SPIE*, Vol. 4182, pp. 316–326, 2000.

[1.3] R. Glasbergen, *Technology bytes!*, CCC Publications, Chatsworth, CA, USA, 1996, ISBN: 1576440141.

[1.4] "Feeling the heat", *The Economist Technology Quarterly*, 15 March 2003, pp. 25–26.

[1.5] Semiconductor Industry Association, International technology roadmap for semiconductors: 2000 update, Austin, TX: International SEMATECH, 2000. (http://public.itrs.net)

[1.6] Semiconductor Industry Association, International technology roadmap for semiconductors: 2001 edition, Austin, TX: International SEMATECH, 2001. (http://public.itrs.net)

[1.7] R. Ronen, A. Mendelson, K. Lai, S-L. Lu, F. Pollack and J.P. Shen, "Coming challenges in microarchitecture and architecture," *Proc. of the IEEE*, Vol. 89, No. 3, pp. 325–340, 2001.

[1.8] R.W. Keyes, "The future of the transistor," *Scientific American Special Issue: the solid-state century*, Vol. 8, No. 1, pp. 46–52, 1997.

[1.9] R. Gonzalez, B. Gordon and M. Horowitz, "Supply and threshold voltage scaling for low-power CMOS," *IEEE J. Solid-State Circuits*, Vol. 32, No. 8, pp. 1210–1216, 1997.

[1.10] S. Borkar, "Design challenges of technology scaling," *IEEE Micro*, Vol. 19, pp. 23–29, 1999.

[1.11] F. Pollack, "New microarchitecture challenges in the coming generations of CMOS process technologies," *32nd Annual International Symposium on Microarchitecture (Micro-32)*, Haifa, Israel, 16–18 November 1999. (http://developer.intel.com/research/mrl/library/micro32keynote.pdf)

[1.12] S.H. Gunther, F. Binns, D.M. Carmean and J.C. Hall, "Managing the impact of increasing microprocessor power consumption," *Intel® Technology Journal*, 2001. (http://developer.intel.com/technology/itj/q12001.htm)

[1.13] A. Allan, D. Edenfeld, W.H. Joyner (Jr.), A.B. Hahng, M. Rodgers and Y. Zorian, "2001 Technology roadmap for semiconductors," *IEEE Computer*, pp. 42–53, 2002.

[1.14] G.D. Hutcheson and J.D. Hutcheson, "Technology and economics in the semiconductor industry," *Scientific American Special issue: the solid-state century*, Vol. 8, No. 1, pp. 66–73, 1997.

[1.15] M.J. Bass and C.M. Christensen, "The future of the microprocessor business," *IEEE Spectrum*, Vol. 39, No. 4, pp. 34–39, 2002.

[1.16] L. Geppert, "Moore's Law mania," *IEEE Spectrum*, Vol. 35, No. 1, pp. 20–21, 1998.

[1.17] K. Bult, "Analog broadband communication circuits in pure digital deep sub-micron CMOS," *Proc. IEEE International Solid-State Circuit Conference*, 15–17 February 1999, San Francisco, CA, USA, pp. 76–77.

[1.18] M. Pelgrom, "Het einde van de éénchip-droom?" *Bits & Chips*, Vol. 4, No. 6, p. 21, 2002 (in Dutch).

[1.19] MEMS Industry Group Annual Report 2001. (http://www.memsindustrygroup.org/)

[1.20] H. Allen and R. van Kampen, "Path to integrated sensors—an industrial perspective," *Proc. 1998 Microsystem Symposium*, 10–11 September 1998, Delft, The Netherlands, pp. 87–91.

[1.21] L. Spangler and C.J. Kemp, "Integrated silicon automotive accelerometer," *Proc. Transducers '95*, Stockholm, Sweden, 25-29 June, 1995, Vol. 1, pp. 585–588.

[1.22] T.L. Michalka, R.C. Varchney and J.D. Meindl, "A discussion on yield modelling with defect clustering, circuit repair and circuit redundancy," *IEEE Trans. Semicond. Manuf.*, Vol. 3, No.3, pp. 116–127, 1990.

[1.23] R.F. Wolffenbuttel, *Silicon sensors and circuits: on-chip compatibility*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995, ISBN: 0-412-70970-8, 313 pp.

[1.24] P.M. Sarro, "Sensor technology strategy in silicon," *Sensors and Actuators A*, Vol. 31, pp. 138–143, 1990.

[1.25] E.H. Klaassen, R.J. Reay, C. Storment and G.T.A. Kovacs, "Micromachined thermally isolated circuits," *Sensors and Actuators A*, Vol. 58, No. 1, 43–50, 1997.

[1.26] J. Puigcorbe, A. Vila, I. Gracia C. Cane and J.R. Morante, "Dielectric micro-hotplate for integrated sensors: an electro-thermo-mechanical analysis," *Proc. Transducers '01 / Eursensors XV*, 10–14 June 2001, Munich, Germany, pp. 312–315.

Chapter 2

THEORETICAL DEVICE PERFORMANCE

The Peltier effect was discovered and interpreted over a century before the first transistor was ever created [2.1]. Nevertheless, the transistor has been analysed in practically every aspect, while the thermoelectric effects remain hard to predict to this very day, even with the most rigorous quantum mechanical treatment. But, once the thermophysical parameters have been determined, the operation of (integrated) Peltier elements can be determined with only a minimal amount of mathematics. This allows an intuitive top-down approach to model the operation of on-chip integrated thermoelectric coolers.

First, an introduction to the thermoelectric effects and their relations is provided in Section 2.1. Next in Section 2.2 the basic set of equations that describes the operation of the conventional Peltier element is defined and consecutively extended in Section 2.4. This is done by identifying and analysing all non-idealities of the on-chip integrated device, after which these are embedded in the basic set of equations. This result provides a fairly accurate prediction of the maximum temperature difference between the hot and cold junctions.

Using a more rigorous treatment by means of Fourier's law, in Section 2.5 a model for the temperature profile in the beam rather than the temperature difference between its ends is derived. Finally, in Section 2.6 finite element analyses (FEA) are presented that merge all existing non-idealities in a single model. These include electrical contact resistance, thermal resistance of the substrate and thermal losses by means of conduction, convection and radiation.

2.1 DEDUCTION OF THE KELVIN RELATIONS

Consider the schematic thermocouple of Figure 2.1a, which is a simplification of the TEC analysed further on. This element is the basic building block of any on-chip integrated thermoelectric device, whether for temperature sensing, thermoelectric power generation or thermal energy transportation. Fundamentally, it consists of two physically dissimilar thermoelements, connected in series. Preferably, these thermoelements are constructed of a semiconductor material, with one of the ther-

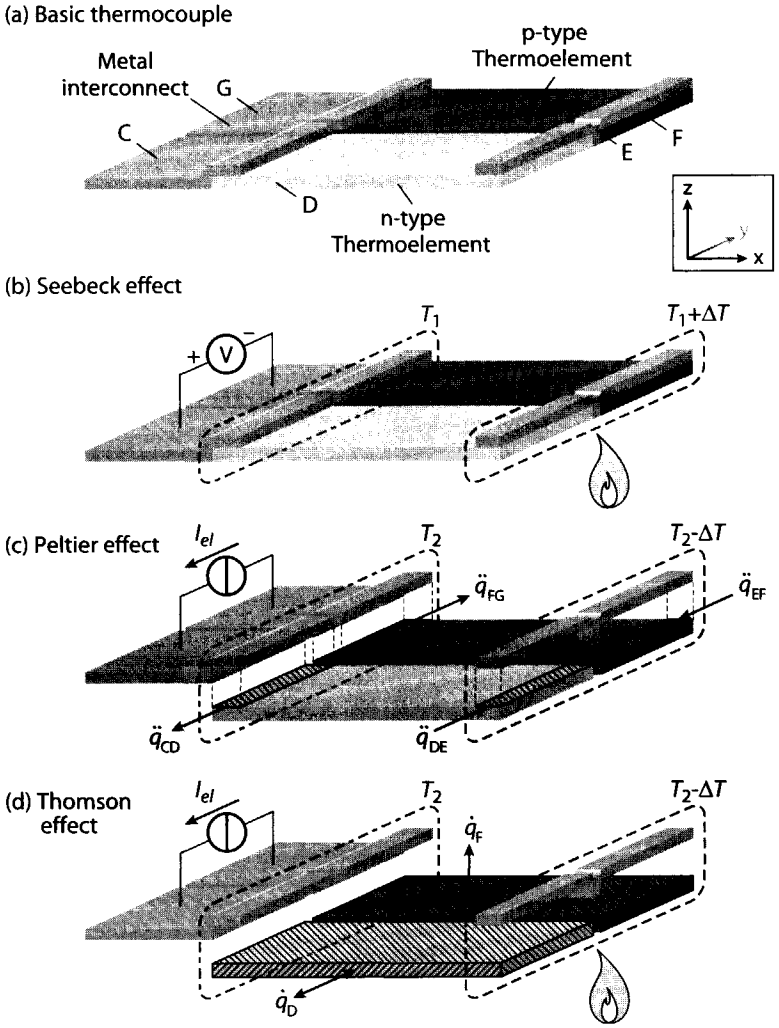


Figure 2.1 (a) Schematic view of a single thermocouple, consisting of a p-type and n-type leg connected in series by metal interconnect. Below that are schematic representations of (b) the Seebeck effect, (c) the Peltier effect and (d) the Thomson effect.

moelements doped with a p-type impurity and the other with an n-type impurity. (An intermediate metallic conductors is required to prevent a direct p-n junction, which prevents the thermocouple from working properly.) The thermoelectric material properties are discussed in more detail in Chapter 3. In any thermoelectric device with a closed electrical loop all three thermoelectric effects introduced in Chapter 1, are present. Their more formal definitions are as follows:

- *Seebeck effect*: the generation of an electrical potential within an electrically conducting material that is subjected to a temperature gradient.
- *Peltier effect*: reversible change in the thermal energy content at an interface between dissimilar electrically conducting materials, that results from the flow of an electrical current across that interface.
- *Thomson effect*: reversible change in thermal energy content within an electrically conducting material, that results from the flow of electrical current through the material, while subjected to a temperature gradient.

In Figure 2.1b, the Seebeck effect is visualised. Heating one side of the thermocouple, a temperature gradient of magnitude ΔT is created. The free carriers at the hot side of the thermocouple now have an increased kinetic energy, which results in a net diffusion of carriers to the cold side. The resulting charge buildup at both ends of the thermoelements creates a reverse emf (electromotive force) that opposes the flow of charge carriers. The open-circuit potential ΔV that arises at the ends of the thermocouple is referred to as the Seebeck voltage. The absolute Seebeck coefficient (ASC) of a material, α, expressed in [VK^{-1}], indicates the rate of change of the Seebeck voltage with respect to a change in temperature, at a given temperature:

$$\alpha = \left.\frac{dV}{dT}\right|_T \tag{2.1}$$

The relative Seebeck coefficient (RSC) of two materials A and B is given by

$$\alpha_{AB} = \alpha_A - \alpha_B = \left.\frac{dV_{AB}}{dT}\right|_T \tag{2.2}$$

If the electrical circuit forms a closed loop, all three thermoelectric effects appear, respectively transferring thermal energy at the junctions and within the thermoelements. For now, it is assumed, that there is no energy transfer with the surroundings, i.e., thermoelectric effects only are present. As these effects are reversible, the power generated by the Seebeck effect has to equal the other two effects:

$$I_{el} \cdot \frac{dV_{AB}}{dT}\Delta T = I_{el} \cdot \left[\left(\pi_{AB}(T+\Delta T) - \pi_{AB}(T)\right) + \left(\beta_B \Delta T - \beta_A \Delta T\right)\right] \tag{2.3}$$

with I_{el} the electrical current, π the Peltier coefficient and β the Thomson coefficient. For unit current, Equation (2.3) may be rewritten as:

$$\begin{aligned} \frac{dV_{AB}}{dT} &= \left[\frac{\left(\pi_{AB}(T+\Delta T) - \pi_{AB}(T)\right)}{\Delta T} + (\beta_B - \beta_A)\right]_{\Delta T \to 0} \\ \rightarrow \alpha_{AB} &= \frac{d\pi_{AB}}{dT} + (\beta_B - \beta_A) \end{aligned} \tag{2.4}$$

The Peltier effect is depicted in Figure 2.1c. If the electrical current I_{el} is consecutively run through the n- and p-type thermoelements (along the path CDEFG as indicated in Figure 2.1a), thermal energy is removed from the system at junctions (DE) and (EF) and is liberated again at the junctions (CD) and (FG). The position of the junctions is indicated by the hatched areas. The absolute Peltier coefficient, π, which indicates the rate of thermal energy removal at unit current at a given temperature, is most conveniently expressed in terms of the Seebeck coefficient (as in [2.2] or [2.3]):

$$\pi \equiv \frac{dV}{dT} T = \alpha T \tag{2.5}$$

Finally, Figure 2.1d provides a graphical interpretation of the Thomson effect. When subjected to both an electrical current and a temperature difference, thermal energy is exchanged within the thermoelements D and F. The difference in Thomson coefficients can again be expressed as a function of the Seebeck coefficient, i.e.,

$$\frac{(\beta_A - \beta_B)}{T} = \frac{d\alpha_{AB}}{dT} = \frac{d^2 V_{AB}}{dT^2} \tag{2.6}$$

Equations (2.5) and (2.6) are generally indicated as the Kelvin relations. For modelling purposes, the description of the thermoelectric effects in the form of Equations (2.1)–(2.6) is sufficient to obtain good results. It should be realised that these equations have a few shortcomings nonetheless, mainly with respect to the temperature dependency of the material properties. These issues will be dealt with at the appropriate moment.

2.2 CONVENTIONAL PELTIER ELEMENT PERFORMANCE

The thermocouple is the basic building block of any conventional (discrete) Peltier device. When operated in cooling mode, thermal energy is removed from the system at junctions (DE) and (EF). These junctions will be referred to as the cold junctions. Similarly thermal energy is released at junctions (CD) and (FG), referred to as the hot junctions. Furthermore, a number of assumptions are made:

- The metal interconnect (parts C, E and G in Figure 2.1a) is assumed to have zero electrical resistance, which implies that no Joule heating is generated within these metallic parts.
- These metallic parts are assumed to have zero thermal resistance, which implies no thermal gradient occur within. Because of this zero temperature gradient, no Seebeck potential occurs across individual metallic conductors.
- The only thermal interaction with the ambient is through the metal interconnect at the hot side (parts C and G). These are assumed attached to a large heat sink, fixed at ambient temperature T_a. ('Large' means large enough to handle any thermal energy flux from the thermocouple, without

a change in the heat sink temperature.) As the electrical and thermal resistance of the metallic parts are assumed to be zero, the temperature of parts C and G is equal to T_a. As a result, the Peltier device's thermal energy is directly released into a body at ambient temperature T_a.

- Finally, thermal energy exchange by means of radiation and convection through the ambient air are ignored, as well as any thermal contact resistance.

Although the above assumptions reduce the modelling accuracy, the models have proven to be robust, under the condition the temperature difference between the hot and cold junctions is reasonably small (up to a few tens of Kelvins).

Parameters of interest

For convenience, Figure 2.2 shows the thermocouple under investigation with the relevant parameters that are required in the thermoelectric cooling performance analysis. From this point on, the subscripts 'n' and 'p' rather than 'A' and 'B' are used to distinguish between the properties of the n- and p-type thermoelements. As such, the material parameters included in the analysis are the thermal conductivities, λ_n and λ_p, the Seebeck coefficients, α_n and α_p, and the electrical resistivities, ρ_n and ρ_p. Furthermore, three geometrical parameters are used, i.e, the lengths L_n and L_p, the widths W_n and W_p, and heights (thicknesses) H_n and H_p of the thermoelements. As stated, the temperature of the hot junctions is fixed at T_h, while the cold junctions are at temperature T_c.

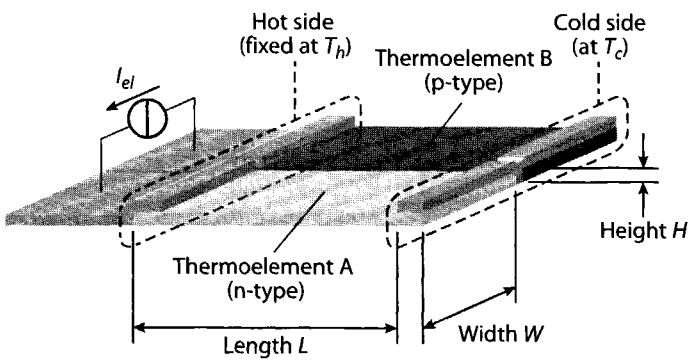


Figure 2.2 Schematic view of a single thermocouple, of which the thermoelectric cooling performance is analysed.

In general, three parameters are of interest in the analysis of Peltier elements:

1. The amount of thermal energy, q_c [W], that is removed from the cold junction.

2. The maximum temperature difference, ΔT_{max} [K], that can be obtained between the hot and cold junction, e.g., $\Delta T_{max} = (T_h - T_c)_{max}$.
3. The efficiency, η, of a Peltier element, defined as the ratio between the total power supplied to the thermocouple, q_{in}, and the amount of thermal energy removed from the cold junctions, q_c, e.g., $\eta = q_c/q_{in}$.

The first two parameter will be derived in subsections 2.2.1 and 2.2.2. At this stage of research, the efficiency of thin-film TEC is of lesser importance. As such, only the work done by other authors is summarised in subsection 2.2.3.

2.2.1 Thermal energy removal rate

The amount of thermal energy removed from the cold junctions by means of the Peltier effect, q_P, can be derived from Equation (2.5) and is equal to:

$$q_P = q_{P,n} + q_{P,p} = -\pi_n I_{el} + \pi_p I_{el} = (\pi_p - \pi_n) I_{el} = (\alpha_p - \alpha_n) T_c I_{el} \tag{2.7}$$

Note that the Peltier and Seebeck coefficients are negative in the case of n-type materials, and positive in the case of p-type materials. The total electrical resistance R of the two thermoelements, electrically in series, is:

$$R = R_n + R_p = \frac{\rho_n L_n}{A_n} + \frac{\rho_p L_p}{A_p} \tag{2.8}$$

where A_n (= $W_n \cdot H_n$) and A_p (= $W_p \cdot H_p$) are the cross-sectional areas of the thermoelements. Hence, the total Joule heating q_J amounts to:

$$q_J = I_{el}^2 R = I_{el}^2 \left(\frac{\rho_n L_n}{A_n} + \frac{\rho_p L_p}{A_p} \right) \tag{2.9}$$

Although not straightforward to see, only half of the Joule heat actually causes the temperature at the cold junction to rise. Even though this effect is often poorly explained or even ignored, this can be explained simply by using the lumped-element model of Figure 2.3.

For unit cross-sectional area, the total thermal resistance K^{-1} of the thermoelement is equal to L/λ [KW^{-1}]. The total Joule heating q_{el} is distributed over the entire length L. To model the distribution, the element is divided into N sections of equal thermal resistance, $(L/\lambda)/N$ (= K^{-1}/N.) The distributed of the Joule heating is modelled by means of (N-1) heat sources of value $q_J/(N-1)$.

Consider the thermal current source of stage (N-1). The only thermal path for the heat generated in this stage to ambient (modelled by the source T_a) is across all thermal resistors of stages (N-1) downto 1. Hence, the temperature rise of the thermal current source in stage (N-1) is equal to $[(N-1)(K^{-1}/N)] \cdot [q_J/(N-1)]$. More general, the temperature rise due to the thermal current source of stage n ($1 \leq n < N$) is equal to

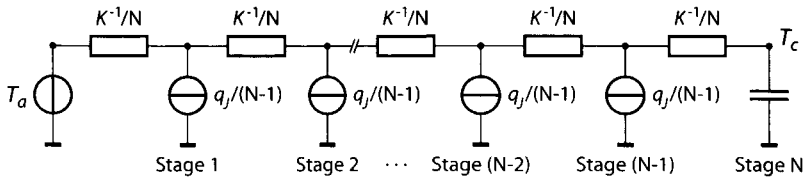


Figure 2.3 Simple lumped-element model of a resistor. The model is split up into N sections, to incorporate the distribution of Joule heating over the entire resistor.

$n{\cdot}(K^{-1}/\mathrm{N}){\cdot}[q_J/(\mathrm{N}\text{-}1)]$. Mathematically, the temperature rise at stage N can now be calculated by summing the individual temperature increase due to each stage:

$$\Delta T = \frac{q_J}{N-1}\frac{K^{-1}}{N}\left[1+\ldots+(N-2)+(N-1)\right] = \frac{q_J K^{-1}}{N^2-N}\sum_{i=1}^{N-1} i \tag{2.10}$$

As $N \to \infty$, this series becomes:

$$\Delta T = \lim_{N\to\infty}\left(\frac{q_J K^{-1}}{N^2-N}\sum_{i=1}^{N} i\right) = K^{-1}\cdot\frac{1}{2}q_J \tag{2.11}$$

Rephrasing Equation (2.11)

$$q = \frac{\Delta T}{K} = \frac{1}{2}q_J = \frac{1}{2}I_{el}^2 R \tag{2.12}$$

Hence, of the total Joule heating, indeed only half contributes to a temperature difference across the thermoelement. In cooling mode, this Joule heating opposes the Peltier cooling effect at the cold junction. Therefore, the heat removal rate at the cold junction, combining the Peltier cooling and Joule heating, now becomes

$$q_c = q_P - q_J = (\alpha_p - \alpha_n)T_c I_{el} - \tfrac{1}{2}I_{el}^2 R \tag{2.13}$$

While electrically in series, the two thermoelements are thermally in parallel. The total thermal conduction, K, of a single thermocouple is:

$$K = K_n + K_p = \frac{\lambda_n A_n}{L_n} + \frac{\lambda_p A_p}{L_p} \tag{2.14}$$

As thermal energy is transported from the hotter to the colder junction, this also opposed the cooling by means of the Peltier effect. Therefore, the final heat removal rate at the cold junction, incorporating Peltier cooling, Joule heating and thermal conduction becomes

$$q_c = (\alpha_p - \alpha_n) T_c I_{el} - K(T_a - T_c) - \tfrac{1}{2} I_{el}^2 R \tag{2.15}$$

Figure 2.4 plots Equation (2.15) as a function of both the electrical current applied and the cold junction temperature, while the hot junction temperature is constant at 300 K. The values of q_c and I_{el} are typical for on-chip integrated polySiGe TECs consisting of a single thermocouple. The material parameters used are those provided for polySiGe in [2.2]. These properties are listed in Table 2.1 on page 51. With respect to the geometrical parameters the following values are used: $H_n = H_p = 0.6$ µm, $L_n = L_p = 250$ µm, $W_n = 45.5$ µm and $W_p = 50$ µm. The reason for the difference in widths is caused by the need to optimise the maximum temperature difference, as will be discussed in Section 2.2.2.

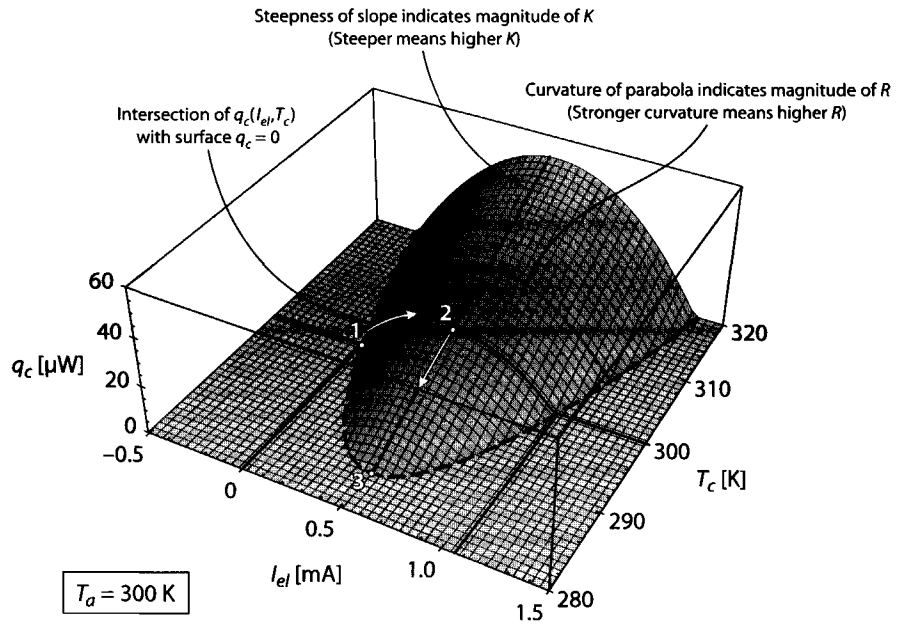


Figure 2.4 Graphical interpretation of the thermal energy removal rate, q_c, as a function of both the electrical current applied, I_{el}, and the cold junction temperature, T_c.

Figure 2.4 should be quite easy to understand. In equilibrium, with zero electrical current applied, the system will be in the state indicated by node 1, with values $(I_{el}, T_c, q_c) = (0,300,0)$. Now, applying a non-zero and positive electrical current, the state of the system will initially move within the $T_c = 300$ K plane, towards node 2 (along the line indicated.) For $0 < I_{el} < 1.06$ mA, q_c is positive, which means that heat is removed from the cold junction. As a result, T_c starts dropping in the (I_{el}, q_c) plane along the surface, until q_c becomes zero. If the optimal current is chosen, T_c will

reach node 3, in which case T_c is at the lowest possible value obtainable, i.e., for this value of I_{el}, the maximum temperature difference $\Delta T_{max} = (T_a\text{-}T_c)_{max}$ is obtained.

Two dashed curves and one dashed line have been included in Figure 2.4, which are quite powerful tools in explaining the performance of the device. The dashed line shows the slope of the surface along the (I_{el}, T_c) plane. The derivative of this slope depends on T_c only and is $dq_c/dT_c = K$. Thus, the steeper the slope of the surface, the higher the thermal conduction in the TEC. Similarly, the second derivative (i.e., the curvature) of the surface for constant T_c directly reveals the magnitude of resistivity R as the second derivative equals $|d^2q_c/dI_{el}^2| = R$. The higher R, the more curved the surface will be. Finally, the dashed curve in the floor of the plot (where $q_c = 0$) is the curve at which T_c is minimised for the specific electrical current applied. The significance of this curve will become clear when the geometrical optimization of the on-chip integrated TEC is discussed, in Section 2.4. As a final remark, it should be noted that the surface actually extends below the surface $q_c = 0$, but this extension would make the figure harder to interpret.

2.2.2 Maximum temperature difference ΔT_{max}

Now that the cooling rate of a Peltier element has been established, the next step is to determine the temperature difference that can be obtained between the hot and cold junctions. This involves two issues:

1. determination of the optimal electrical current, I_{opt}, at which this difference ΔT_{max} (= $T_a\text{-}T_c)_{max}$ is obtained.
2. determination of the optimal geometry for which ΔT_{max} is obtained.

Each of the two issues is discusses in the following subsections.

2.2.2.1 Determination of ΔT_{max}

The first step to determine the value of ΔT_{max} is to calculate at which the optimum electrical current at which this temperature difference is obtained. This is done by calculating the maximum of q_c with respect to I_{el}, i.e., the curve where $dq_c/dI_{el} = 0$:

$$\frac{dq_c}{dI_{el}} = (\alpha_p - \alpha_n)T_c - I_{el}R = 0 \quad \rightarrow I_{opt} = \frac{(\alpha_p - \alpha_n)T_c}{R} \tag{2.16}$$

Next, substituting I_{el} with the above definition for I_{opt}, in Equation (2.15), the maximum thermal energy removal rate, q_{opt}, can be determined:

$$q_{opt} = \frac{(\alpha_p - \alpha_n)^2 T_c^2}{2R} - K(T_a - T_c) \tag{2.17}$$

Graphically, Equation (2.17) indicates the line that runs through points 2 and 3 of Figure 2.4. At the point where q_{opt} equals zero (at node 3 in Figure 2.4,) the maximum temperature difference ΔT_{max} is found:

$$\Delta T_{\max} = (T_a - T_c)_{\max} = \frac{(\alpha_p - \alpha_n)^2 T_c^2}{2KR} \tag{2.18}$$

At this point, most textbooks on thermoelectric refrigeration introduce the figure-of-merit, Z, to simplify this equation. This figure should be interpreted as a measure of thermoelectric performance of that object, i.e. it indicates the suitability of the object for use with specific thermoelectric applications (but not with all applications, as will be discusses in Chapter 5.) The figure-of-merit is defined as:

$$Z = \frac{(\alpha_p - \alpha_n)^2}{KR} \tag{2.19}$$

This equation defines the relative figure-of-merit of two materials. Also, often the figure-of-merit of a single material $z_A = \alpha_A/(\lambda_A \rho_A)$ is encountered. Good values for z are in the order of 10^{-3} K^{-1}. Instead of z, often the dimensionless parameter zT is used, which is ~1 for good thermoelectric materials. Substitution of (2.19) into (2.18) finally yields

$$\Delta T_{\max} = \frac{1}{2} Z T_c^2 \tag{2.20}$$

2.2.2.2 Geometrical optimization

Obviously, from Equation (2.20), $\Delta T_{\max}$ is maximised if Z is maximised. However, modifying the material properties, requires the material fabrication to be altered, which is highly undesirable. Instead, it is preferable to treat the material properties as constants, while the geometrical parameters (L, W and H) may be manipulated so that the product $K{\cdot}R$ is minimised. The product $K{\cdot}R$ is optimised if [2.2]

$$\frac{L_n A_p}{L_p A_n} = \sqrt{\frac{\rho_p \lambda_n}{\rho_n \lambda_p}} \tag{2.21}$$

Until this point in Chapter 2, most issues discussed are straightforward textbook material (except for the proof that only half of the Joule heating counteracts the Peltier cooling effect.) Instead of just referring to this material, it is presented nonetheless, to properly introduce thermoelectrics to the novice reader. Now, the first step towards modelling of integrated Peltier elements is made: Based on issues related to the fabrication process chosen (discussed extensively in Chapter 4), it is preferable to make the lengths and thicknesses of the both thermoelements equal, i.e., $L_n = L_p = L$ and $H_n = H_p = H$. This simplifies Equation (2.21), as this only leaves the optimal width ratio of the thermoelements needs to be determined:

$$\frac{W_p}{W_n} = \sqrt{\frac{\rho_p \lambda_n}{\rho_n \lambda_p}} \tag{2.22}$$

Under these assumptions, minimising $K \cdot R$ has become quite simple. Note that, irrespective of whether Equation (2.21) or (2.22) is used, geometrical optimization of conventional Peltier element depends on material properties only. Consequently, the the maximum temperature difference is determined by the material properties only.

2.2.3 Cooling efficiency

The efficiency of the thin-film TEC is considered of minor importance in this thesis. Only when large temperature differences can be obtained (> 10 K) the efficiency becomes interesting for optimization purposes. The general formulation for the efficiency of a TEC is [2.2]

$$\eta = \frac{q_c}{q_{in}} = \frac{(\alpha_p - \alpha_n) T_c I_{el} - K(T_a - T_c) - \frac{1}{2} I_{el}^2 R}{(\alpha_p - \alpha_n)(T_a - T_c) I_{el} + I_{el}^2 R} \tag{2.23}$$

In that same reference it is shown that the maximum efficiency is

$$\eta_{\max} = \frac{T_c}{(T_h - T_c)} \frac{\sqrt{1 + ZT_{avg}} - (T_h / T_c)}{\sqrt{1 + ZT_{avg}} + 1} \tag{2.24}$$

where T_{avg} is $(T_h + T_c)/2$. Further refinements regarding thin-film TECs is made by Min and Rowe [2.5][2.6], who have investigated the influence of both the thermal and electrical contact resistance on device performance.

To model the influence, they have used the approximation that the thermoelectric properties of both legs in the thermocouple are equal, i.e. $\alpha_n = -\alpha_p$, $\rho_n = \rho_p = \rho$, and $\lambda_n = \lambda_p = \lambda$. Under these conditions the modified efficiency due to the influence of both non-idealities becomes

$$\eta'_{\max} = \frac{L}{L + 2\frac{\lambda}{\lambda_c} H_c} \left(\frac{T_c}{(T_h - T_c)} \frac{\sqrt{1 + ZT_{avg} \frac{L}{L + 2\rho_c / \rho}} - (T_h / T_c)}{1 + \sqrt{1 + ZT_{avg} \frac{L}{L + 2\rho_c / \rho}}} \frac{\lambda L_c}{\lambda_c L} \right) \tag{2.25}$$

with λ_c the thermal contact resistivity, ρ_c the electrical contact resistivity and H_c the thickness of the thermal contact between the junctions and ambient.

2.3 TWO INTEGRATION APPROACHES

Two approaches to on-chip integration of TECs can be distinguished, respectively referred to as transversal (cross-plane) and lateral (in-plane) TECs. These two terms

refer to the direction in which thermal energy is removed, relative to the surface of the device. Figure 3.6 illustrates the difference.

The transversal TEC of Figure 2.5a operates most like the conventional TEC. In the setup depicted, the volume just below the TEC's cold junction (at T_c) has a certain thermal capacitance, C_{th}, in contact with ambient through the thermal resistance of the substrate, K^{-1}. The TEC transports thermal energy from C_{th}, across the integrated TEC to the second body which acts as heat sink (at a rate $q_{P,tr}$ in the z-direction.) The temperature of this second body is indicated as T_b, which is generally unequal to T_a. As with the conventional TEC, the primary contact between the object cooled and the heat sink is through the TEC. Transversal TECs are reported, for example in [2.7] and [2.8].

In comparison to the transversal TEC (or the conventional TEC for that matter,) the lateral TEC as depicted in Figure 2.5b, transports the thermal energy within a single die, from the thermal mass C_{th} at the tip of the thin-film beam to the substrate. Thus, with a lateral TEC, the thermal energy is not removed from the device!

Due to the huge difference in construction, the intended applications for both types of integrated TECs are very different. Compared to the lateral element, the transversal element generally has a large cross-sectional area (~100 × 100 μm^2), with only a very short distance (~2–5 μm) between the hot and cold junctions. This makes these elements specifically suitable to handle large heat fluxes (as a large electrical driving current can be applied) at very high speeds (due to the low thermal resistance). Moreover, the potential application of very high-performance materials like superlattice materials can greatly boost performance. Transversal elements can easily be placed in large two-dimensional arrays along the wafer surface in the (x,y)-plane, further increasing the total thermal energy removal rate.

The transversal TEC has two drawbacks. First, due to its low thermal resistance between the hot and cold junctions, the thermal time constant of such a device is very small, increasing the susceptibility to temperature variations in its surroundings and making a transversal TEC less suitable for thermal stabilisation. Secondly, the hot side temperature is generally ill-defined, due to the way the hot junction is contacted. Consequently the absolute cold junction temperature is also ill-defined. Only when both T_a and T_b can be accurately determined, T_c can be derived. An poor wiring technique was applied in [2.7], where a single-chip solution was attempted by bond wire between the hot junction and the substrate close t the cooled region. Placing such a highly thermally conductive bond wire is in parallel with the TEC caused a significant reduction in the temperature difference between the hot and cold junctions.

The lateral integrated TEC does not have either of the two disadvantages. First, the only thermal (and electrical) path between the hot and cold junction is through the accurately defined thermoelement/dielectric sandwich of the membrane. This improves both the stability and predictability of the temperature difference between the two junctions. Second, as the lateral TEC has a much higher thermal resistance, it

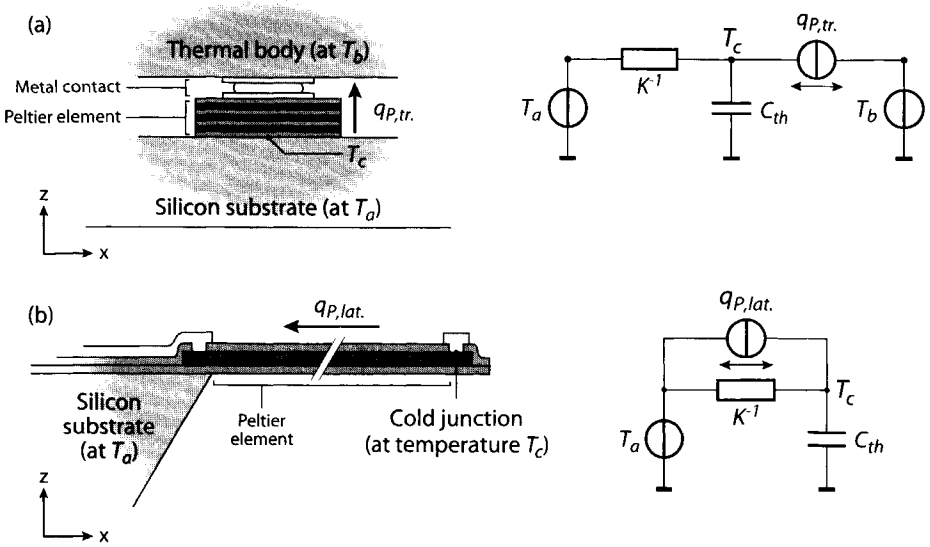


Figure 2.5 Schematic comparison of (a) a transversal on-chip integrated thermoelectric cooler and (b) a lateral on-chip integrated thermoelectric cooler

is less susceptible to rapid changes of the ambient temperature. This high thermal resistance is obtained by a much lower A/L value. Due to the larger distance between the junctions the electrical resistance increases, so only a limited heat flux can be handled by a lateral TEC. Finally, even though it is technically quite possible to increase the cooling power by thin film stacking, the additional fabrication effort required limit the practicality of this appraoch.

2.4 INTEGRATED THERMOELECTRIC COOLER MODELLING

In this section the non-idealities introduced by lateral on-chip integration are discussed and used to extend the conventional model from the previous section. Schematically, the (single-thermocouple) on-chip integrated Peltier element analysed in this thesis looks as indicated in Figure 2.6. Again, the two thermoelements are electrically connected in series (by means of the aluminium interconnect) while thermally in parallel. Additionally, the thermoelements are now sandwiched in between two dielectric layers (respectively of Si_3N_4 and SiO_2), which both have a low thermal conductivity (and zero electrical conductivity.) Furthermore, the hot junction is no longer directly connected to a body at T_a, but to the bulk micromachined silicon substrate of which only the base is at T_a. The top layer of the substrate (i.e., the epi-layer) holds any the electronics with which the TEC is to be co-integrated.

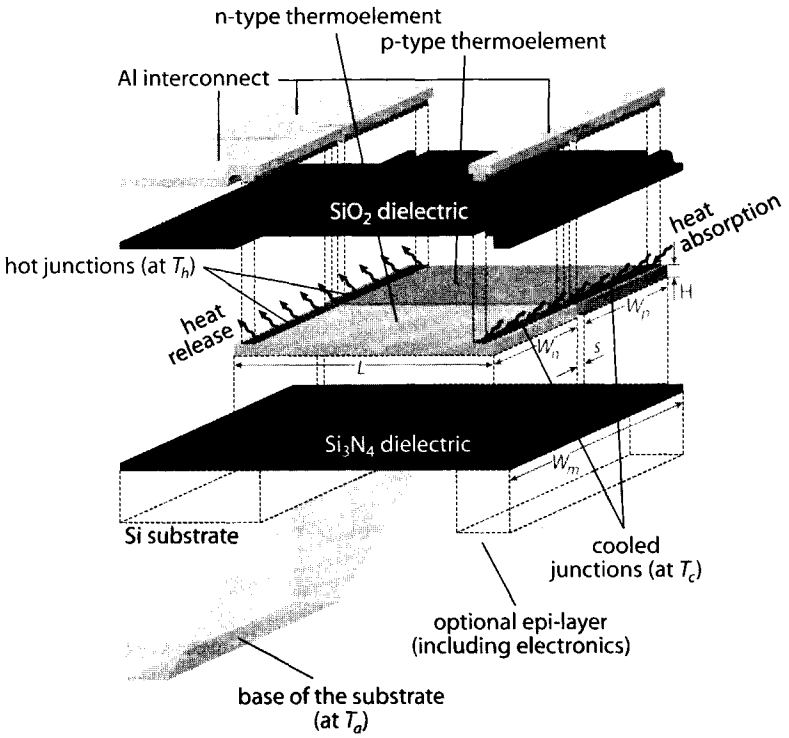


Figure 2.6 Exploded view of a lateral on-chip integrated thermoelectric cooler. The Si_3N_4/polySiGe/SiO_2 stack is thermally isolated by selectively removing the substrate by bulk micromachining.

The ideal model set out in Section 2.2 forms the base for the performance analysis of the on-chip integrated (lateral) TEC as indicated in Figure 2.6. To analyse the operation of such an device, it must be split into two blocks, the thin-film region and the substrate. Conventional, only the thin-film region is of interest, as this is the region in between the hot and cold junctions. What has to be kept in mind though, is that in the case of the planar device the substrate is the only available heat sink. Hence, any heat released from the hot junction has be removed through the substrate. To facilitate the analysis, the (fairly safe) assumption is made that the base of the substrate is in intimate contact with a carrier that has thermal capacity large enough to effectively remain constant, at T_a.

The remainder of Section 2.4 deals solely with the temperature difference across the thin-film region and the influence by the non-idealities in this region. Section 2.5 extends this analysis even further by looking at the temperature profile rather than the maximum temperature difference across the membrane. The thermal resistance

of the substrate does not influence the temperature difference between the junctions, but does influence the absolute temperature of these junctions. This effect will be discussed in detail in Section 2.6.

2.4.1 Thermal leakage through the dielectric membrane (K_m)

In the conventional TEC, the thermoelements form the only considerable thermal contact between the cooled region and the heat sink. In comparison, the on-chip integrated TEC is sandwiched in between two dielectrical layers (see Figure 2.6.) These layers are required due to fabrication constraints, even though these layers are detrimentary to device performance. The bottom Si_3N_4 layer is required as

- mechanical support, to strengthen the membrane
- mechanical etch stop, to halt the bulk etch when removing the substrate below the membrane, and
- electrical insulator, to ensure proper device operation.

Similarly, the top SiO_2 dielectric layer serves as both mechanical support and electrical insulator. The additional thermal conduction through these dielectric layers (thermally in parallel with the thermoelements) can be taken into account by adding a term into Equation (2.14) that describes the total conduction K_m through the additional layers:

$$K' = K + K_m \tag{2.26}$$

where

$$K_m = \sum_i K_{m,i} = \sum_i \left(\frac{\lambda_{m,i} A_{m,i}}{L_{m,i}} \right) \tag{2.27}$$

In the case of the device in Figure 2.6, $i = 2$, as two dielectric layers are present (i.e., the Si_3N_4 and SiO_2 layers.) The decrease in q_c (through (2.15)) due to the substitution $K \rightarrow K'$ should be obvious. Less obvious is the change in ΔT_{max}. As ΔT_{max} is expressed as a function of the relative figure-of-merit Z, it is easiest to regard the change to Z first:

$$Z_{K'} = \frac{(\alpha_p - \alpha_n)^2}{K'R} = \frac{K}{K'} \frac{(\alpha_p - \alpha_n)^2}{KR} = \frac{K}{K'} Z(K_m) \tag{2.28}$$

In this equation, $Z(K_m)$ has to be used instead of Z, as the product $K{\cdot}R$ slightly changes due to the influence of K_m. Following the substitution $K \rightarrow K'$, ΔT_{max} changes to

$$\Delta T_{max,K'} = \frac{1}{2} Z_{K'} T_c^2 = \frac{K}{K'} \left(\frac{1}{2} Z(K_m) T_c^2 \right) \tag{2.29}$$

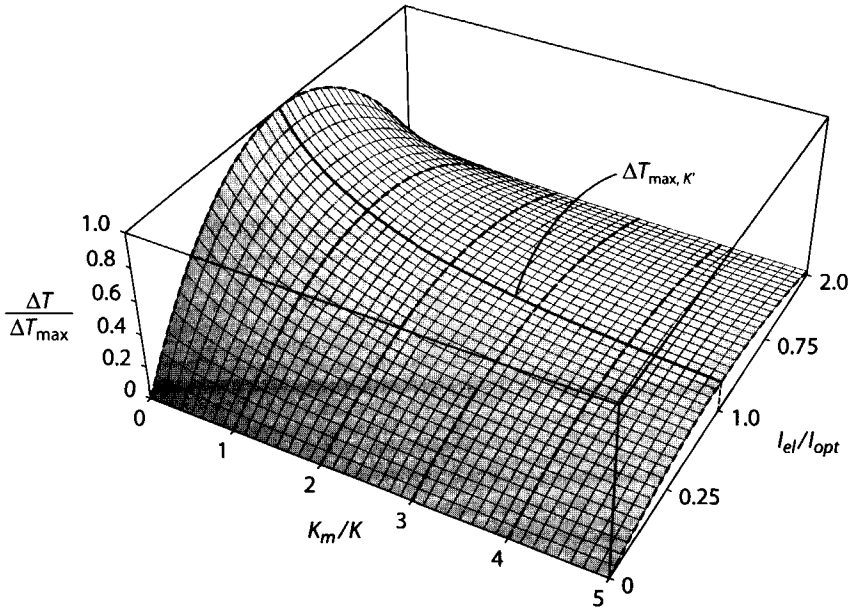


Figure 2.7 Typical plot of the relative decrease in temperature difference ΔT, as a function of the relative thermal conduction through the membrane, K_m, and the relative electrical current applied, I_{el}.

The reason that Z slightly changes is due to the change in the ratio W_p/W_n (for which ΔT_{max} is optimised.) In general, the width of the membrane $W_m = W_n + W_p$, as the membrane needs to support both thermoelements. In the generic case (when $L_n \neq L_p$ or $H_n \neq H_p$) the optimal ratio is

$$\left.\frac{W_p}{W_n}\right|_{K'} = \frac{H_n L_p}{H_p L_n}\sqrt{\frac{\rho_p(\lambda_n L_m + \lambda_m L_n H_m / H_n)}{\rho_n(\lambda_p L_m + \lambda_m L_n H_m / H_p)}} \tag{2.30}$$

Here, λ_m, L_m and H_m are the respective average thermal conductivity, length and total thickness of all non-thermoelectric layers in the membrane. For the integrated lateral TEC, $L_n = L_p = L_m = L$ and $H_n = H_p = H_m = H$, so Equation (2.30) reduces to

$$\left.\frac{W_p}{W_n}\right|_{K',lt} = \sqrt{\frac{\rho_p(\lambda_n + \lambda_m)}{\rho_n(\lambda_p + \lambda_m)}} \tag{2.31}$$

Still the optimal width ratio depends on material properties only, although the average thermal conductivity of the supporting membrane needs to be included.

Figure 2.7 plots the reduction in the temperature difference (T_a-T_c) relative to the maximum temperature difference ΔT_{max}, as a function of the thermal conduction

through the membrane K_m, and the electrical current applied, I_{el}. (K_m is displayed in proportion to the thermal conduction through the thermoelements, K.) To get an idea of the influence by K_m, when $K_m = K$, ΔT_{max} is already reduced by half. For clarity, Equation (2.29) has been indicated by the solid curve in Figure 2.7.

2.4.2 Electrical contact resistance (R_c)

A second issue with on-chip integrated TECs that is commonly ignored, is the electrical contact resistance of the metal-semiconductor interface, R_c. For a conventional TEC this is generally allowed, as the contact between the two conductors is significantly smaller than the electrical resistance of the thermoelements. Nevertheless, already in 1950 (when fabrication techniques were less sophisticated) Gehlhoff *et al.* were forced to investigate the influence of electrical contact resistance, as their devices performed less than expected due to this effect [2.9].

In contrast to a discrete TEC, the influence of electrical contact resistance can not be ignored with integrated TECs. This is primarily due to the size of the contact (as the area:side length ratio scales down linearly with the perimeter size.) For comparison, the conventional thermoelement has a contact area in the order of $(1\ \text{mm})^2$, while the integrated thermoelements has one in the order of $(10\ \mu\text{m})^2$, i.e., some 10^4 times smaller. The electrical contact resistance manifests itself at both the hot and cold junction of the device. However, as the hot junctions are considered to be in direct contact with ambient, the Joule heat generated in these junctions is absorbed immediately, and as such does not contribute to a change in ΔT_{max}. All of the Joule heat generated by the electrical contact resistance of cold junctions has to travel trough the entire length of the device to reach ambient. Thus (with similar reasoning as applied in Section 2.2.1,) not half but all of the Joule heat generated by R_c at the cold junctions contributes to a rise in T_c. Hence, the total Joule heat generated due to electrical resistance (and electrical contact resistance) becomes

$$q_{J'} = q_J + q_{R_c} = \frac{1}{2} I_{el}^2 R + I_{el}^2 R_c = \frac{1}{2} I_{el}^2 R' \tag{2.32}$$

with

$$R' = R + 2R_c = \frac{\rho_n L_n}{A_n} + \frac{\rho_p L_p}{A_p} + 2(R_{c,n} + R_{c,p}) \tag{2.33}$$

To analyse the influence of R_c only on device performance, K_m is disregarded in this subsection. In similarity to the substitution $K \rightarrow K'$, the substitution $R \rightarrow R'$ is applied to model the behaviour of the contact resistance on the figure-of-merit:

$$R' = R + 2R_c = \frac{\rho_n L_n}{A_n} + \frac{\rho_p L_p}{A_p} + 2(R_{c,n} + R_{c,p}) \tag{2.34}$$

Here, $Z(R_c)$ refers to the figure-of-merit optimised for electrical contact resistance, i.e.,

$$Z_{R'} = \frac{(\alpha_p - \alpha_n)^2}{KR'} = \frac{R}{R'}\frac{(\alpha_p - \alpha_n)^2}{KR} = \frac{R}{R'}Z(R_c) \tag{2.35}$$

In the general case, R_c changes the optimal geometrical ratio W_p/W_n to

$$\left.\frac{W_p}{W_n}\right|_{R'} = \frac{H_n L_p}{H_p L_n}\sqrt{\frac{\lambda_n \rho_p}{\lambda_p \rho_n} + \frac{2R_c A_p}{\lambda_p \rho_n L_p}} \tag{2.36}$$

while this reduces to

$$\left.\frac{W_p}{W_n}\right|_{R',lt} = \sqrt{\frac{\lambda_n \rho_p + 2R_c(A_p/L)}{\lambda_p \rho_n}} \tag{2.37}$$

for the lateral on-chip integrated TEC. This results differs significantly from either of Equations (2.22) or (2.31), as the optimal width ratio no longer is a function of material properties only, but of the device dimensions as well. As such, the maximum temperature difference, considering the electrical contact resistance, equals

$$\Delta T_{\max,K'} = \frac{1}{2} Z_{R'} T_c^2 = \frac{R}{R'}\left(\frac{1}{2} Z(R_c)\, T_c^2\right) \tag{2.38}$$

Figure 2.8 plots the relative decrease in the temperature difference as a function of the relative electrical contact resistance R_c/R and the relative electrical current, I_{el}/I_{opt}. For convenience only the values where $\Delta T > 0$ K are displayed. At the point where $R_c = R$, the maximum temperature difference is almost reduced by two thirds. Thus, relatively speaking, the electrical contact resistance has a much more negative impact on device performance than the thermal conductivity through the dielectric layers. Besides the difference in the severity of performance degradation, Figure 2.8 reveals another major difference between the behaviour of R_c and K_m. Along the curve $\Delta T_{\max,K'}$ in Figure 2.7 the value of I_{el} hardly changes with K_m. In contrast, with increasing values of R_c, the electrical current needs to be decreased significantly, to maintain on the curve $\Delta T_{\max,R'}$. The observation will be discussed further in Section 2.4.3. Finally, the solid curve in Figure 2.8 is the graphical interpretation of Equation (2.38), which shows $\Delta T_{\max}$ as a function of R_c.

2.4.3 Quantitative comparison of $\Delta T_{\max,K'}$ and $\Delta T_{\max,R'}$

Figure 2.9 provides a direct comparison of the influence by both K_m and R_c on ΔT, with and without the optimization of the width ratios. The values K_{ideal} and R_{ideal} are the values found for K and R from the ideal model, using the non-optimised width ratio as defined in Equations (2.22). The functions $K' = K_{ideal}+K_m$ and $R' = R_{ideal}+R_c$

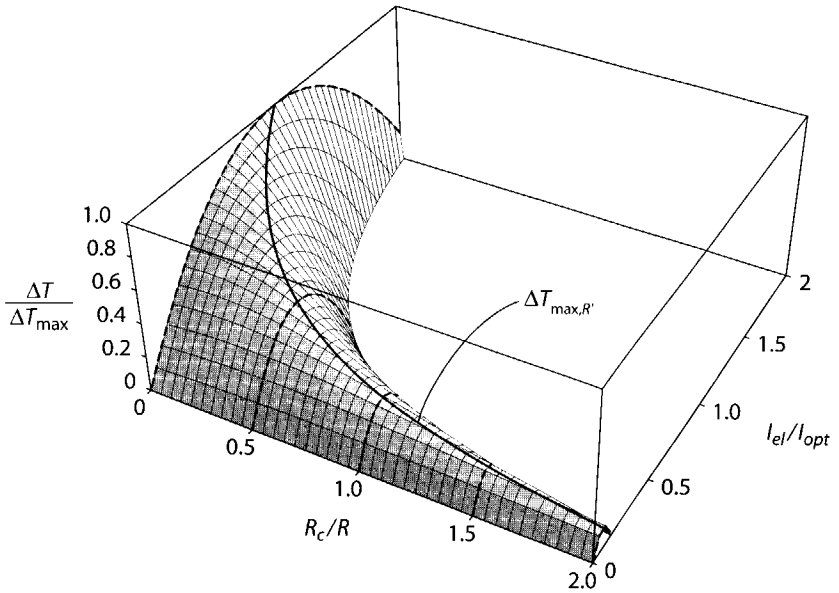


Figure 2.8 Typical plot of the relative decrease in temperature difference, ΔT, obtainable, as a function of the relative electrical contact resistance, R_c, and the electrical current applied, I_{el}. Only the portion of the surface, where $\Delta T > 0$ K is shown.

are based on the same non-optimised width ratio. The dashed curves indicate the decrease in temperature difference as a function of the respective non-ideality. The functions plotted are $K_{ideal}/K' = (1+(K_m/K_{ideal}))^{-1}$ and $R_{ideal}/R' = (1+2(R_c/R_{ideal}))^{-1}$. The latter function is exactly what Gehlhoff *et al.* had derived in 1950 [2.9]. Even though the result from the two dashed curves is entirely correct, they modelled the influence of K_m and R_c as purely parasitic rather than as an integral part of the device. Thus, K' and R' serve as correction factors only. This always leads to a suboptimal design.

In contrast to the above, the curves $\Delta T_{max,K'}$ and $\Delta T_{max,R'}$ as defined respectively by Equations (2.29) and (2.38). For these equations, the width ratios have been optimised, respectively through Equations (2.30) and (2.36). This way, K_m and R_c have become an integral part of the design: Determining the values of λ_m, $R_{c,n}$ and $R_{c,p}$ beforehand, these can be anticipated for in the design and, as such, be partially compensated. Both optimised functions are indicated in Figure 2.9 by the solid curves.

The underlying explanation of the improvement by the integral approach is subtle, yet simple. Recall Eqns. (2.31) and (2.37). In both cases the optimal ratio of W_p/W_n varies with the non-ideality under investigation. Due to the required change in widths to optimise this ratio, the values for K and R, and consequently for Z, change accordingly. The terms K_{ideal}/K' and R_{ideal}/R' simply do not reflect this poten-

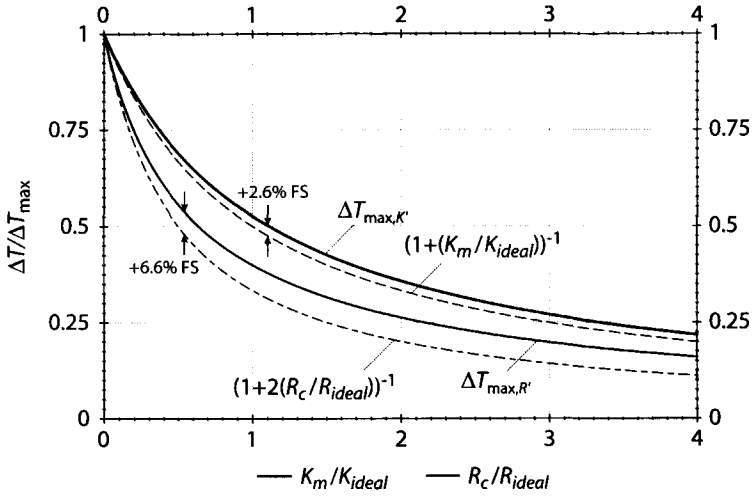


Figure 2.9 Plots of the relative maximum temperature difference obtainable, $\Delta T/\Delta T_{max}$, in relation to the relative electrical contact resistance and relative thermal conduction through the membrane.

tial optimisation, unlike $\Delta T_{max,K'}$ and $\Delta T_{max,R'}$. This was the reason why reference is made to the optimised figure-of-merit, respectively $Z(K_m)$ and $Z(R_c)$, rather than the conventional figure-of-merit Z. The individual optimisation of R and K is simple. Fixing the dimensions of the n-type leg, then R should be equal to $R_n+R_p/(W_p/W_n)|_{R'}$ and K should be equal to $K_n+K_p\cdot(W_p/W_n)|_{K'}$. This way, ΔT may be improved by up to 2.6% when considering K_m during the design phase, while an improvement of up to 6.6% can be obtained when considering R_c in the design phase.

It was already noted in Section 2.4.2 that the optimal driving current needs to decreased as R_c increases, while this hardly is the case for a change in K_m. Figure 2.10 plots these optimised electrical currents, $I_{opt}(K_m)$ and $I_{opt}(R_c)$, relative to the optimal current of the ideal device, $I_{opt,ideal}$. The figure clearly endorses the previously made observation: Going from $K_m = 0$ to $K_m = K$ only increases I_{opt} by just over 3%, while going from $R_c = 0$ to $R_c = R$ decreases I_{opt} by over two thirds! The electrical contact resistance can be made small with respect to the electric resistance of the other components. Nevertheless, R_c will generally show a large spread, as it is a process parameter rather than a material parameter. As a result, I_{opt} can change significantly from device to device, which complicates the design.

2.4.4 A new derivation of the optimal current I_{opt}

The major problem with the conventional expression for I_{opt} is that it can't be solved directly. Looking at this expression—see Equation (2.16)—it is noticed immediately that $I_{opt} \propto T_c$. Unfortunately, this requires T_c (= T_a–ΔT_{max}) to be determined first. In-

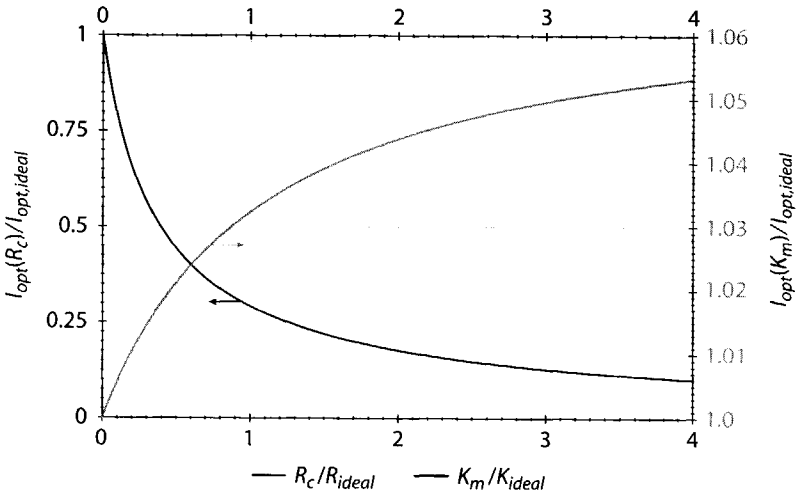


Figure 2.10 Plot of the optimised electrical currents, $I_{opt}(R_c)$ and $I_{opt}(K_m)$, respectively as a function of the electrical contact resistance (thin black curve) and thermal conduction through the membrane (thick grey curve).

stead, it is preferable to express I_{opt} as a function of T_a only (rather than T_c or ΔT_{max}.) This requires I_{opt} to be derived in a way completely different from the derivation found in the standard works on thermoelectrics. At this point the dashed curve in the floor of Figure 2.4 comes into play. This curve plots the solutions to $q_c = 0$, which has the solution

$$T_c = \frac{-\frac{1}{2} I_{el}^2 R - KT_a}{K + I_{el}(\alpha_p - \alpha_n)} \tag{2.39}$$

Next, from this function $dT_c/dI_{el} = 0$ is calculated, which is the point of optimal electrical current:

$$i_{opt} = \frac{KR\sqrt{KR(KR + 2(\alpha_n - \alpha_p)^2 T_a)}}{R(\alpha_n - \alpha_p)} \tag{2.40}$$

Even though this equation is more complex than Equation (2.16), it has a significant advantage in that it 'only' depends on the device geometry, material properties and ambient temperature and, hence, can be solved directly.

2.4.5 Combined influence of K_m and R_c on ΔT_{max}

In a practical device, both K_m and R_c are present. Therefore, the final step in Section 2.4 is to combine the influence of both non-idealities in a single model. In this situation, both substitutions $K \rightarrow K'$ and $R \rightarrow R'$ are applied, in which case the optimised figure-of-merit becomes

$$Z_{K'R'} = \frac{(\alpha_p - \alpha_n)^2}{K'R'} = \frac{KR}{KR + (2KR_c + K_m R + 2K_m R_c)} Z(K_m, R_c) \tag{2.41}$$

where $Z(K_m,R_c)$ is the figure-of-merit optimised for both K_m and R_c and as such again is unequal to the ideal figure-of-merit Z. Using $Z(K_m,R_c)$, the maximum temperature difference becomes

$$\Delta T_{max,K'R'} = \frac{1}{2} Z_{K'R'} T_c^2 = \frac{KR}{K'R'} \left(\frac{1}{2} Z(K_m, R_c) T_c^2 \right) \tag{2.42}$$

which is reached at the optimum width ratio of

$$\left. \frac{W_p}{W_n} \right|_{K'R'} = \frac{H_n L_p}{H_p L_n} \sqrt{\frac{\left(\lambda_n L_m + \lambda_m L_n (H_m / H_n)\right)\left(\rho_p + 2R_c (A_p / L_p)\right)}{\rho_n \left(\lambda_p L_m + \lambda_m L_p (H_m / H_p)\right)}} \tag{2.43}$$

Again, in the case of the integrated lateral TEC this equation can be simplified to:

$$\left. \frac{W_p}{W_n} \right|_{K'R',lt} = \sqrt{\frac{\left(\lambda_n + \lambda_m\right)\left(\rho_p + 2R_c (A_p / L)\right)}{(\lambda_p + \lambda_m)\rho_n}} \tag{2.44}$$

As will be discussed in the following chapters, a well-designed integrated polySiGe TEC (consisting of a single couple of thermoelements) has an electrical contact resistance in the order of a few percent at most, while the thermal conduction through the dielectric layers will be in the range of 10%–30%. As a result, the combined temperature decrease can be as much as 25% of ΔT_{max}.

2.5 TEMPERATURE PROFILE ACROSS THE PELTIER DEVICE

In the analysis presented in the previous section, only the electrical contact resistance and thermal conduction through the dielectic membrane were taken into account. These two are not the only non-idealities in the membrane region of the thermoelectric cooler. Thermal energy with its surroundings is also exchanged through radiation and conduction/convection through the surrounding air. The radiant heat flux from a surface with area A_s (at temperature T_s) can be written as

$$q_{rad} = A_s \varepsilon \sigma_B \left((T_s + T_a)(T_s^2 + T_a^2) \right) (T_s - T_a) = A_s h_{rad} (T_s - T_a) \tag{2.45}$$

with σ_B the Stefan-Boltzmann constant ($5.67 \cdot 10^{-8}$ $Wm^{-2}K^{-4}$), ε the emissivity of the surface and h_{rad} [$Wm^{-1}K^{-1}$] the radiation heat transfer coefficient. Use of the above expression will yield a very non-linear equation to be solved, so a linearised approximation is used for the radiant heat transfer coefficient:

$$h_{rad} \cong 4 \varepsilon \sigma_B T_a^3 \tag{2.46}$$

It has to be kept in mind that this is a good approximation only if $T_s \approx T_a$. The convective heat exchange from that surface is defined as

$$q_{conv} = h_{conv} A_s (T_s - T_\infty) \tag{2.47}$$

where T_∞ is the temperature far away from the surface. In the analysis that follows the (safe) assumption is made that that $T_\infty = T_a$.

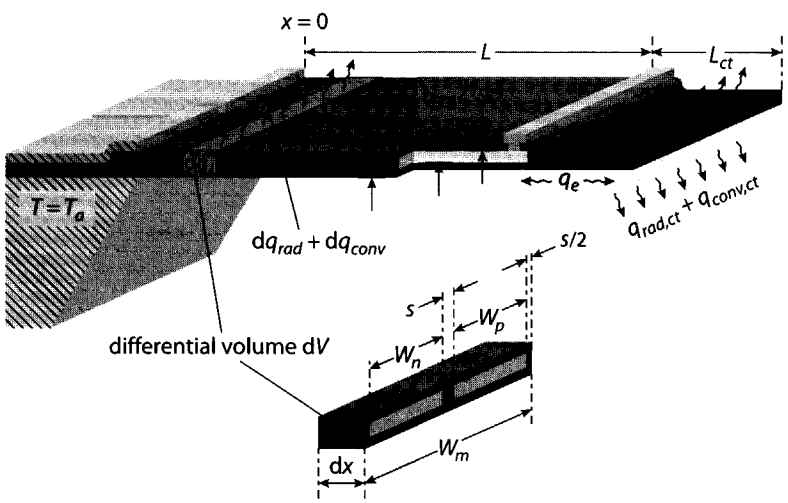


Figure 2.11 Model used to derive the temperature profile across the thin-film TEC.

Both the radiant and convective heat exchange can be evaluated from the model shown in Figure 2.11. The TEC itself stretches from the hot junction at $x = 0$ (at temperature T_a) to the cold junctions at $x = L$ (at T_c). The inset reveals that the thermoelements (which have a thickness H_{TE}) are fully encapsulated; On the bottom by a Si_3N_4 dielectric layer of thickness H_{SiN}, on the top by a SiO_2 dielectric layer of thickness H_{SiO}. The spacing between the thermoelements (of width s) is filled with SiO_2 and has the same thickness as the thermoelements. The nitride and oxide films

extend over the entire width W_m of the cantilever, which is equal to the widths of both thermoelements plus twice the spacing between the thermoelements, i.e.,

$$W_m = W_n + W_p + 2s \tag{2.48}$$

The difficulty of modelling the radiant and convective heat exchange between the on-chip integrated TEC and its ambient is that the temperature across the surface of the TEC is not constant. Therefore, the heat exchange depends on the position along the surface. Consider the differential volume dV indicated in Figure 2.11 (which has a differenential length $dx = \Delta x \to 0$.) Assuming no thermal energy is stored in this volume, the heat transported across the volume in the x-direction plus the heat generated within the volume have to balance the heat exchanged with the surrounding by means of convection and conduction, i.e.,

$$\frac{dq_x}{dx} + \frac{dq_J}{dx} = \frac{dq_{rad}}{dx} + \frac{dq_{conv}}{dx} \tag{2.49}$$

The conduction term is

$$\begin{aligned} \frac{dq_x}{dx} &= \left[\begin{array}{l} W_m(H_{SiO}\lambda_{SiO} + H_{SiN}\lambda_{SiN}) \\ \qquad + H_{TE}(W_n\lambda_n + W_p\lambda_p + 2s\lambda_{SiN}) \end{array} \right] \frac{d^2\left(T(x) - T_a\right)}{dx^2} \\ &= A\frac{d^2\left(T(x) - T_a\right)}{dx^2} \end{aligned} \tag{2.50}$$

As both the radiant and convective heat exchange are functions of $(T(x)\text{-}T_a)$, these can be combined into one expression, i.e.,

$$\begin{aligned} \frac{dq_{rad}}{dx} + \frac{dq_{conv}}{dx} &= \left[2W_m(4\varepsilon\sigma_B T_a^3 + h_{conv})\right]\left(T(x) - T_a\right) \\ &= B\left(T(x) - T_a\right) \end{aligned} \tag{2.51}$$

Finally, the Joule heating term is equal to

$$\frac{dq_J}{dx} = \frac{I_{el}^2}{H_{TE}}\left(\frac{\rho_n}{W_n} + \frac{\rho_p}{W_p}\right) = C \tag{2.52}$$

By substituting Equations (2.50)–(2.52) into Equation (2.49), the following differential equation is obtained:

$$A\frac{d^2\left(T(x) - T_a\right)}{dx^2} - B\left(T(x) - T_a\right) + C = 0 \tag{2.53}$$

Using the substitution $\theta(x) \equiv T(x)\text{-}T_a$, this equation simplifies to

$$A\frac{d^2\theta}{dx^2} - B\,\theta + C = 0 \tag{2.54}$$

However, before this differential equation can be solved, the boudary conditions need to be established. The first one is that the hot junction temperature is equal to ambient temperature, i.e., $T(0) = T_a$ at $x = 0$:

$$\theta(0) = 0 \tag{2.55}$$

The second boundary condition is determined by the region at the tip of the cantilever, from $x = L$ and beyond. (This would be the region to be thermall stabilised.) Across the entire surface of the cantilever tip, A_{ct} (= $W_m \cdot L_{ct}$), heat is exhanged with the surroundings by means of radiation and conduction. For simplicity, it is assumed that the entire surface area of the cantilever tip is at a uniform temperature, equal to $T(L) = T_c$. Thus, the radiant / convective heat exchange away from A_{ct} equals

$$q_{rad,ct} + q_{conv,ct} = 2A_{ct}\left(h_{rad} + h_{conv}\right)\theta\big|_{x=L} \tag{2.56}$$

Two more power sources manifest themselves at the position of the cold junctions ($x = L$; shown best in Figure 2.6); The Peltier cooling power q_P removing heat from these junctions, while the electrical contact resistance adds Joule heata at a rate $q_{J,Rc}$ (see Equation (2.32)). The last power source is the heat load from power-dissipating electronic devices located on top of the membrane or inside the epi-layer mass. This heat load is referred to as q_e and, for simplicity, is considered to be distributed uniformly across the volume at the cantilever tip. To correctly model the Peltier effect, the absolute temperature dependency of this effect should be taken into consideration. This can be done by substituting $T(L) \rightarrow T_a + \theta$:

$$q_P = (\alpha_p - \alpha_n) I_{el} T(L) = (\alpha_p - \alpha_n) I_{el} T_a + (\alpha_p - \alpha_n) I_{el}\theta\big|_{x=L} \tag{2.57}$$

The total net power generated in the cantilever tip, q_{ct} equals $q_{rad,ct} + q_{conv,ct} + q_P + q_{Rc} + q_e$, where q_e is an external heat load at the tip of the cantilever:

$$\begin{aligned} q_{ct} &= \left(2A_{ct}\left(h_{rad} + h_{conv}\right) + (\alpha_p - \alpha_n) I_{el}\right)\theta\big|_{x=L} + \left((\alpha_p - \alpha_n) I_{el} T_a - I_{el}^2 R_c - q_e\right) \\ &= D\,\theta\big|_{x=L} + E \end{aligned} \tag{2.58}$$

The second and final boundary condition states that the the power deficit / surplus of q_{ct} can only be transported to ambient through the TEC. Therefore, at $x = L$, the heat flux through the cross-section of the TEC has to equal q_{ct}:

$$A\frac{d\theta}{dx}\bigg|_{x=L} = -\left(D\,\theta\big|_{x=L} + E\right) \tag{2.59}$$

The generic solution of Equation (2.54), using the boundary conditions of Equations (2.55) and (2.59) is

$$\theta(x) = \frac{\left[\begin{array}{r} 2C\sinh\left(\sqrt{\frac{B}{A}}\frac{x}{2}\right)\left(D\cosh\left(\sqrt{\frac{B}{A}}\frac{(x-2L)}{2}\right)-\sqrt{AB}\sinh\left(\sqrt{\frac{B}{A}}\frac{(x-2L)}{2}\right)\right) \\ -(CD+BE)\sinh\left(\sqrt{\frac{B}{A}}L\right) \end{array}\right]}{\sqrt{A}B^{3/2}\cosh\left(\sqrt{\frac{B}{A}}L\right)+BD\sinh\left(\sqrt{\frac{B}{A}}L\right)} \tag{2.60}$$

The difference between the cold junction temperature and ambient temperature, $T_c - T_a = T(L) - T_a = \theta(L)$ can be directly found by substituting x with L in Equation (2.60):

$$\theta(L) = \frac{2\sqrt{A}C\sinh\left(\sqrt{\frac{B}{A}}\frac{L}{2}\right)^2 - \sqrt{B}E\sinh\left(\sqrt{\frac{B}{A}}L\right)}{\sqrt{A}B\cosh\left(\sqrt{\frac{B}{A}}L\right)+\sqrt{B}D\cosh\left(\sqrt{\frac{B}{A}}L\right)} \tag{2.61}$$

To plot a temperature profile of the on-chip integrated TEC, incorporating the radiant and convective heat exchange, it is impossible to avoid specifying material parameters as well as device geometries. The values used are listed in Table 2.1. Only three parameters are varied, the heat load q_e, the emissivity ε and spacing s. Furthermore, the TEC under investigation is assumed to work in vacuum ($h_{conv} = 0$), L_{ct} is set to zero. For the spacing s, two values are considered, $s = 0$ µm (which approaches the ideal case) and $s = 3$ µm (which is a practical value.)

When $\varepsilon \to 0$, Equation (2.60) no longer produces a meaningful result. Therefore, to plot the temperature profile of the TEC without radiant and convective heat exchange, but with R_c and K_m (so $s = 0$ µm, $q_e = 0$ µW and $\varepsilon = 0$), Equation (2.54) has to be solved again, this time without the term $B\theta$. The resulting expression for the temperature profile becomes

$$\theta(x)\big|_{B=0} = -\frac{E - CL - (CDL^2)/(2A)}{A + DL}x - \frac{C}{2A}x^2 \tag{2.62}$$

and is indicated by the thick solid line in Figure 2.12. The value for $\theta(L)$ obtained from this modified solution is identical to the value for $(T_c - T_a)$ as obtained from Equation (2.42).

In the next step there is still zero spacing between the thermoelements (so the thermoelements are not really encapsulated as in Figure 2.11.) and there is still no heat load q_e at the cantilever tip. For these conditions, the temperature profile result-

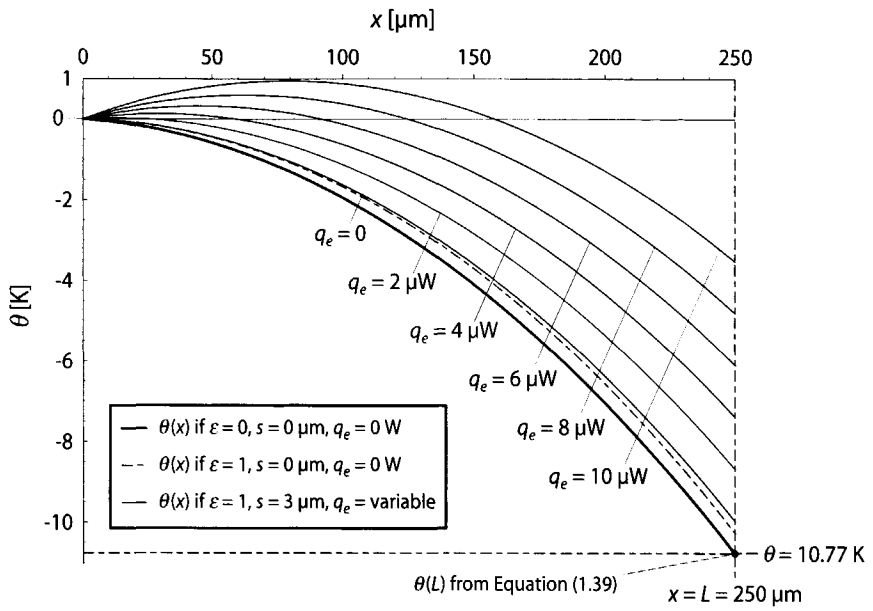


Figure 2.12 Various plots of the temperature profile of an on-chip integrated TEC, using the material properties listed in Table 1.1.

Table 2.1 Parameters used for calculating the temperature profile of a TEC.

Material properties			
Material	α [μVK^{-1}]	ρ [μΩm]	λ [Wm^{-1}K^{-1}]
n-polySiGe	−136	10.1	4.45
p-polySiGe	144	13.2	4.80
SiO_2	--	--	1.2
Si_3N_4	--	--	1.5
Device geometries			
L	250 μm	H_{TE}	600 nm
L_{ct}	0 μm	H_{SiO}	300 nm
W_n	41.4 μm	H_{SiN}	300 nm
W_p	50.0 μm		
Other parameters			
R_c	10 Ω	h_{conv}	0 Wm^{-2}K^{-1}

ing from Equation (2.60) is indicated in Figure 2.12 by the dashed curve. This is the exact same result as obtained from the analysis performed by Völklein *et al.* [2.10], when their differential equation is adapted for the geometries of Figure 2.11. The disadvantage of the analysis performed by Völklein *et al.* is that, rather than an explicit solution like Equation (2.60), an implicit solution is derived for $\theta(x)$. That is, in their case $\theta(x) \propto \theta(L)$, so $\theta(L)$ needs to be solved first.

The thin solid curves in Figure 2.12 include the effect of the spacing, using $s = 3$ μm. In case q_e is zero, a slight decrease of 0.3 K is observed in $\theta(x)$, compared to the case where $s = 0$ μm. The remaining set of thin solid curves indicates the influence of a heat load at the tip of the cantilever, for different values of q_e. Note that for the material properties and geometries specified in Table 2.1, a heat load of only 8 μW is already sufficient to cut $\theta(L)$ in half.

2.5.1 Influence of the Thomson effect and temperature-invariant material properties

In all of the models in this thesis, the material properties have been kept constant and the Thomson effect has been ignored, which decreases the accuracy of the models. In this subsection, a simple qualitative estimation of the discrepancy between the models in this thesis and the real performance is provided.

The influence of the Thomson effect

Besides the Peltier effect, thermal energy is also absorbed by the Thomson effect. Over an entire element the rate of energy absorption amounts to [2.3]

$$q_T = I_{el} T_c \int T d\alpha \cong I_{el} T_c (\alpha_h - \alpha_c) \tag{2.63}$$

where α_c and α_h are the respective Seebeck coefficients at the cold and hot junction temperatures. The combination of the cooling rates by the Peltier and Thomson effects now yields

$$q_{PT} = I_{el} T_c \alpha_c + \frac{1}{2} I_{el} T_c (\alpha_h - \alpha_c) = I_{el} T_c \frac{(\alpha_h + \alpha_c)}{2} \tag{2.64}$$

The material values applied in the models throughout this thesis are the values obtained at room temperature (typically the hot junction temperature.) Instead, Equation (2.64) shows it would be better to use the average thermal value rather than the hot junction temperature.

The influence of constant-parameter modelling

In Chapter 14 of [2.2], the influence of temperature dependency of material properties on device performance is analysed, among others with respect to the maximum temperature difference ΔT_{max}. This is gone about by comparing the results from different frequently encountered temperature-independent approximations to an FEA model that does include the thermal dependency of the materials (and which is esti-

mated to be accurate to within 1% of the actual TEC performance). Two of the temperature-independent models have already been introduced in this thesis. In the so-called $q(T_h)$ model the values are determined similarly as in this thesis, i.e., from the values at the hot junction temperature T_h. With the $q(T_{avg})$ method the material properties are determined as proposed in the previous subsection, i.e., from the average temperature, T_{avg}.

The most interesting remark from this study is that all constant-material value approaches underestimate the magnitude of ΔT_{max}, even the $q(T_h)$ method. A striking detail is that the $Q(T_h)$ method appears to be the most accurate model in terms of the maximum temperature difference. From the above observations I conclude that—if all relevant non-idealities are considered—the models developed in this thesis can be regarded as worst-case models that underestimate the actual device performance.

2.6 THERMAL RESISTANCE OF THE SUBSTRATE

In the few references of lateral on-chip integrated thermoelectric coolers that exist, the thermal resistance of the substrate is persistently ignored. Nevertheless, the substrate is the TEC's only available heat sink. To achieve a high thermal sensitivity the on-chip TEC must be placed in a vacuum. Only the base of the substrate surface is remains in intimate contact with the ambient (i.e., the chip carrier). As a result, the heat released at the hot junctions of the TEC can only be transferred off-chip by traversing the substrate (to the carrier).

If the thermal resistance of the substrate between it's base and the hot junctions is considerable, the hot junction temperature T_h will rise in relationship to the ambient (carrier) temperature T_a. As the ratio $\Delta T = (T_h - T_c)$ remains nearly unaffected by this rise, the cold junction temperature T_c will vary equally with T_h. Therefore, without knowing the thermal resistance of the substrate it is not possible to talk about an absolute cold junction temperature.

The thermal resistance of the substrate not only depends on the thermal conductivity of the substrate, but also on the volume available for heat removal. Therefore, the choice of etch technique (to thermally isolate the thin-film TEC) influences the thermal resistance of the substrate. In total, four etch techniques are evaluated with respect to their relative thermal resistance. These are, in order of increasing volume:

- Back side etching. This techniques removes the substrate through an opening created in the backside of the wafer, resulting is the typical substrate profile as indicated in the upper three schematic drawings in Figure 2.13a. With thist technique the characteristic angle φ of the wafer surface with the wall of the substrate is 54.74°. (See Chapter 4 for more information on etching.)
- Deep reactive ion etching (DRIE). Even though this technique will never be really used for bulk micromaching, it is included for modelling purposes.

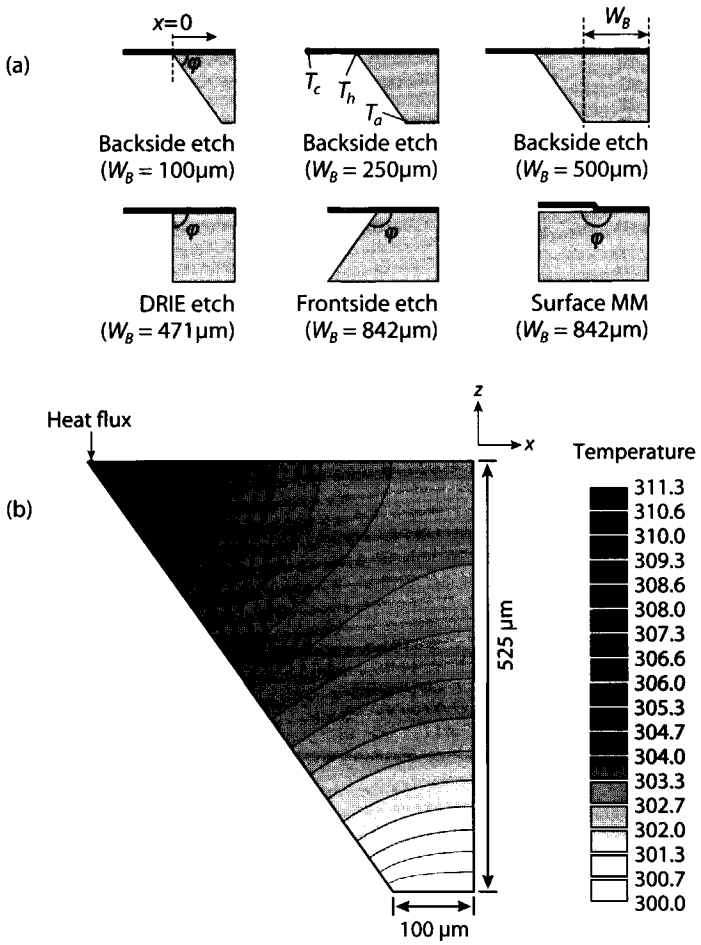


Figure 2.13 (a) All six two-dimensional substrate geometries investigated for thermal resistance. (b) Typical result of a single FEA simulation. As the heat flux as well as the temperature at the 'hot' junction are known, the thermal resistance can be directly calculated.

Again the substrate is removed through an opening in the backside. With DRIE the etching mechanism is such, that the substrate would be etched in the direction perpendicular to the wafer surface only, hence $\varphi = 90°$. The resulting structure is shown in the lower left schematic of Figure 2.13a.

- Front side etching. Instead of etching through an opening in the back side of the wafer, the wafer is opened from the front side. The etchant can be the same as for the back side etch. With front side etching, $\varphi = 125.26°$. The resulting shape is shown in the lower middle drawing of Figure 2.13a.

- Surface micromachining. This is not really an substrate etching technique, as the substrate is not etched. Instead, some sacrificial layers in between the substrate and the thin-film TEC are removed. This way, a thin-film structure can be constructed that merely floats a few micrometres above the substrate surface. In this situation $\varphi = 180°$. The schematic shape is shown in the lower right drawing Figure 2.13a.

2.6.1 Parameters analysed

In the construction of the actual devices, back side etching has been used to thermally isolate the thin-film structures. As can be seen directly from Figure 2.13a, this etch technique creates the longest and narrowest path from T_h to T_a (and, thus, will have the highest thermal resistance.) Therefore, it is investigated to what extent an increase in the base width W_B of the substrate helps reduce the thermal resistance. Together with these three variations for W_B (i.e., W_B = 100, 250 and 500 μm,) six other two-dimensional substrate configurations are analysed using finite element analysis (FEA); In this case ANSYS® was used.

The two-dimensional analysis is valid only if the heat flow in the y-direction (perpendicular to the cross-section of the substrate) is zero. This configuration resembles the situation in the middle of a (near-)infinite row of thermoelements in the y-direction. As heat can only flow in the (x,z)-plane, this automatically becomes the worst-case scenario.

In constrast to the two-dimensional case, at the corners created through a bulk etch the substrate volume available for heat sinking increases significantly as a large amount of heat can now flow in the y-direction as well. Three three-dimensional configurations have been evaluated, which are are shown in Figure 2.14a (the back side etch, DRIE etch and front side etch.) Moreover, for the back side etch the same three variations in W_B as in the two-dimensional analysis are evaluated (i.e., W_B = 100, 250 and 500 μm.)

The third issue investigated is the influence of a dielectric layer on top of the silicon substrate. As can be seen clearly from the schematic of Figure 2.6, the thermal energy released at the hot junction has to travel through the Si_3N_4 dielectric layer before it enters into the substrate. As the thermal conductivity of a dielectric is generally much lower than that of the silicon substrate, the contribution of the dielectric layer to the thermal resistance may be considerable. (In the FEA analyses performed, a thermal conductivity of 1.5 $Wm^{-1}K^{-1}$ was used for the Si_3N_4 dielectric, while monocrystalline silicon has a thermal conductivity of 145 $Wm^{-1}K^{-1}$.)

In total 13 configurations are investigated, eight two-dimensional and five three-dimensional. The typical result of a two- and three-dimensional simulations are respectively shown in Figures 1.13b and 1.14b. To draw a meaningful conclusion on these 2-D and 3-D simulation results, the thermal resistance encountered at various locations along the top surface is determined. To do so, in the 3-D geometry, first an 8 × 25 μm rectangle is defined in the top surface of the substrate, on which a certain

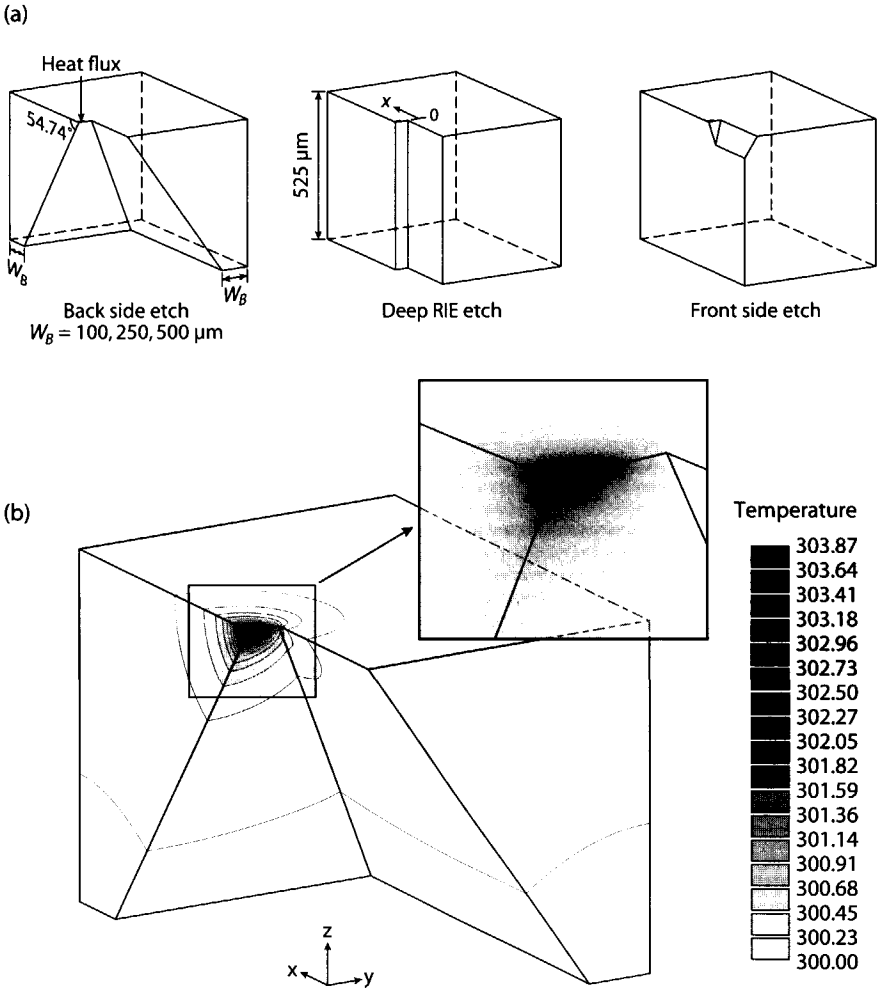


Figure 2.14 (a) The five three-dimensional substrate geometries investigated for thermal resistance. (b) Typical three-dimensional result of a single FEA simulation.

heat flux is placed. (This area closely resembles the total hot junction area of the average thermoelement.) Furthermore, the bottom surface is fixed at T_a, while all other surfaces are adiabatic. Thus, thermal energy entering the substrate through the heat flux at the top surface can only leave the substrate through the bottom surface. For the 2-D shape similar conditions apply, although the heat flux is placed on an 8 µm line instead, while the bottom line remains fixed at T_a.

With each simulation, the thermal resistance of one specific position along the wafer surface is determined. To see how the thermal resistance varies across the surface, the position of the heat flux area is moved in the x-direction with each simula-

tion (while keeping the magnitude of this heat flux constant.) The thermal resistance K^{-1} [Km^2W^{-1}] encountered at that point is found from the relationship $K^{-1} = \Delta T/q''$, where q'' [Wm^{-2}] is the heat flux and ΔT is the temperature between the hot junction and the substrate base.

2.6.2 2-D and 3-D simulation results

Figure 2.15 shows the result of the FEA simulations. One curve results for each of the 13 configurations described above. Each dot (•) or cross (×) represents a single simulation. As the heat flux is released over an 8 μm area, the value indicated on the x-axis refers to the middle of this region. (That is, the heat is released at the x value ±4 μm.)

If well fabricated, not the entire range for x plotted in Figure 2.15 is of interest. The range below $x = 20$ μm normally should not need to be considered as the hot junction is designed to be at least 20 μm away from the point where the membrane merges into the substrate. However, the thickness of the wafer can vary by ±25 μm. As a result, the origin of the x-axis can shift by ±25/tan(54.74°) ≈ ±18 μm. Moreover, it frequently occurs that the substrate is overetched a few microns, for example due to a poor wafer quality. As a result, a total overetch of up to 20 μm can occur. For this reason, 20 μm is the minimum distance the hot junction has to be away from $x = 0$ (as calculated for a wafer of exactly 525 μm thick, without any overetch occurring.) If the maximum overetch occurs the entire 'safety margin' is consumed, so the hot junction is located exactly at what then becomes $x = 0$. Under this circumstance, the range below $x = 20$ μm does need to be considered.

For the fabrication of the on-chip integrated TECs as presented in this thesis, back side etching has been applied, so the base width W_B is the only available 'degree-of-freedom'. Curves (1), (2) and (3) show the 2-D result for respective base widths of 100, 250 and 500 μm. For the hot junction position designed for, $x = 20$ μm, going from $W_B = 100$ μm to $W_B = 500$ μm, the thermal resistance drops from 1499 KW^{-1} to 1215 KW^{-1}, which is a reduction of 19%. Therefore, for a long array of thermoelements, increasing the base width can really help reduce the thermal resistance of the silicon substrate. Still when going from $W_B = 250$ μm to 500 μm, the improvement is limited, so a base width beyond 500 μm is not meaningful.

Comparing the back side etch to curves (4), (5) and (6), which respectively are the FEA results for the DRIE etch, front side etch and surface micromachining, the thermal resistance can be decreased significantly. The DRIE etch as well as surface micromachining have been included as hypothetical etching techniques only,. Only front side etching is a technique that can actually be applied. For $x = 20$ μm, the thermal resistance using front side etching can be reduced to 650 KW^{-1}, which is only 43% of that obtained from a back side etch with $W_B = 100$ μm. Hypothetically, the thermal resistance obtained through front side etching can be reduced further by using surface micromachining instead. (from curve (5) to (6), the thermal resistance

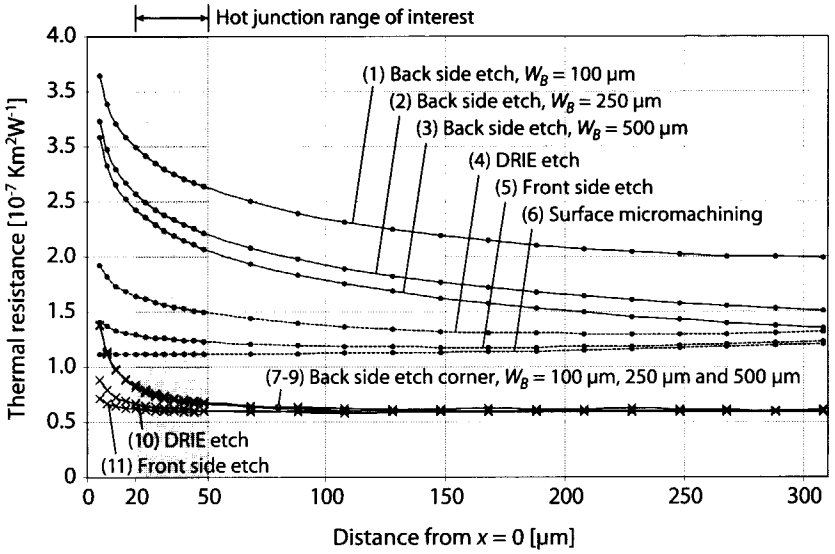


Figure 2.15 Calculated thermal resistance along the x-axis, for the eleven different substrate configurations shown in Figure 2.13 and Figure 2.14. The dots (•) represent the 2-D simulation results, while the crosses (×) represent the 3-D simulation results.

can be decreased by some 14% for x = 20 µm.) However, with surface micromachining, the membrane floats only a few micron above the substrate at best. This increases the thermal losses to ambient by means of convection and conduction through any residual air, which is very likely to outweigh the decrease in the thermal resistance of the substrate. Therefore, for long arrays, front side etching is the preferred bulk micromachining technique.

For the 3-D simulations, resembling the corner created by the specific bulk etch, the situation is quite different. Curves (7) to (9) are overlapping almost perfectly, which implies that for the back side etch an increase of the base width near the corner does not reduce the thermal resistance of the silicon substrate. For x = 20 µm, the 'corner configuration' for a back side etch with any base width has a thermal resistance of approximately 418 KW^{-1}, which is about 28% of the result from the 2-D case where W_B = 100 µm. Again the DRIE and front side etches, indicated by curves (10) and (11), can further decrease the thermal resistance. Going from curve (7) to (11), the thermal resistance decreases by some 20%. So, even when close to a corner in the substrate, the front side etch is remains preferred, although the improvement is much less distinct as in the 2-D case.

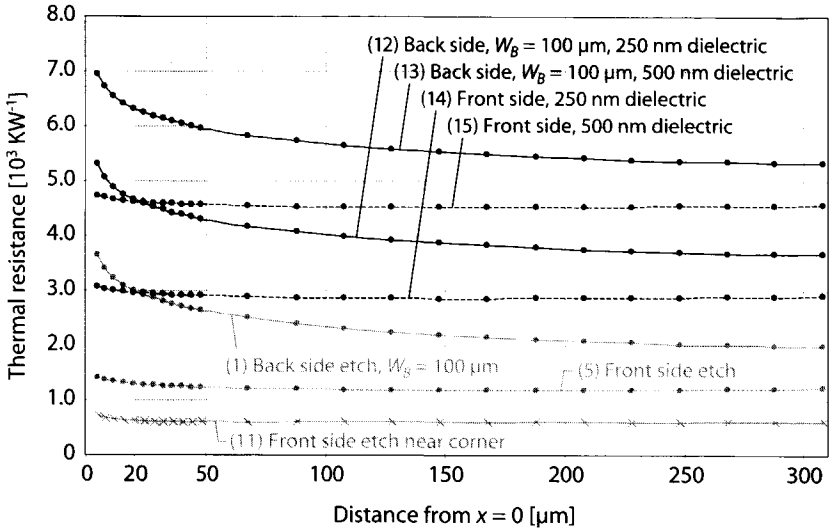


Figure 2.16 FEA simulation result that compares some of the silicon substrate-only configurations (n grey),to some that do include a thin-film dielectric layer.

2.6.3 Thin film dielectric influence

In Figure 2.16 curves (12) until (15) are the results from a substrate with a thin-film dielectric layer of either 250 nm or 500 nm on top (i.e., in between the hot junction and the substrate.) The thermal conductivity of the thin-film dielectric was set at 1.5 $Wm^{-1}K^{-1}$, some 10% of the 145 $Wm^{-1}K^{-1}$ for the silicon substrate.

For reference, three curves are listed in grey; The 2-D back side etch (with W_B = 100 µm) and the 2-D (long array) and 3-D (corner) front side etch. The numbers of these curves correspond to those listed in Figure 2.15.

What is most striking about the the thin-film dielectric is that its thickness translates into an offset in terms of thermal resistance. Staying with the location x = 20 µm, for the back side etch, from curve (1) to curves (12) and (13), the thermal resistance respectively increases by 56% and 111%. For the front side etch, going from curve (5) to curves (14) and (15), the effect is even worse, as the the offset increases the thermal resistance by 128% and even 256%, respectively. In conclusion it is not the substrate but rather the thin-film dielectric between the thermoelement and the substrate that causes the largest thermal resistance.

2.7 2-D FEA ANALYSIS OF COMPLETE STRUCTURE

So far, the simulations performed have only served to compare the thermal resistance of the various substrate geometries (both 2-D and 3-D) as well as the influence of a thin-film dielectric layer on the thermal resistance. This does not say anything

about the actual temperature profile in the the substrate or the actual hot junction temperature.

The modelling performed in Section 2.6 only provides the thermal resistance at one particular spot at a time, i.e., as 'seen' by the hot junction. In contrast, Joule heat generated within the TEC is distributed over the entire cantilever region. As a result, the heat flowing from the membrane into the substrate will be distributed over an area much larger than just the hot junction. The resulting temperature profile can only be verified by executing an FEA of the entire device, incorporating the thermoelements, the metal interconnect, all dielectric films and the substrate.

The main problem with a simulation of the entire structure lies in how to include the Peltier effect in the finite element analysis. Equation (2.7) shows the Peltier effect behaves as a heat source controlled by its own temperature, i.e., $q_P \propto T_c$. Hardly any FEA solvers can handle this sort of load. An extensive survey of all FEA solvers in 2000—ANSYS®, CFDRC®, Femlab®, Flotherm®, COSMOS/M®, Memcad®, Memscap®, FastFlo, PDEasy®—revealed that only the last two have the versatility to allow direct definition of a temperature controlled heat source. Of these two, FastFlo® 3.0, designed by the Numerical Algorithms Group, is designed for use with Femap® from SDRC. The latter is a powerful the pre- and postprocessor, i.e., for mesh generation as well as displaying of the results. At the time of performing the survey, this combination provided the most potential way to deal with the simulation at hand.

For people who would like to redo or expand the simulation described, it is worthwhile to reconsider the available solvers, as there are now two more good solvers available. PDEasy® has been revamped by PDE Solutions into FlexPDE® 3.0, and has become significantly more powerful and easier to use. Also, Harvard Thermal has developed Thermal Analysis System® (TAS) which at the time of writing must be considered the most powerful all-in-one thermal solver that is capable of handling temperature controlled heat sources.

Even though FastFlo® 3.0 is designed for use with Femap®, the process of running a single simulation remained elaborate. The most serious problems came from the Femap-to-FastFlo and FastFlo-to-Femap conversion filters supplied, as these simply did not work. Intead, dedicated conversion filters had to be coded. After these filters had been written and verified, the FEA procedure went along fine.

Only the Peltier effect, Joule heating, thermal conduction and electrical contact resistance are included in the FEA. For the same reasons as before, convection, conduction through air and radiant heat exchange are ignored. The same properties as listed in Table 2.1 on page 51 have been used. Furthermore, for the silicon (Si) substrate a thermal conductivity of $\lambda_{Si} = 150\ \mathrm{Wm^{-1}K^{-1}}$ is used. For the aluminium interconnect, a thermal conductivity $\lambda_{Al} = 220\ \mathrm{Wm^{-1}K^{-1}}$ and an electrical resistivity $\rho_{Al} = 3.08 \cdot 10^{-8}\ \Omega\mathrm{m}$ are assumed.

During the analysis, the following three boundary conditions are applied: 1) at the cold junction heat is removed by the Peltier effect and heat is generated by the electrical contact resistance 2) at the hot junction heat is generated due to both the

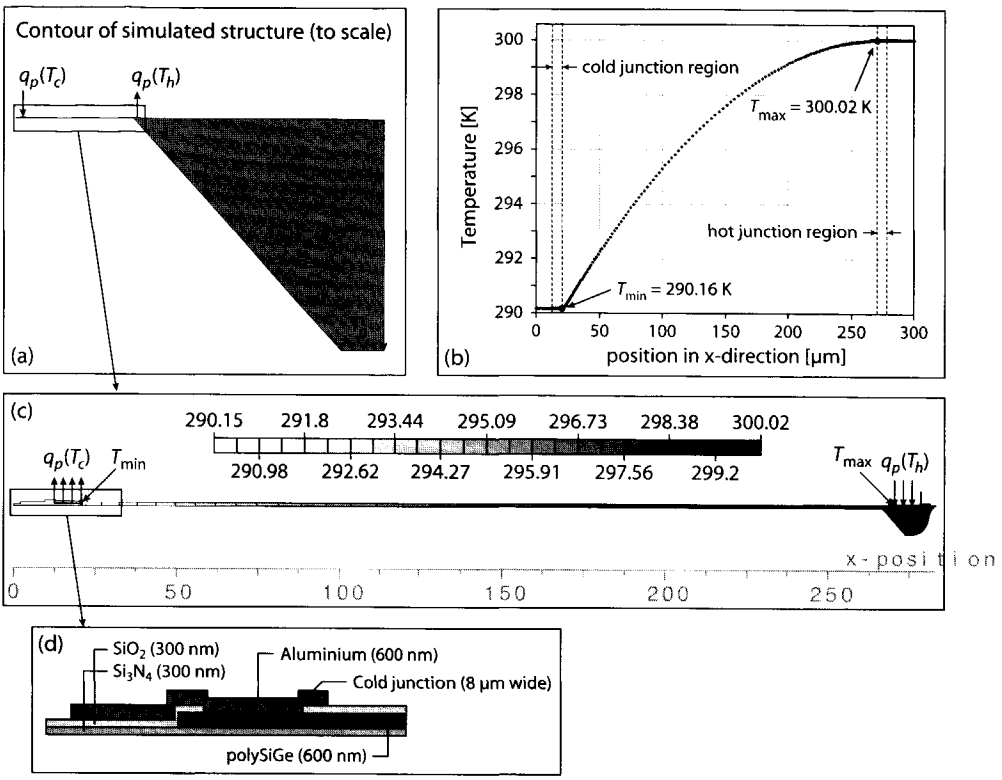


Figure 2.17 (a) Contour plot to scale clearly showing the size difference between the thin-film TEC and the substrate. (b) The resulting temperature profile in the thin-film region. (c) Graph of the temperature profile in the thin film region. (d) A schematic plot to scale of the layers used in the thin-film region.

Peltier effect and the electrical contact resistance, and 3) the temperature at the bottom of the substrate is fixed at 300 K. The locations of these three sources are indicated in Figure 2.17a, designated by $q_p(T_c)$, $q_p(T_h)$ and T_a. Furthermore, Joule heating in the thermoelements as well as in the aluminium interconnect is included. In Figure 2.17d the configuration near the tip of the cantilever is shown to scale.

Figure 2.17b and Figure 2.17c shows the results of the FEA simulation. The temperature difference between the hot and cold junctions is $T_{max}-T_{min}$ = 9.86 K. This is slighly less than would be expected from the solution provided in Section 2.5. Most likely this is due to the fact that in the FEA simulation the cold junction is 8 µm wide, unlike with the solution in Section 2.5, where the junction is modelled as a point source. Nevertheless, the result is less than 1 K off, which is quite good considering the simplifications made in both Sections 2.5 and 2.7.

The second observation from Figure 2.17b is that over 99% of the temperature gradient in the device is located across the cantilever. The temperature rise at the hot junction is a mere 0.02 K. Thus, for the power levels anticipated from a polySiGe on-chip integrated TEC (typically a few tens of μW per thermoelement) the temperature gradient across the substrate may be ignored. In particular as the simulation performed is a worst-case simulation. (In Section 2.6 it is explained why a 2-D simulation may be considered a worst-case simulation.) Needless to say, for substrates with a lower thermal conductivity than silicon or when the dielectric film between the thermoelement and the substrate is thicker or has a lower thermal conductivity, the result should be reconsidered, as such changes significantly increase the thermal resistance.

2.8 REFERENCES

[2.1] M. Peltier, "Nouvelles expériences sur la caloricité des courans électriques," *Annales de Chimie et de Physique*, LVI (56), pp. 371-386, 1834.

[2.2] D.M. Rowe, ed., *CRC Handbook of thermoelectrics*, CRC Press, Boca Raton, FL, USA, 1995.

[2.3] H.J. Goldsmid, *Applications of thermoelectricity*, Methuen, London, UK, 1960.

[2.4] W. Thomson, "On a mechanical theory of thermoelectric currents," *Proc. of the Royal Society of Edinburgh*, pp. 91-98, 1851.

[2.5] G. Min and D.M. Rowe, "Cooling performance of integrated thermoelectric microcooler," *Solid-State Electronics*, Vol. 43, pp. 923–929, 1999.

[2.6] G. Min and D.M. Rowe, "Improved model for calculating the coefficient of performance of a Peltier module," *Energy conversion & Management*, Vol. 41, pp. 163–171, 2000.

[2.7] C. LaBounty, A. Shakouri and J.E. Bower, "Design and characterization of thin film microcoolers," *J. Appl. Phys.*, Vol. 89, No. 7, pp. 4059-4064, 2001.

[2.8] X. Fan, G. Zeng, C. LaBounty, J.E. Bowers, E. Croke, C.C. Ahn, S. Huxtable, A. Majumdar and A. Shakouri, "SiGeC/Si superlattice microcoolers," *Appl. Phys. Lett.*, Vol. 78, No. 11, pp. 1580-1582, 2001.

[2.9] P.O. Gehlhoff, E. Justi and M. Kohler, "Verfeinte Theorie der elektrothermischen Kälteerzeugung," *Abh. braun. wiss. Ges.*, Vol. 2, pp. 149-164, 1950.

[2.10] F. Völklein, G. Min and D.M. Rowe, "Modelling of a microelectromechanical thermoelectric cooler," *Sensors and Actuators*, Vol. A75, pp. 95–101, 1999.

Chapter 3

MATERIAL CHARACTERISATION

The thermoelectric material forms the foundation of the on-chip integrated thermoelectric cooler. In Chapter 3 it was already shown that the maximum temperature difference between the hot and cold junctions roughly increases linearly with the figure-of-merit, which includes the Seebeck coefficient, thermal conductivity and electrical resistivity. (In fact, $\Delta T_{max} \propto \{\alpha^2, \lambda^{-1}, \rho^{-1}\}$.)

As these material properties are so dominantly present in the device performance, a good characterisation of these material parameters is indispensable and is treated in this chapter. In Section 3.1 an overview is provided of the current state of research on thin-film materials for thermoelectrics. In Section 3.2 the characterisation of thin-film poly$Si_{0.7}Ge_{0.3}$ is presented.

3.1 THIN-FILM THERMOELECTRIC MATERIALS

Less revolutionary, but of much more practical value is the research of thin-film thermoelectric materials, as these can be readily grown or deposited using common cleanroom facilities. On the one hand the phonon mean free path cannot be reduced to the magnitude of superlattice structures. On the other hand, compared to bulk fabrication technologies, crystal growth can be controlled meticulously, and smaller crystals can be formed more easily. This has to potential to reduce the lattice thermal conductivity, e.g., through boundary scattering. In particular Si-Ge alloys appear to benefit from this effect.

Running through literature, two categories of compounds and one materials have been studied in particular. These three are 1) the $(Bi_{1-x}Sb_x)_2Te_3$ compounds 2) the $Si_{1-x}Ge_x$ compounds and 3) polySi. Even though thin-film materials may have material properties that deviate from bulk properties, the latter do provide a good estimate. An alternative and intuitive approach to grading the potential of a thermoelectric materials is from the $q_c(I_{el}, T_c)$ diagram. For polySiGe such a plot was already drawn in Figure 2.4. In Table 3.1, typical values for Bi_2Te_3, polySiGe, polySi and two well-known metals are listed. Using the same geometrical values as used in

Section 2.2.1, the plots in Figure 3.1 are constructed, from Equation (2.15). The graphs are clipped at $q_c = 0$ mW and at $q_c = 5$ mW for clarity.

For a decent cooling performance, the parameter of interest must be ΔT_{max}. Thus, recalling $\Delta T_{max} \approx zT_c^2/2$, the figure-of-merit z is the decisive factor. Not surprisingly, Table 3.1 clearly reveals that Bi_2Te_3 outperforms both polySiGe and polySi by an ample margin. (In fact, polySi only performs on par with the best thermoelectric metal, Ni.) This table also indicates why: the difference in the figure-of-merit is primarily caused by the difference in the thermal conductivities.

The same conclusion can be derived graphically from Figure 3.1. The lowest value for T_c—and, hence, the largest ΔT_{max}—is found for Bi_2Te_3, in Figure 3.1a. In contrast, ΔT_{max} for polySi, plotted in Figure 3.1b, is in the order of a few Kelvin at best. PolySiGe has a ΔT_{max} in between. Even when the exact material data as in Table 3.1 is not available, the dominant material parameter can identified directly from the $q(I_{el}, T_c)$ diagram: As was already explained in Section 2.2.1, the curvature dq/dI_{el} and slope dq/dT_c of the surface, are indicators for the magnitude of the thermal conductivity and electrical resistivity. The curvature of the surface is approximately equal for all three materials, so ρ is about equal. However, going from Bi_2Te_3 to polySiGe and finally to polySi, the slope increases drastically, which indicates that λ differs significantly.

As a rule-of-thumb regarding the suitability of a thermoelectric material for cooling and power generation, the dominating performance parameter of a thermoelectric material is the thermal conductivity λ, not the power factor $\alpha^2\rho$. In line with these findings, only the thin-film $(Bi_{1-x}Sb_x)_2Te_3$ and $Si_{1-x}Ge_x$ compounds have received attention.

Thin-film bismuthides and antimonides

Of all bulk materials, tellurides and antimonides are still the best thermoelectric materials near 300 K. Downscaling to thin-film dimensions (over a few hundred nm thick) should not change the material properties too drastically. Consequently, tellurides are the most logical candidates for thin-film materials. This is reflected in amount of research that has been performed. Thin-film deposition is performed using sputtering [3.1], co-evaporation [3.2]–[3.5], flash evaporation [3.6][3.7], molecular beam epitaxy (MBE) [3.8] and metal organic chemical vapour deposition (MOCVD) [3.9]–[3.12]. The material values obtained using the various techniques are summarised in Table 3.2. Below, the essential processing parameters for each of the deposition techniques are outlined.

Thin-film bismuthide and antimonide deposition techniques

Sputtering: RF diode sputtering has been applied to deposit both Bi_2Te_3 and PbTe films. Sputtering is done directly from (99.999% pure) stoichiometric targets in an Ar atmosphere, although it is unclear which Ar gas pressure is used. The RF power applied is 100–400 W. This way a deposition rate of 1 nm s^{-1} is established. While the resulting Bi_2Te_3 composition is non-stoichiometric (44.5% Bi and 55.5% Te) the

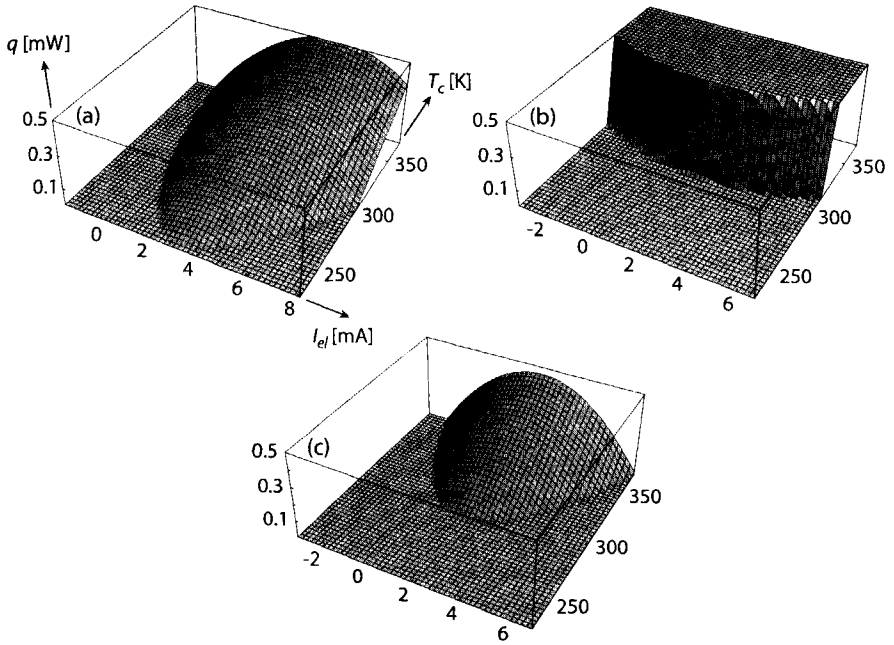


Figure 3.1 Comparison of the three thin-film materials considered for thermoelectric applications: (a) Bi_2Te_3, (b) polySi and (c) polySiGe.

Table 3.1 Comparison of bulk values for some popular thermoelectric materials.

Material		α [μVK^{-1}]	ρ [$\mu\Omega m$]	λ [$Wm^{-1}K^{-1}$]	z [$10^{-3}K^{-1}$]	Ref.
Bi_2Te_3	n	-240	10	2.02	2.89	[2.2]
	p	162	5.5	2.06	2.32	
polySiGe	n	-136	10.1	4.45	0.328	[2.2]
	p	144	13.2	4.80	0.413	
polySi	n	-120	8.5	24	0.071	[3.13]
	p	190	58	27	0.037	
	p	219.3	28.3	33.6	0.060	[3.14]
Metals	Ni	-19.5	$7.20 \cdot 10^{-2}$	90.5	0.059	
	Al	-1.66	$2.73 \cdot 10^{-2}$	237	0.00043	

PbTe composition is. Moreover, when PbTe is 1.5% Sn-doped, an improvement of the power factor can be observed. When annealed at temperature above 250 °C, both films experienced a decrease in thermoelectric properties.

Co-evaporation: Direct evaporation of Bi_2Te_3 is not considered appropriate due to the large differences in the vapour pressures of Bi and Te. In general Bi_2Te_3 films were deposited by co-evaporation from high-purity (99.999%) Bi and Te targets. Bismuth is evaporated from a molybdenum boat, while tellurium is evaporated from either a tantalum crucible or a tungsten boat. The deposition rates need to be controlled and monitored independently. To ensure the proper Bi:Te ratio is deposited, the substrate holder has to be stabilised at a temperature T_s above ambient temperature. Excess Te concentrations are observed for $T_s < 250$ °C, while Bi_2Te_3 is deficient of Te for $T_s > 310$ °C.

Zou *et al.* [3.4][3.5] performed deposition at a pressure of $3{\cdot}10^{-6}$ Torr. For p-type Bi_2Te_3 they used a flux ratio F[Bi]:F[Te] equal to 3:1, with $T_s = 300$ °C. Similarly, for n-type Bi_2Te_3 a flux ratio F[Bi]:F[Te] equal to 1:2 is used, with $T_s = 260$ °C. Finally, for p-type Sb_2Te_3 a flux ratio F[Sb]:F[Te] of 1:2, with $T_s = 230$ °C is used. Shafai *et al.* [3.2] used slightly different setting for Bi_2Te_3, i.e., $p = 2{\cdot}10^{-6}$ Torr, F[Bi]:F[Te] = 1:2.3 and $T_s = 270$ °C. The typical growth rate is in the order of 1.5 μm h^{-1}.

Flash evaporation: Unlike co-evaporation, flash evaporation is suitable for deposition of multiple elements with different vapour pressures. This makes it suitable to evaporate from stoichiometric targets. Both Völklein *et al.* [3.6] and Foucaran *et al.* [3.7] made their own evaporation targets.

The constituents are first mixed in stoichiometric proportions in a quartz ampoule. This ampoule is sealed in an H_2 atmosphere at 1 Pa. A homogeneous alloy is created by heating the mixed powder. Foucaran *et al.* used a temperature of 800 °C for 24 h for (Bi_2Te_3–Bi_2Se_3) and (Bi_2Te_3–Sb_2Te_3). Instead, Völklein *et al.* used a temperature of 950 °C for 24 h. After quenching to room temperature, the alloys are pulverised into grains several hundreds of micrometer in diameter. Next, flash evaporation is performed at a typical pressure of 10^{-3} Pa and a substrate temperature of 480 °C. The growth rate is in the order of 1 nm s^{-1}. After deposition, the films are annealed. Foucaran *et al.* found an anneal of 250 °C for 2 h was sufficient to stabilise the thermoelectric properties. Völklein *et al.* found somewhat different values; They required an anneal at 573 K during 4 h.

Molecular beam epitaxy (MBE): The MBE process description provided by Mzerd *et al.* [3.8] is very meagre. Deposition is performed from condensation of the Bi, Sb and Te molecular beams. During growth, the substrate was kept at a temperature of 260 °C. For both Bi_2Te_3 and $Bi_{0.1}Sb_{1.9}Te_3$, the respective flux ratios F[Te]:F[Bi] and F[Te]:F[Bi+Sb] were kept above 3.

Metal organic vapour deposition (MOCVD): MOCVD has been applied to fabricate both n-type and p-type Bi_2Te_3. The bismuth precursor used is trimethylbismuth (TMBi). As tellurium precursor, three materials are suitable, dimethyltelluride (DMTe), diethyltellure (DETe) and diisopropyltelluride (DIPTe). DETe seems to be

Table 3.2 Comparison of thermoelectric properties of various bismuthides and antimonides and PbTe.

Material		n [cm^{-3}]	α [μVK^{-1}]	ρ [$\mu\Omega m$]	λ [$Wm^{-1}K^{-1}$]	ψ [$10^{-3}\,Wm^{-1}K^{-2}$]	z [$10^{-3}\,K^{-1}$]	Reference	Remarks
Bi_2Te_3	n	$7.2 \cdot 10^{20}$	-56	15.4	–	0.20	–	[3.1]	(1)
Bi_2Te_3	n	–	-187	11.5	–	3.0	–	[3.2]	(2)
Bi_2Te_3	p	$1.6 \cdot 10^{21}$	81	3.2	–	2.1	–	[3.4]	(2)
Sb_2Te_3	p	$3.4 \cdot 10^{19}$	185	31	–	1.1	–	[3.5]	(2)
Bi_2Te_3	n	$6.5 \cdot 10^{19}$	-228	13	–	4.0	–	[3.4]	(2)
$Bi_{0.5}Sb_{1.5}Te_3$	p	$1.4 \cdot 10^{19}$	230	17	1.1	3.1	2.9	[3.6]	(3),(6)
$(Bi_2Te_3)_{0.9}(Bi_2Se_3)_{0.1}$	n	–	-90	30	–	0.58	0.17	[3.7]	(3),(7)
$(Bi_2Te_3)_{0.25}(Sb_2Te_3)_{0.75}$	p	–	40	50	–	0.27	0.0021	[3.7]	(3),(7)
Bi_2Te_3	n	–	-164	29	–	0.93	–	[3.8]	(4)
$Bi_{0.1}Sb_{1.9}Te_3$	p	–	164	19	–	1.4	–	[3.8]	(4)
Bi_2Te_3	n	$4.5 \cdot 10^{19}$	-210	12	–	3.7	2.48	[3.9]	(5),(7),(8)
Bi_2Te_3	n	$1.6 \cdot 10^{19}$	-94	66	–	0.13	0.5	[3.10]	(5),(7),(9)
Bi_2Te_3	p	$1.4 \cdot 10^{19}$	190	78	–	0.46	2.5	[3.10]	(5),(7),(9)
Bi_2Te_3	n	$1.6 \cdot 10^{20}$	-218	6.92	–	6.9	–	[3.11]	(5)
PbTe		$5.6 \cdot 10^{19}$	175	28.3	–	1.3	0.060	[3.1]	(1)

(1) Sputter deposition, (2) co-evaporation, (3) Flash evaporation, (4) MBE, (5) MOCVD
(6) α, λ and ρ estimated from figures in reference
(7) z estimated by author from assumed value for λ
(8) Doping concentration estimated from figure in reference
(9) ρ derived from μ and n.

the most suitable precursor, due to the low cracking temperature (390 °C for DETe versus 500 °C for DMTe; The latter temperature is very close to the melting temperature of Bi_2Te_3.) Both TMBi and DETe gasses are stored at controlled temperatures, respectively 5 °C and 20 °C. Hydrogen (3 slm to 6 slm) is used as a carrier gas. To prevent early reactions, both gasses remain separated until the reactor head is reached.

To ensure 100% cracking of the precursors, the growth temperature is 450 °C. The pressure during growth is kept constant at about 10^{-4} atm. The minimum DETe:TMBi partial pressure ratio to obtain stoichiometric alloys is 1.5:1. Below a ratio of 2.5, p-Bi_2Te_3 is obtained, above a ratio of 2.5, n-Bi_2Te_3 is obtained. The optimum thermoelectric properties are obtained for a DETe:TMBi partial pressure ratio of about 7:1. For ratios varying from 2 to 15, the growth rate was practically constant, with about 0.5 μm h^{-1} for 3 slm H_2 and 1.2 μm h^{-1} for 6 slm H_2.

Thin-film polycrystalline silicon germanium

Thin-film polycrystalline silicon germanium has received less academic attention than bismuthides and tellurides due to the inferior thermoelectric performance. To my knowledge, only three groups have actively worked on the polySiGe for thermoelectric purposes, including the work in this thesis. The most elaborate data set has been collected at IMEC in Belgium [3.14]–[3.17]. This data set not only includes thermoelectrical data, but also data on the mechanical stress and strain in thin-film polySiGe. The next-largest data set is presented in this thesis. Finally, Strasser *et al.* [3.18] at Infineon, Germany, has shown the major benefit of polySiGe over Bi_2Te_3: it can be directly integrated with a standard BiCMOS process (or any other microelectronic process.) In Table 3.3 the different data sets are compared. Also included in this table are the best bulk SiGe properties reported, provided in[2.2].

Thin-film polySiGe preparation techniques

In general (e.g., also for SiGe transistors) the most favourite deposition technique for polySiGe is chemical vapour deposition. In particular atmospheric pressure CVD (APCVD) and reduced pressure CVD (RPCVD) and ultra low pressure CVD (ULPCVD) are applied frequently. Furthermore, interesting work has been performed on reactive thermal CVD (RTCVD) [3.19] and co-sputtering[3.20]. Both techniques allow polySiGe to be grown at much lower temperature than possible with APCVD or RPCVD. This furthers the compatibility between processing of TEC and microelectronics.

Work performed at IMEC: In [3.16], Sedky *et at.* have performed an in-depth study on the structural and mechanical properties of polySiGe. In particular the grain structure and stress treatment are covered extensively. Both APCVD and RPCVD (at a pressure of 40 Torr) were investigated, as well as ULPCVD polySiGe doped *in situ* to a concentration of 10^{20} cm^{-3} [3.14].

Deposition is performed in an ASM Epsilon CVD reactor, from a mixture of germane (GeH_4) and dichlorosilane (Si_2Cl_6). APCVD and RPCVD samples were grown,

Table 3.3 Comparison of thermoelectric properties of different polySiGe samples.

n [cm^{-3}] or S [cm^{-2}]		Technique	T_{dep}, T_{ann} [°C, °C]	α [μVK^{-1}]	ρ [$\mu\Omega m$]	λ [$Wm^{-1}K^{-1}$]	ψ [$Wm^{-1}K^{-2}$]	z [$10^{-3}\ K^{-1}$]	Ref.
$\pm 2.0 \cdot 10^{20}$ cm^{-3}	n	sintering	–	-136	10.1	4.45	$1.8 \cdot 10^{-3}$	0.328	
$\pm 3.0 \cdot 10^{20}$ cm^{-3}	p	sintering	–	144	13.2	4.80	$1.6 \cdot 10^{-3}$	0.413	
$7.5 \cdot 10^{15}$ cm^{-2}	n	APCVD	700, 1000	-174	29.2	5.1	$1.0 \cdot 10^{-3}$	0.203	our work
$5.0 \cdot 10^{15}$ cm^{-2}	p	APCVD	700, 1000	129	28.9	4.7	$5.8 \cdot 10^{-4}$	0.123	
$2.5 \cdot 10^{20}$ cm^{-3}	n	RPCVD	670, –	-77	23.7	9.4	$2.5 \cdot 10^{-4}$	0.026	[3.18]
$2.5 \cdot 10^{20}$ cm^{-3}	p	RPCVD	670, –	59	18.7	11.1	$1.9 \cdot 10^{-4}$	0.016	
$5.0 \cdot 10^{15}$ cm^{-2}	p	APCVD	650, 950	133.6	33.8	3.5	$5.3 \cdot 10^{-4}$	0.152	[3.14]
$1.0 \cdot 10^{20}$ cm^{-3}	p	ULPCVD	600, –	76.6	28.3	33.6	$2.1 \cdot 10^{-4}$	0.060	

at a respective temperature of 650 °C and 625 °C, and a respective growth rate of 37 nm min^{-1} and 17 nm min^{-1}. Measurements indicate that both deposition techniques provide a homogeneous SiGe alloy (rather than a polySi film with Ge-rich clusters.) Transmission electron microscopy (TEM) and X-ray diffraction (XRD) revealed that the as-deposited APCVD material has a columnar structure (on average with columns 200 nm in diameter.) As-deposited RPCVD material is more randomly oriented. This is explained from the low deposition temperature of 625 °C, which is close to the amorphous/polycrystalline transition temperature.

After deposition the RPCVD polySiGe has a slight tensile stress (+60 MPa). Further annealing at temperatures up to 1050 °C has a negligible effect on both stress and grain size. In contrast, APCVD polySiGe is significantly influenced by an anneal at temperature of 700 °C and above. First, the structure goes from columnar to more randomly oriented grains with an average diameter of 400 nm. Second, at higher temperatures the initially compressive stress (-145 MPa) is reduced. At a temperature of approximately 950 °C the stress even becomes tensile. In Figure 3.2 the change is stress level against annealing temperature is plotted.

To be able to measure the thermal conductivity of the thin film layer, a structure similar to that discussed in Section 3.2.4 was constructed. Instead of a 300 nm Si_3N_4 supporting membrane as in our case, a 1 μm polyimide layer ($\lambda \approx 0.2$ $Wm^{-1}K^{-1}$) was spun onto the wafer before polySiGe deposition. The required cure at 350 °C did cause the polyimide to buckle. The best results were obtained with APCVD polySiGe, annealed for 2 hours at 950 °C and a B implant with a dose of $5 \cdot 10^{15}$ cm^{-2}. Under these conditions, a figure-of-merit of $0.153 \cdot 10^{-3}$ K^{-1} is obtained.

Work performed at Infineon: At Infineon, Strasser *et al.* have investigated the suitability of polySiGe for 100% IC compatible TEGs. As such, they have compared polySi TEGs to $polySi_{0.7}Ge_{0.3}$ TEGs. Indeed polySiGe outperformed polySi, although—not surprisingly—the largest performance gain came from improving thermal isolation. Like at IMEC, a gas mixture of dichlorosilane and germane (1% GeH_4 in H_2) is used. A 400 nm polySiGe layer was deposited by CVD at a pressure of 16 Torr and a temperature of 670 °C. The growth rate was 100 nm min^{-1}. Again a polySi nucleation layer is applied, in this case with a thickness of 5 nm. Unfortunately, no mention is made of any anneal. Considering the low figure-of-merit obtained (from $16 \cdot 10^{-6}$ K^{-1} to $26 \cdot 10^{-6}$ K^{-1}) it is likely that no such step has been applied.

Work performed at the Delft University of Technology: The work presented in this thesis concerns the fabrication of TECs from $polySi_{0.7}Ge_{0.3}$. After deposition of a 10 nm polySi nucleation layer, a 600 nm $polySi_{0.7}Ge_{0.3}$ layer is grown by APCVD. The growth took place at a temperature of 700 °C. A gas mixture of hydrogen (40 slm), with 5 % germanium (51 sccm) and dichlorosilane (20 sccm) was used to establish a growth rate of 35 nm/min.

After introduction of the B and P impurity atoms, and after a TEOS layer is deposited over the polySiGe, a 55 min anneal is performed; 20min at 600°C followed by

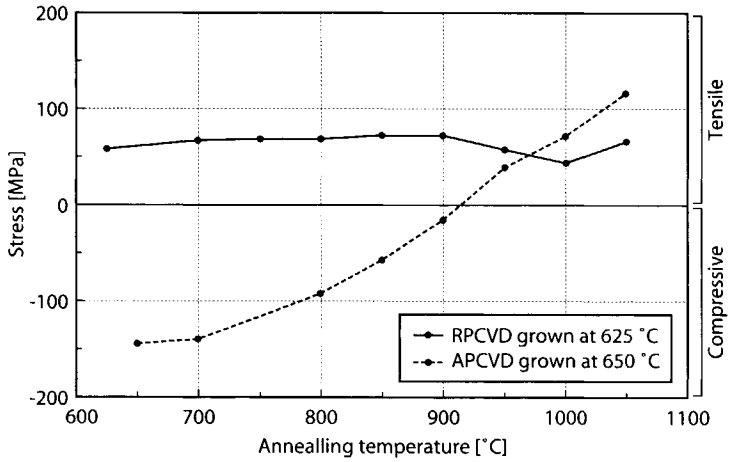


Figure 3.2 Stress in polySiGe versus the annealing temperature. (From Sedky *et al.* [3.16].)

35min at 1000°C. The thermoelectric material properties thus obtained are listed in Table 3.3.

3.2 MATERIAL CHARACTERISATION

To reliably estimate the performance of an on-chip integrated TEC, it is essential to determine all relevant material properties. First and foremost, this applies to the properties of the thermoelectric materials, as these are the fundamental building block of any integrated TEC. The essential material properties are the electrical resistivity, ρ, the Seebeck coefficient, α, and the thermal conductivity, λ. Recall that the ratio of these three properties determines the figure-of-merit, z. As the TEC is designed to operate over a large temperature range, whilst subjected to large temperature gradients, it is insufficient to determine the properties at room temperature only. Instead, the corresponding temperature coefficients (TCs) need to be determined as well. If possible, the properties are determined over the temperature range from -25 °C to +125 °C, which is limitated by the measurement setup.

Besides the material properties of the thermoelements, Chapter 3 revealed that the properties of a number of performance-limiting effects need to be taken into account as well. These are the electrical contact resistance between the thermoelements and the metal interconnect, and the thermal conductivities of the SiO_2 and Si_3N_4 dielectric layers.

3.2.1 Electrical resistivity

The electrical resistivity is one of the most frequently monitored parameters during and after the fabrication of integrated circuits and devices. Over the years a number

of different techniques have been developed to measure this parameter [3.21]. By far the best known techniques are the four-point measurements like the bridge structure and the Van der Pauw structure. However, contactless measurements by means of eddy current and confocal resonators are also of encountered. The thermoelectric material characterisation presented is this thesis used contact measurements, so only the four-point measurement techniques are discussed in more detail.

Four-point collinear probe

After the epitaxial growth and impurity implantation of the polySiGe layer (but before the anneal) the resistivity of the polySiGe samples is determined using a four-point probe, similar to the collinear four-point probe shown in Figure 3.3. An electrical current I_{el} is passed through probes 1 and 4, while the voltage difference across probes 2 and 3 is recorded. The advantage of sensing the voltage from these two separate probes rather than directly from the current-carrying probes is that non-linear effects like the probe resistance, contact resistance between the probe and the sample and current spreading resistance near the probe tip are avoided. For an arbitrarily shaped sample, with a constant probe spacing $s_1 = s_2 = s_3 = s$ the resistivity ρ is

$$\rho = 2\pi s F \frac{V_{32}}{I_{14}} \tag{3.1}$$

where $F = F_1F_2F_3$ is a correction factor, $V_{23} = V_3 - V_2$ and I_{14} is the current running from probe 1 to probe 4. The factor F_1 is used to correct for the sample thickness, F_2 corrects for the lateral sample dimensions and F_3 corrects for the probe placement in relation to the sample edges. As the polySiGe sample is in fact a separate test wafer entirely covered with polySiGe, F_2 and F_3 are close to unity, so only the influence of the sample thickness H needs to be considered [3.21]:

$$F_1 = \frac{H/s}{2\ln\left([\sinh(H/s)]/[\sinh(H/2s)]\right)} \cong \left.\frac{H/s}{2\ln 2}\right|_{H \le s/2} \tag{3.2}$$

The latter approximation is valid for very thin samples, when $H \le s/2$, as is the case for the polySiGe layer grown. As such, the resistivity from the collinear four-point probe for thin film samples becomes

$$\rho = H \frac{\pi}{\ln 2} \frac{V_{34}}{I_{12}} \tag{3.3}$$

Van der Pauw structures

When the resistivity of a thin film is an important design parameter (as is the case with the on-chip integrated TEC,) dedicated test structures are incorporated into the design. Bases on the theory of Van der Pauw [3.22], various four-point configurations have been designed of which some are shown in Figure 3.3b. The current is

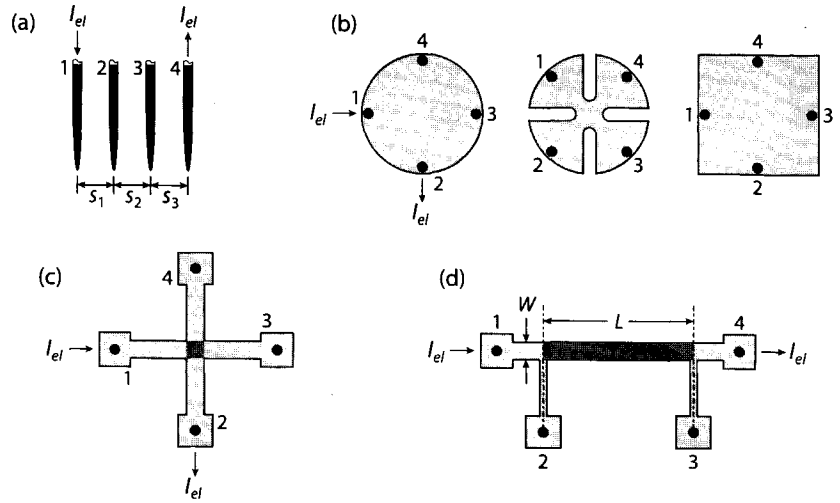


Figure 3.3 Various four-point measurement configurations: (a) the collinear probe, (b) Van der Pauw configurations, (c) the Greek cross and (d) the bridge configuration.

forced through points 1 and 2, while the voltage is measured across points 3 and 4. For a perfectly symmetrical Van der Pauw structure, the resistivity equals

$$\rho = H \frac{\pi}{\ln 2} \frac{V_{34}}{I_{12}} \tag{3.4}$$

Often the designer is not interested in the resistivity of the material, as he/she needs to know the thickness of a layer in order to determine the resistance of a strip of material he has drawn. Instead, most foundries specify the sheet resistance of a material, R_{sh}, which simply is the resistivity of the layer divided by its thickness. Thus, the sheet resistance of the Van der Pauw structure is

$$R_{sh} = \frac{\rho}{H} = \frac{\pi}{\ln 2} \frac{V_{34}}{I_{12}} \tag{3.5}$$

In particular for shapes without notches it is very difficult to assure that all four contact points with the shape are small and are placed exactly symmetrical with respect to the centre of the shape. Therefore, in IC technology the Greek cross, shown in Figure 3.3c, is preferred, as alignment of the probe on each of the four bond pads is much less critical. Equations (3.4) and (3.5) remain valid for this structure, and can be used to calculate the resistivity and sheet resistance of the darkened square in the middle of the structure.

Bridge structure

Besides the Greek cross, the bridge structure as indicated in Figure 3.3d is often encountered and is preferred by me. The electrical current runs from bondpad 1 to 4, while the voltage over bondpads 2 and 3 is measured to determine the resistance / resistivity of the dark gray area. These two values respectively are

$$R_{sh} = \frac{V_{23}}{I_{14}} \frac{W}{L} \tag{3.6}$$

and

$$\rho = R_{sh} H = \frac{V_{23}}{I_{14}} \frac{HW}{L} \tag{3.7}$$

where L and W are the respectively are the length and width of the resistor.

There are two advantages of using the bridge structure over the Greek cross. First, instead of measuring the voltage over a single square, the length:width ratio can be chosen arbitrarily. In particular if the resistivity of a layer is small, a bridge can be made long to obtain a reasonable voltage drop over the resistor. This is important to minimise self-heating of the resistor. A second advantage is that the bridge structure is more suitable to calculate any deviations in the width W of the resistor. As the resistor is either patterned by means of an etch or a lift-off, there are variations in the exact geometry of the resistor. This does not influence L, but does influence W. For example, an overetch can make W smaller, while an improper lift-off can leave more resistor material than anticipated, effectively widening W. If the width has a variation $\pm\Delta W$, then the resistance $R \propto L/(W \pm \Delta W)$. Assuming ΔW is the same for each device on a specific wafer, the deviation in the width can be calculated from the derivative of R to L:

$$\frac{dR}{dL} = \frac{R_1 - R_2}{L_1 - L_2} = \frac{1}{W \pm \Delta W} \tag{3.8}$$

By measuring R for multiple values of L, the derivative and, hence, ΔW can be directly determined.

One issue that is important when dealing with folded bridge resistors (as shown in Figure 3.4e) is the weighting of the corner resistance, as the current experiences a different resistance at a corner [3.23]. In Figure 3.4 the most common corner configurations are displayed, with the corresponding correction value.

3.2.1.1 Thermal issues

One issue that has to be regarded is the prevention of dissimilar thermal gradients across the leads from the resistor to the measurement device. When using a four-point measurement, this applies particularly to the leads over which the voltages are

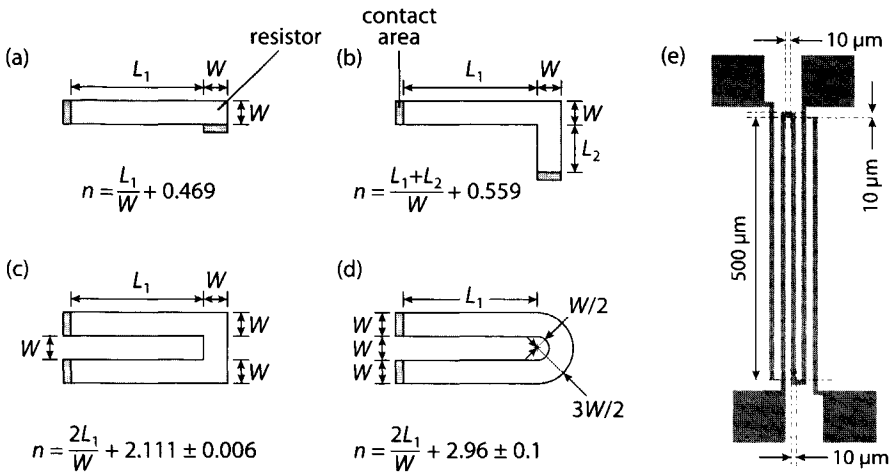


Figure 3.4 (a)–(d) Different correction values for specific corner configurations. (e) Drawing to scale of an folded bridge resistor used to determined electrical resistivity and TCR.

measured. If the thermal gradients are dissimilar, the resulting Seebeck potential will manifest itself as a DC potential.

Therefore, unless the thermal gradients can be accurately controlled, a DC four-wire measurement is not the most suitable resistance measurement. Instead, it is better to use an AC measurement or a pulsed DC measurement. The latter type of measurement uses the finite thermal diffusivity of the resistor and package, so that thermal gradients do not get sufficient time to build up and cause a Seebeck potential. Moreover, performing a pulsed DC measurement largely prevents self-heating of the resistor, which is of particular importance in micromachined and thermally isolated structures. If the polarity of each consecutive pulse is inverted, any Peltier effect that occurs at the metal-semiconductor junction will average out over time.

3.2.1.2 Measurement results

Figure 3.4e shows a folded bridge structure (to scale) as it has been used to determine the resistivity and TCR of the aluminium interconnect and the n-type and p-type polySiGe. This structure has two corner configurations as shown in Figure 3.4c. Therefore, the total number of squares of the folded bridge is equal to $3 \cdot (500/10) + 2 \cdot (2.111 \pm 0.006) \approx 154.22$ squares. As such, $R_{sh} = R/154.22$.

The die holding the various bridge structures is glued to a ceramic chip carrier and electrical connections established by wire bonding. The resistance of the meandered structure is measured using a digital multimeter.

To ensure the environment is thermally stabilised and to allow the temperature to be controlled over a range from −25 °C to +125 °C, the packaged devices are placed inside a climate chamber. The temperature inside the climate chamber is de-

termined using a Pt-100 resistor. By stabilising at these different temperatures, the temperature coefficient of resistance (TCR) of the resistors can be determined. The TCR β is defined as

$$\beta = R(T)^{-1} \frac{\mathrm{d}R(T)}{\mathrm{d}T} \tag{3.9}$$

As $\mathrm{d}R/\mathrm{d}T$ cannot be measured directly, the TCR is determined from the difference quotient, i.e.,

$$\beta(T + \tfrac{1}{2}\Delta T) \cong R(T + \tfrac{1}{2}\Delta T)^{-1} \frac{R(T + \Delta T) - R(T)}{\Delta T} \tag{3.10}$$

Over the temperature range of interest, -25 °C to +125 °C, the TCR is practically constant, so only the first order coefficient needs to be determined. The results obtained are listed in Table 3.4. At first glance, the positive TCR of the n-type polySiGe seems surprising. However, due the high doping concentration (near degeneracy), a positive TCR is commonly obtained.

Table 3.4 Electrical resistivity and TCR measured.

Material	Resistivity [Ωm]	TCR [°C^{-1}]
Aluminium ($Al_{0.98}Si_{0.02}$)	$(30.8 \pm 0.5)\cdot 10^{-8}$	–
p-poly$Si_{0.7}Ge_{0.3}$	$(28.9 \pm 0.2)\cdot 10^{-6}$	$(10.6 \pm 0.7)\cdot 10^{-4}$
n-poly$Si_{0.7}Ge_{0.3}$	$(29.3 \pm 0.3)\cdot 10^{-6}$	$(5.6 \pm 0.4)\cdot 10^{-4}$

3.2.2 Electrical contact resistance

As was established in Chapter 3, a second electrical parameter that needs to be determined is the electrical (interfacial) contact resistance R_c, which is capable of significantly degrading TEC performance.

For example, for a rectifying (metal to n-type semiconductor) junction with a low to moderate doping concentration the thermionic emission dominates [3.24]. In such a case the specific contact resistance is

$$R_c = \frac{k_B T}{e} \frac{\exp\left(\frac{e\phi_{Bn}}{k_B T}\right)}{A^* T^2} \tag{3.11}$$

where

$$A^* = \frac{4\pi e m_n^* k_B^2}{h^3} \tag{3.12}$$

is the effective Richardson constant for thermionic emission in $AK^{-2}cm^2$. Thus, the contact resistance decreases rapidly as the barrier height (ϕ_{Bn}) decreases.

When the semiconductor has a higher doping concentration, the field emission will dominate thermionic emission, Under this circumstance, the relationship between specific contact resistance and doping concentration [3.21][3.24]

$$R_c \propto \exp\left[\frac{2\sqrt{\varepsilon_s m^*}}{\hbar} \cdot \frac{\phi_B}{\sqrt{N}}\right] \tag{3.13}$$

where $\hbar = h/2\pi$ is the modified Planck's constant and ε_s is the semiconductor's permittivity. This clearly reveals that the specific contact resistance is a fairly strong function of semiconductor doping.

3.2.2.1 Measurement method

Electrical contact resistance can be determined directly by means of a Kelvin structure as shown in Figure 3.5. This structure is based on a model developed in the 1970's [3.25], which bridged the incomplete models used until then, the Kennedy–Murley Model (KMM) and the Transmission Line Model (TLM). From this model, the interfacial contact resistance can only be determined indirectly from the so-called end contact resistance [3.26].

In 1983, Proctor *et al.* improved a four-terminal test structure reported by Anderson and Reith in 1975 [3.27] such that both the end contact resistance and interfacial contact resistance can be measured directly. This is the structure shown in Figure 3.5.

By running an electrical current I_{el} from terminal 1 to terminal 3, and measuring the potential difference across terminals 6 and 2. The electrical interfacial contact resistance R_c is directly determined from

$$R_c = \frac{V_6 - V_2}{I_{el}} = \frac{V_{62}}{I_{el}}. \tag{3.14}$$

The specific electrical contact resistance, ρ_c directly follows as

$$\rho_c = A_c R_c. \tag{3.15}$$

where A_c is the contact area.

3.2.2.2 Measurement results

Based on the Kelvin structure, a series of measurements have been performed in order to establish the (specific) electrical contact resistance. Both the contact resistance from $Al_{98.5}Si_{1.5}$ to p-type as well as n-type poly$Si_{0.7}Ge_{0.3}$ has been determined. InTable 3.5 and Table 3.6 the results are summarised in tabular form. In Figure 3.6 on page 79, the individual R_c measurement values are compared graphically.

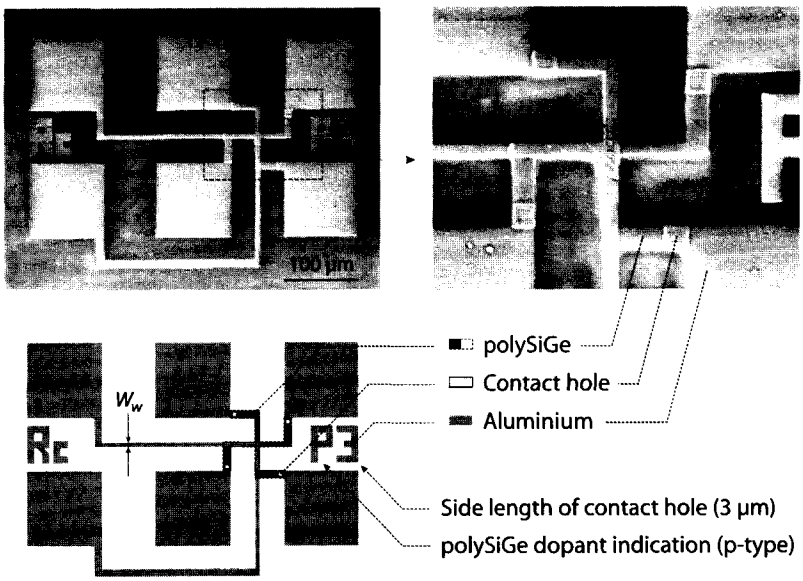


Figure 3.5 (a) A SEM image with (b) close-up, as well as (c) a schematic drawing of a six-terminal Kelvin structure to measure the interfacial contact resistance R_c (after [3.27].)

Table 3.5 Electrical contact resistance for the $Al_{98.5}Si_{1.5}$/p-type $polySi_{0.7}Ge_{0.3}$ interface.

A_c [µm²]	$\langle R_c \rangle$ [Ω]	$\langle \rho_c \rangle$ [Ω·µm²]	σ [Ω·µm²]
4	8.72	34.9	2.03
9	5.08	45.7	3.06
16	2.89	46.2	2.65
25	1.65	41.3	3.00
100	0.402	40.2	9.10

Table 3.6 Electrical contact resistance for the $Al_{98.5}Si_{1.5}$/n-type $polySi_{0.7}Ge_{0.3}$ interface.

A_c [µm²]	$\langle R_c \rangle$ [Ω]	$\langle \rho_c \rangle$ [Ω·µm²]	σ [Ω·µm²]
4	14.9	59.7	22.1
9	12.7	114	34.3
16	10.3	158	49.4
25	9.00	215	74.4

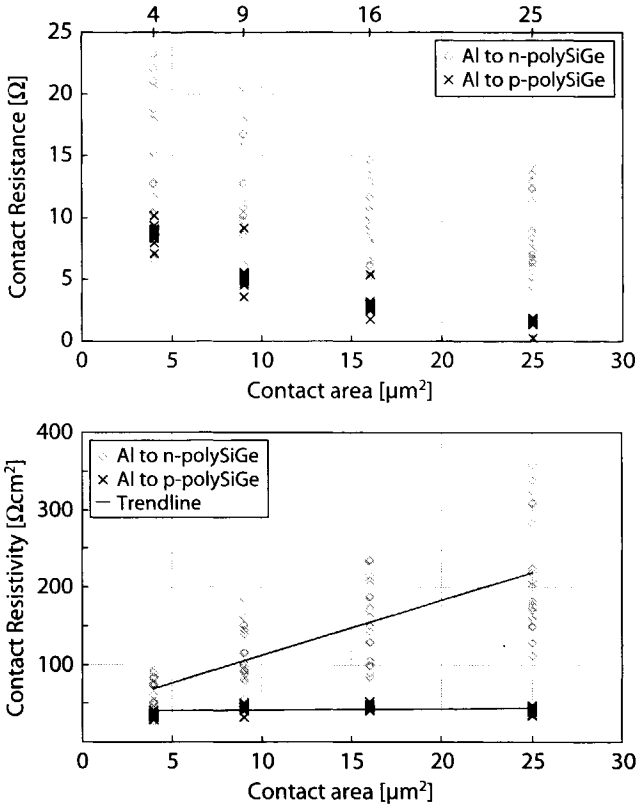


Figure 3.6 (a) The electrical contact resistance R_c and (b) the specific contact resistance of the $Al_{0.85}Si_{0.15}$–$polySi_{0.7}Ge_{0.3}$ junction.

The first columns of Table 3.5 and Table 3.6 indicates the surface area of the square metal-semiconductor contact. In the second column, the average contact resistance values measured are listed. In the third and fourth column, the mean specific contact resistance and the corresponding standard deviation are listed.

The first and most obvious observation from the values reported is that the contact to the n-type polySiGe is significantly higher than to the p-type polySiGe. Moreover, the spread in the values obtained is also much larger for n-type than p-type contacts. This result is directly in line with the general rule that an n-type contact is much harder to establish than a p-type contact.

What has to be noted for all of the n-type contact resistance measurements is that a number of samples were outside the $\pm 3\sigma$ range, so these were discarded. As a result, the measurement population dropped to below 20 samples, indicating the mean and standard deviation values have to be used with reservation. This is illustrated in

particular for the 5μm × 5 μm n-type contacts, where the 3σ-value (= 223 Ωμm^2) is larger than the mean value $\langle \rho_c \rangle$ of 215 Ωμm^2.

Focusing on the more reliable values for the p-type contact, there still is a significant difference in $\langle \rho_c \rangle$ across different contact areas. A plausible explanation to this variation was provided in 1985, by Loh *et al.* [3.28]. They performed a thorough analysis of the Kelvin structure. Some interesting conclusions were drawn, in particular with respect to the width ratio between the side length of a contact, L_c, and the width of the wires carrying the electrical current, W_w. (These values are indicated in Figure 3.5.)

The major conclusion from their work is that W_w should match L_c as closely as possible, in order to remove the effect of current crowding near the contacts. In the case W_w and L_c are not equal, this can lead to a significant measurement error. For example, for the 9 μm^2 contact, the ratio L_c:W_w is 3:6. Estimating from the values reported in [3.28], it would appear the measured value of ρ_c is 20% off the actual value. This error becomes less as the contact area decreases and the L_c:W_w ratio gets closer to unity. Keeping an error of 20% in mind, the various mean values do no longer contradict.

3.2.3 Seebeck coefficient

The second material parameter embedded in the figure-of-merit is the Seebeck coefficient of thin films. Recalling from Chapter 3, the Seebeck coefficient α is defined as the rate of change of the Seebeck voltage with respect to a change in temperature, at a given temperature, i.e.

$$\alpha = \left.\frac{\mathrm{d}V}{\mathrm{d}T}\right|_T \tag{3.16}$$

In essence, this is the equation describes how the Seebeck coefficient is measured: In essence, if the voltage difference over a strip of material can be determined, which is subjected to a known temperature difference, the Seebeck coefficient follows automatically.

3.2.3.1 Measurement method

Two measurement structures can be used to measure the Seebeck coefficient [3.29]. They differ in that one structure requires micromachining, while the other does not. The structures are constructed from the same thin-film layers, using processing steps similar to the thin-film Peltier devices discussed in this thesis (See the next chapter for processing details). In Figure 3.7 a schematic top view of the micromachined structure is shown, as it was used to obtain the results presented in this section. This structure is directly derived from the micromachined structure presented in [3.29].

A single heater is used to raise the temperature at one end of the (thermoelectric) material sample under investigation. The measurement of the temperatures as both the hot and cold end are made using either two-wire or four-wire resistors. To meas-

ure ΔV, an electrical connection is established to both ends of the sample, to measure the potential difference when subjected to a temperature gradient. From these three measurements (T_h, T_c and ΔV), the Seebeck coefficient can be determined.

As dV/dT cannot be measured directly, the difference quotient is once again used to measure the Seebeck voltage. This implies that the parameter measured is not that of Equation (3.16), but rather the Seebeck coefficient at the temperature half way in between T_h and T_c:

$$\alpha\left(T+\tfrac{1}{2}\Delta T\right)=\frac{\Delta V}{T_h-T_c}=\frac{\Delta V}{\Delta T} \tag{3.17}$$

This approximation is valid if the temperature difference ΔT ($=T_h-T_c$) is sufficiently small to allow ΔV to be linearised. It has been established empirically that a temperature difference of around 10 K is acceptable.

3.2.3.2 Measurement results

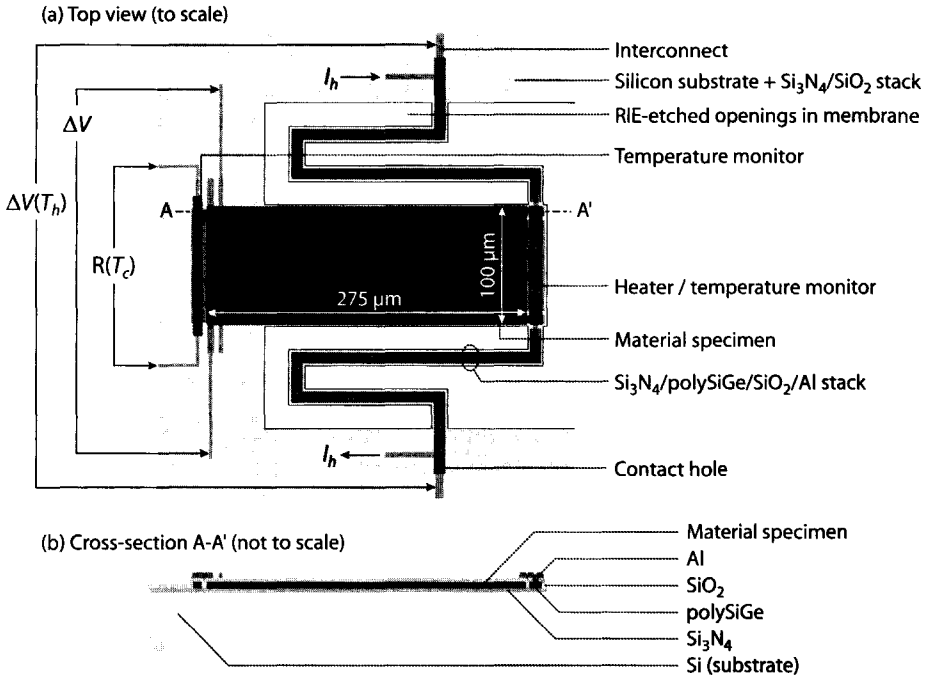
Unlike in [3.29], the choice was made to use a single resistor to both heat the sample and monitor T_h. For this purpose, the heater has been configured as a four-terminal resistor. (In [3.29] a separate two-terminal resistor was used to monitor the temperature at the hot end of the micromachined structure.) To measure T_c, a two-terminal resistor was used which, in a more recent design, was changed to a four-terminal resistor. In order to use a resistor as temperature monitor (similar to a Pt-100 resistor) it is necessary to know the (R, T)-relationship of the specific resistor. Thus, for each resistor used, both the resistivity and TCR need to be determined beforehand. This is done in the manner described in Section 3.2.1.

Most dimensions of the micromachined structure are not very critical. The only geometrical parameters that are of real interest are the width and length of the heater as well as the spacing between the resistors and the material specimen. The specimen itself was 275 µm × 100 µm. As heater, a 10 µm-wide p-type polySiGe strip was used with a length of 100 µm. As the effective length in between the two contacts is 88 µm, the heater has a resistance in the order of 250 Ω Similarly, the resistance of the resistor used to measure T_c is in the order of 450 Ω As the (R, T)-characteristic of the p-type polySiGe resistors is practically linear over the temperature range from -25 °C to 125 °C, the resistance value can be easily converted to a temperature by using

$$\Delta T=\frac{1}{\beta(T)\,R(T_a)}\Delta R(T) \tag{3.18}$$

where $R(T_a)$ is the resistance measured at the ambient temperature and $\beta(T)$ is the TCR. When a significant current is run through the heating resistor, the corresponding resistance change $\Delta R(T)$ can be converted directly to a change in temperature.

As the temperature gradients are largest near the heater, it is critical to minimise the distance between the heater and the material specimen (See section 3.2.4 for FEA



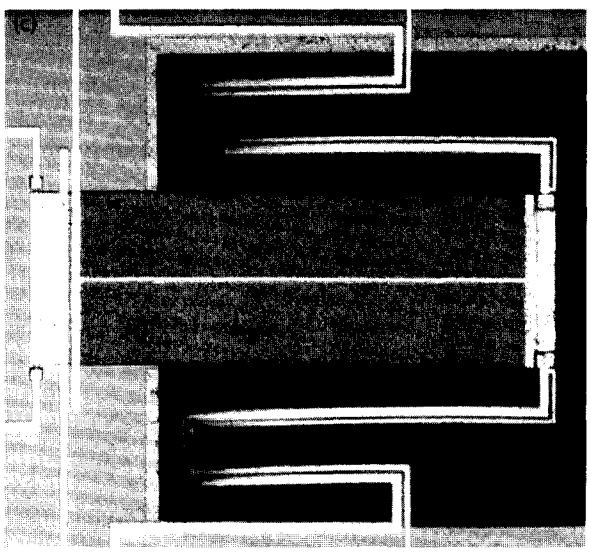


Figure 3.7 (a) Top view and (b) cross-sectional view of a micromachined structure for measurement of the Seebeck coefficient. (c) Photograph of an earlier-version measurement structure (the stress from the interconnect layer is clearly visible, causing the arms to bend out of focus of the camera.)

simulations on similar structures). A spacing of 3.5 µm was chosen, to ensure that no short-circuit between the heater and the material specimen occurs. Furthermore, to minimise the thermal gradient between the heater and material specimen, the Al interconnect was laid across the heater as well as a part of the material specimen. This is clearly visible in Figure 3.7b.

Table 3.7 list the measurement results. The values are in good agreement with the data reported in [3.30].

Table 3.7 Average Seebeck coefficient for poly$Si_{0.7}Ge_{0.3}$ from -25 °C to 125 °C.

Material	Seebeck coefficient [µV°C^{-1}]	Temperature coefficient [µV°C^{-2}]
Al/p-poly$Si_{0.7}Ge_{0.3}$	175	0.2
Al/n-poly$Si_{0.7}Ge_{0.3}$	-129	-0.1

3.2.4 Thermal conductivity

Whether performed on a discrete sample or an on-chip integrated sample, of the three parameters that make up the figure-of-merit the thermal conductivity is the hardest parameter to determine. The reason for this can be derived from the statement by Lasance [3.31]: *"In many cases of practical importance, there is no such thing as purely one-dimensional heat conduction in the thermal world, in contrast to the electrical world."* Where the electrical conductivity stretches across 20 orders of magnitude the thermal conductivity only varies about 3 orders of magnitude at most. For this reason, regions of high thermal conductivity cannot be separated as good from regions of low thermal conductivity as is possible in the electrical domain. The resulting increase in thermal losses makes the test structure behaviour difficult to predict and control. This reduces the accuracy with which the measurement can be performed.

3.2.4.1 Conventional measurement methods

Thermal conductivity measurement techniques can be divided into static and dynamic measuments. The advantages and disadvantages of either category is outlined below.

Static method

The static method most frequently applied is referred to as the longitudinal steadstate method [3.32]. This technique should be considered the thermal equivalent of the bridge structure used for resistivity measurements (see Section 3.2.1.), i.e.,

$$q = \lambda \frac{A_{th}}{L_{th}} \Delta T \quad \Leftrightarrow \quad I = \rho^{-1} \frac{A_{el}}{L_{el}} \Delta V \tag{3.19}$$

where the subscripts '*th*' and '*el*' are used to distinguish between the thermal and electrical domains, and A and L refer to the cross-sectional area and length of the thermal / electrical resistor. InFigure 3.8 the similarity between the structures is expressed graphically.

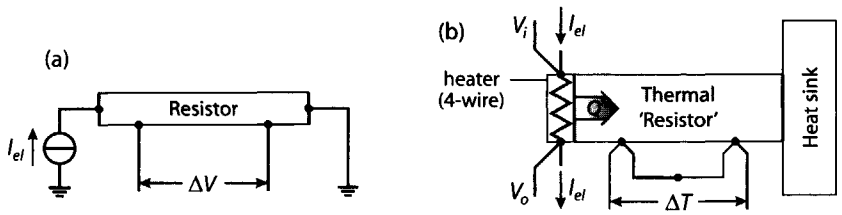


Figure 3.8 (a) The electrical 4-terminal resistor versus the (b) thermal 4-terminal resistor.

For the static method to be successful, two major requirements have to be fulfilled. First, the heater and heat sink have to be in intimate contact with the specimen (in Figure 3.8 referred to as the thermal resistor.) If an intimate contact cannot be established, the thermal contact should at least be well-defined, so that it may be corrected for.

The second requirement is that thermal losses to the ambient should be reduced to an absolute minimum. This is particularly critical for specimen with a low thermal conductivity, where the thermal losses through the electrical leads become significant. Compared to the electrical resistor that requires only four leads, the thermal resistor (plus heater) requires at least six. Moreover, convective and conductive losses through air need to be circumvented. These can be reduced to negligible values by performing the measurement in a good vacuum. Radiant heat exchange can't be eliminated. Still, the influence can be significantly limited by keeping the temperature difference between the material under test and the objects in its field of view small (typically less than 10 K.) Under these circumstances, only the thermally conductive flow through the specimen and the leads connected needs to be considered. When all of the requirements are met, the thermal conductivity follows from rearranging (3.19):

$$\lambda = \frac{q}{\Delta T}\frac{L_{th}}{A_{th}} \tag{3.20}$$

Nolas [3.32] argues that, for bulk samples where λ is not too low, it is possible to keep the measurement error within 1%, even when a large ΔT is applied. For very low values of the thermal conductivity (< 2 $Wm^{-1}K^{-1}$), the measurement error increases, but can still be kept below 3%. To limit the errors for such low thermal conductivities, it is essential to limit the length of the specimen.

Dynamic methods

Because of the lengthy time period to reach equilibrium as well as the high number of parameters that need to be measured and controlled, dynamic methods are often preferred. In general, with a dynamic measurement method, the thermal conductivity is derived from the thermal diffusivity δ_{th}

$$\delta_{th} = \frac{\lambda}{c_v} \tag{3.21}$$

Three thermal actuating principles are distinguished [3.33]. The laser flash methods uses pulsed actuation [3.32], while both the hot strip and hot wire methods use as step-wise power change [3.34][3.35] and the 3ω method uses sinusoidal actuation [3.36]. In my opinion the main drawback of dynamic methods is that the specific heat of the material under investigation generally is not accurately known and needs to be determined before the thermal conductivity can be calculated.

Of these dynamic measurement methods only the 3ω method has been proven to yield accurate results on thin-film samples. Therefore, the 3ω technique is the only transient measurement method discussed in more detail here. The name stems from the fundamental parameter measured, which is a voltage oscillation at three times the frequency of the electrical current used for heating the sample.

In essence, the measurement structure is no more than the bridge-type 4-terminal resistor presented in Section3.2.1. The two outer leads are used to drive an AC current through the resistor, while the inner two leads measure the voltage across the resistor.

Assume the electrical current applied is defined as $I_{el} = I_0 \sin(\omega)$, then the Joule heat generated in this resistor equals

$$q_J = I_{el}^2 R = \left(I_0^2 \sin^2(\omega)\right) R = \frac{1}{2} I_0^2 R \left(1 - \cos(2\omega)\right) \tag{3.22}$$

This equation clearly shows that the heat generated alternates at a frequency 2ω The thermal variation of the heater $\Delta T = z_{th} q_J$, where the term z_{th} resembles the thermal impedance of the connection between the heater and ambient. Therefore, the temperature of the heater will also vary at a frequency 2ω:

$$\Delta T = z_{th} q_J = C_1 \left(1 - \cos(2\omega)\right) \tag{3.23}$$

with

$$C_1 = \frac{z_{th} I_0^2 R}{2} \tag{3.24}$$

When the heater has a non-zero TCR, say C_2, the resistance of the heater will vary along with the thermal fluctuations, also at a frequency 2ω:

$$R = R_0(1 + C_2\Delta T) = R_0\left(1 + C_2C_1(1 - \cos(2\omega))\right) \tag{3.25}$$

This equation ignores the non-linear heat generation due to the TCR of the resistor. However, this will generally yield an error of less than 1%, which is more than ample to illustrate the 3ω technique. In the final step, the voltage measured across the two inner leads, V (= $I_{el}R$), equals.

$$\begin{aligned} V &= I_0 \sin(\omega) R_0 \left(1 + C_1C_2 (1 - \cos(2\omega))\right) \\ &= I_0R_0\left(\left(1 + \frac{3C_1C_2}{2}\right)\sin(\omega) - \frac{C_1C_2}{2}\sin(3\omega)\right) = V_\omega + V_{3\omega} \end{aligned} \tag{3.26}$$

where V_ω and $V_{3\omega}$ are the respective voltage components at frequencies ω, and 3ω

The thermal conductivity λ is hidden in the parameter C_2, so it may be determined from both the ω the 3ω components. However, as $C_1C_2 \ll 1$, measurement of the ω component is inaccurate. In amplitude the 3ω component is much smaller than the ω component (Cahill indicates a factor 1000 at least [3.36].) However, as it has no offset component, it is more accurate and can be isolated from the other frequency components, by detecting it with a lock-in amplifier. By measuring the 3ω voltage at two (or more) frequencies, the thermal conductivity is calculated from

$$\lambda = \frac{C_2}{4\pi LR_0^2}\frac{V_\omega^3 \ln(f_2/f_1)}{(V_{3\omega,1} - V_{3\omega,2})} \tag{3.27}$$

with L is the length of the heater.

3.2.4.2 Thin-film measurement methods

The methods described above are intended to measure bulk samples only, under well controlled circumstances. Direct application of these methods on thin film samples is not a very suitable way to measure the thermal conductivity:

1. The thermal load of the actuating and measuring entities is much larger in thin film structures. In particular for low-λ samples, the heat flux through electrical leads away to ambient will outweigh the heat flux through the specimen under test several times.
2. The design rules of most thin-film devices do not allow specimen, heaters and sensors to be in intimate contact. Mostly these are separated by one or more dielectric layers.
3. Unlike with the conventional discrete measurement methods, it is very difficult to establish a controlled environment in which every measurement parameter is accurately known. For example, it is practically impossible to know the exact geometrical profile of each layer used in the test structure.

Still, four thin-film thermal conductivity measurement techniques are encountered in literature. These are outlined below.

Modified 3ω technique

The 3ω technique is the only transient bulk measurement method that is directly applicable to thin-film structures. Illustrated clearly in [3.37], the thermal conductivity of a thin film on top of a thick substrate can ben accurately determined, even with a thermal conductivity below 1 $Wm^{-1}K^{-1}$.

The thermal conductivity of a thin film on top of a thick substrate cannot be determined directly using the 3ω technique. Instead a differential measurement is performed. The first step is determination of the thermal conductivity and specific heat of the substrate. Next, the second measurement is performed on an identical substrate that holds the thin-film. Use is made of the fact that the total amplitude of the thermal oscillation, ΔT_{tot}, is the sum of the amplitude in the thin film, ΔT_{tf} and the amplitude in the substrate, ΔT_{sub}. As the thin-film thickness H_{tf} is generally much less than the width of the heater W_h. Thus, ΔT_{tf} is practically constant for low frequencies (< 10 kHz). This allows ΔT_{tf} to be determined directly, from the relationship $\Delta T_{tf} = \Delta T - \Delta T_{sub}$. Once ΔT_{tf} is determined, the thermal conductivity of the thin film, λ_{tf}, is calculated using

$$\lambda_{tf} = \frac{P}{\Delta T_{tf}} \frac{H_{tf}}{W_h} \tag{3.28}$$

where P is the amplitude of the power generated by the heater.

Thermal Van der Pauw structure

Recently an entirely new measurement structure was developed, the thermal Van der Pauw structure [3.38]. As the name already suggests, this structure is the thermal equivalent of the Van der Pauw structure discussed in Section 3.2.1. Unlike the 3ω method, this structure does allow the thermal conductivity of micromachined membranes to be determined directly.

The structures looks similar to the Greek cross geometry shown in Figure 3.3. On each of the four legs a single heater/temperature monitor is placed. Instead of an current flow, the heaters are used to establish a thermal energy flow along two adjacent legs (e.g. legs 1 and 2 in Figure 3.3.) The other two legs are used to measure the resulting temperature difference (e.g. legs 3 and 4 in Figure 3.3.)

Even though good results have been obtained using the thermal Van der Pauw structure, it has a considerable disadvantage in the sense that four individual heaters/temperature monitors have to be operated simultaneously: In the electrical Van der Pauw structure, the voltage probes are entirely passive; Without any bias applied, no electrical current will flow away through the voltage probes. In contrast, with the thermal structure, at least three of four legs need to be heated (read: biased) at all times. Each leg that is not heated will remain at the temperature of the substrate.

To ensure efficient operation, the structure needs to be thermallyisolated from the ambient. This is achieved by micromachining, similar to the fabrication of the thin film TECs. In equilibrium, each heater generates an amount of power q, so that a temperature difference ΔT above ambient temperatureis established. The power dissipation in the temperature monitors in legs 3 and 4 is kept constant throughout the measurement. Now, if the power q_1 in leg 1 is increased, the power q_2 in leg 2 needs to be decreased accordingly, to fulfil the requirement that the total power generated within the structure remains constant. The resulting temperature difference across legs 3 and 4, ΔT_{34}, at a given power level q_{12} (= q_1–q_2) is a direct measure for the thermal sheet resistance / resistivity:

$$R_{sh,th} = \frac{\pi}{\ln 2} \frac{\Delta T_{34}}{q_{12}} \tag{3.29}$$

Membrane-based lateral measurement methods

The third thin-film measurement technique, has the simplest geometry of any structure used to measure the thermal conductivity of thin films. Only a membrane is required, on which a single large heater is placed in the middle and one or more temperature monitors are placed at predetermined positions on the membrane.

In literature, two variants of this concept are found. Zhang *et al.* [3.33] position a four-wire resistor in parallel with the heater, which acts as an absolute temperature monitor. Instead, Irace and Sarro [3.39] used thermocouples to measure the temperature at various points on the membrane, in relation to the substrate temperature. Both use

$$\lambda = \frac{q'}{\Delta T_{sa}} (L_m - x) \tag{3.30}$$

to calculate the thermal conductivity. Here, q' is the heat flux from the heater towards ambient. Along the direction of the flux, L_m is the distance from the centre of the membrane (where the heater is located) to the edge of the membrane, x is the distance between the heater and the temperature monitor and ΔT_{sa} is the temperature difference between the sensor and ambient.

In both [3.33] and [3.39] it was shown that the heater can be driven using either a dc current, or using a transient electrical current, either in ac mode or pulsed mode.

Cantilever-based lateral steady state method

The fourth thin-film measurement technique closest resembles the longitudinal steady-state method discussed in Section 3.2.4.1. A membrane is micromachined so that a beam-type or cantilever-type structure results. This thermal insulation enhances the control and predictability of the lateral heat flow. Not coincidentally, due to the well-defined shape (and resulting ease of calculation), this technique is en-

countered very often in literature, e.g., [3.40][3.41] and has also been applied to obtain the results presented in this Chapter.

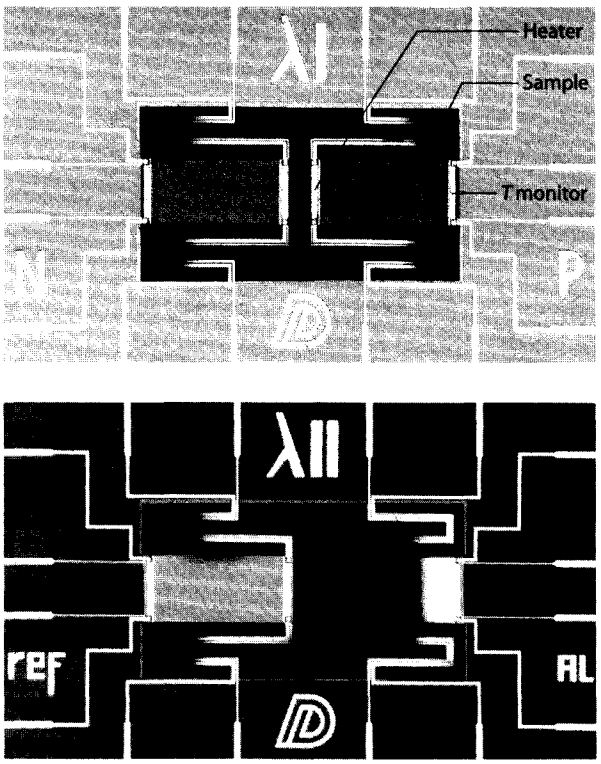


Figure 3.9 Cantilever type structures to measure the thermal conductivities of (a) the n- and p-type polySiGe and (b) the reference Si_3N_4/SiO_2 membrane and the Al interconnect.

To help explain the basic operation of this technique, some fabricated devices are shown in Figure 3.9. The reference membrane consists of a 300 nm Si_3N_4 and a 300 nm SiO_2 stack. This is the cantilever structure marked 'ref' in Figure 3.9b. The 600 nm polySiGe specimen is located between these two layers, either doped or undoped. Two structures, respectively with n- and p-type doping are shown in Figure 3.9b. Finally, the interconnect layer (600 nm $Al_{0.99}Si_{0.01}$) can be deposited on top of the Si_3N_4/SiO_2 membrane, which is the right structure in Figure 3.9b.

Like with the Seebeck measurement structure, a p-type polySiGe heater is placed at the tip of the cantilever that simultaneously serves as temperature monitor. To simultaneously monitor the temperature of the heater and the thermal energy q generated within, a 4-terminal configuration is used. A second resistive temperature monitor is placed on the beam close to the substrate.

From the three parameters, the two heater temperatures T_h, T_c and the power q, as well as all geometrical parameters, the total thermal conductance K_{tot} of the membrane plus specimen follows from

$$K_{tot} = \sum_{m=1}^{n} K_m = \frac{q}{T_h - T_c} \tag{3.31}$$

with K_m the thermal conductance of the m^{th} layer in the stack:

$$K_m = \lambda_m \frac{A_m}{L_m} \tag{3.32}$$

The parameters A_m and L_m are the respective cross-sectional area and length of the specimen, seen in the direction of thermal flow.

For mechanical reasons, it is generally impossible to construct a beam or cantilever from only one of these layers, (as such a submicron layer is too fragile to process.) Instead, a differential measurement method needs to be applied in which the results from different stack configurations are subtracted, according to:

$$K_m = \sum_{n=1}^{z} K_n - \left(\sum_{n=1}^{m-1} K_n \Big|_{m\geq 1} + \sum_{n=m+1}^{z} K_n \Big|_{m\leq z} \right) \tag{3.33}$$

Once K_m has been determined, λ_m can be determined from (3.32).

The thermal conductivity can only be determined accurately if all thermal conductances $K_1 \ldots K_z$ are determined accurately first. This requires the geometry of each thin film component (e.g., resistors, interconnect, membranes, specimen, etc.) to be accurately known. Moreover, no energy losses may occur due to convection and conduction of the surrounding air. This implies that, like with the other thermal conductivity measurement methods, the specimen must be placed in a vacuum. As stated earlier, radiant heat exchange cannot be avoided. Keeping the temperature difference between the heater and the ambient below 10 K, the error due to radiant heat exchange is kept well below 1%, which is far better than the total inaccuracy in the measurement.

3.2.4.3 Measurement setup

At the start of our research on lateral thin-film TECs, no suitable vacuum chamber was available that allowed the measurement to be performed in. The requirements for this chamber were:

- A large number of electrical feedthroughs (20+) to the vacuum chamber are required to allow all electrical components to be connected simultaneously.
- The vacuum clock and pump should have a small footprint and be portable.

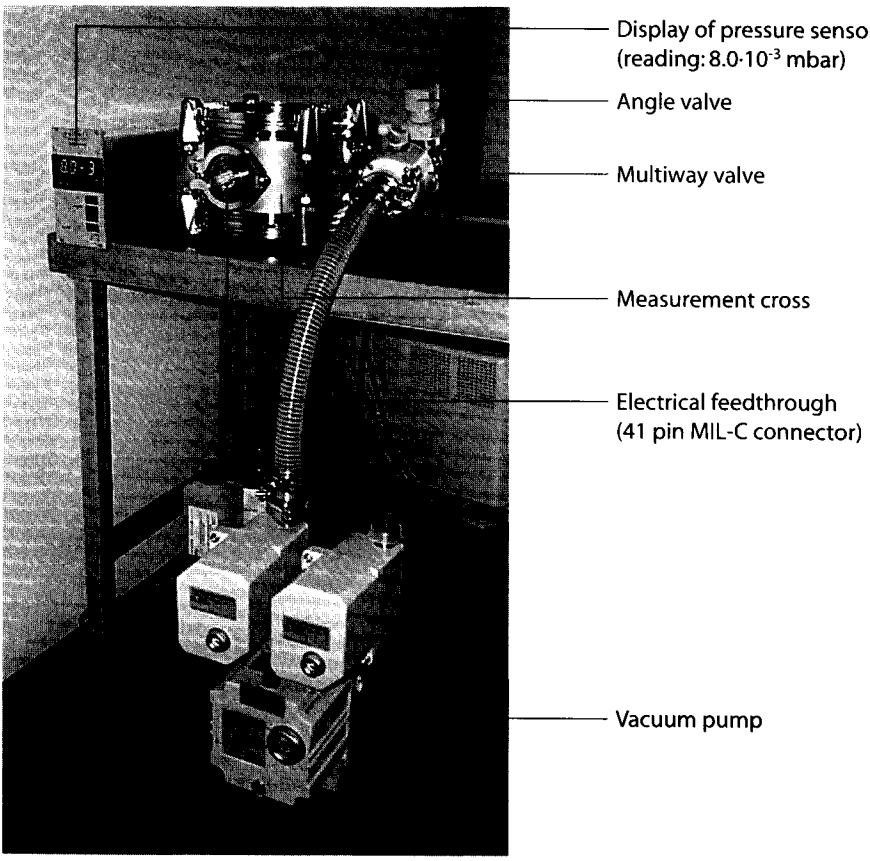


Figure 3.10 Photograph of the vacuum setup for measurement of the thermal conductivity.

- For the thermal losses through air to be negligible, the end pressure in the vacuum chamber should be < 1 Pa [3.41][3.42]. The actual end pressure observed was slightly higher than 1 Pa. From [3.43] the losses are estimated to be less than 1% still.
- The temperature of the substrate must be adjustable to any temperature in between ambient temperature (outside the vacuum chamber) and 125 °C.
- The thermostat inside the vacuum chamber should be able in intimate contact with the package.

Based on the above requirements, the setup shown in Figure 3.10 was assembled. As vacuum chamber a DN 160 ISO-K measurement cross was selected. To seal the top and bottom of this chamber stainless steel DN 160 ISO-K blanking flanges are mounted. For visual inspection during testing, a DN 160 ISO-K borosilicate glass viewport can be mounted instead.

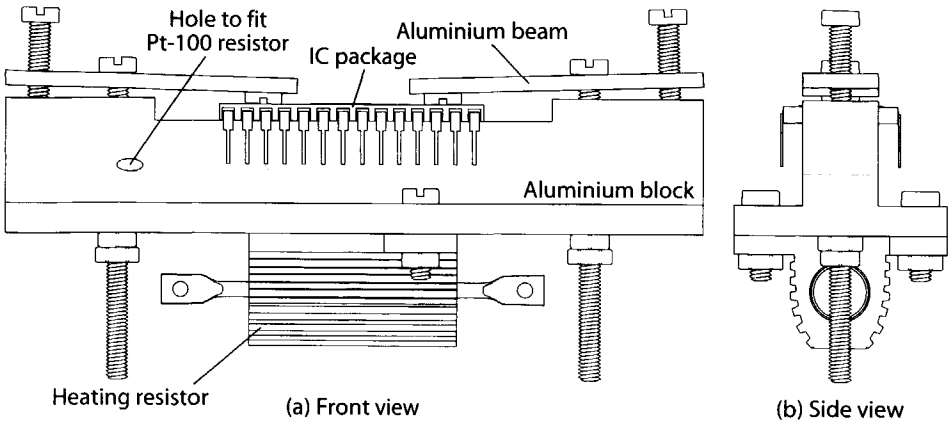
The measurement cross has three lateral connection flanges, with respective diameters of 16, 25 and 40 mm (DN 16, DN 25 and DN40 ISO-KF). The DN 40 flange is used for electrical feedthroughs by means of a 41 pin MIL-C type flange. To the DN 25 flange, a Pfeiffer Vacuum PKR 250/251 FullRange™ gauge is connected. This cold cathode/Pirani gauge is capable of measuring pressures from $5 \cdot 10^{-9}$ to $1 \cdot 10^{-9}$ mbar. By connecting a Balzers VVA016H multiway valve to the DN 16 flange of the measurement cross, a selection between air intake and air outlet can be made. On one of the two flanges of the multiway valve an angle valve is placed to control the air intake speed (to prevent a pressure shock in the chamber.) To the other valve flange the vacuum pump is connected. The vacuum pump is a Pfeiffer Vacuum Duo 2.5 two-stage rotary vane pump. This pump has a theoretical end pressure of less than 0.6 Pa ($6 \cdot 10^{-3}$ mbar). In practise, values < 1 Pa are rarely obtained, depending on the thoroughness of the cleaning procedure before assembly as well as the length of vacuum tubes between the pump and the chamber.

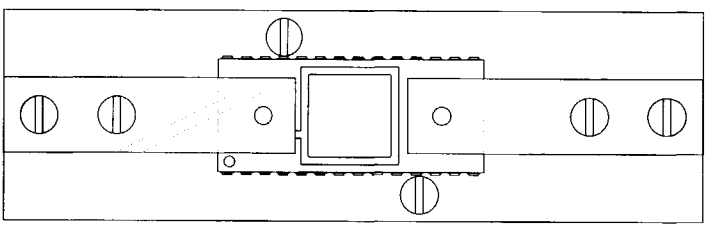
To control the temperature of the chip inside the vacuum chamber, a simple thermostat was developed that could hold DIL20...DIL40 packages. This structure is shown in Figure 3.11. The schematic drawings are to scale. For clarity, Figure 3.11d displays an artist impression of this simple unit.

The chip package is clamped onto a inverted T-shaped aluminium block in which a recess is made. (The recess is large enough to hold a DIL40 package.) Clamping is obtained by two levers that are hinged by a nylon screw in the middle of the lever. By driving the outer nylon screws further into the lever, pressure on the package is increased. The nylon screws act as a force limiter for the IC package as the screw thread will be torn away before the pressure buildup could crack the package. To further reduce the risk of damage to the package as well as to provide a well-defined pressure point, nylon screwheads on the bottom side of the lever tip are used to transfer the pressure from the lever to the package.

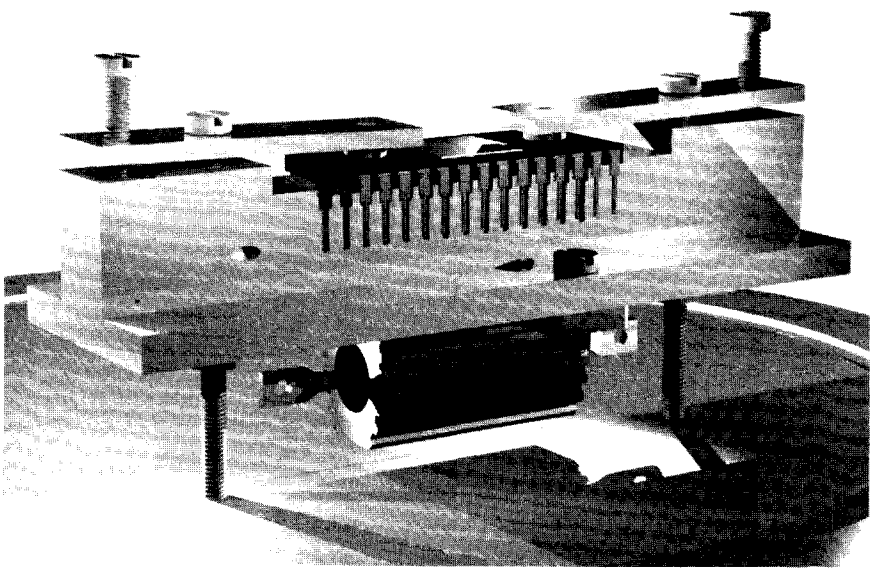
To make the aluminium block into a thermostat, the temperature must be measured and the heat generation controlled. To measure the temperature, a hole is drilled into the aluminium block to accommodate a Pt-100 resistor. The resistor type is Bauart M-K 20 15, which has a length of 25 mm and a diameter of 1.5 mm. To heat the aluminium block, a 25W wirewound resistor (Welwyn WH25) was fixed to the bottom of the aluminium block. The aluminium block including the heater is suspended from two metallic screws that are fixed into the stainless steel blanking flange. These screws form the only connection between the thermostat and the flange. A power of only a few Watt is already sufficient to heat the aluminium block up several tens of degrees when a vacuum is applied.

Active cooling of the thermostat was also considered. A Peltier cooler between the heater and the vacuum flange is an option, but simple calculations indicated too much of the power from the heater would directly flow away through the Peltier cooler into the flange. This would cause a serious reduction in the maximum temperature that can be obtained. Therefore, it was decided to use passive cooling.





(c) Top view



(d) Artist's impression

Figure 3.11 (a)–(c) Schematic drawings to scale and (d) an artist impression of the thermostat used to control the temperature of a chip package between ambient temperature and 125 °C.

Table 3.8 Materials values used in the FEA simulation of the λ test structures.

Material	λ [$Wm^{-1}K^{-1}$]	ρ [μΩm]	H [nm]
Si_3N_4	1.55	–	300
SiO_2	1.17	–	300
p-polySiGe	4.8	$3.08 \cdot 10^{-2}$	600
n-polySiGe	4.45	–	600
Al	216.5	13.2	600

3.2.4.4 FEA inspection

To assess 1) the influence of the specimen on the temperature profile in the cantilever/beam structure and 2) the thermal flux concentration near the heater, FEA simulations were performed. The material parameters used in the FEA are listed in Table 3.8 Typical simulation results are shown in Figure 3.12. Due to the high width:thickness ratio, only the lateral (two-dimensional) heat flow needs to be analysed. To ensure this 2-D approach is valid, the potential heat flux concentrations around the heater still need to be considered. Unlike the structure reported in [3.41] a small strip of the Si_3N_4/SiO_2 stack is located in between the heater and the specimen, which forces the flow through the specimen to be all but two-dimensional around the heater. Moreover, far less layers are stacked than described in [3.41]. Therefore, the assumption that only the lateral heat flow needs to be considered is justified.

Two parameters have to be defined for the FEA simulation. First, the average thermal conductivity across all layers of the regions in Figure 3.12a need to be specified. This is done according to the thermal conductivity λ_m and thickness H_m of each individual layer. The average thermal conductivity is calculated from

$$\lambda_{avg} = \frac{\sum_{m=1}^{n} \lambda_m H_m}{\sum_{m=1}^{n} H_m} \tag{3.34}$$

This values has to be corrected by a term C to account for the relative thicknesses of the stacks, i.e.,

$$\lambda'_{avg} = C \cdot \lambda_{avg} \tag{3.35}$$

In the simulations, five layers stacks are encountered (labelled in Figure 3.12):

- Region '1': Si_3N_4/SiO_2 stack: $\lambda_{avg} = 1.36\ Wm^{-1}K^{-1}$, $C = 1$.
- Region '2': Si_3N_4/polySiGe/SiO_2 stack: $\lambda_{avg} = 3.08\ Wm^{-1}K^{-1}$, $C = 2$.
- Region '3': Si_3N_4/p-polySiGe/Al stack: $\lambda_{avg} = 88.83\ Wm^{-1}K^{-1}$, $C = 2.5$.

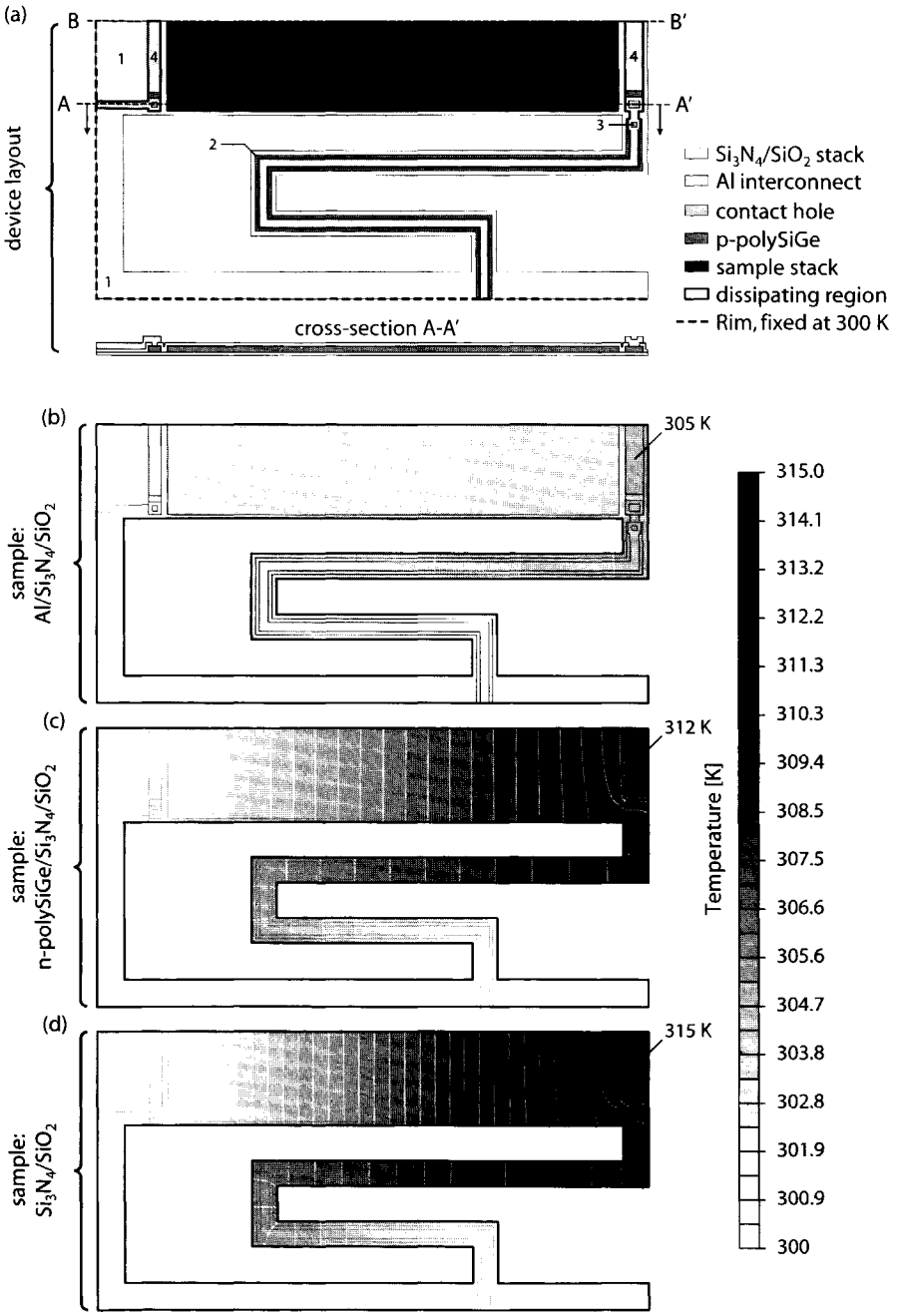


Figure 3.12 (a) Overview and FEA simulation results of (b) an Al specimen, (c) a polySiGe specimen and (d) the reference structure without any specimen.

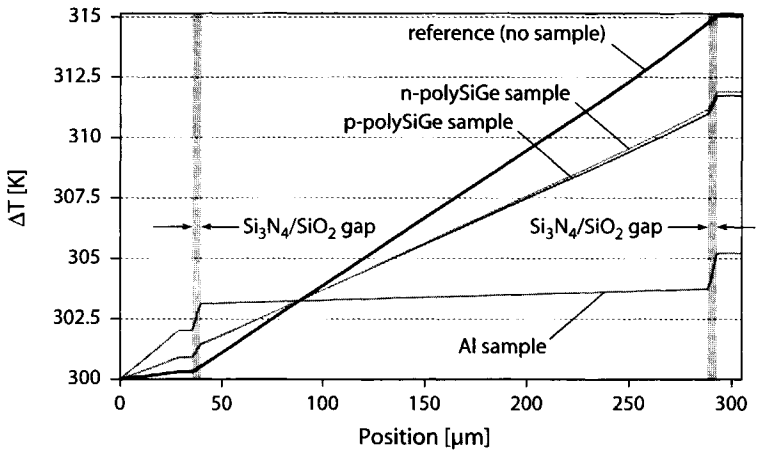


Figure 3.13 Temperature profile along the path B-B' in Figure 3.12, for the reference structure as well as different material samples.

- Region '4': Si_3N_4/p-polySiGe/SiO_2/Al stack: $\lambda_{avg} = 74.21$ $Wm^{-1}K^{-1}$, $C = 3$.
- Region '5': Si_3N_4/SiO_2 stack plus the material specimen.

In the regions indicated with a thicker border in Figure 3.12, Joule heat is generated. (Electrical power is required for heating and resistor read-out.) For these regions, the electrical sheet resistance together with the electrical current is used to calculate the heat flux in these regions. In the aluminium wires running to the heater at the tip of the cantilever, the heat flux is $2.06 \cdot 10^3$ Wm^{-2}. In the p-polySiGe heater this is $4.40 \cdot 10^4$ Wm^{-2}. Finally, the thick dashed line in Figure 3.12a indicates the rim of the device which is implemented as a heat sink, at a constant temperature of 300 K.

The first observation from the FEA simulation results shown in Figure 3.12b-Figure 3.12d, is that an Al layer over the heater indeed acts as a very effective heat spreader. This significantly improves the temperature homogeneity of the heaters, which improves the accuracy with which the average temperature of the heater can be calculated. The effect is clearest in the FEA simulation of the reference structure, of Figure 3.12d. the maximum temperature difference across the heater is less than 0.5 K, equivalent to less than 3% of the total temperature gradient in the measurement structure. This Al strip was implemented based on the findings of Paul *et al.* [3.41]. Unlike them, I have chosen not to extend the Al strip over the specimen to be characterised. The idea of such an overlap is to make the temperature of the specimen equal to the heater temperature. I have chosen for a well defined Si_3N_4/SiO_2 membrane in between the heater and the specimen, and to correct for this thermal resistance afterwards.

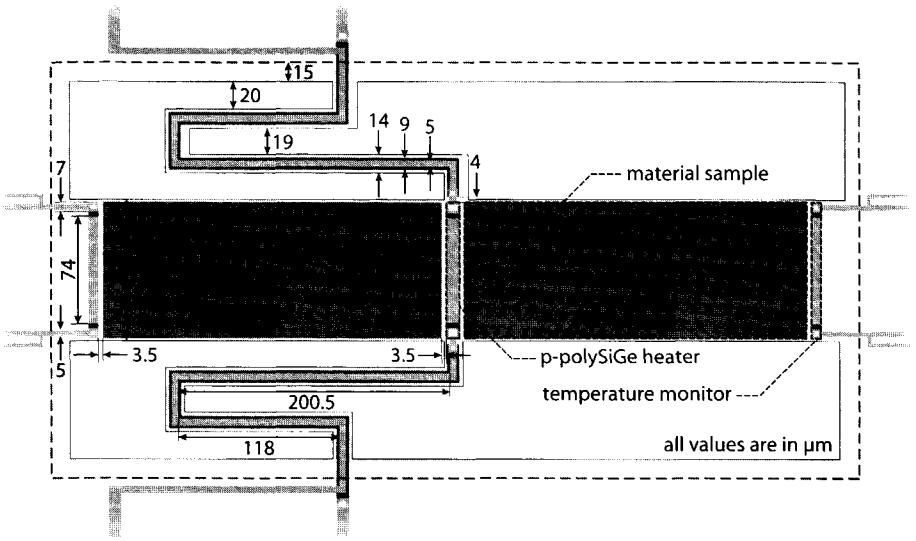


Figure 3.14 Layout of the thin-film structure used to measure the thermal conductivity of a material. All values are dimensions in μm.

A second observation is made from the equithermal line density in Figure 3.12. In the reference structure, the lines are spaced practically equal. As a material sample is added and the thermal conductivity of the sample increases, the line density builds up in the Si_3N_4/SiO_2 membrane 'gaps' between the heaters / temperature monitors and the sample. These resulting temperature drop across these 'gaps' is shown more clearly in Figure 3.13. The curves in this figure represent the temperature profile across the central line of the structure, line B-B' of Figure 3.12a, for each of the configuration of different sample materials. In the case the Al interconnect is used as specimen, the temperature drop across the dielectric 'gap' is nearly 50% of the total thermal gradient across the measurement structure. This clearly indicates the steady state measurement method should be used cautiously on specimen with a high thermal conductivity, as the slightest deviation in the geometry of such a specimen (due to an overetch, for example) has a very pronounced effect on the temperature gradient measured.

3.2.4.5 Measurement results

Determination of the thermal conductivity of a thin-film polySiGe sample located in between the nitride/oxide membrane consists of two steps:

1. Determination of the Q–ΔT relationships for both the structure without a sample (i.e. the reference structure) and the structure holding the polySiGe sample.

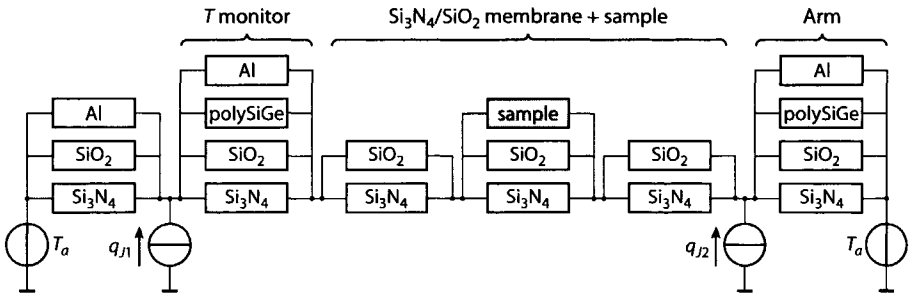


Figure 3.15 Equivalent thermal network of the structure in Figure 3.14.

2. Determination of λ from the ratios between the difference between Q–ΔT relationships (incorporating the difference in construction between the different structures.)

Unlike the cantilever structure used in the simulations, a beam structure was used for the actual measurements, for two reasons. First, a beam structure provides improved mechanical strength, so the yield is improved (especially after wafer dicing.) Second, as a beam allows two material samples to be placed on each side of the heater, the parasitic thermal conduction through the arms leading to/from the heater is effectively reduced by a factor of two. The layout of this beam-type structure is shown in Figure 3.14, together with the most important dimensions.

Based on the FEA simulation results the dimensions of the structure provided in Figure 3.14 were converted into an equivalent thermal network, which is used to calculate the thermal conductivity of the sample from measurement data. For symmetry reasons only half of the structure of Figure 3.14 needs to be modelled. The resulting equivalent thermal network is shown in Figure 3.15.

The majority of thermal energy is generated in the heater, modelled as a thermal energy source q_{J1}. From there the thermal energy flows back to the substrate through the arms and the beam. As the arms are sufficiently long and consists of a uniform Si_3N_4/p-polySiGe/SiO_2/Al stack, the thermal resistance of each individual layer can be modelled by a single resistor, in a parallel configuration. For the beam, the configuration is slightly more complex. a small strip of the Si_3N_4/SiO_2 is present between the sample under investigation (shown in grey in Figure 3.15) and both the heater and temperature monitor. Finally, the temperature monitor that dissipates a small amount of power q_{J2} is located in between the beam and the rim of the device.

The value of each thermal resistor is determined from

$$R_{th} = K_{th}^{-1} = \frac{N}{\lambda \cdot H} \tag{3.36}$$

where H is the height of the layer and N is the number of squares (in the direction of the 1-D flow.) For clarity, the thermal network is split into two parts, the beam (consisting of all thermal resistors to the left of q_{J2}) and the arm (consisting of the resistors to the right of q_{J2}). Filling in the geometrical parameters from Figure 3.14 gives

$$R_{\text{th,beam}} = \frac{2.70 \cdot 10^5}{2.15\lambda_{\text{Al}} + 2.31\lambda_{\text{p-SiGe}} + 1.2\lambda_{\text{SiN}} + \lambda_{\text{SiO}}} + \frac{7.71 \cdot 10^6}{\lambda_{\text{Al}} + 13.2\lambda_{\text{SiN}} + 11.0\lambda_{\text{SiO}}} + \frac{9.88 \cdot 10^6}{1.20\lambda_{\text{SiN}} + \lambda_{\text{SiO}}} \tag{3.37}$$

and

$$R_{\text{th,arm}} = \frac{1 \cdot 10^9}{7.11\lambda_{\text{Al}} + 12.2\lambda_{\text{p-SiGe}} + 10.4\lambda_{\text{SiN}} + 8.66\lambda_{\text{SiO}}} \tag{3.38}$$

The subscripts 'SiN', 'SiO', 'p-SiGe' and 'Al' are used to distinguish the thermal conductivities of the respective layers. Already discounted in the above expressions are the thicknesses of the four layers (H_{SiN} = 300 nm, H_{SiO} = 250 nm, $H_{\text{p-SiGe}}$ = 600 nm and H_{Al} = 600 nm.) With a sample in place, Equation (3.37) changes to

$$R'_{\text{th,beam}} = \frac{2.70 \cdot 10^5}{2.15\lambda_{\text{Al}} + 2.31\lambda_{\text{p-SiGe}} + 1.2\lambda_{\text{SiN}} + \lambda_{\text{SiO}}} + \frac{7.71 \cdot 10^6}{\lambda_{\text{Al}} + 13.2\lambda_{\text{SiN}} + 11.0\lambda_{\text{SiO}}} + \frac{2.70 \cdot 10^5}{1.20\lambda_{\text{SiN}} + \lambda_{\text{SiO}}} + \frac{6.25 \cdot 10^8}{2.50 \cdot 10^8 H_{\text{Sample}}\lambda_{\text{Sample}} + 3.12\lambda_{\text{SiN}} + 2.60\lambda_{\text{SiO}}} \tag{3.39}$$

where only the third terms is modified compared to Equation (3.37) and a fourth term is added, which incorporates the thermal conductivity of the specimen.

First, the q–ΔT relationship of the reference structure is determined from the q–R measurements. Next, the q–ΔT relationship of the reference plus sample is determined. The resulting curves for the reference structure as well as the structure with a p-polySiGe sample are shown in Figure 3.16. The experiments have been performed at different temperatures, in the range from 25 °C to 85 °C.

As Equations (3.37) and (3.38) reveal, the thermal resistance of the beam and arm do not just depend on the thermal conductivities of the Si_3N_4 and SiO_2 layers, but also on the $Al_{0.99}Si_{0.01}$ and p-type $polySi_{0.7}Ge_{0.3}$. As such, to derive the thermal conductivity of the p-type $polySi_{0.7}Ge_{0.3}$ has to be derived first. This can only be done by assigning estimate values to the Si_3N_4, SiO_2 and $Al_{0.99}Si_{0.01}$. Table 3.9 lists the values that are used in the calculation.

From the difference in slope of the Q-ΔT curves the thermal conductivity is determined. (See Figure 3.16b.) An automated Mathematica® script was created to do the calculation. Table 3.10 lists the values obtained from the script. According to this

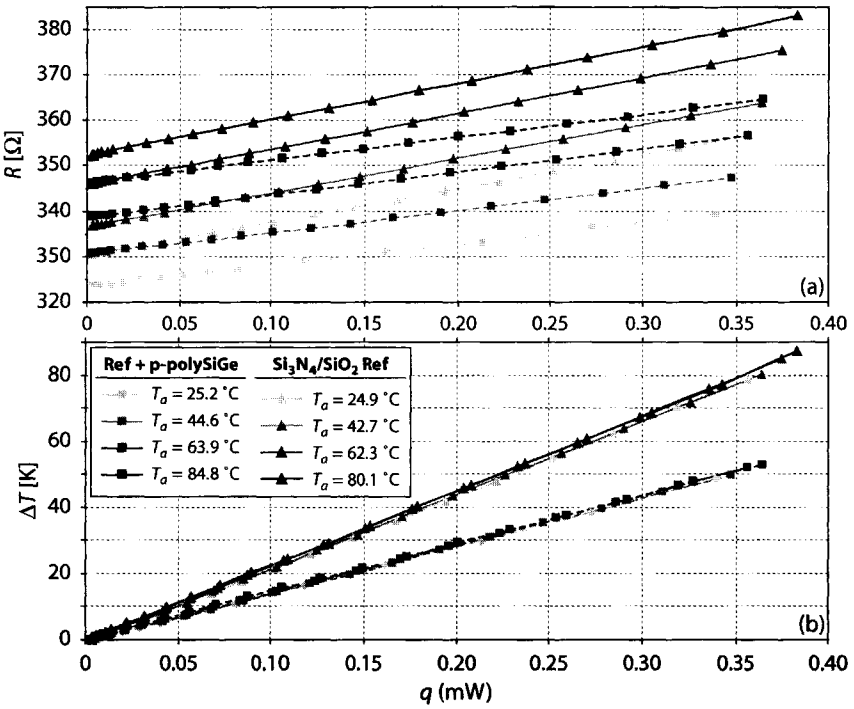


Figure 3.16 (a) Ppower input–resistance output curves as measured from the heater located in the middle of the beam of Figure 3.14. (b) When the R-ΔT relationship of the resistor has been measured, the Q-ΔT curves can be plotted

Table 3.9 Material values used to calculate λ of polySiGe.

Material	λ [$Wm^{-1}K^{-1}$]	Reference
Si_3N_4	1.55	[3.39]
SiO_2	1.17	[3.39]
$Al_{0.99}Si_{0.01}$	180	[3.41]

Table 3.10 PolySiGe thermal conductivity measurement results.

Material	Energy [keV]	Dose [atom, cm^{-2}]	λ [$Wm^{-1}K^{-1}$]
$polySi_{0.7}Ge_{0.3}$	undoped		5.0 (±20%)
$polySi_{0.7}Ge_{0.3}$	40	B, $5.0 \cdot 10^{15}$	4.7 (±20%)
$polySi_{0.7}Ge_{0.3}$	40	P, $7.5 \cdot 10^{15}$	5.1 (±20%)

table the thermal conductivity of p-type polySiGe is lower than that of undoped polySiGe, which should not be the case. This has to be attributed to the measurement errors made in determination of the values. Sources of error are:

- Power dissipation in the arms of the structure. The power density in the Al wires that carry the electrical current to the heater is 20 times less than the power density in the heater itself. Like with the Peltier element, of all power dissipated in the arm, only half contributes to a rise temperature. As such, the error by power dissipation in the arms is around 2%–3%.
- The non-uniformity of the temperature profile in the heater. Figure 3.12 clearly shows the aluminium strip over the heater very effectively reduces thermal gradients in the heater. As a result, the temperature drop over the arms is practically equal to the temperature drop over the beam. Both the arms and beam are connected to the rim of the heater, where the temperature of the heater is lowest. As such, the thermal non-uniformity across the heater will causes an average ΔT to be read that is slightly higher than the actual ΔT from the edge of the heater to ambient. The thermal gradient across the heater is most noticeable when an aluminium sample is in place. In this situation, the average heater temperature will be 5%–10% higher than the temperature at the edge of the heater. When a polySiGe heater is in place, this error reduces by roughly 50% to 3%–5%.
- Flow of thermal energy close to the heater can't be modelled entirely accurate in a 1-D thermal network or a 2-D FEA. It is hard to predict what the error of these approximations is. For a similar structure, Paul *et al.* [3.41] reported that the error between the measured and actual thermal conductivity value (for polySi) is about 5%. Our structure uses less and thinner layers so the error is not expected to be larger than the 5% reported in [3.41].
- Thickness variations in the various layers (in particular across structures.) Unlike the other errors, this is not a systematic error. The thickness of any given layer will vary from wafer to wafer. Even across a single wafer, significant thickness variations can occur. To give an indication of the kind of error thickness variations may cause, consider the aluminium layer. If the thickness is reduced from 600 nm to 550 nm, the calculated value changes by 8%. It it hard to predict how much the actual variation is, but a corresponding error of 5% (= 30 nm on 600 nm) may not be excluded; Even though the values for Si_3N_4 and SiO_2 come from measurements performed on layers processed at DIMES.
- The values in Table 3.9 are only estimates. For example, the values for the thermal conductivity of silicon nitride at room temperature range from 1 $Wm^{-1}K^{-1}$ to 16 $Wm^{-1}K^{-1}$ [3.37][3.39][3.44][3.45]. Similarly, values for the thermal conductivity of $Al_{0.99}Si_{0.01}$ are known to range from 180 $Wm^{-1}K^{-1}$ to

236 $Wm^{-1}K^{-1}$. These errors will most likely be even larger than the error due to thickness variations.

In conclusion, an error margin of 20% in the measurement must be applied.

As Figure 3.16 shows, the measurements have been performed at various temperatures in the range of 25 °C to 85 °C. No significant change in the thermal conductivity could be observed in this temperature range. This result is in accordance with [3.46], for example.

3.3 REFERENCES

[3.1] Y.H. Shing, Y. Chang, A. Mirshafii, L. Hayashi, S.S. Roberts, J.Y. Josefowicz and N. Tran, "Sputtered Bi_2Te_3 and PbTe thin films," *J. Vac. Sci. Technol. A*, Vol. 1, No. 2, pp. 503–506, 1983.

[3.2] C. Shafai and M.J. Brett, "Optimization of Bi_2Te_3 thin films for microintegrated Peltier heat pumps," *J. Vac. Sci. Technol. A*, Vol. 15, No. 5, pp. 2798–2801, 1997.

[3.3] C. Shafai, M.J. Brett and J.N. Broughton, "Passivation and performance analysis of thin film Peltier heat pumps," *J. Vac. Sci. Technol. A*, Vol. 17, No. 1, pp. 305–309, 1999.

[3.4] H. Zou, D.M. Rowe and G. Min "Growth of p- and n-type bismuth telluride thin films by co-evaporation," *Journal of Crystal Growth*, Vol. 222, pp. 82–87, 2001.

[3.5] H. Zou, D.M. Rowe and G. Min, "Preparation and characterization of p-type Sb_2Te_3 and n-type Bi_2Te_3 thin films grown by co-evaporation," *J. Vac. Sci. Technol. A*, Vol. 19, No. 3, pp. 899–903, 2001.

[3.6] F. Völklein, V. Baier, U. Dillner and E. Kessler, "Transport properties of flash-evaporated $(Bi_{1-x}Sb_x)_2Te_3$ films I: optimization of film properties," *Thin Solid Films*, Vol. 187, pp. 253–162, 1990.

[3.7] A. Foucaran, A. Sackda, A. Giani, F. Pascal-Delannoy and A. Boyer, "Flash evaporated layers of $(Bi_2Te_3–Bi_2Se_3)$(N) and $(Bi_2Te_3–Sb_2Te_3)$ (P)," *Materials Science and Engineering*, Vol. B52, pp. 154–161, 1998.

[3.8] A. Mzerd, F. Tcheliebou, A. Sackda and A. Boyer, "Improvement of thermal sensors based on Bi_2Te_3, Sb_2Te_3 and $Bi_{0.1}Sb_{1.9}Te_3$," *Sensors and Actuators*, Vol. A46–47, pp. 387–390, 1995.

[3.9] A. Boulouz, A. Giani, F. Pascal-Delannoy, A. Foucaran and A. Boyer, "Bi_2Te_3 thin films grown by MOCVD," *Proc. 16th International Conference on Thermoelectrics*, 26–29 August 1997, Dresden, Germany, pp. 167–170.

[3.10] ,A. Giani, F. Pascal-Delannoy, A. Boyer, A. Foucaran, M. Gschwind and P. Ancey, "Elaboration of Bi_2Te_3 by metal organic chemical vapor deposition," *Thin Solid Films*, Vol. 303, pp. 1–3, 1997.

[3.11] A. Boulouz, A. Giani, F. Pascal-Delannoy, M. Boulouz, A. Foucaran and A. Boyer, "Preparation and characterization of MOCVD bismuth telluride thin films," *Journal of Crystal Growth*, Vol. 194, pp. 336–341, 1998.

[3.12] A. Giani, A. Boulouz, F. Pascal-Delannoy, A. Foucaran and A. Boyer, "MOCVD growth of Bi_2Te_3 layers using diethyltellurium as a precursor," *Thin Solid Films*, Vol. 315, pp. 99–103, 1998.

[3.13] O. Paul, M. von Arx and H. Baltes, "Process-dependent thermophysical properties of CMOS IC thin films," *Proc. Transducers '95 / Eurosensors IX*, 25–29 June 1995, Stockholm, Sweden, pp. 178–181.

[3.14] K. Baert, P. Mertens and P. van Gerwen, "Thin film poly-$Si_{70\%}Ge_{30\%}$ for thermopiles," *Proc. Micro System Technologies '98*, 1–3 December 1998, Potsdam, Germany, pp. 655–658.

[3.15] P. van Gerwen, T. Slater, J.B. Chévrier, K. Baert and R. Mertens, "Thin-film boron-doped polycrystalline $silicon_{70\%}$–$germanium_{30\%}$ for thermopiles," *Sensors and Actuators*, Vol. A53, pp. 325–329, 1996.

[3.16] S. Sedky, P. Fiorini, M. Caymax, S. Loreti, K. Baert, L. Hermans and R. Mertens, "Structural and mechanical properties of polycrystalline silicon germanium for micromachining applications," *J. of Microelectromechanical systems*, Vol. 7, No. 4, pp. 365–372, 1998.

[3.17] S. Sedky, P. Fiorini, K. Baert, L. Hermans and R. Mertens, "Characterization and optimization of infrared poly SiGe bolometers," *IEEE Trans. on Electronc Devices*, Vol. 46, No. 4, p. 675–682, 1999.

[3.18] Strasser, M., R. Aigner, M. Franosch, and G. Wachutka, "Miniaturized thermoelectric generators based on poly-Si and poly-SiGe surface micromachining," *Sensors and Actuators*, Vol. A 97–98, pp. 535–542, 2002.

[3.19] K .Shiota, D. Inoue, K. Minami and J. Hanna, "Structural and electrical properties of poly-SiGe thin films prepared by reactive thermal CVD," *J. Non-Crystalline Solids*, Vol. 227–230, pp. 1074–1078, 1998.

[3.20] E.V. Jelenkovic, K.Y. Tong, W.Y. Cheung, I.H. Wilson, S.P. Wong and M.C. Poon, "Low temperature doping of poly-SiGe films with boron by co-sputtering," *Thin Solid Films*, Vol. 368, pp. 55–60, 2000.

[3.21] D.K. Schroder, *Semiconductor material and device characterization*, second edition, John Wiley & Sons, New York, NY, USA, 1998, ISBN: 0-471-24139-3.

[3.22] L.J. van der Pauw, "A method of measuring specific resistivity and Hall effect of discs of arbitrary shape," *Philips Research Report*, Vol. 13, pp. 1–9, 1958.

[3.23] A.B. Glaser and G.E. Subak-Sharpe, *Integrated circuit engineering*, Addison-Wesley, Reading, MA, USA, 1977.

[3.24] D.A. Neamen, *Semiconductor physics and devices: basic principles*, Irwin, Homewood, IL, USA, 1992, ISBN: 0-256-08405-X.

[3.25] H.H. Berger, "Models for contacts to planar devices," *Solid-State Electronics*, Vol. 15, pp. 145–158, 1972.

[3.26] J.G.J. Chern and W.G. Oldham, "Determining specific contact resistivity from contact end resistance measurements," *IEEE Electron Device Letters*, Vol. 5, No. 5, pp. 178–180, 1984.

[3.27] S.J. Proctor, L.W. Linholm and J.A. Mazer, "Direct measurement of interfacial contact resistance, end contact resistance, and interfacial contact layer uniformity," *IEEE Trans. on Electron Devices*, Vol. 30, No. 11, pp. 1535–1542, 1983.

[3.28] W.M. Loh, K. Saraswat and R.W. Dutton, "Analysis and scaling of Kelvin resistors for extraction of specific contat resistivity," *IEEE Electron Device Letters*, Vol. 6, No. 3, pp. 105–108, 1985.

[3.29] M. von Arx, O. Paul and H. Baltes, "Test structures to measure the Seebeck coefficient of CMOS IC polysilicon," *IEEE Trans. on Semiconductor Manufacturing*, Vol. 10, No. 2, pp. 201–208, 1997.

[3.30] F. Edelman, M. Stölzer, T. Raz, Y. Komem, C.B. Vining, H. Zeindl and P. Zaumseil, "Structure and transport properties of microcrystalline SiGe films," *Proc. 6th International Conference on Thermoelectrics*, 26–29 August 1997, Dresden, Germany, pp. 232–235.

[3.31] C. Lasance: "Thermal resistance: an oxymoron?" *Electronics Cooling Magazine*, Vol. 3, No. 2, pp. 20–24, 1997.

[3.32] G.S. Nolas, I. Sharp and H.J. Goldsmid, *Thermoelectrics: basic principles and new material developments*, Springer-Verlag, Heidelberg, Germany, 2001, ISBN: 3-540-41245-X.

[3.33] X. Zhang and C.P. Grigoropoulos, "Thermal conductivity and diffusivity of free-standing silicon nitride thin films," *Rev. Sci. Instrum.*, Vol. 66, No. 2, pp. 1115–1120, 1995.

[3.34] S.E. Gustafsson and E. Karawacki, "Transient hot-strip probe for measuring thermal properties of insulating solids and liquids," *Rev. Sci. Instrum.*, Vol. 54, No. 6, pp. 744–747, 1983.

[3.35] P. Andersson and G. Bäckström, "Thermal conductivity of solids under pressure by the transient hot wire method," *Rev. Sci. Instrum.*, Vol. 47, No. 2, pp. 205–209, 1976.

[3.36] D.G. Cahill, "Thermal conductivity measurement from 30 to 750 K: the 3ω method," *Rev. Sci. Instrum.*, Vol. 61, No. 2, pp. 802–808, 1990.

[3.37] S.-M. Lee and D.G. Cahill, "Heat transport in thin dielectric films," *J. Appl. Phys.*, Vol. 81, No. 6, pp. 2590–2595, 1997.

[3.38] O. Paul, P. Ruther, L. Plattner and H. Baltes, "A thermal Van der Pauw test structure," *IEEE Trans. on Semiconductor Manufacturing*, Vol. 13, No. 2, pp. 159–168, 2000.

[3.39] A. Irace and P.M. Sarro, "Measurement fo thermal conductivity and diffusivity of single and multilayer membranes," *Sensor and Actuators*, Vol. A76, pp. 323–328, 1999.

[3.40] F. Völklein and H. Baltes, "A microstructure for measurement of thermal conductivity of polysilicon thin films," *J. Microelectromechanical Systems*, Vol. 1, pp. 193–196, 1993.

[3.41] O.M. Paul, J. Korvink and H. Baltes, "Determination of the thermal conductivity of CMOS IC polysilicon," *Sensors and Actuators*, Vol. A41–42, pp. 161–164, 1994.

[3.42] H. Baltes and D. Moser, "CMOS vacuum sensors and other applications of CMOS thermopiles," *Proc. Transducers '93*, 7–10 June 1993, Yokohama, Japan, pp. 736–741.

[3.43] A.W. van Herwaarden, *Thermal vacuum sensors based on integrated silicon thermopiles*, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands.

[3.44] P. Eriksson, J.Y. Andersson and G. Stemme, "Thermal characterization of surface-micromachined silicon nitride membranes for thermal infrared detectors," *J. Microelectromechanical Systems*, Vol. 6, No. 1, pp. 55–61, 1997.

[3.45] S.H. Kong, *Infrared micro-spectrometer based on a multi-slit grating*, Ph.D. Thesis, Eburon Publishers, Delft, The Netherlands, 2002, ISBN 90-5166925-9.

[3.46] S.-M. Lee and D.G. Cahill, "Thermal conductivity of Si–Ge superlattices," *Appl. Phys. Lett.*, Vol. 70, No. 22, 1997, pp. 2957–2959.

Chapter 4

DEVICE FABRICATION

While Chapters 2 and 3 focus on rigorous characterisation of lateral thin-film thermoelectric elements, this chapter is focuses on the practical implementation of such devices. In Section 4.1, the prior conditions for fabrication are established, in particular regarding fabrication compatibility. Sections 4.2 and 4.3 respectively provide a high-level classification of fabrication technologies and an overview of relevant fabrication technologies. Then, in Section 4.4 the actual TEC fabrication process is discussed, while Section 4.5 addresses the fabrication difficulties and potential improvements to the current process. Finally, Section 4.6 presents a versatile technique for the deposition of thin metal films, which is particularly useful for the fabrication of temperature-independent resistive heaters.

4.1 FABRICATION-COMPATIBLE PROCESSING

Integrated silicon transducer technology has enabled both high-volume batch fabrication of integrated monolithic transducers as well as low-volume to medium-volume fabrication of microelectromechanical systems (MEMS). The combination of conventional IC technologies with this new class of integrated devices is by no means trivial, as was already discussed in Chapter 1. In the following subsections, the major consideration, regarding material/process invasiveness and co-integration are discussed.

4.1.1 The need for non-standard materials

For analogue electronics, the susceptibility of silicon to both temperature and light is hardly described as a virtue. Nevertheless these susceptibilities, which are clearly be categorised as parasitic to normal IC operation, have been the driving force behind, the recent generations of solid-state temperature sensors [4.1][4.2] and CMOS imagers [4.3][4.4], for example. Combined with the availability of the advanced photolithographic capabilities from the IC industry, the structural dimensions of the devices—which largely determine the sensing characteristics—can be controlled accurately and reproducibly.

As always, there is a catch: silicon isn't always the best sensing material, and the materials that are superior in performance often can't be co-integrated with electronics. This generally leads to the approach of a system-in-a-package (SiP). Still, the more complex system-on-a-chip (SoC / co-integration) approach was chosen for the fabrication of the lateral on-chip TECs. This choice is elaborated below.

4.1.2 Invasiveness of a transducer processing step

What has to be determined before the development of a transducer fabrication process is to what extent the required materials and processing steps conflict with the IC processing, in case co-integration is pursued. (The concept of co-integration was introduced in Chapter 1.) That is, the level of invasiveness of the transducer fabrication process needs to be determined. I distinguish four categories:

1. *Non-invasive transducer processing.* This applies to all materials and fabrication steps that result directly from a well-established conventional IC process. As all materials and processes are standard, these are fabrication compatible by definition. The simplest examples of this category are the optical, thermal and magnetic sensors. This class of sensors can be made 'smart' easily, and is a logical candidate for on-chip co-integration.
2. *Minimally invasive transducer processing.* Minimally invasive materials and processing steps are those that are directly derived from standard IC fabrication processes, but require modifications to the standard process. Examples are selective epitaxial growth, oxide thinning and non-standard implantations. Devices using these technologies are, among others, phototransistors [4.4][4.5] magnetotransistors [4.6] and optical interference filters [4.7][4.8]. This category of transducer processing is a good candidate for co-integration, although additional and/or deviant masks and fabrication steps are required.
3. *Moderately invasive transducer processing.* Moderately invasive are those materials and processing steps that go beyond conventional IC processing, but can be implemented with the infrastructure that is already in place (i.e. the implementation is in the margin of what is possible.) Polycrystalline silicon germanium belongs to this class of materials. It can be processes instead of polySiGe, and is allowed inside the cleanroom, but does contaminate the epitaxial reactor (non-destructively). Though not compromising material compatibility, this category of integration often interferes with the thermal budget of the process. This requires careful planning of all processing steps required. This level of invasiveness requires careful consideration of whether or not to pursue on-chip co-integration.
4. *Highly invasive transducer processing.* If processing and contamination are such that the standard IC process will be significantly degraded by the transducer fabrication process, on-chip co-integration should be avoided

when possible. Nonetheless, many of the advanced transducer materials and processes are poorly compatible, often requiring special etchants, deposition techniques or sacrificial layers. Most micromachining techniques belong to this category. Even though KOH and TMAH are widely used bulk etchants, they attack aluminium [4.9][4.10]. More seriously, KOH can easily (and destructively) contaminate a cleanroom by means of the K^+ ions. Unless special post-processing techniques and equipment are available, highly invasive materials and processes cannot be used for on-chip co-integration.

4.1.3 The choice for co-integration

From the classification in the subsection above and the process flow in Section 4.4 it becomes obvious that the thermoelectric cooler presented in this thesis requires moderately invasive materials (i.e. polySiGe) as well as highly invasive processing techniques (high-temperature anneal, KOH etch, RIE etch.) It has therefore been given considerable thought whether or not to pursue on-chip co-integration. The motivation to do so is provided below.

Motivation for co-integration

In chapter 1 the concepts of hybrid integration, semi-hybrid integration and co-integration were introduced. Compared to hybrid or semi-hybrid integration (i.e. the system-in-a-package approach), on-chip co-integration (i.e. the system-on-a-chip approach) is more complex, has potential yield issues and optimization of both transducer and electronics remains difficult. Moreover, steady progress is made in hybrid integration regarding to total size of the multi-chip solution. It is routinely possible to both dice a wafer no further than 2 μm from the active area, and to position individual dies as little as 5 μm apart. Even though individual components need to be wire-bonded, the total device area remains in the same order of magnitude as can be achieved with co-integration (if the latter is at all possible.)

Even though development of a thermoelectric cooler suitable for co-integration is more complex than a a thermoelectric cooler suitable for hybrid integration, it was still decided to pursue on-chip co-integration. The main reason is the applications targeted: An integrated thermoelectric cooler should be capable of thermally stabilising on-chip electronic circuits like bipolar or CMOS references and input stages *within* a chip. In such a case co-integration is obligatory.

Motivation for back-end processing

In Chapter 1 it was already pointed out that for a device to be suitable for co-integration, it needs to be material compatible and process compatible (see Section 1.3.3.) In many cases, such 'rules' can be bent. Platinum, for example, is not to be considered material compatible as it adheres insufficiently to oxide surfaces [4.11]. Titanium oxide does adhere very well to SiO_2 as well as Pt [4.12]. Therefore, if TiO_2 is deposited directly before Pt a Pt-100 resistors can be integrate on-chip.

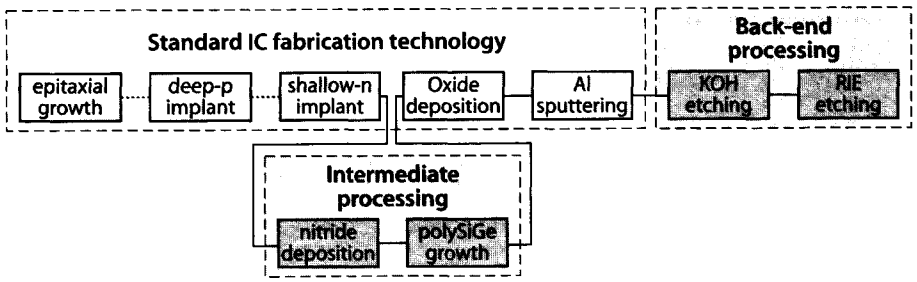


Figure 4.1 The schematic fabrications sequence of a lateral on-chip integrated TEC.

The largest challenge for co-integration is the modularity of the design. As explained in Chapter 1, modularity is essential if fabrication compatibility is to be achieved: 1) Fabrication of transducer and electronics remain separated in time and space. 2) In turn, the separate processing steps ensure the microelectronics processing infrastructure already in place may be employed. This minimises the need to invest in new processing equipment. 3) It prevents the need to redesign the entire process flow when a small step has been changes as with a fully customised process.

Looking back at Figure 1.8, three modular design approaches were distinguished. Frond-end processing, where the transducer is processed before the electronics, intermediate processing, where the transducer is processed in between electronics processing steps, and back-end processing, where the transducer is processed after the electronics are finished.

Front-end transducer processing is only restrictedly applicable. If pre-processed wafers are to be accepted in a microelectronic processing facility, these have to be identical to prime wafers in terms of thickness, planarity and contaminants present. Although this is possible, pre-processed transducers are hardly ever reported. For obvious reasons, intermediate processing should also be avoided whenever possible, as this quickly leads to a fully customised process. One major exceptions exists, which is fabrication of a transducer directly in advance of the final metallisation. Many foundries have equipment outside their class-100 and class-1000 facilities, for metallisation of 'dirty' wafers.

Back-end processing is preferred, as this saves the most dirty and intrusive processing steps until after the standard IC fabrication process. In Figure 4.1 the (simplified) process flow schematic for a lateral on-chip integrated TEC is shown. Contrary to the previous remarks, intermediate transducer processing steps are required. However, both nitride deposition (using low pressure chemical vapour deposition—LPCVD) and epitaxial growth of polySiGe (on a polySi seed layer) are fabrication compatible. For this reason it is allowed to return the wafer into the IC line after these deposition steps. The 'dirty' processing steps, KOH etching in particular, are done at the end, outside the class-100 and class-1000 cleanrooms.

4.2 CLASSIFICATION OF TRANSDUCER FABRICATION TECHNOLOGIES

Three categories of transducer fabrication technologies can be distinguished, bulk micromachining, surface micromachining and thin-film technology (although the latter technique is never recognised as such.) The basic characteristics of the three categories are outlined below.

Bulk micromachining

Put simply, bulk micromachining (BMM) is used to remove well-defined parts of the silicon substrate, through etch windows on either the front- or backside of the wafer. This technology is discussed in more detail in Section 4.3.3, as this is an essential step in the fabrication of integrated thermoelectric coolers.

Surface micromachining

Surface micromachining is applied to create freestanding structures using layers near the surface of the wafer. Instead of removing the entire substrate as with bulk micromachining, free-standing structures are created by selective removal of thin sacrificial layers. Such a layer is patterned first, after which a structural layer is deposited and patterned to partially overlap the sacrificial structures. The sacrificial layer is completely removed, leaving a freestanding structure. This process is shown schematically in Figure 4.2. For a nice introduction into surface micromachining techniques, the reader is referred to [4.13].

A common problem in surface micromachining is the large ratio between the lateral and the vertical dimensions, which results in 1) long underetch times, hampered by the need for high directional selectivity and 2) sticking of freestanding structures after the sacrificial etch [4.14].

For a fabrication compatible surface micromachining process, use should preferably be made of the materials available in the microelectronic fabrication process. As standard materials and processing steps are designed for maximum microelectronic performance these are not necessarily suitable for the construction of micromechanical structures. Nevertheless, this option has the interesting economic advantage of using the microelectronic industry as subcontractor for realising on-chip microsystems. A wide range of potential sacrificial layers are available like aluminium, polysilicon, polymers and oxides. An nice implementation of a fabrication compatible strategy is presented in [4.15].

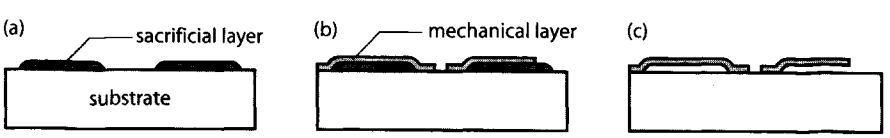


Figure 4.2 Basic surface micromachining process. (a) sacrificial layer deposition and patterning. (b) Mechanical layer deposition and pattering. (c) Sacrificial layer removal.

Thin-film technology

Thin-film technology is not categorised as a transducer fabrication technology as such. This technology does not necessarily have to be an enabling technology (to enable a transduction effect) but is often used as an enhancing technology. In the latter case, think of 1) polySiGe replacing polySi to obtain improved thermoelectric performance, 2) deposition of ferromagnetic materials for magnetic flux concentration or 3) special metallic films for zero-TCR heaters. Without these modifications and additions, the devices would still work, but with far inferior performance.

TIn most cases, when the material itself is altered, material compatibility—and, thus, fabrication compatibility—is lost. In this sense, a material like polySiGe is a positive exception, as it is deposited using the same technique and equipment as used for polySi, with the modification of some of the precursors.

Still, in most other cases, the thin film material is not compatible with the IC fabrication process, so a special (back-end) fabrication module has to be developed. A good example presented at the end of this chapter, when the deposition of thin metal films is addressed.

4.3 RELEVANT FABRICATION TECHNOLOGIES

In the previous section, the general transducer fabrication technologies were outlined. Not all technologies are required for the fabrication of the lateral on-chip integrated TECs. In this section only the technologies relevant to the fabrication of the integrated TECs are introduced.

4.3.1 Chemical vapour deposition and epitaxial growth

Sputtering and chemical vapour deposition (CVD) are some of the most frequently applied thin-film deposition techniques. Sputtering is used mostly for metal films, whereas CVD is used for dielectric and polycrystalline (silicon) thin films. Silicon nitride, polySiGe and TEOS are required for the fabrication of the lateral thin-film TEC and are all deposited using LPCVD (low pressure chemical vapour deposition.)

Chemical vapour deposition is the formation of solids through decomposition of gaseous precursors. The CVD process consists of a number steps. After decomposition of the source gases, the reactants diffuse to the surface of the wafer. Next, the reactants completely decompose and either start nucleation or are integrated in an existing crystal matrix. Finally, the reaction products are desorbed from the surface and diffuse away to be removed from the reactor.

Generally, an energy source (such as heat or a plasma) is required for the decomposition. The process of pyrolitic breakdown of molecules (between 800 °C and 1200 °C) is generally referred to as thermal chemical vapour deposition (TCVD). Growth rates up to 5 μm/min can be obtained. Alternatively, hot wire chemical vapour deposition (HWCVD) can be used, where a heated (W or Ta) filament initiates a catalytic rather than pyrolitic decomposition. At around 0.3 μm/min, the deposition rates of HWCVD are lower than for TCVD. A third CVD technique is

plasma spraying, where a powder is lead across a plasma torch to be deposited on a substrate. This technique allows deposition rates of up to 10 μm/min. The high deposition rate is offset by the high level of impurities encountered in plasma sprayed thin films, either form the torch of from an impure powder [4.16].

Compared to evaporated or sputtered films, CVD films can be deposited more uniformly, generally with a better step coverage. Depending on the temperature and deposition rate, either a monocrystalline, polycrystalline or even an amorphous film is deposited. As a rule of thumb, lower temperatures (around 650 °C and below) will produce polycrystalline films (if slowly deposited) or amorphous films (if rapidly deposited.) Monocrystalline films can only be deposited at a low rate, with temperatures of 1000 °C and above [4.17]. To recrystallise a film or to repair a damaged monocrystalline crystal, or to convert amorphous films into polycrystalline films, a high-temperature anneal is required.

Plasma enhanced CVD (PECVD) reactions are commonly used during the fabrication of microsensors, where gases are decomposed and ionised by means of plasma. PECVD has the advantage of a higher deposition rate yet a lower deposition temperature than LPCVD films. This makes PECVD more suitable for co-integration of sensors and electronics. Unfortunately, due to the large amount of hydrogen incorporated in the deposited films [4.18] PECVD films are generally inferior to LPCVD layers.

Finally, epitaxial growth (or epitaxy for short) is a special CVD process where the crystal structure of the thin film exactly matches that of the underlying (seed) layer. This technique is often applied for surface micromachining of silicon transducers [4.13][4.17]. For example, with epitaxy it is possible to grow films with a totally different doping concentration than the substrate resulting in much sharper junctions than obtainable through common doping techniques [4.17].

CVD of silicon nitride

Two deposition methods are used to deposit silicon nitride, LPCVD and PECVD. PECVD has the potential to deposit nearly stress-free nitride films. Unfortunately due to the high concentrations of hydrogen, PECVD nitride is rather porous, making it vulnerable to etchants. In contrast, LPCVD nitride is highly resistant to chemical attack.

Typical deposition temperatures and pressures respectively range from 700 °C to 900°C and from 200 mTorr to 500 mTorr. LPCVD nitride films deposited between these temperatures are amorphous. As precursor gas, SiH_2Cl_2 is preferred over SiH_4 as the first provides a better thickness uniformity. Because nitride deposition is reaction-limited rather than diffusion limited, films on both sides of a wafer are of equal thickness.

The residual stress in stoichiometric Si_3N_4 is tensile and quite large. To reduce the residual stress, non-stoichiometric silicon nitride films can be LPCVD deposited. These low-stress films (also referred to as Si-rich films), are deposited with an excess

of Si, by decreasing the NH_3:SiH_2Cl_2 precursor ratio. For a 1:6 ratio (at a temperature of 850 °C and a pressure of 500 mTorr) the as-deposited films are practically stress free [4.19].

CVD of silicon dioxide through TEOS (tetraethylorthosilicate)

The largest difference between TEOS and other precursors for the formation of most other oxides is that in TEOS the silicon atom is already oxidized (i.e. the silicon atom is directly surrounded by four oxygen atoms.) As a result, TEOS is converted into silicon dioxide by adsorption rather than by means of an oxidising reaction. Due to the native presence of oxygen atoms in the TEOS molecule this process can take place in a inert atmosphere, although addition of oxygen increases the deposition rate.

For the development of lateral TECs, TEOS is also preferred over other oxides for a second reason: TEOS diffuses well into trenches and deposits well on steep slopes, providing excellent step coverage characteristics. Moreover, when deposited at temperatures over 650 °C, the structure is densified during deposition, improving stability and allowing a thinner film to be deposited. Both characteristics (better step coverage and densification) are essential to minimise the thermal losses through the Si_3N_4/poly$Si_{0.7}Ge_{0.3}$/SiO_2 stack. A layer as thin as 100 nm is sufficient.
The main restriction of TEOS is the high processing temperature required. For this reason TEOS deposition has to take place before metal deposition. After deposition, TEOS films will generally exhibit a modest compressive stress. More information on the deposition of TEOS can be found in [4.20]–[4.22], for example.

4.3.2 Dopant introduction

In general, there are three ways through which dopant atoms can be introduced in a material. In case a constant doping concentration throughout the entire thickness of the material is required, *in situ* doping is preferred, where the dopant atoms are introduced during deposition of a material. The drawback is that if higher doping concentrations are required (~10^{20} cm^{-3}), the deposition becomes increasingly slow. This limits the practical thickness of an *in-situ* doped thin film.

Diffusion

For reasonable dopant introduction speeds, the other two techniques, diffusion and implantation are employed. Diffusion is typically performed by placing a patterned wafer in a furnace and flowing an inert gas carrying the dopant atoms across the wafer surface. Dopants travel into the wafer by means of either vacancy diffusion or interstitial diffusion. As a result, the dopant concentration will decrease monotonously from the wafer surface. In case of a constant surface doping concentration C_s, the dopant concentration at position x is given by

$$C(x,t) = C_s \mathrm{erfc}\left(\frac{x}{2\sqrt{Dt}}\right) \tag{4.1}$$

where $x = 0$ corresponds to the wafer surface, erfc is the complementary error function and $(Dt)^{1/2}$ is the diffusion length with

$$D = D_0 C_s \exp\left(\frac{-E_a}{k_B T}\right) \tag{4.2}$$

In this equation, D_0 [cm^2s^{-1}] is the diffusion coefficient, E_a [eV] is the activation energy and C_s is the surface dopant concentration.

Implantation

In the lateral TECs discussed in this thesis, only the third dopant introduction technique, implantation, has been applied. With this technique the dopants are driven into the surface by means of a high-energy ion beam. Unlike with diffusion, the doping concentration will not peak at the wafer surface, but inside the wafer. The advantage over diffusion is the improved control over the amount of dopants and the improved reproducibility of the doping profile.

To estimate the implantation profile, two parameters are of importance, the implantation energy [keV] and dose S [cm^{-2}]. The doping profile cannot be calculated directly from these two values. Based on the implantation energy, the projected range R_p (i.e. the depth of the peak concentration inside the substrate) can be derived from tabular or graphical data. The relevant curves (reproduced from [4.23]) have been reproduced in Figure 4.3. Also of importance is the spread on the projected range, in the implantation direction referred to as the projected straggle ΔR_p and perpendicular to the implantation direction the lateral straggle $\Delta R_\perp$. Based on these values a relatively accurate Gaussian description of the implantation profile $n(x)$ is

$$n(x) = \frac{S}{\sqrt{2\pi}\Delta R_p} \exp\left[-\frac{(x - R_p)^2}{2\Delta R_p^2}\right] \tag{4.3}$$

Annealing

After the dopants have been implanted, the lattice has to recrystallise, for three reasons. First, the dopants have to be activated. Second, during ion implantation the crystal lattice has been damaged due to the dopant atoms that collided with the lattice. This damage has to be repaired. Third, the SiGe deposited might have the wrong crystallographic structure. The grain size and shape become very irregular or even have become amorphous for deposition temperatures around 600 °C to 650 °C, or when the phosphorous dose was greater than 10^{15} cm^{-2} [4.24].

During the anneal, the doping profile will broaden due to diffusion of the dopants. According to Sze [4.24], for monocrystalline materials, this limited-source diffusion can be captured in Equation (4.3) by substituting ΔR_p^2 by $(\Delta R_p^2 + 2Dt)$. In [4.25] the diffusion behaviour in polycrystalline materials is discussed extensively, which is very different from the diffusion in monocrystalline materials. Even though

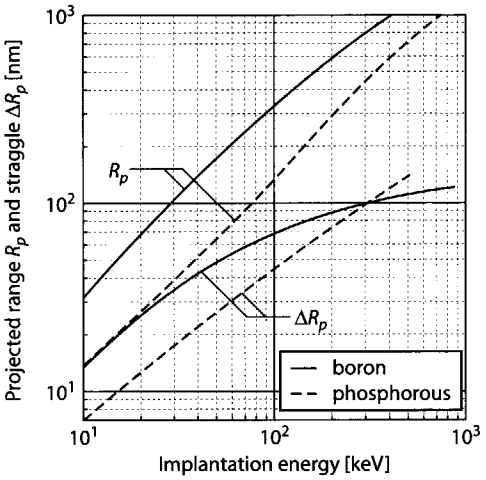


Figure 4.3 Projected range and projected straggle in Si according to [4.23].

the grain boundaries of the polycrystalline material only make up a small fraction of the volume, this disorder at these boundaries causes the overall diffusion into the material to increase significantly. As a result, dopants will first move along these grain boundaries before diffusing into the grain. This effect is most noticeable with heavier dopants.

For example, when the distance between dislocations is small, they will eventually form a slab of high diffusivity. The columnar structure frequently observed for polySiGe approximates this condition. It can be shown that the dopant concentration may be expressed as

$$n(x,y) = C_s \exp\left(\frac{-x\sqrt{2}\sqrt[4]{D}}{\sqrt{\delta D_g}\sqrt[4]{\pi t}}\right)\operatorname{erfc}\left(\frac{y}{2\sqrt{Dt}}\right) \tag{4.4}$$

Here x is the position along the grain boundary, y is the position perpendicular to the grain boundary (into the grain), δ is the width of the grain boundary and D_g is the diffusivity along the grain boundary.

The result of the increased diffusivity along the grain boundary is that the doping profile becomes irregular. As such, secondary ion mass spectroscopy (SIMS) or similar analysis tools for measurement of the doping concentration do no longer provide reliable results. Instead, voltage-contrast scanning electron microscopy and electron-beam-induced-current measurements are required to determine the amount of active dopants at the grain boundary and inside the grain.

4.3.3 (Wet) bulk micromachining

For many transducers, there is a need to structure or remove the substrate, either to create free-standing mechanical structures or membranes of a specific thickness, or—as is the case for thermal applications like the integrated TEC—to increase the thermal insulation. Two classes of etching techniques are available:

- Dry etching. Either a plasma etch or a deep reactive ion etch (DRIE) is used for this purpose. The main advantage of a dry etch is that a very good directionality can be obtained. On the downside, the equipment for a dry etch is rather expensive and the etch process requires a very delicate control. More on dry etching techniques can be found in [4.26], for example.
- The majority of bulk etches are done by means of wet etching. The main advantage over dry etching is the higher cost-effectiveness. Much less (if any) expensive equipment is required, and the processes are quite easy to control. In [4.27] an extensive review of the various wet etching techniques is provided.

In relation to the integrated TEC, wet etching has been applied to remove the substrate below the thermoelements, in order to minimise thermal losses. Moreover, regarding device performance, it was already established through modelling that bulk micromachining can have a profound effect on device operation; In particular when large amounts of power have to be removed through the substrate. In the remainder of this subsection the option regarding wet etching are outlined. This concerns both the bulk etchant properties as well as the etch stop, in order to accurately control the thickness of the remaining structure.

Wet bulk etchants

Four wet bulk etchants are commonly encountered. These are KOH, TMAH (tetramethyl ammonium hydroxide), EDP (ethylenediamine pyrocatechol) and hydrazine. All of these are alkaline etchants, which anisotropically etch monocrystalline silicon (i.e. faster in some crystallographic directions than others.)

The (111) plane has the slowest etch rate, while the (100) plane is the fastest etching plane. As a result, the etch front is bound by the (111) planes, initially resulting in an inverted pyramid with the sides always aligned to the (110) planes. The angle between the <100> and <111> directions is 54.74 °. To obtain a pit with an etch depth D, the mask opening on the backside of the wafer needs to have a side length of $2D{\cdot}\cot(54.74\,°)$. This is illustrated inFigure 4.4. If D is more than the wafer thickness—which is what is required for the lateral TECs—the etch front will reach the other side of wafer, after which an inverted pyramid structure results.

The main drawbacks of hydrazine and EDP is that these etchants are very toxic. Moreover, hydrazine is explosive, making it a delicate material to handle. The main drawback of KOH is that it is not cleanroom compatible, as it easily causes K^+ contamination, even though it is easy to dispose of. On the positive side, KOH has the

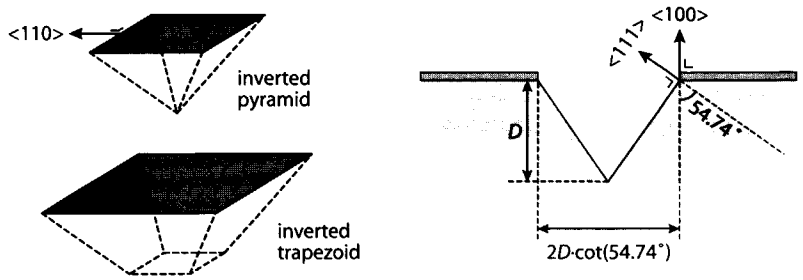


Figure 4.4 Basic etch shape resulting from an alkaline etch.

highest anisotropy of all etchants, as the (100) plane etches around 400 times faster than the (111) plane. Moreover, Si_3N_4 is hardly attacked (< 1 nm h^{-1}) making it an excellent mask / barrier material.

TMAH offers an interesting alternative to KOH. First, the TMAH ion does not diffuse into the silicon lattice as easily as the K^+ ion, making this etchant cleanroom compatible. Moreover, it does not attack SiO_2 nearly as much as KOH. This enable SiO_2 to be used as masking material instead of Si_3N_4. On the downside, TMAH is an organic solution, which is not so easy to dispose of as anorganic solutions like KOH. What is interesting to note about TMAH is that different dielectric materials are etched at very different rates [4.27]. In a 5% TMAH solution and at 80 °C, PECVD silicon oxide, PECVD silicon nitride, thermal silicon oxide and LPCVD silicon nitride respectively have etch rates of 70, 45, 15 and < 5 nm h^{-1}.

Of the four, KOH is the most popular etchant and is the etchant applied in the fabrication of the on-chip integrated TECs. The high anisotropy is crucial for the development of on-chip integrated TECs, as this minimises underetching of a membrane, making the size of the membrane and free-standing thermoelements more controllable. In Table 4.1 the most important properties of the four wet etchants are outlined.

Undercutting

One of the difficulties of a bulk etch is accurate control over the vertical dimension. To illustrate the potential problems, Figure 4.5 shows the basic progress of a KOH etch. After the etch window has been opened (step 1), the etch will rapidly proceed in the <100> direction (steps 2 - 3), until the front side of the wafer is reached (step 4.)

At step 4, the inverted pyramid shape of Figure 4.4 is obtained. Unfortunately, the etch does not stop when the front side of the wafer is reached. Even though the anisotropy for KOH and THAM is high, the (100):(111) etch rate ratio is not infinite. Thus, silicon will still be removed in the <111> direction. Therefore, in time, the masking material will be undercut. This effect is shown in step 5 of Figure 4.5. It worsens with decreasing wafer quality: An increase in lattice defects and other crystal

Table 4.1 Main properties of the most conventional wet etchants (after [4.13].)

Etchant	Masking material	Anisotropy (100)/(111)	SiO_2 etch rate [nm h^{-1}]	Si_3N_4 etch rate [nm h^{-1}]	Comment
Hydrazine	SiO_2, Si_3N_4, most metals	16:1	100	< 10	Highly toxic, explosive
EDP	SiO_2, Si_3N_4, Cr, Au	35:1	120	60	Highly toxic
TMAH	SiO_2, Si_3N_4, Cr, Au	300:1	< 100	< 20	Cleanroom compatible
KOH	SiO_2, Si_3N_4, Cr, Au	400:1	1700–3600	< 1	Not cleanroom compatible

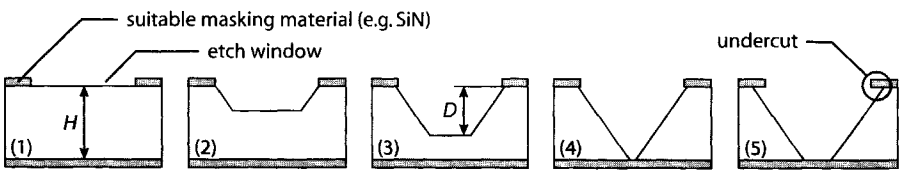


Figure 4.5 The basic progression of a KOH etch.

irregularities that appear in the (111) plane generally create new etch planes, which etch significantly faster than the <111> direction. This way, undercutting of up to 10 μm has been observed.

For many transducers, in particular mechanical ones like pressure sensors, the resulting device structure should be as indicated in step 3 of Figure 4.5. The silicon layer of thickness (*H-D*), usually the epi-layer, is either required for mechanical strength or contains entities like strain gauges thermopiles. If the etch is not stopped, it will continue etching until this layer is entirely removed. To prevent this, various etch stop techniques have been developed. One of these, the electrochemically controlled etching (ECE) has been applied to create 'epi islands'. Such epi-layer islands are suspended from the membrane and contain can electronics. Due to the good thermal isolation of these islands, the temperature can be accurately controlled.

In [4.28] a very clear distinction between etch stop techniques is made, intrinsic etch stop techniques and extrinsic etch stop techniques. The intrinsic etch stop techniques (referred to as chemically limiting) do no require an external power source. Extrinsic etch techniques do require an external power source, like an electrical current source or a light source. Below the most frequently used etch stop techniques are outlined. A more extensive review of etch stop techniques is provided in [4.27].

Intrinsic etch stop techniques

The best known intrinsic etch stop techniques are:

1. The timed etch stop. This is the most straightforward and cheapest method. The main problem associated with this approach is that the etch rate and wafer thickness can only be determined very roughly, which limits the accuracy and uniformity of the layer thickness. In particular for TMAH solutions, the etch rate changes significantly during etching, as Si dissolves into the solution [4.29].
2. The p^+ etch stop, also called the high-boron etch stop. KOH etching stops when a boron concentration exceeding ~$5 \cdot 10^{19}$ cm^{-3} (or ~10^{20} cm^{-3} for TMAH) is reached. Due to the excellent control over implantation depth and dose, a thin and uniform layer of a well-defined thickness is obtained. The main disadvantage is that the high level of background doping makes it impossible to fabricate electronics over the region where the etch stop has been created.
3. The buried dielectric etch stop. In this technique etching proceeds until a buried dielectric layer is reached. The disadvantage lies in the requirement to create the dielectric layer before processing of the electronics. Even though the use of SOI wafers is increasingly commonplace, these wafers are still expensive. Alternatively, epitaxial lateral overgrowth [4.17] or silicon oxide-epipoly [4.13] can be applied to grow an epitaxial layer on top of a dielectric layer.

Extrinsic etch stop techniques

The above-mentioned intrinsic etch stop techniques are easy to apply, yet all have difficulties that limit the applicability or accuracy. The extrinsic etch stops eliminate most of these difficulties, yielding better control over the thickness of the silicon layer that is to remain.

Regarding the work in this thesis, electrochemically controlled etching (ECE) is the extrinsic etch stop technique used to create the epi islands. This is the most commonly applied extrinsic etch stop technique (first presented in 1970 [4.30].) For this technique, an n-type epitaxial layer is grown on top of a p-type substrate (forming a pn junction.) The epi layer has to be biased positively with respect to the Pt counter electrode in the solution, to which the p-type silicon substrate is exposed. This reverse-biased pn junction ensures the etch will stop once the depletion layer is reached.

The appeal of the ECE is that no layers are required in addition to those provided by the microelectronic fabrication process. Moreover, the regions of the epitaxial layer that are required can be preserved selectively, by implanting deep p-implantation, stretching across the epi-layer. This technique is discussed in some more detail in Section 4.4.2.

4.3.4 Dry etching

Dry etching, or plasma etching, is another frequently applied etching technique. A plasma simply is a partially ionised gas, created from breaking up nonreactive molecules, often freons, into highly reactive species. Unlike the wet alkaline etch, which is direction dependent, the basic plasma etch is isotropic (see Figure 4.6a) and relies on the diffusion of ions and radicals to the wafer surface. Once the surface is reached, the reactive particle forms a chemical bond with the surface after which a reaction product is formed that is desorbed from the surface and eventually removed from the reactor.

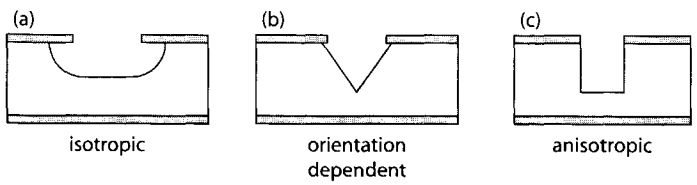


Figure 4.6 The three fundamental etch profiles.

Alternatively, an ion-assisted plasma etch, commonly referred to as a reactive ion etch (RIE), can be tuned to provide a very high level of anisotropy (see Figure 4.6c). Unlike the above-mentioned plasma etch, the ions are accelerated towards the wafer surface, by means of dc, RF, microwave, or inductively coupled energy. Depending on the acceleration potential, three different processes may occur: a) isotropic chemical etching, b) sputtering by physically bombarding the surface and c) reactive ion etching in which the ion bombardment enhances the chemical etching rate. Inductively coupled plasma (ICP) sources are finding more frequent use as these achieve higher plasma densities at lower pressures. This results in a better uniformity and a higher selectivity with lower ion energy [4.31]. For further details on dry etching, the reader is referred to [4.32][4.33].

4.4 LATERAL TEC FABRICATION PROCESSES

For clarity, the fabrication of the co-integrated lateral thermoelectric cooler is divided into two subsection. In the first subsection, the fabrication of the TEC itself is discussed. In the second subsection integration of the TEC within a microelectronics process is discussed.

4.4.1 Lateral thin-film TEC process flow

In literature many different process descriptions are found for the fabrication of CMOS compatible thin-film thermopile-based structures [4.34]-[4.36], mostly based on front-side etching. The basic process flow used for the structures presented in this thesis is an extension of the fabrication sequence previously developed by

Sarro *et al.* [4.37]. In total, six masks are required, including one full-wafer mask for the alkaline etch. A graphical representation of the process flow is shown in Figure 4.7. The most important processing steps are outlined below.

Nitride deposition

From the viewpoints of fabrication compatibility and mechanical stress reduction it would be best to limit the dielectric layers used to SiO_2 of equal thickness only. However, as KOH is the bulk etchant, SiO_2 can not be used. Instead, a 300 nm layer of low-stress silicon nitride is deposited by means of LPCVD in a Tempress CVD oven. The precursors are SiH_2Cl_2 (340 sccm) and NH_3 (60 sccm). At a deposition temperature of 850 °C and a pressure of 150 mTorr, the growth rate is 7.5 nm/min. The layer is deposited on both sides of the wafer.

If the epi layer is not biased, the alkaline etch will proceed straight to the nitride layer. So the primary function on the front side of the wafer is to act as alkaline etch stop. The nitride film on the backside of the wafer acts at as masking material for the alkaline etch, after it is patterned in a later stage of the process.

Polycrystalline silicon germanium preparation

The second step is the preparation of the thermoelectric material, poly$Si_{0.7}Ge_{0.3}$. Due to the presence of chlorine-based compounds (from the dichlorosilane) nucleation on the nitride layer can be difficult (resulting in a discontinuous film of isolated grains). For this purpose, a 15 nm polySi wetting/nucleation layer is LPCVD deposited (45 sccm SiH_4, at 570 °C and 150 mTorr, with a growth rate 1.7 nm/min.) After this layer has been deposited, a 600 nm poly$Si_{0.7}Ge_{0.3}$ layer is epitaxially grown. The growth took place at a temperature of 700 °C, under atmospheric pressure. A gas mixture of hydrogen (40 slm), with 5 % germanium (51 sccm) and dichlorosilane (20 sccm) was used to establish a growth rate of 35 nm/min.

Introduction of impurity atoms into the polySiGe is performed using implantation. For the p-type regions (Mask 1, called 'PP'), boron was used, while phosphorous was used to obtain n-type polySiGe (Mask 2, 'NP'). All structures characterised in this thesis were implanted using an energy of 40 keV. In the case of boron, a dose of $5 \cdot 10^{15}$ ions/cm^2 is applied while the dose is $7.5 \cdot 10^{15}$ ions/cm^2 for phosphorous.

To complete the thermoelectric material preparation, the polySiGe is patterned (Mask 3, 'PE'). This is done by means of an anisotropic plasma etch, for which either a gas mixture of CF_4 (70 sccm), SF_4 (10 sccm) and O_2 (10 sccm) is used, or a gas mixture of 36 sccm C_2F_6 and 144 sccm CHF_3 at a pressure of 180 mTorr. The RF power for the latter etch recipe is 300 W. Typical etch rates are in the order of 9 to 10 nm/s.

Oxide deposition (from TEOS)

After poly$Si_{0.7}Ge_{0.3}$ preparation is completed, TEOS (tetraethylorthosilicate) is used for LPCVD of SiO_2. The TEOS bubbler is kept at 40 °C while N_2 is used as a carrier gas. Deposition takes place at a temperature of 700 °C and a pressure of 250 mTorr. The thinnest layer deposited is 100 nm, which takes about 13 min.

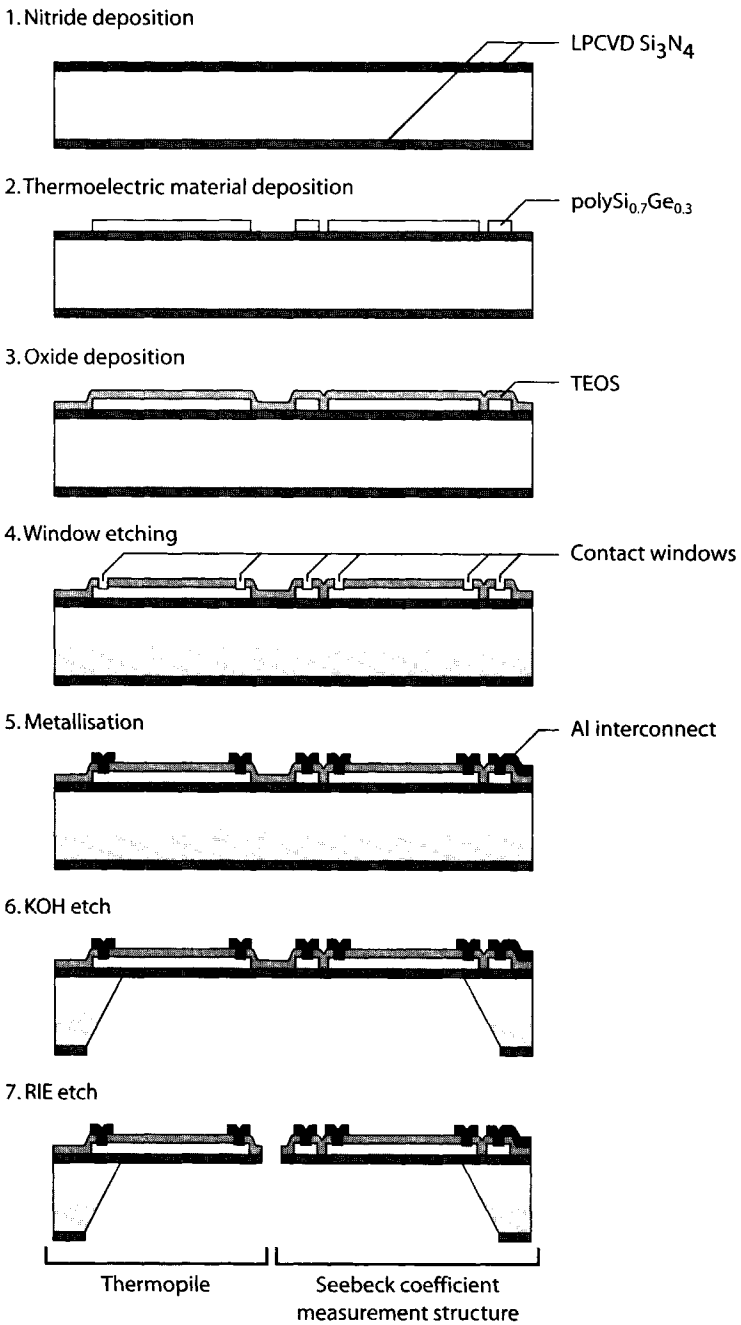


Figure 4.7 Graphical representation of the process flow for lateral polySiGe TECs.

Anneal

After the SiO_2 is deposited, a 55 min anneal is performed; 20 min at 600 °C and 35 min at 1000 °C. This temperature is chosen to recrystallise the polySiGe such that the residual stress is minimised. Furthermore, at this temperature it activates the dopant atoms are activated andthe quality of the oxide layer is improved through densification.

Metallisation

To establish an electrical connection to the thermoelectric material, contact windows are etched into the SiO_2 by means of dry etching (Mask 4, 'CO'). Thereafter $Al_{0.99}Si_{0.01}$ is sputtered onto the wafer at 350 °C, which is patterned and etched by means of dry etching (Mask 5, 'IC').

Bulk and surface etch

The final step is to thermally isolate the thin-film device. This is done in two steps. First, the bulk is removed by an alkaline etch with 33% KOH at 80 °C–85 °C (Mask 6, 'KOH'; full wafer mask). At this temperature and concentration the average etching speed is 1.3 µm/min. Finally, the unwanted thermal losses through the thin film are minimised by removing any unnecessary membrane regions. This is done using a reactive ion etch (Mask 7, 'RIE'). Note that as the membranes are too thin to go through photoresist spinning and pattering, the layer for the 'RIE' mask must be deposited and patterned prior to the wet etch.

Two of the resulting structures prepared by means of the above process are shown in Figure 4.8. These are two cantilever-type thermoelectric coolers.

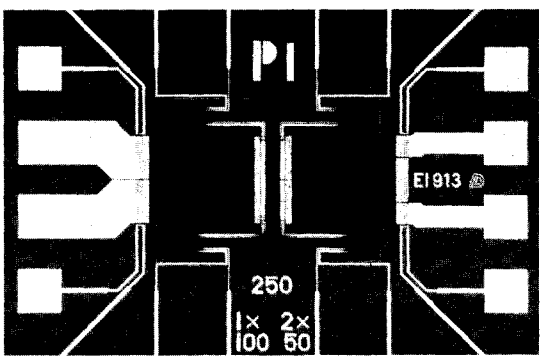


Figure 4.8 Lateral thin-film thermoelectric coolers. The thermoelement lengths are 250 µm, while the respective widths are 100 µm and 50 µm.

4.4.2 Lateral TEC co-integration

The process flow presented in the previous subsection is still of limited use, as only a part of all applications targeted can be served, i.e., for cooling and thermal stabilisa-

tion of devices that can be constructed in the TEC fabrication process (like IR detectors and humidity detectors.) If electronic components need to be cooled or thermally stabilised, electrochemically controlled etching has to be applied to thermally isolate these electronics. The thermal conductivity of the silicon substrate is too high for the lateral TEC to be effective in otherwise.

To establish an electrical connection between the electronics and the thermoelements, and enable an electrochemically controlled etch, the process flow presented in Section 4.4.1 needs to be expanded. In Figure 4.9 on page 124, this modified process flow is outlined. A 2 μm bipolar process 'DIMES01' was used for the electronics. Due to relatively large feature size in the process, sufficient robustness with respect to the thermal budget is created. Some of the standard layers from this bipolar process are also required for the ECE:

1. The electrochemically controlled etch needs to stop at a pn junction. This junction is generally formed by the p-substrate / n-epi layer interface. Thus, the epitaxial layer is required.
2. To ensure the etch stops at the interface between the epitaxial layer and the substrate, a bias potential has to be applied across this interface. For this purpose, each epi island that has to remain after the ECE needs to be electrically contacted. For a good contact to the n-type epi layer, a shallow n^+ contact is implanted. Such a contact is encircled in step 6 of Figure 4.9.
3. Finally, the contours of the epi islands are defined by properly implanting deep p^+ barriers that stretch across the thickness of the epitaxial layer. This way a second pn junction is created, at which the ECE stops. (The regions in between the p^+ barriers that are not electrically biased will be etched away.)

So, in total three processing steps of the bipolar process are used, i.e., 1) the epi layer growth, 2) the p^+ implant and 3) the n^+ implant. All are standard DIMES01 processing steps. Below the steps of the combined process are outlined.

Epitaxial layer growth

Processing is started off by growth of the n-type epitaxial layer (step 1 in Figure 4.9). The growth takes place at a temperature of 1050 °C and a pressure of 60 Torr. Arsenic is used with a concentration of $1{\cdot}10^{16}$ cm^{-3}, which gives a sheet resistivity of 1200–1500 Ω/sq. The final epitaxial layer thickness is 4.0±0.1 μm.

Deep p^+ and shallow n^+ diffusion

Along the course of the standard bipolar process, the p^+ barrier (Mask 1, 'DP') and the shallow n^+ layer (Mask 2, 'SN') are introduced. Both dopants are introduced by means of diffusion. For the p^+ barrier, boron is deposited on the wafer prior to drive-in. The resulting layer has a thickness of 4.7±0.2 μm. The shallow n^+ layer is created by phosphorous drive-in, which is deposited on the wafer surface from liquid $POCl_3$. The thickness of this layer is 2.4–3.4 μm.

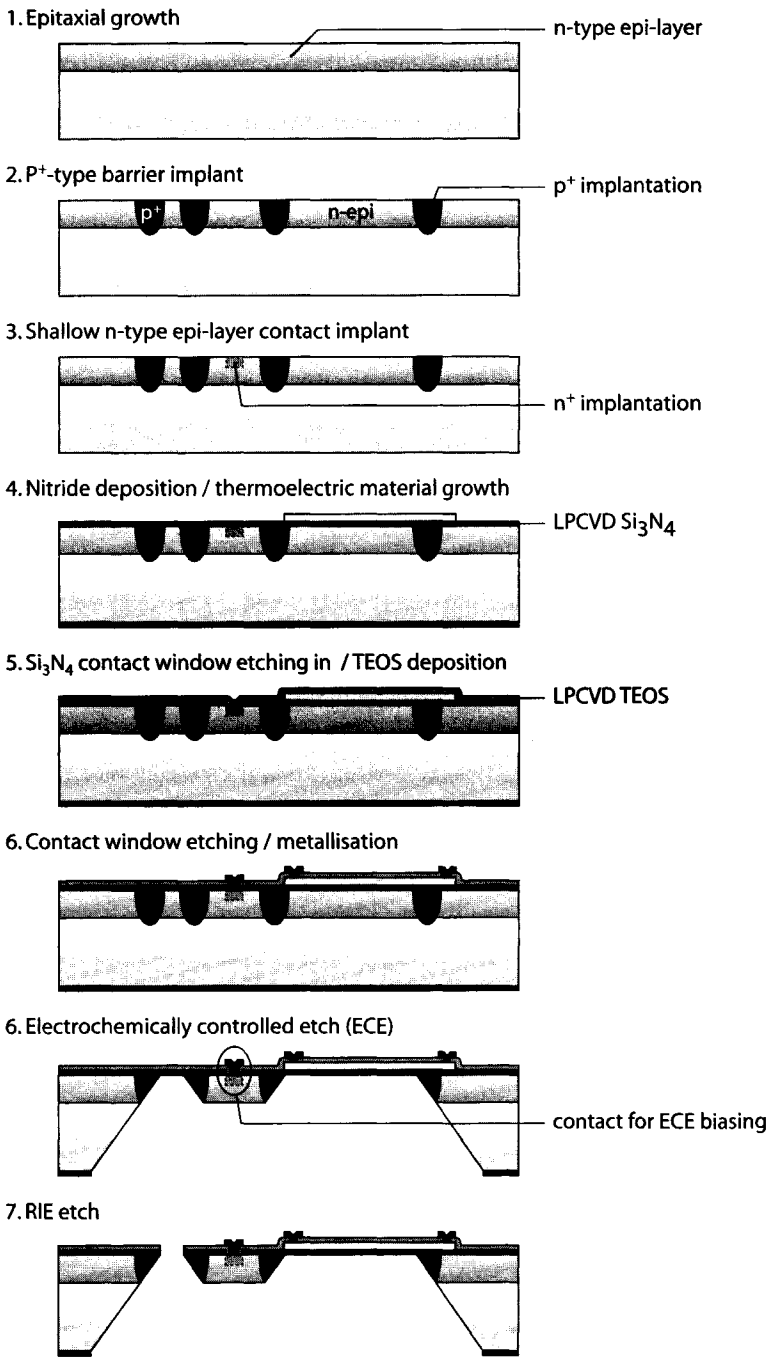


Figure 4.9 Graphical representation of the process flow required to integrate lateral TECs with microelectronic circuits.

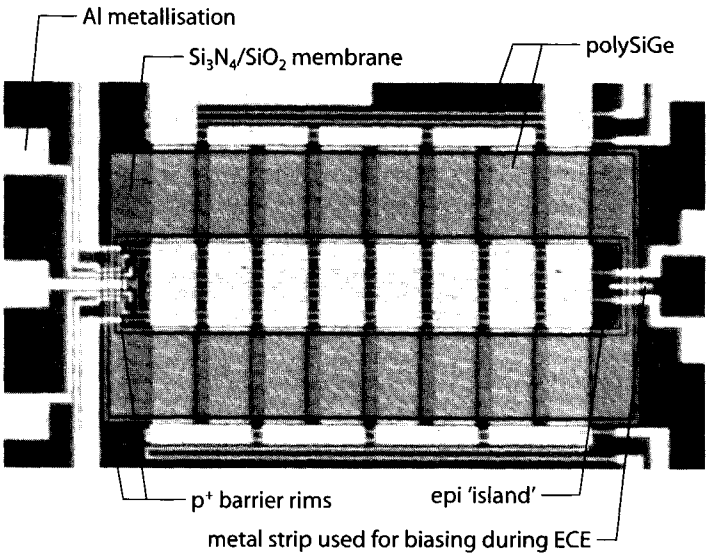


Figure 4.10 Thin-film thermoelectric cooler, fabricated using electrochemically controlled etching. The dark epi island in the middle (underneath the aluminium interconnect) is suspended from the Si_3N_4/SiO_2 membrane.

Lateral TEC fabrication steps - part 1

The first deviation from the standard bipolar process flow is introduced only after all doping steps have been performed, when the silicon nitride layer is LPCVD deposited. (This is step 4 in Figure 4.9 and corresponds to step 1 in Figure 4.7.) The thin film TEC fabrication process is continued by deposition of the polySi nucleation layer and subsequent polySiGe layer. After the polySiGe has been implanted and patterned (using the 3 masks, 'NP', 'PP' and 'PE') an additional step is required that was not discussed in the TEC process description in the previous subsection.

Contact window etch in nitride layer

Unlike with the thin film TEC process, an electrical connection with the substrate needs to be established. For this purpose, a contact window needs to be etched into the silicon nitride layer. This can be the same mask as used for the shallow n^+ implant as the position of the contact hole corresponds with the position of the n^+ implant. (This has been experimentally verified to work.)

Lateral TEC fabrication steps - part 2

After the contact window has been etched, the TEC fabrication process is continued. The TEOS layer is grown and the polySiGe/TEOS are annealed. Subsequently, contact windows are etched into the TEOS to establish an electrical connections to be-

tween the interconnect and the electronics in the substrate as well as between the interconnect and the thermoelements. (This is shown in step 6 of Figure 4.9.)

Alkaline etch with ECE

After all thin-film layers are in place, the two steps that still remain are the ECE of the silicon substrate, followed by a RIE etch to minimise thermal losses through unnecessary regions of the thin-film membrane. The ECE is again performed in a 33% KOH solution, of which the temperature is kept constant at 80 °C–85 °C. The RIE etch is the same as discussed in the previous subsection. After the process has been finalised, structures similar to that shown in Figure 4.10 result. In this figure, the RIE etch has been left out.

4.5 TECHNOLOGICAL IMPROVEMENTS OF YIELD AND PERFORMANCE

Even though the basics of the fabrication process were established in 1994 by Sarro *et al.* [4.37], the actual fabrication of the on-chip integrated TECs turned out to be anything but trivial. To provide an indication, of the 35 to 40 wafers used for experimenting, only four survived to be diced. Of these thin-film structures, on average only 30% remained after dicing. So, in the end only a few tens of structures survived; Sufficient for academic purposes, but far from sufficient for commercial purposes.

Below, the major 'yield-killers' are discussed together with the solutions on how to deal with these. Thereafter, directions for future improvements to both yield and performance are discussed. In fact, if properly implemented, yield and performance can be improved simultaneously, at the cost of a few masks only.

4.5.1 Yield issues encountered

A number of annoying fabrication issues were encountered, most of which were easy to solve. Most of the problems originated from 1) bad electrical connections, 2) wet etching, 3) plasma etching and 4) dicing. What can be seen from these issues is that most of them occur at the end of the fabrication process and they are irreversible. This is why so many wafers were lost.

4.5.1.1 Open circuits

The first yield issue encountered was an improper Al/polySiGe contact. It turned out the TEOS plasma etch was timed too critically, while the TEOS layer was slightly thicker than anticipated. This caused the contact window not to penetrate the TEOS layer. The solution was simple. By increasing the time of the dry etch, an overetch situation was created[5]. The polySiGe was sufficiently thick to survive this overetch.

Open circuits also cause the shallow n-well to be improperly biased during ECE. This leads to the effect observed in Figure 4.11. The etch does not stop at the deep-p/shallow-n junction, and eventually reaches the nitride layer, which is all but pene-

5. After a suggestion made by Sander van Herwaarden (Xensor Integration, The Netherlands).

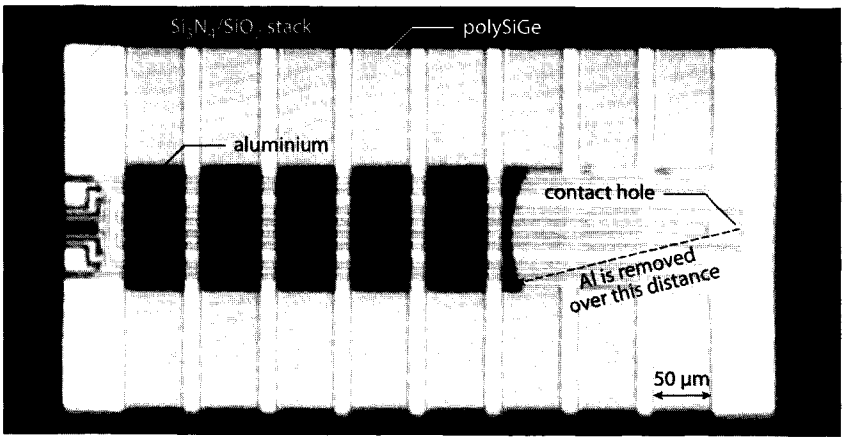


Figure 4.11 Aluminium removal due to KOH seeping through the contact hole, as a result of failure to properly bias the n-type epi-layer.

trated during the contact hole etch. Even though the etch rate of Si_3N_4 is slow this thin layer is rapidly removed so the front side of the wafer becomes exposed. Figure 4.11 clearly reveals to which distance from the contact hole the aluminium has been removed (that is, nearly 200 μm.)

4.5.1.2 Short circuits

The next problem encountered was the formation of random yet frequent short circuits throughout the wafer. Careful measurements indicated the problem occurred where the aluminium interconnect and the polySiGe layers overlapped. First it was though that the TEOS did not properly cover the sides of the polySiGe, so that the aluminium interconnect causes short-circuits between the polySiGe entities. Further investigations were carrie out, by rearranging the pinning of the various test structures so configurations as in Figure 4.12 could be examined.

These investigations ruled out the possibility that the side walls of the polySiGe were uncovered: In such a case, interconnect overlapping multiple polySiGe strips or two metal strips on different sides of a polySiGe strip (as in Figure 4.12a) should have been observed. However, if metallic strips ran in parallel and did cross over a polySiGe strip, as in Figure 4.12b, short circuits were observed.

This left only one possible explanation for the short-circuits: some sort of stringer formation occurred in between the two metal lines. A closer inspection of the process chart revealed the problem. In the course of the foregoing months, a wet dip etch after the Al plasma etch was removed from the process chart. After renewed introduction of this wet dip etch, the short circuits were encountered only seldomly.

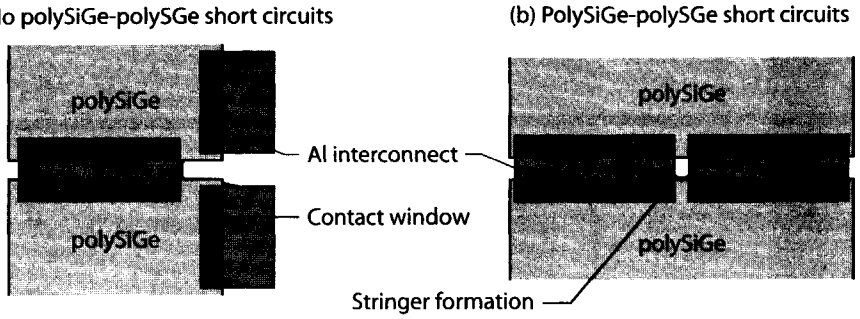


Figure 4.12 Two interconnect configurations examined to determine the source of short circuits that were encountered during the development of the TEC fabrication process.

4.5.1.3 Wet etching issues

Mechanical failure during KOH etching was the single most significant cause for the loss of wafers. Nearly 50% of all wafers were destroyed during this step. As soon as the substrate is removed, the thin-film TEC loses its mechanical support at which point the mechanical forces within the wafer resettle. Careless clamping of the wafer in the wafer holder (possibly combined with a wafer containing a lot of defects) is sufficient to pre-stress a wafer so severely, that it can no longer withstand the temperature and pressure differences occurring during the KOH etch.

In the least severe case the stress lead to snapping of a number of thin-film membranes, as most of the stress is released near the corners of these membranes. However, of a more severe nature, entire wafer were observed to break in half. In both cases KOH seeps through to the otherwise protected front side of the wafer, so that the aluminium interconnect is removed.

The problem originates from three sources. The first, careless clamping, can be avoided easily. The second, conceivably poor quality wafers, can't be remedied. The third source simply is the thickness of the thin-film structure. The thinnest structure had a nitride layer of 250 nm, a polySiGe layer of 300 nm and an oxide layer of 100 nm, which gives a total stack thickness of 650 nm. It is very easy for stresses to build up in such thin layers.

The first solution to this problem is to improve the protection of the front side of the wafer. This can be done by covering the front side of the wafer in wax [4.38] or PDMS [4.39], for example. Even when a membrane breaks, the KOH cannot seep through to the front side of the wafer if such a layer is in place. It is unclear whether wax or PDMS can be removed from the wafer without leaving traces that might influence the performance of the device. As the wax is used to protect thin-film thermoelectric structures and the PDMS is used to protect thin-film deformable mirrors, such remnants should already have been noticed.

The second solution is to increase the thickness of the membrane stack. The latter can be achieved by increasing the individual layer thickness or by increasing the number of layers. In fact, increasing the number of layers will—if properly implemented—simultaneously increase device performance and mechanical strength. This modified process will be discussed in some more detail in Section 4.5.3.2.

4.5.1.4 Dry etching issues

In an early research stage an effort was made to reduce the number of masks to an absolute minimum. It was concluded that, of the seven masks used (NP, PP, PE, CO, IC, KOH and RIE), two masks (CO and RIE) could be merged. The combined mask (labelled 'CR') then replaces the CO mask.

Using the CR mask, the SiO_2 is patterned using dry etching. Next, the aluminium interconnect deposition and wet bulk etch are performed as usual. As final step, the already patterned SiO_2 layer serves as a mask for the dry etch used to shape the beams and cantilevers. During the nitride etch, the oxide is also attacked, as the SiO_2:Si_3N_4 etch ratio was 1:1.35. Therefore, depending on the thickness of the nitride layer, the thickness of the oxide layer has to be increased accordingly, in comparison to the 7-mask process.

Even though this 6-mask approach is functioning, it was experimentally concluded that the timing of the nitride etch was too critical. As such, this approach was eventually abandoned and the 7-mask process was used instead. Directly after the backside nitride layer is patterned (for the KOH etch), the RIE mask is applied on the front side of the wafer. Photoresist is used as masking material.

To ensure the photoresist would not be damages from the pressure applied by the the wafer holder, the photoresist needs to be baked. The initial bake was at a 130 °C for 30 min. However, the additional deep UV cure most likely 'overbaked' the protoresist. The result is shown in Figure 4.13. Even though the dry etch is successful, a very thin layer of the photoresist remains. Nor a dry etch nor exposure to ozone could successfully remove this photoresist film. Even through such a thin layer is not expected to have a huge effect on device performance, the layer covers the bond pads, so it is impossible to bond the structure.

4.5.1.5 Dicing concerns

After wet etching, wafer dicing is the largest cause for a low yield. A high-pressure water jet is required to both cool the saw and the spray the silicon powder (generated during dicing) off the wafer. When leaving the thin-film structures unprotected from the jet, the structures are simply blown away. To prevent this, the structure on the front side are first covered by a (non-sticking) plastic foil. This foil is fixed to the wafer (outside the inner 60 x 60 mm) by means of two layers of sticky foil. This is a sufficient protection from the water jet. Still, the power of the saw causes mechanical shocks in the wafer that are sufficient to fracture the thin-film devices (most likely in combination with the stress already present in the structures.) Moreover, the structures that do survive are covered in a layer of silicon dust that is very hard to remove.

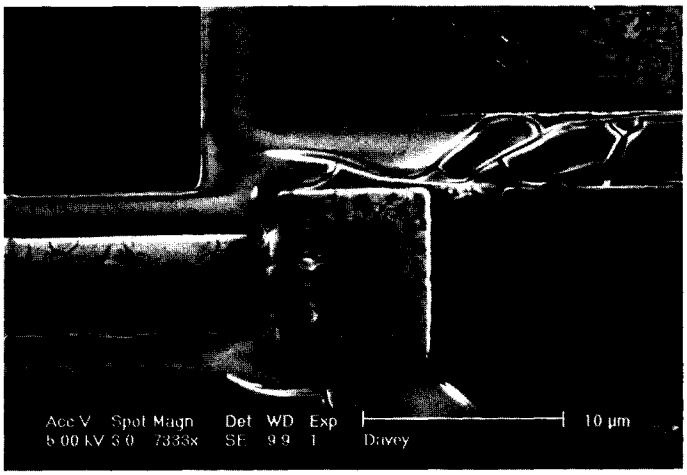


Figure 4.13 Very thin layer of cracked photoresist that, due to excessive hardening, could not be stripped from the wafer.

Like with the wet etch, the structures can be fixed in position by means of wax, PDMS or even photoresist. The advantage is that by removing these layers after dicing, any silicon powder stuck to the wax/PDMS/photoresist is also removed. Another solution can come from scribe-and-break separation of the individual structures.

4.5.2 Yield improvements

Each layer in the TEC has different deposition requirement,s in particular concerning the thermal treatment. Therefore, mechanical stress is impossible to avoid. Still, this stress might be reduced or compensated. Both approaches are addressed below.

4.5.2.1 Reduction of stress in polySiGe

Next to properly annealing the polySiGe (see Section 3.1), stress can be avoided by introducing carbon in the SiGe lattice. The underlying idea is that difference between the Si atoms and larger Ge atoms causes an increased strain level in the SiGe lattice. This can be compensating by incorporating smaller-sized C atoms into the lattice [4.40]–[4.43]. Moreover, as the melting point of SiGeC is lower than the melting point of SiGe, diffusion, dopant activation, recrystallisation and stress reduction by annealing are expected to occur at lower temperature [4.40].

For example, Hsieh *et al.* [4.43] have performed extensive research on stress in $Si_{0.68}Ge_{0.31}C_{0.01}$. A 120 nm $Si_{0.68}Ge_{0.31}C_{0.01}$ layer is deposited on a silicon nitride layer by means of rapid thermal chemical vapour deposition (RTCVD), and is patterned by means of lift-off. Deposition, took place at a pressure of 2 Torr, using H_2, SiH_4, GeH_4, and SiH_3CH_3 (methylsilane) as precursors. using respective flow rates were of 1.2 lpm, 60 sccm, 5 sccm, and 0.6 sccm, with which a growth rate of about 120 Å/min

was established. Their research revealed that the strain reduces with increasing SiH_3CH_3 flow rate, up to 0.8 sccm.

4.5.2.2 Stress compensation

The current Si_3N_4/poly$Si_{0.7}Ge_{0.3}$/SiO_2 layer stack is not preferred in terms of stress, as each layer has a different level of pre-stress and a different coefficient of expansion. the situation could be improved by replacing the 100 nm–250 nm SiO_2 layer by a nitride layer of equal thickness as the nitride layer already present in the device (i.e. 300 nm to 500 nm).

Such a modification was not implemented as this would decrease device performance. In Section 4.5.3.2 an alternative process flow with 2 polySiGe layers and 3 dielectric layers. There the layer stack can be built up symmetrically, while still improving the performance of the current triple-layer stack.

4.5.2.3 Avoidance of Al

This appears to be an oversimplified statement. Yet, observing Figure 3.7 and Figure 3.9, in particular the arms that contain Al wires are severely stressed. (the cantilever with the Al specimen in Figure 3.9b is even bent by nearly 90 °.) By limiting the volume of the Al interconnect to a minimum, this stress can be largely avoided. This improvement was implemented in the beam structure, when going from the earlier design EI913 to the newer design EI989:

The structure shown in Figure 4.14 is no longer subjected so the large levels of stress previously observed. Mostly, aluminium wires can only be left out at the cost of reduced functionality. For example, the heating resistor at the tip of the cantilever still has to be configured as a four-terminal resistor. However, by replacing the narrow Al voltage-sensing wires with much wider p-polySiGe wires, the four-wire configuration can be preserved, compensating the stress in the Al layer by the polySiGe and simultaneously lowering the thermal conduction through the arms. Such a replacement is only allowed for arms that are not meant to carry large current (i.e. for the sensing of a voltage or driving the base or gate of a transistor.)

4.5.3 Technological performance improvements

If research on the lateral thin-film TECs is continued, a few technological improvements can still be implement to improve the performance of the device. These are outlined below.

4.5.3.1 Replacement of Al by polySiGe

Although already implemented, and discussed in Section 4.5.2.3, the replacement of Al by polySiGe not only improves the mechanical properties of the device, but also improves the performance for two reasons.

First, the Al has a thermal conductivity of 180–236 $Wm^{-1}K^{-1}$, which is replaced by a polySiGe film with a thermal conductivity of 5–10 $Wm^{-1}K^{-1}$. Thus, depending on the actual thermal conductivities, and assuming the Al and polySiGe layers are of

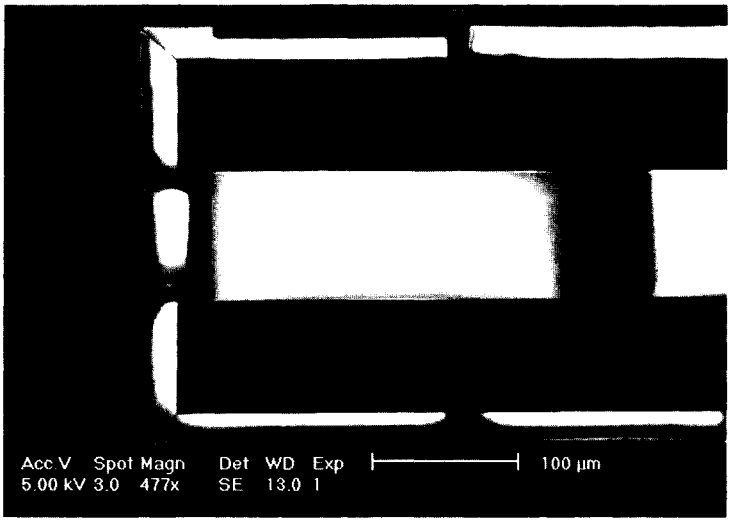


Figure 4.14 An improved reference structure for the thermal conductivity measurements, in which stress due to the Al interconnect was reduced by partially replacing Al with polySiGe.

equal thickness, a polySiGe wire may be 10 times to 45 times wider than an Al wire, without compromising the total thermal conductance.

Second, looking at the structure in Figure 4.14, the current-carrying Al wire is stacked on top of the polySiGe wire, in contrast to the older structures in Figure 3.9, where the current-carrying and voltage-sensing Al arms have to be placed next to each other. As a result, the Si_3N_4/SiO_2 supporting layer can be made narrower. This means a further reduction in the thermal conductance can be achieved.

4.5.3.2 Dual polySiGe layer process

The thin-film TECs discussed in this thesis are constructed using a minimal number of masks. With the application of two additional masks a thin-film TEC can be created that has both a better thermoelectric performance and improved mechanical characteristics. This 9-mask process is displayed in Figure 4.15.

The start of the process is similar to that of the thin-film TEC discussed earlier. First a nitride layer is deposited. Next, a polySiGe layer is deposited, implanted (mask 'NP') and patterned (mask 'PE1'). Note that this polySiGe layer contains only one kind of dopant. On top of this layer either a nitride or oxide layer is deposited. Thereafter a second polySiGe layer is deposited, implanted (mask 'PP') and patterned (mask 'PE2'). Again only one kind of dopant is implanted in this layer. Next a nitride layer of equal thickness as the first nitride layer is deposited and patterned (mask 'NE'). Thereafter the contact windows are etched (mask 'CO') and electrical connections are established by the aluminium interconnect (mask 'IC'). Finally, the structures are released by means of a wet etch followed by a dry etch.

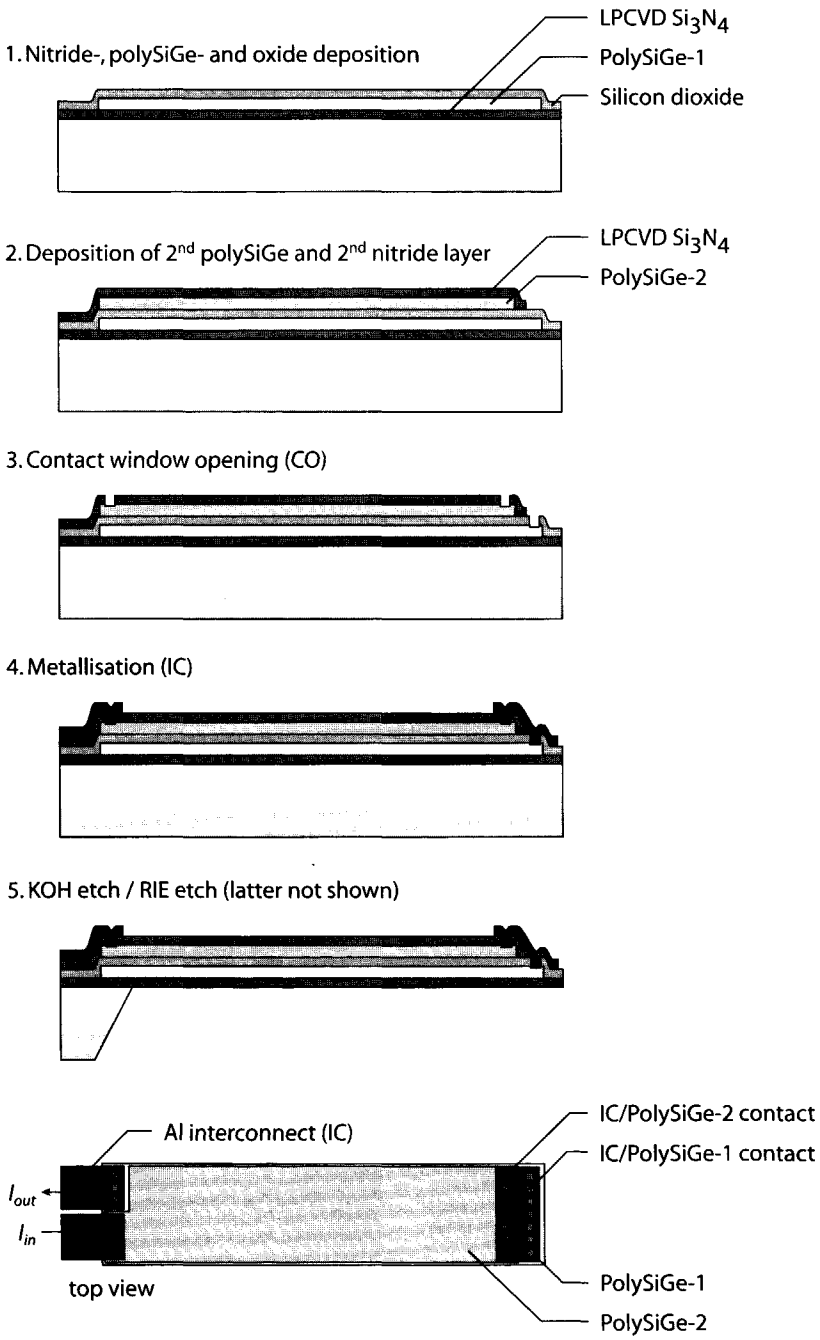


Figure 4.15 An improved lateral thin-film TEC fabrication sequence, which both improves the mechanical stability and the thermoelectric performance.

The first additional mask is required for patterning of second polySiGe layer (mask 'PE2'.) The second additional mask is required for patterning of dielectric layer in between the two polySiGe layers (mask 'NE').

The bottom drawing in Figure 4.15 shows the schematical top view of a thermocouple that can be created using the alternative process flow. As the two polySiGe layers are stacked, the footprint is reduced by 50%, which has significant advantages. Using the same surface area, twice the amount of thermal energy can be removed from the cold junction per unit area. Moreover, the ratio between the number of polySiGe layers and dielectric layers is shifted from 1:2 to 2:3, which means the parasitic thermal conductance through the membrane can be reduced by 20%.

4.5.3.3 Front-side wet etching

In Chapter 2, the effect of different bulk etch techniques and configurations was investigated. In particular the backside and front side alkaline etches are frequently encountered in practise. As such, these two are of most interest. Compared to the front side alkaline etch, the backside alkaline etch leaves only about 1/3 of the substrate volume to sink the heat from the hot junction. As such, in a new design, a front-side alkaline etch is preferable.

Unless a gold film is deposited over the Al interconnect, KOH will attack the interconnect and as such is not considered suitable for a front side etch. Instead, either TMAH or EDP are used, as these etchants can be used in such a solution that a high etch selectivity between the nitride, oxide and aluminium is obtained.

The typical EDP-based etching solution is the so-called S-type EDP, which is composition of 1000 ml ethylenediamine, 160 g pyrocatechol, 6 g pyrazine and 133 ml deionised (DI) water at 95 °C. Using such a solution, hydroxides will form on the aluminium surface, preventing proper bonding. A 2 min dip in 5% ascorbic acid ($C_6H_8O_6$) followed by a 2 min rinse in DI water removes the OH^- ions. A 1 min dip in hexane (C_6H_{14}) ensures the beams and cantilevers will dry without distortion.

As EDP is rather dangerous, TMAH is generally preferred. When the TMAH solution is doped with silicon or silicid acid ($Si(OH)_4$) the aluminium will not be attacked. Typical solution encountered are 5 wt.% TMAH with 16 g l^{-1} of Si or 16 g l^{-1} of $Si(OH)_4$ [4.34], and 10 wt.% TMAH with 36 g l^{-1} of dissolved Si [4.35]. To prevent hillock formation during the etch (that reduces the etch rate significantly) ammonium peroxydisulfate can be added to the solution [4.35].

4.5.3.4 Multi-step implantation

In the thin-film TECs discussed in this thesis, dopants have been introduced through a single implantation. For both B and P, the implantation energy was 40 keV. The thick lines in Figure 4.16 indicate the corresponding implantation profiles, which are derived from Equation (4.3) and the projected range and straggle in Figure 4.3.

With a single implant, the energy should not be chosen too high as this will result in a poor-quality contact. The consequence is that the implant doesn't stretch across the entire height of the polySiGe film. Due to the high dopant diffusion rate along

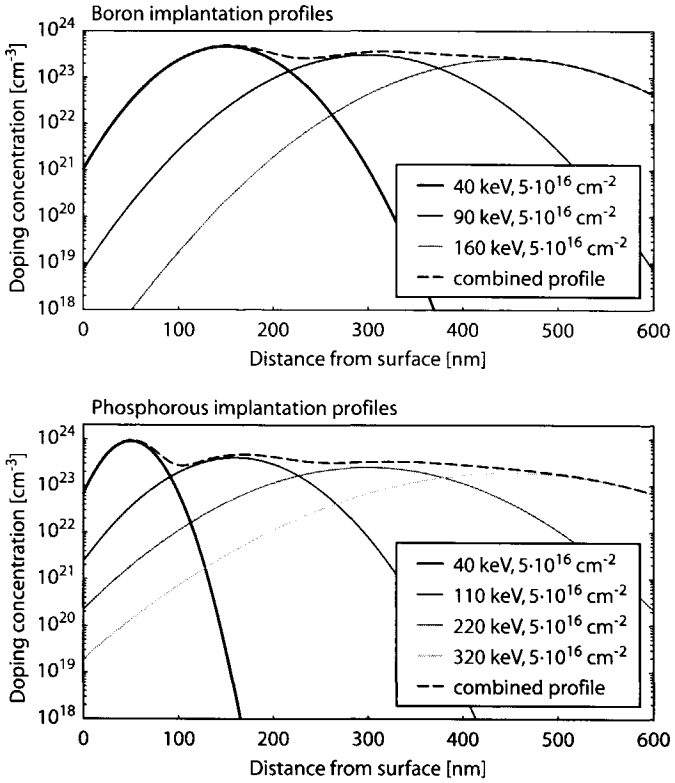


Figure 4.16 An example of doping profile smoothing by means of a multi-step implantation.

the crystal boundaries, an anneal will drive the dopants deeper into the layer. It depends on the anneal temperature and time how well the final doping concentration is homogenized over the entire thickness of the layer.

If a multi-step implantation is used, the dopants can be spread much better over the entire thickness of the film, thereby providing a more uniform doping profile to start with. As examples, a B doping profile based on a three-step implant and a P profile based on a four-step implant are provided in Figure 4.16. As the doping gradients are much smaller with a multi-step implant, the dose should be altered accordingly, toreach the optimal polySiGe doping range of $1–4{\cdot}10^{20}$ cm^{-3}.

4.6 THIN-FILM NICKEL-CHROMIUM DEPOSITION

Before the development of the lateral thin-film TECs, work was performed on the development of near-zero-TCR resistors [4.44]. Such resistors are indispensable for high-accuracy applications as the thermal susceptibility can be removed entirely.

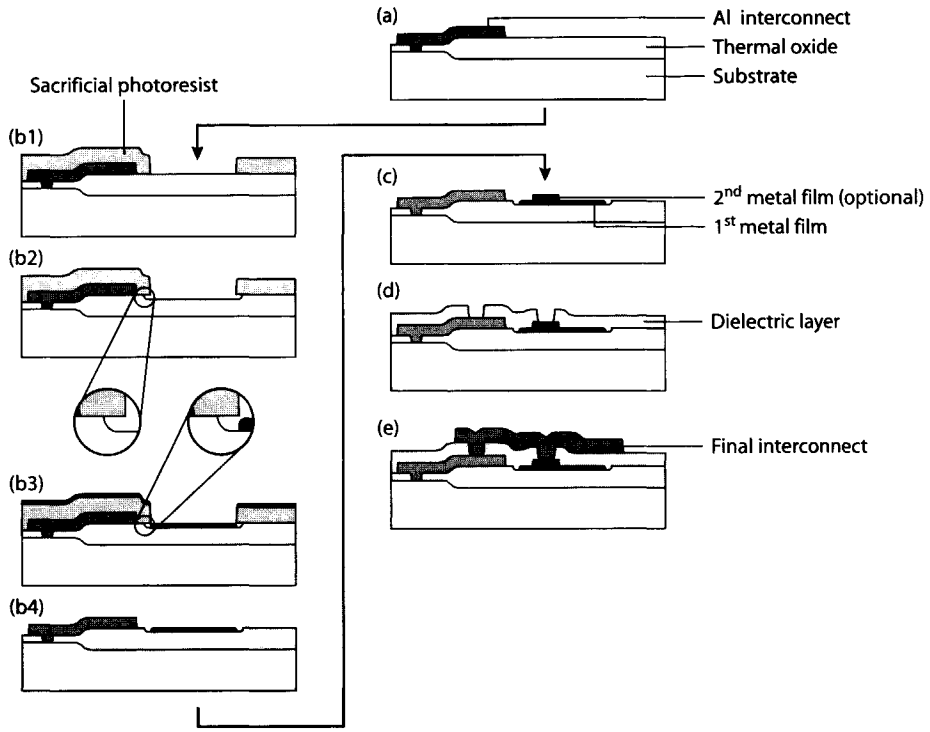


Figure 4.17 Schematic overview of the thin-film metal deposition module in case two thin metal films are deposited.

The work on thin-film NiCr resistors was initially targeted towards high-accuracy D/A converters based on a weighted resistor network (i.e., the Kelvin-Varley attenuator scheme) [4.45]–[4.47]. In the context of this thesis, such temperature-independent resistors are also preferred as heating resistor, as they have a near-constant power output over a very large temperature range. When looking at one of the applications targeted, the micro-thermostat, temperature independent (zero-TCR) resistors are ideally suited as heater.

4.6.1 Thin-film deposition module

The main consideration regarding the deposition and patterning of metallic thin films was that these should be independent of the metal itself. In particular the need specific etchants should be avoided. Therefore, patterning is based on lift-off, with photoresist as sacrificial material. The choice for photoresist automatically implies the process temperature has to be kept low, which has the added advantage that the thin metal films can be deposited after deposition of the aluminium interconnect.

Table 4.2 Metal alloys with a low temperature coefficient of resistance [4.48].

Alloy	Resistivity μΩcm	TCR (0-100°C) 10^{-6}
$Cu_{84\%}Mn_{12\%}Ni_{4\%}$ (Manganin)	44	< 5
$Ni_{75\%}Cr_{20\%}Al_{2.5\%}Cu_{2.5\%}$ (Evanohm)	133	< 5
$Ni_{0.75\%}Cr_{20\%}Al_{2.5\%}Fe_{2.5\%}$ (Karma)	134	< 5
$Cu_{88\%}Mn_{9\%}Al_{3\%}$	38	10
$Cu_{86\%}Mn_{9\%}Al_{5\%}$	42	12
$Cu_{85\%}Mn_{12\%}Al_{3\%}$	48	-5
$Cu_{86\pm1\%}Mn_{12\%}In_{2\pm1\%}$	42.5	10
$Cu_{70.6\%}Mn_{23.2\%}Fe_{6.2\%}$	77	10
$Cu_{55\%}Ni_{45\%}$	49	20
$Au_{97.65\%}Cr_{2.1\%}Co_{0.25\%}$	39.8	20
$Au_{94.4\%}Cr_{2.1\%}Pd_{3.5\%}$	36.9	20
$Au_{88.9\%}Cr_{2.1\%}Pd_{9\%}$	38.7	20

Figure 4.17 shows the basic thin film deposition module, in case two metallic films are deposited. Step (a) indicates a wafer after completion of a standard microelectronic process (including the aluminium metallisation.) Steps (b1) to (b4) show the actual deposition module for a single metallic thin film that is deposited on a dielectric layer. In step (b1) the sacrificial photoresist is deposited and patterned. Step (b2), undercutting of the photoresist, is ignored at this point. After photoresist patterning, the thin metal film is deposited (step (b3)). Thereafter, in step (b4) the sacrificial photoresist is stripped. Step (c) serves to illustrate that the module can be repeated for multiple additional films, for example when an adhesive film or barrier layer is required. If the metal film is not deposited over the aluminium interconnect, a second interconnect layer can be used to bridge the first interconnect layer and the thin metal film. This is shown in steps (d) and (e).

The deposition process displayed in Figure 4.17 is relatively elaborate. In most cases of practical importance, a single thin-film will be deposited. In such a case, the contacts of the metallic film will be deposited over the contacts of the aluminium or polySi interconnect, in which case a direct connection is established. If required, for example, to prevent corrosion or oxidation, the thin film can be protected by means of a passivation layer.

4.6.2 Results

In [4.48] an extensive overview of the TCR of metal alloys is provided. The alloys with the lowest TCR are listed in Table 4.2. This table clearly reveals that in particular

CuMn-based and NiCr-based alloys are very suitable for temperature-independent resistors. Based on the availability of targets within DIMES, a $Ni_{0.5}Cr_{0.5}$ target was used. Even though NiCr as such is not listed in Table 4.2, work performed prior to our work indicated a TCR below 20 ppm should be obtainable [4.49][4.50].

Deposition characteristics

The $Ni_{0.5}Cr_{0.5}$ films have been deposited using e-beam evaporation from the $Ni_{0.5}Cr_{0.5}$ target, at a pressure of $2{\cdot}10^{-3}$ Pa. The deposition rate was approximately 0.5 nm/s. Measurements using an Alpha stepper indicated a thickness uniformity of better than 5% across an entire wafer. This was confirmed from electrical resistivity measurements. The latter measurements also revealed that the worst-case difference between resistors that are spaced 300 μm apart is no more than 0.2%. The latter corresponds to the $\cos[\varphi]^{-1}$ deposition profile from a point source onto a flat surface.

Step coverage

The excellent step coverage of the 35 nm to 75 nm thin NiCr films was the major difficulty to be dealt with. Figure 4.18 illustrates exactly how good this step coverage was. At that point, lift-off was performed in an acetone bath without any ultrasonic agitation. Even though the photoresist is obviously removed, the NiCr that was deposited on the sidewalls of the photoresist was strong enough to survive.

To improve this situation, ultrasonic agitation was introduced. This already improved the situation, as shown on the left side of Figure 4.19. The Bridges observed in Figure 4.18 no longer occurred. Still, frays of the NiCr deposited on the sidewalls remained. This makes it hard to design a resistor with an absolute value. Moreover, the frays are mechanically unstable and may cause erratic and sudden alterations of the resistance in time. Thus, mere ultrasonic agitation is insufficient to yield high-quality resistive films.

To yield high-quality resistors, the shape needs to be precisely defined. This can only be achieved if there is no direct contact between the metal film on the wafer surface and the photoresist. The simplest way to achieve this is by means of a photoresist with a negative slope (i.e. a photoresist that gets narrower near the bottom.) A second way is to undercut the photoresist. This is displayed in steps (b2) and (b3) of Figure 4.17. After the photoresist is spun onto the wafer and patterned, the photoresist is used for isotropic etching of the underlying dielectric film, thereby undercutting the photoresist. If the depth of the etch is more than the thickness of the film to be deposited, the metal film deposited on the dielectric film never touches the photoresist.

When the etch depth and film thickness do not vary much, the trenches etched are practically refilled by the metal film, which all but planarises the surface. This is clearly observed in Figure 4.19b. The major restriction to undercutting is that it can not be applied in the case the metal film is in direct contact with interconnect layers as the etch would attack the interconnect.

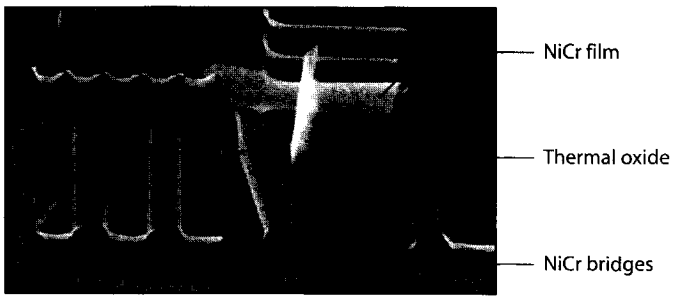


Figure 4.18 Step coverage by NiCr over HPR504 photoresist.

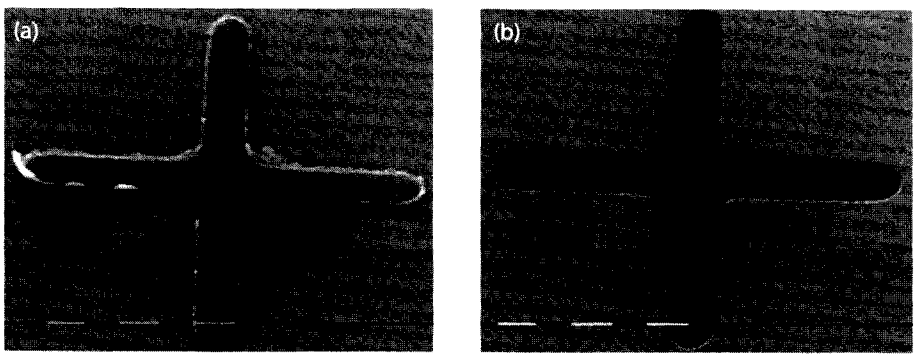


Figure 4.19 NiCr films after lift-off (a) using ultrasonic agitation and (b) using undercutting of the photoresist.

Measurement results

The measurement performed indicate a $Ni_{0.5}Cr_{0.5}$ resistivity of 120 μΩcm. As no passivation layer was used, thermal cycling was applied to promote ageing. After ten cycles between 25 °C to 145 °C, with 2 h per cycle, a TCR of $4.0 \cdot 10^{-5}$ K^{-1} was obtained. This is factor of 2 higher than reported elsewhere. The most likely explanation is that ageing is not yet completed. In [4.50] a thermal treatment for the duration of 24 h at 150°C is proposed. This puts the resistor at a higher temperature for a much longer period of time.

The work described above clearly establishes the applicability of the thin-film deposition module. An improved thermal anneal should reduce the TCR to less than $2.0 \cdot 10^{-5}$ K^{-1}. Klonz *et al.* [4.51] have even described the deposition of $Ni_{0.45}Cr_{0.5}Si_{0.05}$ films with a TCR below $1 \cdot 10^{-6}$ K^{-1}.

4.7 REFERENCES

[4.1] G.C.M. Meijer, G. Wang and F. Fruett, "Temperature sensors and voltage references implemented in CMOS Technology," *IEEE Sensors Journal*, Vol. 1, No. 3, pp. 225–234, 2001.

[4.2] A. Bakker, "High-accuracy CMOS smart temperature sensors," *Ph.D. Thesis*, Delft University of Technology, 2000, 132 pp. ISBN: 90-901-3643-6.

[4.3] www.foveon.com/X3_tech.html

[4.4] G. de Graaf and R.F. Wolffenbuttel, "CMOS sensor system for measuring environmental light conditions,"*Proc. Eurosensors XVI*, 15–18 September 2002, Prague, Chech Republic, pp. 415–418.

[4.5] J.F.L. Goosen, P. French and P.M. Sarro, "Pressure, flow and oxygen saturation sensors on one chip for use in catheters," *Proc. MEMS 2000*, 23–27 January 2000. Miyasaki, Japan, pp. 537-540.

[4.6] S. Kordic, *Offset reduction and three-dimensional field sensing with magnetotransistors*, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 1987.

[4.7] R.F. Wolffenbuttel, "Silicon photodetectors with a selective spectral response," *Sensors Update*, Vol. 9, Iss. 1, pp. 69–101, 2001.

[4.8] J.H. Correia, G. de Graaf, M. Bartek and R.F. Wolffenbuttel, "A CMOS optical microspectrometer with light-to-frequency converter, bus interface, and stray-light compensation," *IEEE Trans. on Instrumentation & Measurement*, Vol. 50, No. 6, 2001, pp. 1530–1537.

[4.9] J.F.L. Goosen, B.P. van Drieënhuizen, P.J. French and R.F. Wolffenbuttel, "Etching problems in the presence of aluminium interconnect," *Proc. Eurosensors X*, 8–11 September, 1996, Leuven, Belgium, pp. 271-274.

[4.10] J.F.L. Goosen, *Design and fabrication of a surface micromachined positioning device*, Ph.D. Thesis, Delft University Press, Delft, The Netherlands, 1996.

[4.11] J.A. Foerster, *Integrated micro vacuum tubes in silicon*, Ph.D. Thesis, Delft University Press, Delft, The Netherlands, 1996.

[4.12] I. Kondo, T. Yoneyama and O. Takenaka, "Formation of high adhesive and pure Pt layers on TiO_2," *J. Vac. Sci. Technol. A*, Vol. 10, No. 6, pp. 3456–3459, 1992.

[4.13] P.T.J. Gennissen, *Micromachining techniques using layers grown in an epitaxial reactor*, Ph.D. Thesis, Delft University Press, Delft, The Netherlands, 1999, ISBN 90-407-1843-1.

[4.14] N. Tas, T. Sonnenberg, H. Jansen, R. Legtenberg and M. Elwenspoek, "Stiction in surface micromachining," *J. Micromech. Microeng.*, Vol. 6, pp. 385–397, 1996.

[4.15] G.K. Fedder, S. Santhanam, M.L. Reed, S.C. Eagle, D.F. Guillou, M.-C. Lu and L.R. Carley, "Laminated high aspect ratio microstructures in a conventional CMOS process," *Proc. MEMS 1996*, 11–15 February 1996, San Diego, CA, USA, pp. 13–18.

[4.16] A.J.M.M. van Zutphen, *High-temperature CVD silicon on ceramic substrates for solar cell applications*, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 2001, ISBN 90-90144-77-3.

[4.17] M. Bartek, *Selective epitaxial growth for smart silicon sensor applications*, Ph.D. Thesis, Delft University Press, Delft, The Netherlands, 1995, ISBN 90-407-1178-X.

[4.18] D.P. Poenar, Thin film colour sensors, Ph.D. Thesis, Delft University Press, Delft, The Netherlands, 1996, 90-407-1343-X.

[4.19] M. Sekimoto, H. Yoshihara, and T. Ohkubo, "Silicon nitride single-layer X-ray mask," *J. Vac. Sci. Technol.*, Vol. 21, pp. 1017–1021, 1982.

[4.20] S. Rojas, A. Modelli, W. Wu, A. Borghesi and B. Pivac, "Properties of silicon dioxide films prepared by low-pressure chemical vapor deposition from tetraethylorthosilicate," *J. Vac. Sci. Technol.*, Vol. B8, Iss. 6, pp. 1177–1184, 1990.

[4.21] L. Tedder, G. Lu an J. Crowell, "Mechanistic studies of dielectric thin film growth by low pressure chemical vapor deposition: the reaction of tetraethylorthosilicate with SiO_2 surfaces," *J. Appl. Phys.*, Vol. 69, Iss. 10 pp. 7037–7049, 1991.

[4.22] M. IslamRaja, C. Chang, J. McVittie, M. Capelli and K. Saraswat, "Two precursor model for low-pressure chemical vapor deposition of silicon dioxide from tetraehtylorthosilicate," *J. Vac. Sci. Technol.*, Vol. B11, pp. 720–726, 1993.

[4.23] K.A. Pickar "Ion implantation in silicon," in R. Wolfe, ed., *Applied solid state science*, Vol. 5, Academic Press, New York, NY, USA, 1975.

[4.24] S.M. Sze, *Semiconductor devices, physics and technology*, John Wiley & Sons, New York, NY, USA, 1985, ISBN: 0-471-83704-0.

[4.25] T. Kamins, *Polycrystalline silicon for integrated circuit applications*, Kluwer Academic Publishers, Boston, MA, USA, ISBN 0-89838-259-9.

[4.26] H. Ohji, *Macroporous silicon based micromachining*, Ph.D. Thesis, Eburon Publishers, Delft, The Netherlands, 2002, ISBN 90-5166-901-1.

[4.27] C.M.A. Ashruf, *Galvanic etching of silicon*, Ph.D. Thesis, Delft University Press, Delft, The Netherlands, 2000, ISBN 90-407-2001-0.

[4.28] S.D. Collins, "Etch stop techniques for micromachining," *J. Electrochem. Soc.*, Vol. 144, 1997, pp. 2242–2262.

[4.29] P.M. Sarro, S. Brida, C.M.A. Ashruf, W. van der Vlist and H. van Zeijl, "Anisotropic etching of silicon in saturated TMAHW solutions for IC-compatible micromachining, *Sensors and Materials*, Vol. 10, pp. 201–212, 1998.

[4.30] H.A. Waggener, "Electrochemically controlled thinning of silicon," *Bell Sys. Tech. J.*, Vol. 49, 473–475, 1970.

[4.31] J. Hopwood, "Review of inductively coupled plasmas for plasma processing," *Plasma Sources Sci. Tech*, Vol. 1, pp. 109–116, 1992.

[4.32] D.M. Manos and D.L. Flamm, *Plasma Etching: An Introduction*, Academic Press, New York, NY, USA, 2000, ISBN 0-124-69370-9.

[4.33] M. Sugawara, *Plasma Etching*, Oxford Science Publications, New York, NY, USA, 1998, ISBN: 0-198-56287-X.

[4.34] E.H. Klaassen, R.J. Reay, C. Storment and G.T.A. Kovacs, "Micromachined thermally isolated structures," *Sensors and Actuators*, Vol. A58, pp. 43–50, 1997.

[4.35] T. Akin, Z. Olgun, O. Akar and H. Kulah, "An integrated thermopile structure with high responsivity using any standard CMOS process," *Sensor and Actuators*, Vol. A66, pp. 218–224, 1998.

[4.36] C.-H. Du and C. Lee, "3D thermoelectric structures derived from a new mixed micromachining process," *Jpn. J. Appl. Phys.*, Vol. 39, No. 12B, pp. 7125–7129, 2000.

[4.37] P.M. Sarro, A.W. van Herwaarden and W. van der Vlist, "A silicon–silicon nitride membrane fabrication process for smart thermal sensors," *Sensors and Actuators*, Vol. A41–42, pp. 666–671, 1994.

[4.38] P. van Gerwen, *Microsystems technology for integrated biochemical analysis systems*, Ph.D. Thesis, Katholieke Universiteit Leuven, 2000, ISBN 90-5682-249-7.

[4.39] S. Sakarya, *Micromachining techniques for fabrication of integrated light modulating devices*, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 2003, ISBN: 90-6734-196-7.

[4.40] J.L. Regolini, F. Gisbert, G. Dollino, and P. Boucaud, "Growth and characterization of strain compensated $Si_{1-x-y}Ge_xC_y$ epitaxial layers," *Mater. Lett.*, Vol. 18, Iss. 1–2, p. 57, 1993.

[4.41] P. Warren, M. Dutoit, P. Boucaud, J.-M. Lourtioz, and F. H. Julien, "RTCVD growth and characterization of SiGeC multi-quantum well," *Thin Solid Films*, Vol. 294, Iss. 1–2, pp. 125–128, 1997.

[4.42] C. Guedj, X. Portier, A. Hairie, D. Bouchier, G. Calvarin and B. Piriou, "Carbon and germanium distribution in $Si_{1-x-y}Ge_xC_y$ layers epitaxially grown on Si(001) by RTCVD," *Thin Solid Films*, Vol. 294, Iss. 1–2, pp. 129–132, 1997.

[4.43] M.-C. Hsieh, Y.-K. Fang, P.-M. Wu, C.-C. Yang, Y.-C. Lin, W.-D. Wang, S.-F. Ting and and J.-J. Ho, "Design and fabrication of a novel crystal SiGeC far infrared sensor with wavelength 8–14 micrometer," *IEEE Sensors Journal*, Vol. 2, No. 4, pp. 360–365, 2002.

[4.44] D. Wijngaards, M. Bartek and R.F. Wolffenbuttel, "Silicon IC process compatible thin metal film post-processing module," *Sensors and Actuators*, Vol. A68, pp. 419–428, 1998.

[4.45] B.F. Field, "A sub-PPM automated 1-10 Volt DC measuring system," *IEEE Trans. on Instr. and Meas.*, Vol. 34, No. 2, pp. 327–330, 1985.

[4.46] Y. Sakamoto, T. Endo and T. Sakuraba, "A 10 V calibrating system using an automated voltage divider and a 1V Josephson array," *IEEE Trans. on Instr. and Meas.*, Vol. 42, No. 2, pp. 583–587, 1993.

[4.47] S. H. Tsao, "A 25-bit reference resistive voltage divider," *IEEE Trans. on Instr. and Meas.*, Vol. 36, No. 2, pp. 285–290, 1987.

[4.48] C.J. Smithells, *Metals reference book*, Butterworth Scientific Publications, London, 5th edn., 1979.

[4.49] F. Hegner, "The industrial production of high quality nickel-chromium resistors with controlled temperature coefficient of resistance," *Thin solid films*, Vol. 57, pp. 359–362, 1979.

[4.50] A. García-Alonso, J. García, E. Castaño, I. Obieta and F.J. Gracia, "Strain sensitivity and temperature influence on sputtered thin films for pierzoresistive sensors," *Sensors and Actuators*, Vol. A37–38 pp. 784–789, 1993.

[4.51] M. Klonz, H. Laiz and E. Kessler, "Development of thin-film multijunction thermal converters at PTB/IPHT," *IEEE Trans. on Instrum. and Meas.*, Vol. 50, No. 6, pp. 1490–1498, 2001.

Chapter 5

APPLICATIONS

In this fifth chapter, thermoelectric applications are classified in a somewhat fundamental way, which allows a better distinction between the different modes of operation, and between the parameters of importance for that particular mode of operation. First, the macroscopic model of the on-chip integrated thermoelectric cooler is introduced. Such a thermal network allows the electrical and thermal domains to be directly merged. Thereafter, the generic mode of operation are discussed. These are cooling (Sections 5.2), thermal stabilisation / active thermal feedback (Section 5.3) and heat spreading (Section 5.4). Finally, in Section 5.5 the use of polySiGe in thermal sensor applications is discussed.

5.1 MACRO-SCALE OPERATION OF AN INTEGRATED TEC

When going through literature, the Peltier device is most frequently referred to as a thermoelectric cooler (also in this thesis.) This emphasises the thermal energy removal at the cold junction and treats the rest of the system as parasitic. Other sources justifiably argue that the Peltier device should be seen as a regulated thermal impedance [5.1]. The term 'controlled thermal impedance' suggests the thermal resistance K^{-1} is the dominant factor in the device, where the total number of 'K/W' can be controlled by means of the Peltier effect. Finally, the Peltier device is regularly referred to as a heat pump or thermal energy pump [5.2]. The term 'heat pump' emphasis the reversibility of the Peltier effect (and the device), which I believe is the most important aspect of the system. As such, when one term must be identified as generic, the term 'heat pump' would be the most correct of the three. Still, all three designation are correct. In the end it depends on the application which one suits best.

5.1.1 Lumped-element approach

The easiest way to understand how a Peltier device interfaces with its surroundings is by means of a lumped-element model, also referred to as a thermal network. Recalling Equation (2.15), the rate with which thermal energy is removed from the cold junctions of a conventional Peltier element is given by

$$q_c = (\alpha_p - \alpha_n) I_{el} T_c - K(T_h - T_c) - \tfrac{1}{2} I_{el}^2 R \tag{5.1}$$

Thus, three effects are incorporated:

- Thermal energy is removed from the cold junction at a rate q_c, equal to $(\alpha_p - \alpha_n) I_{el} T_c$. This parameter behaves as a thermal 'current' source controlled by both the electrical current as well as the cold junction temperature. The Seebeck coefficients merely behaves like a scaling factor.
- Thermal energy is transferred from the hot junction area at a rate $K(T_h - T_c)$. This passive effect is modelled most easily by its reciprocal, the thermal resistance.
- Within the device, thermal energy is generated at a rate $I_{el}^2 R$. Half of the energy generated diffuses to the cold junctions while the other half diffuses to the hot junctions. (See Section 2.2.1 for an explanation of this behaviour.) As energy is generated by this effect, it should be modelled by a thermal current source, controlled by both I_{el} and R.

A similar thermal energy removal rate q_h is obtained at the hot junction. The three effects are combined in the simple model of Figure 5.1. This way, the Peltier device can be modelled as a two-port element, with I_{el} as input and $\Delta T = (T_h - T_c)$ as output. This model still has a few shortcomings. To obtain a model that approximates the real physical behaviour of the integrated TEC discussed in Chapter 2, the following aspects need to be addressed:

- Within the Peltier device, the electrical contact resistance at both the hot and cold junctions causes Joule heating q_{Rc} at a rate $I_{el}^2 R_c$.
- The regions in the vicinity of the cold and hot junctions have a certain thermal capacitance. In the integrated TEC, the thermal capacity near the cold junction $C_{th,c}$ will be rather small, (i.e. from a part of the volume of the TEC plus the region at the tip of the beam, like an epi-island.) The hot junction is located directly over the substrate, which has a much larger volume and, as a result, will also have a much larger thermal capacitance $C_{th,h}$.
- There is a thermal (contact) resistance between the cold junction and the region to be cooled/stabilised, R_{p-c}, and the hot junction and the ambient, R_{p-h}. Thermal energy generated at the junctions first needs to cross both the polySiGe and the Si_3N_4 before reaching the respective areas. Similarly, there is a certain thermal resistance from both the cooled region and the region below the hot junction to ambient. These are referred to as R_{c-a} and R_{h-a}.
- There is radiant heat exchange between the membrane and the surroundings and, if not operated in vacuum, there will be thermal conduction and convection through the surrounding air.This makes R_{c-a} and R_{h-a} non-linear.

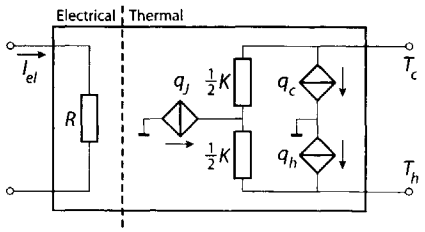


Figure 5.1 Simple electro-thermal network of a Peltier device.

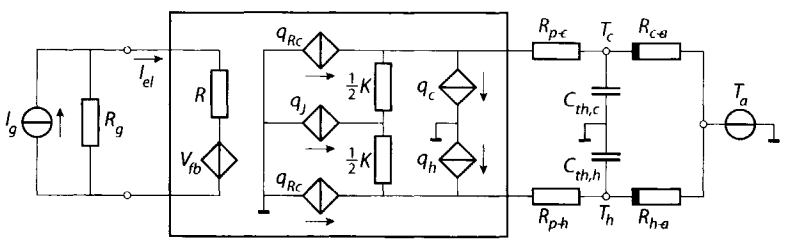


Figure 5.2 Enhanced electro-thermal network of a Peltier device.

Incorporating these effects into the model of Figure 5.1, the enhanced model of Figure 5.2 arises. In this enhanced model, T_a is the ambient temperature, modelled by a 'thermal' voltage source. This above model was published [5.3] (in slightly different form) at virtually the same time as the model by Chávez *et al.* [5.4]. The resemblance between the models is striking. The main difference lies in the 1) the inclusion of electrical contact resistance in our model and 2) the way of incorporating electrical Joule heating q_J into the model. I have chosen a single power source, which power flows away equally to the nodes at T_c and T_h by means of a thermal conductance $K/2$. They chose to use a single conductance K in between T_c and T_h, while using two power sources of magnitude $q_J/2$.

In both [5.3] and [5.4], the models were depicted as a three-port element. The above models are a modification in the sense that the TEC is now modelled as a two-port element. This is strictly a matter of interpretation. In the above model the thermal 'earths' are kept inside the model; In the older models the 'earths' are included as separate terminals. The philosophy behind this modification is that thermal energy transferred at the junctions comes from or goes to within the device.

The work performed by Chávez *et al.* was directed towards modelling of conventional Peltier elements. This gave them the opportunity to quickly verify the model using off-the-shelf TECs. Cycling over a temperature range of nearly 32 °C, their model deviated less than 0.3 °C. This is a clear indication of the accuracy that can be obtained with a thermal network.

Further thermal macro-models of thermoelectric coolers are found in [5.5] and [5.6], for example. In particular the latter reference is of some interest as it refers to influence by the Thomson effect and temperature dependency of thermoelectric material parameters on the performance of planar integrated TECs.

5.1.2 Thermoelectric feedback

In the model of Figure 5.2, one component has remained undiscussed. Due to the temperature difference generated by the TEC, a Seebeck voltage V_{fb} is fed back across the input terminals of the TEC. This V_{fb} has a magnitude equal to $(\alpha_p-\alpha_n)(T_h-T_c)$. If a good electrical current source I_g is applied (i.e. $R_g \rightarrow \infty$) all of the electrical current is forced through the TEC so V_{fb} does not influence the input current. In this case, a power $I_g \cdot V_{fb}$ ($= I_{el} \cdot V_{fb}$) is dissipated, which is exactly equal to q_h-q_c:

$$\begin{aligned} q_h - q_c &= (\alpha_p - \alpha_n) I_{el} T_h - (\alpha_p - \alpha_n) I_{el} T_h \\ &= I_{el}(\alpha_p - \alpha_n)(T_h - T_c) = I_{el} V_{fb} \end{aligned} \tag{5.2}$$

In Figure 5.3 the typical influence of the current source's output impedance on device performance is displayed. This R_g is in parallel with the resistance R of the TEC. As such, the relationship between I_{el} and function of I_g is

$$I_{el} = \frac{R_g}{R+R_g} I_g \tag{5.3}$$

Thus, for very small values of R_g, where the current source behaves more like a voltage source, the required shift in I_g to maintain the right I_e is considerable. This can be seen from Figure 5.3 as follows: when R_g increases, the value for I_g where $(T_h-T_c)/T_h$ peaks shifts back towards the value expected for an idealised integrated TEC. (The numbers along the axes are based on typical material values.)

In conclusion, for the Seebeck voltage at the input of the TEC to be suppressed, any TEC must be driven by a current source, never by a voltage source (using R to estimate the current I_{el}.)

5.2 COOLING (NEAR AND BELOW AMBIENT TEMPERATURE)

The primary role of a Peltier device undoubtedly is that of a cooler. Classical applications range from cooling of beer cans and electronics [5.7] (using single-stage TECs) to the freezing of warts and other dermatological conditions [5.8] (using triple-stage TECs). Regarding the lateral on-chip integrated Peltier devices the obvious application is cooling of single or multiple (micromachined) volumes within a single chip [This thesis][5.9][5.10]. The transversal TEC (like the cross-plane superlattice structures discussed in Section 2.3 and Appendix B) is primarily targeted at cooling of non-micromachined volumes, e.g. the VCSELs (Vertical Cavity Surface Emitting Lasers) discussed in [5.11]–[5.13].

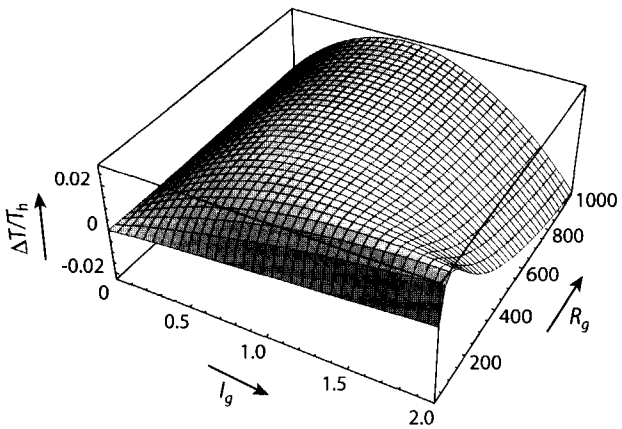


Figure 5.3 Influence of a current source with finite output impedance on the optimal electrical driving current.

In this section the actual cooling performance of a polySiGe-based lateral TEC is assessed. Figure 5.4 shows a rendered image of the devices in question. The beam holds two pairs of n-polySiGe/p-polySiGe thermoelectric legs. These pairs can be operated independently (in case the characteristics of both pairs are different.) Between the thermoelectric legs lie 2 resistors. The first, a two-terminal resistor lies exactly in the centre of the beam and can be used to bias the temperature at the cold junction. The second resistor consists of two resistors in series, both in between the 2-terminal resistor and the thermoelectric legs. This resistor is configured as a 4-terminal resistor, so that this resistor can be used to read out the temperature in the central region of the device.

The steady-state temperature reduction obtained at the centre of the beam is plotted in Figure 5.5. At an electrical current of a mere 145 μA, cooling to 2.1 K below ambient temperature is obtained, at a reduced pressure of 1.7 Pa. This value is still rather low as, during the design of the structure the error was made to ignore the Al wires. The current carrying arms to the resistors are necessarily made of aluminium. The proportionally large thermal conductance through the arms accounts for about 60% of the total thermal conduction from the central region of the beam to the surrounding substrate. Thus, if the arms including the metal wires are removed, the temperature difference would increase from 2.1 K to 5.3 K. For the device in Figure 5.4, the analytical model presented in Chapter 2 predicts a ΔT_{max} of 5.5 K. This is very close to what is expected when the arms were to be removed.

The measurement results clearly indicate that thermal leakage through metallic interconnect is the major source of performance degradation in lateral on-chip integrated TECs.

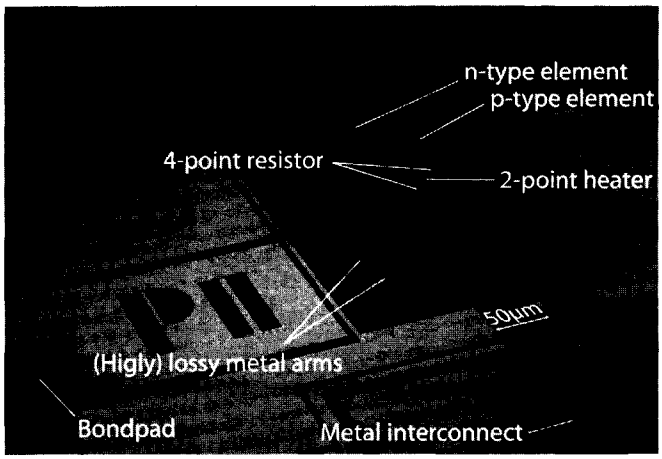


Figure 5.4 Rendered image of the lateral on-chip integrated structure used to characterise the cooling performance of polySiGe TECs.

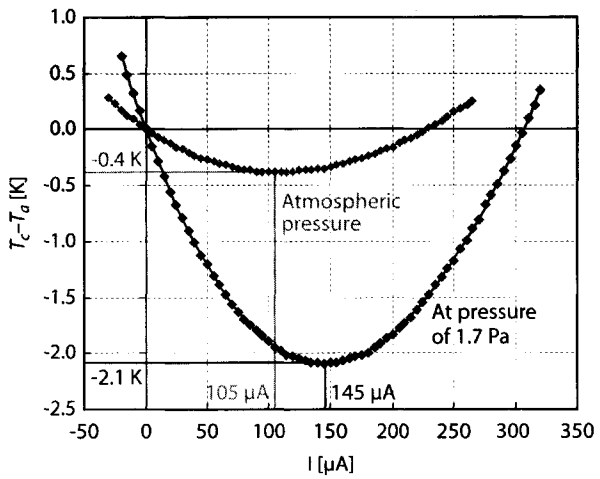


Figure 5.5 Plots of the cooling measured using the structure in Figure 5.4. In a vacuum, cooling to 2.1 K below ambient temperature is established. This value is expected to increase to 5.3 K if thermal leakage of the aluminium wires is eliminated.

5.3 THERMAL STABILISATION AND ACTIVE FEEDBACK

Temperature control is one of the most fundamental and most essential operations encountered in many physical systems. Active temperature control, as compared to passive, is capable of accurately and rapidly bringing the intended temperature

above, below or in between preset limits. These are set to either ensure safety (e.g., above 42 °C the proteins in the human body start to disintegrate) or to optimise the performance of a system (like the speed of a chemical reaction.) In particular in biomedicine thermal control systems are encountered frequently, for example in electrophysiology [5.14] radiopharmaceutical synthesis [5.15], microbial studies [5.16], rapid thermal cycling of cells [5.17] as well as DNA sequencing [5.18].

Most of these applications will benefit from miniaturisation if accurate micro-scale thermal control systems can be realised. At elevated temperatures (90 °C and above) such systems have to rely on passive cooling. These systems, generally referred to as micro-hotplates have already been implemented [5.19]–[5.22], e.g. for thermal stabilisation of bandgap references. However, in most applications, temperatures need to be controlled in the 10 °C – 50 °C range, for which integrated micro-thermostats are not yet available. Such applications not only include the biomedical examples mentioned above, but also dew-point sensors [5.23]–[5.25] and even electronics [5.7][5.26].

A special class of sensors that can benefit from the availability of a micro-thermostat are the thermally based sensors that can be operated in an active thermal feedback configuration. These include humidity sensors [5.27], IR detectors [5.28], micro-calorimeters [5.29] and flow sensors [5.30]–[5.32]. Conventionally, the temperature of or temperature difference in such a device is used as indicator for the magnitude of the effect observed. Using active thermal feedback, it is not the temperature, but the amount of power dissipated by the device to keep the temperature or temperature difference constant that serves as indicator. A thermal feedback measurement can be made more accurate than a direct measurement, as the first can be configured as a null-balance circuit. The work in this thesis is aimed at these applications, which require thermal stabilisation near ambient temperature.

Thermal stabilisation of micro-scale and/or micromachined applications near ambient temperature only becomes feasible when an acceptable response time is obtained. This can only be achieved if a micro-hotplate is also equipped with a cooler, to actively ramp down the temperature. Of the different cooling techniques available, the most suitable cooling device to fulfil this roll is the TEC. Due to its solid-state operation and single-wafer construction there are very few packaging restrictions in comparison to fluidic solutions. Moreover, the work in this thesis shows that a micro-thermostat can be constructed using fabrication compatible thin-film technology only, so that it can be readily co-integrated with electronics.

5.3.1 Transient response

Besides the steady-state response measured in the previous section, it is important to understand the transient behaviour of the thermoelectric cooler. Again, a lumped-element model is indispensable for such an analysis. As all structures are designed to be highly symmetrical, the one-dimensional model of Figure 5.6 is adequate.

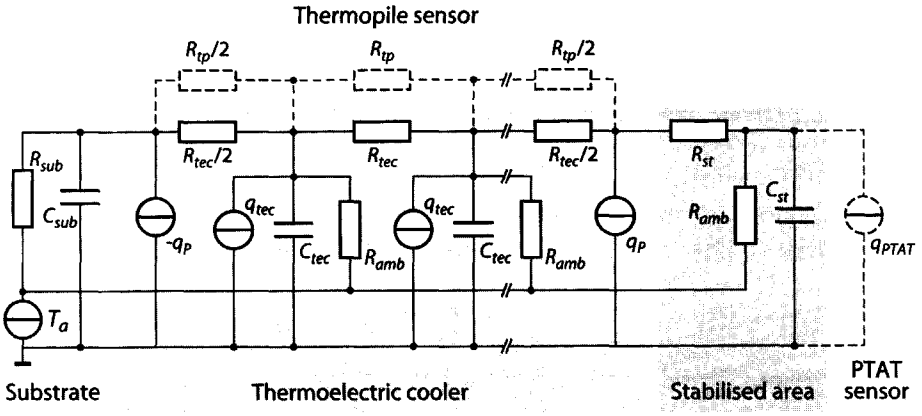


Figure 5.6 Simple thermal network to model the transient behaviour of lateral on-chip integrated thermoelectric coolers.

As the substrate has a large volume compared to the thin-film section, it is sufficient to model the substrate with a single RC-section (R_{sub}, C_{sub}). According to [5.33], C_{sub} may be left out altogether. Similarly, the assumption is made that the volume to be thermally stabilised has such a shape that this area can be approximated using R_{st} and C_{st} only. The exact error introduced by such an assumption can only be verified by means of FEA.

With this model, in particular the influence of the thermoelectric element is studied. Its thermal conduction is modelled by resistors designated R_{TEC}. Besides this, the thermal capacitance of and Joule heating in the TEC are distributed by means of the capacitors C_{TEC} and thermal current sources q_{TEC}. The Peltier effect, which manifests itself at the hot and cold junctions (see Fig. 2) is modelled by the sources q_P

This model can be extended easily, for example if a PTAT circuit or any other power-consuming circuits are integrated in the cooled / thermally stabilised region. Such a source will thermally load the stabilised region, as indicated by the thermal current source q_{PTAT} (plus a correction term to R_{TEC}, as metallic wires have to b used to contact such devices.) In contrast, a thermopile sensor to sense the temperature difference between the cooled region and the substrate does not dissipate any power, and can be modelled using the resistors designated R_{TP} (together with a correction term to C_{TEC}.) Finally, to model the losses to ambient by means of radiant and convective heat exchange, (non-linear) the resistors R_{amb} can be employed.

The transient response to a stepwise change in electrical current through the heater at the centre of the beam is plotted in Figure 5.7. Besides the thermal conductivities of the material applied, also the volumetric heat capacity needs to be known. For Si_3N_4, the heat capacity varies from $1.7 \cdot 10^6$ $Jm^{-3}K^{-1}$ to $2.0 \cdot 10^6$ $Jm^{-3}K^{-1}$, while this property varies between $1.5 \cdot 10^6$ $Jm^{-3}K^{-1}$ and $2.3 \cdot 10^6$ $Jm^{-3}K^{-1}$ for SiO_2; See references

[3.44][5.34][5.35]. For pure Al, the volumetric heat capacity is around $2.4 \cdot 10^6$ $Jm^{-3}K^{-1}$ [5.36] while the reported heat capacity for polySiGe is $1.7 \cdot 10^6$ $Jm^{-3}K^{-1}$ [5.37]. As these thermophysical parameters show a significant spread, an average thermal conductivity of 3 $Wm^{-1}K^{-1}$ and an average volumetric heat capacity of $2 \cdot 10^6$ $Jm^{-3}K^{-1}$ [5.33] are used. The model predicts the device as shown in Figure 5.4 has a time constant of needs only 7 ms to settle within a 1% inaccuracy of the final temperature. This shows the micro-thermostat has the potential to be extremely fast. Furthermore, the influence of dividing the TEC into multiple RC-sections has been investigated. This only has some effect directly after the stepwise change is applied, but hardly influences the time constant at all.

Some simple experiments have been performed to determine the time constant of an actual on-chip integrated TEC. In the experiment, the two-terminal resistor is used to raise the temperature of the central region of the structure in Figure 5.4 by means of Joule heating. The Seebeck voltage that builds up across the arms of the Peltier device as a result is monitored using an oscilloscope. The response for the lateral integrated TEC in air is depicted in Figure 5.8. The time constant of the TEC is only 2 ms. The time to settle within 1% of the final temperature is 9 ms, slightly larger than predicted from the thermal network of Figure 5.6.

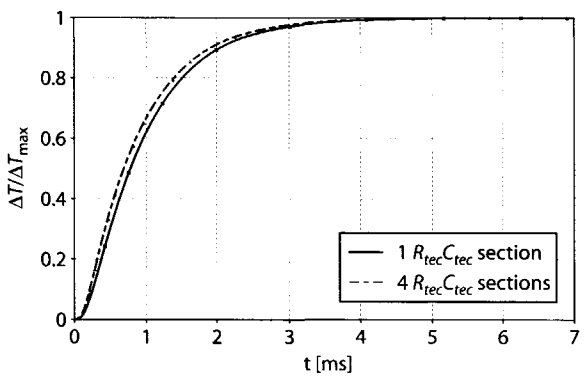


Figure 5.7 Simulated transient response of the device in Figure 5.4, without the Al wires.

5.3.2 Overall micro-thermostat design

In Figure 5.9, the schematic configuration of an on-chip integrated micro-thermostat is displayed. The actual temperature of the thermally isolated structure is read out and compared to the required temperature. In case of a discrepancy between the two values, heat is either added or removed from the isolated volume, based on the magnitude and sign of the discrepancy.

Three of four major system components of the micro-thermostat are located on the thermally isolated membrane: the cooler, heater and temperature sensor. The

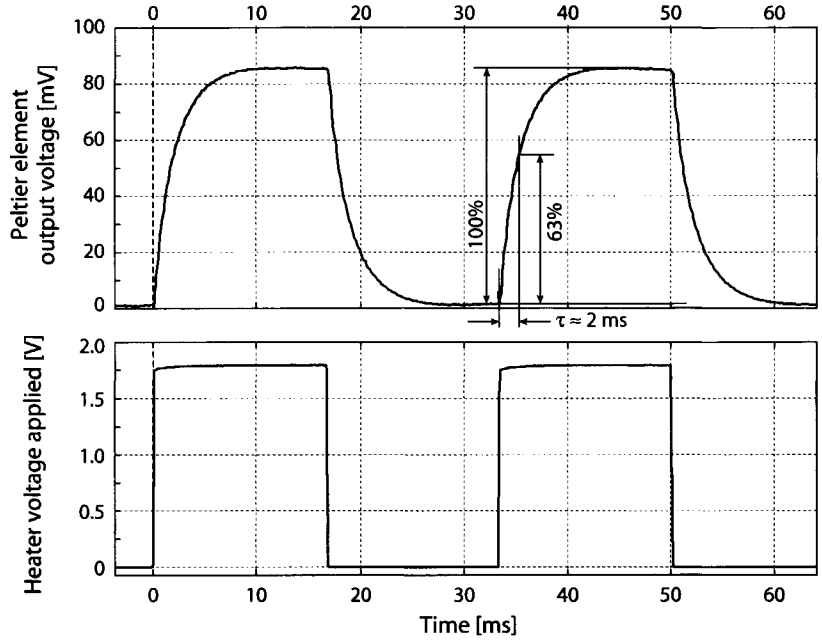


Figure 5.8 Transient response of the lateral on-chip integrated TEC shown in Figure 5.4.

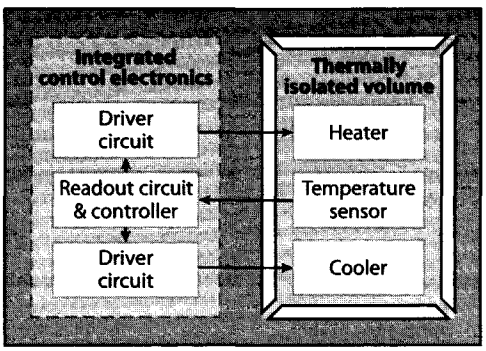


Figure 5.9 Schematic diagram of the basic configuration of a micro-thermostat.

choice of cooler is undisputed, i.e., a TEC will be used to implement this functionality. The remaining concerns are how to implement the heater and the temperature sensor. These considerations are discussed directly below. The implementation of read-out and driving electronics of the micro-thermostat system will be discussed in following subsection.

Choice of heater

The obvious choice for a heater is a (thin-film) resistor. Such a resistor can be integrated directly on top of (or in) the region to be thermally stabilised, so all of the Joule heat generated is used to increase the temperature of that region. So, in terms of power efficiency resistors are a good choice.

Either an integrated resistor (embedded in an epi-island suspended from a membrane) or a deposited thin-film resistor can be applied. Thin film resistors—in particular metallic ones—are preferred for two reasons. First, the properties of a metal resistor are practically process-independent. Second, the temperature coefficient of resistance (TCR) of thin metal-film alloys can be brought back sub-ppm values, similar to the NiCr resistors discussed in the previous chapter. Due to the extremely low variation in resistance over a very large temperature range, the heating power is virtually constant over temperature. Moreover, even though alloys are non-standard in terms of IC fabrication, they can be readily co-integrated with electronics using simple post-processing, as discussed in Section 4.4.

The main drawback of a resistor is that it requires low-Ohmic—and thus metallic—leads. As was pointed out in the previous section, the interconnect is the largest source of thermal losses. The resulting increase in parasitic thermal conduction significantly decreases the temperature range of the micro-thermostat as the lower temperature boundary shifts upward.

A second, less conventional heating technique is application of the TEC by reversing the direction of the electric current, operating the device as a thermoelectric heater. The appeal of this solution is that a single element is used for active heating and active cooling alike. This not only reduces the number of components required, but most importantly, no additional metal leads are required, so thermal leakage is reduced to an absolute minimum.

The main drawback of heating by means of a thermoelectric element is that the power efficiency is lower. In Chapter 2 it was proven that only half of the Joule heat generated within a thermoelement contributes to a rise in temperature at the membrane. This implies that together with the heat transported by the Peltier effect (at the junctions) only 60-70% of all power is used for heating. So, in terms of power efficiency, a resistor is preferable over thermoelectric heating. However, this is greatly outweighed by the fact that the thermoelectric cooler/heater does not increase the parasitic thermal conductance, so the lower operating range of the micro-thermostat remains intact. As the operating range of the thermostat is the dominant design criterion, the TEC is the preferred heater element.

Choice of temperature sensor

In order to control the amount of power released at or withdrawn from the thermally stabilised region, an accurate temperature sensor is required. For the micro-thermostat to be flexible, the thermally stabilised volume should be allowed to settle at any user-specified temperature within the operating range. For this purpose, the

system must be capable of measuring the absolute temperature of the thermally stabilised region.

When considering an absolute temperature measurement, a widely employed sensor system is the proportional to absolute temperature (PTAT) circuit. When used in conjunction with advanced offset- and noise-reduction techniques, an inaccuracy as low as 0.7 K can be achieved, with a repeatability and long-term drift of 0.1 K, and a noise level as low as $5{\cdot}10^{-3}$ K [5.38]. The major advantage of using a PTAT circuit directly at the thermally stabilised region is that this provides a spot-on absolute temperature reading. On the other hand, the PTAT structure has to carry an electrical current and, therefore, dissipates power. Combined with the fact that a minimum of three metal connections are required to read out the circuit, a significant decrease in the operating range is the result. This does not favour a dedicated PTAT circuit (unless metal leads are unavoidable anyway, e.g., in the case a similar electronic circuit, like a bandgap reference is the component to be thermally stabilised.)

A second absolute temperature sensing technique is by means of resistors, which can be read out using a 4-terminal configuration. However, for a high sensitivity, semiconductor resistors are required, which are non-linear and show a strong process-dependency. In contrast, a metal resistor has a more reproducible response, but due to the smaller TCR and resistance large electrical currents are required for a good signal level. This may induce significant self-heating, specifically as the resistor is located on a thin-film membrane. Again, metallic arms are required for read-out, decreasing the temperature range of the micro-thermostat.

A third option is the application of thermopiles. These are self-generating, have zero self-heating (if read out properly) and exhibit zero offset. A minor drawback is found in the fact that thermopiles sense a temperature difference, rather than an absolute temperature. Therefore, the absolute temperature of the stabilised region can only be determined in conjunction with an absolute temperature sensor in the substrate. Thermocouples are placed in parallel with the thermoelectric cooler so there will thermal leakage. Still, the thermal losses through a thermopile are nowhere near as severe as through metallic wires.

As fourth and final option, the thermoelectric cooler can also be used as thermopile. This has the advantage that there are no additional thermal losses between the membrane and ambient, so the thermal operating range remains uncompromised. On the other hand, the thermoelements are limited in number and are distributed all over the rim of the stabilised region. Therefore, the output signal will be rather small and can only provide the average temperature of the membrane (i.e., it is difficult to sense thermal gradients across the thermally stabilised region.) Finally, as the thermoelements cannot be used for sensing and heating/cooling at the same time, some form of time multiplexing is required. Depending on the magnitude of the duty cycle required for temperature sensing, a decrease in the operating range results, even though this will be the smallest of all four temperature-sensing techniques.

Concluding, from the standpoint of maximising the operating range, it is unfavourable to use an absolute temperature sensor within the stabilised region, due to the large thermal leakage. Instead, a two-stage measurement technique is preferred, where the absolute temperature is measured within the substrate and a differential temperature measurement is performed between the substrate and the stabilised region. This differential measurement can be performed either by a dedicated thermopile or by using the thermoelectric cooler/heater as thermopile.

As our aim is to extend the temperature range of operation to the lowest possible temperature, the TEC is chosen as sensing element. As a result, in the case a micro-thermostat with active cooling is designed, it is preferable to use the TEC for cooling, heating and sensing alike. Due to a lack of time, co-integration of the TEC with electronics has not yet been implemented. Nevertheless, the micro-thermostat design is finalised and some of the critical electronic components have already been fabricated and tested. The electronic part of the micro-thermostat is discussed below.

5.3.3 Driving- and read-out electronics

Based on the choices made in previous subsections, the schematic configuration shown in Figure 5.10 is chosen. On the left side of the figure, the top view of the micromachined thin-film structure is shown as it has actually been designed. The length and width of the thermoelectric arms suspending the central membrane are 250 μm and 50 μm respectively, which provides a central section with an area of 200×200 μm^2.

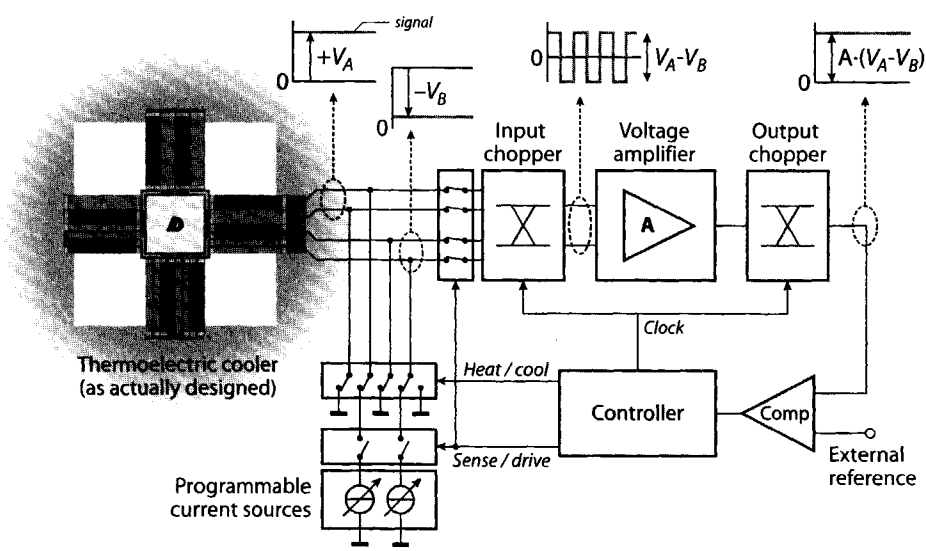


Figure 5.10 Graphical interpretation of the the micro-thermostat implementation targeted.

To implement heating, cooling and sensing with an array of Peltier elements only, two control signals are used in a feedback configuration. The first signal is used to enable switching between driving and sensing mode. The second signal is used in driving mode, to select whether the suspended membrane needs to be heated or cooled. The latter decision is made based on the signal coming from the comparator, which compares the output signal of the thermopiles to an external signal. This external signal is such that it provides the difference between the absolute temperature measured in the substrate and the thermal set point as specified by the user. Programmable current sources to drive the thermoelectric elements are preferable. First, this makes the electronic design independent of the thermoelectric element configuration. Second, this allows any fluctuation in the optimal driving current between similar devices to be compensated.

From the values found in Chapter 3, it follows a single Peltier element has a sensitivity of about 300 μVK^{-1}. The electronics are designed in the DIMES01 bipolar process, as only this process has been verified to work in conjunction with ECE. As a consequence of using a bipolar process (which generally shows a large voltage offset,) a special configuration is required for accurate read-out of the relatively small Seebeck voltages. First, the voltage from the TECs has to be chopped at the amplifier input to reduce the noise level. The resulting input signal, a square wave with an amplitude equal to the sum of all eight Peltier elements (depicted in Figure 5.10) is amplified, demodulated by the second chopper and low-pass filtered, thereby removing the amplifier offset. The resulting dc voltage is compared to the external signal in order to determine whether the suspended membrane needs to be heated or cooled.

In the following subsections a short description of the input voltage chopper, amplifier and comparator is provided. These circuits have been designed by Ger de Graaf, who I would like to thank at this point for his efforts.

Chopper circuit

Voltage chopping is best implemented using MOS transistors. However, as only bipolar transistors are available, chopping is done by means of a so-called balanced voltage chopper as described by Riedijk [5.39]. When a bipolar transistor is used in the normal mode, the saturation voltage (between the collector and emitter) can only decrease to some tens of millivolts. This voltage adds to the thermopile / TEC output signal. However, when driven in reverse mode, the saturation voltage can be reduced to as low as 0.3 mV. In reverse mode, the saturation voltage can be expressed as

$$V_{sat} = I_{base} R_{coll} - \frac{k_B T}{q} \ln\left(\frac{1}{1+\beta_F}\right) \tag{5.4}$$

with I_{base} the base current, R_{coll} the collector resistance and β_F the forward beta of the transistor. Injection of the base current into the signal path is avoided by always switching to ground. However, this requires a floating source, which generally is the case with a thermopile or TEC.

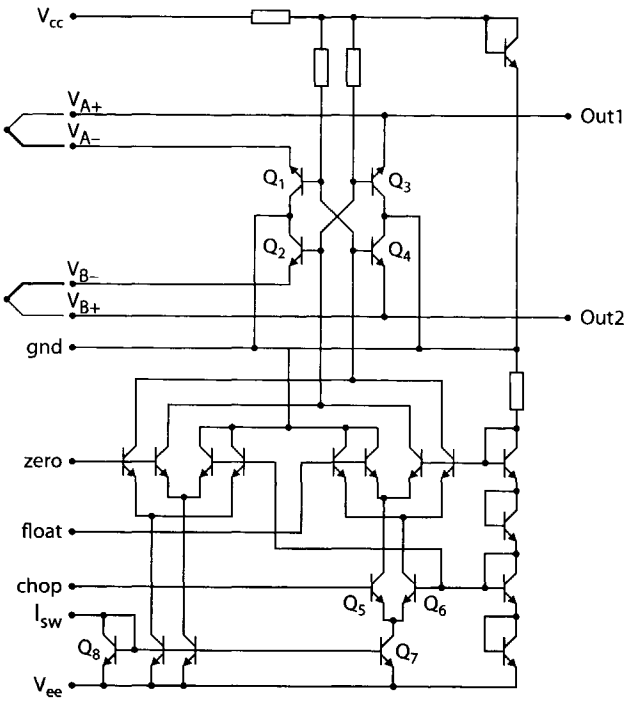


Figure 5.11 Circuit diagram of the balanced voltage chopper.

The details of the implemented circuit are shown in Figure 5.11. The signal 'I_{sw}' is used to switch the chopper on and off. (It needs to be switched off when the Peltier elements are electrically driven.) The terminals 'zero' and 'float' are respectively used for auto-zeroing and making the chopper float. the supplies rails V_{cc} and V_{ee} respectively are at a potential of +5 V and –5 V.

Due to the somewhat unusual operating principle of the chopper, two thermopile output voltages are presented at the input of the chopper. The chopper itself consists of transistors Q_1 to Q_4 that all have a quad layout to compensate for thermal gradients and to improve matching. When the 'chop' signal is high, transistor Q_5 conducts and only leaves transistors Q_1 and Q_4 in conduction. The potential across the output (Out1–Out2) now equals the positive Seebeck potential (V_{A+}–V_{A-}) across thermopile A. When Q_5 is turned off, the current mirrored by Q_7 (from transistor Q_8) runs entirely through Q_6, turning on Q_2 and Q_3, so the potential across the output terminals equals the negative Seebeck potential (V_{B-}–V_{B+}). By applying a square wave to the terminal 'chop' the voltage between the two output pins is sequentially alternated between the Seebeck potentials (V_{A+}–V_{A-}) and (V_{B-}–V_{B+}). The amplitude of the chopped signal is the sum of both Seebeck voltages. (For clarity, see Figure 5.10.)

Measurements indicate that transistors Q_1 to Q_4 have an average saturation voltage of 4 mV. However, as the circuit is balanced, only the mismatch between the transistors remains which leaves a typical residual saturation voltage of 350 μV. The typical chopped frequency is 5kHz.

Amplifier circuit

The amplifier circuit shown in Figure 5.12 is used as a voltage amplifier. Except for the input stage, the circuit is very straightforward. This stage is designed for low noise, using very large transistors. To obtain a high gain and low offset, these transistors are biased at high currents. The input bias current is reduced by current mirrors at the base of the input transistors. The offset is further reduced by using a cross-coupled quad transistor layout.

The signal coming from the terminals Out1 and Out2 of the chopper are connected to the input terminals 'In+' and 'In−' of a differential pair. The amplified signal is fed into a level shift stage. The resulting signal is used to drive the output terminal using a class AB push-pull stage. Measurements show a unity-gain bandwidth of 7 MHz, with a β_F of 150. At 1.5 mV the input offset is rather large. The noise level inside the circuit is extremely low. Calculations indicate the equivalent noise voltage and current should be $2nVHz^{-1/2}$ and 50 $pAHz^{-1/2}$. Unfortunately this could not be verified experimentally, as the noise in the measurement setup was too high, i.e., equal to15 $nVHz^{-1/2}$.

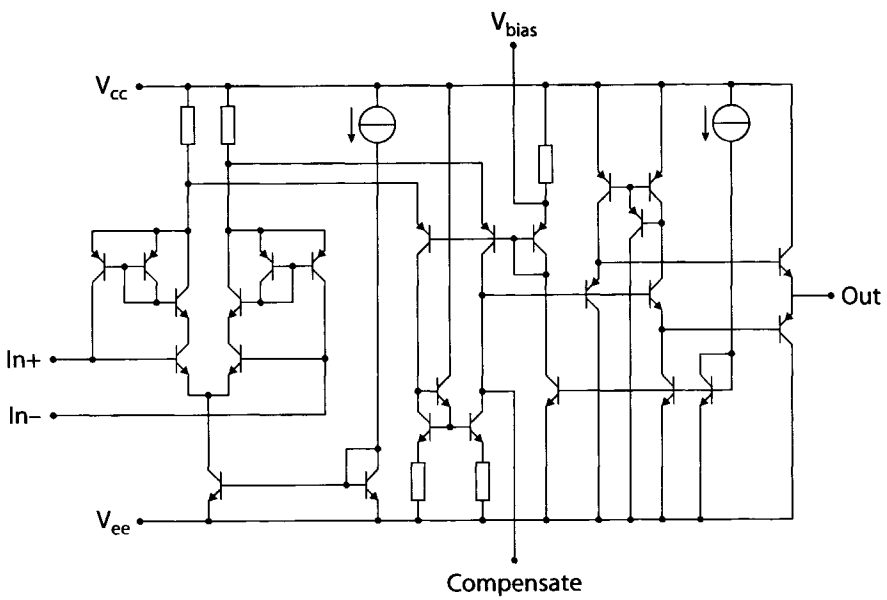


Figure 5.12 Circuit diagram of a low-noise operation amplifier.

Comparator

The comparator circuit that compares the measured and amplified voltage to the user-specified temperature is shown in Figure 5.13. It basically is a differential pair with active current mirror load to increase the gain.An emitter follower is used as output buffer.

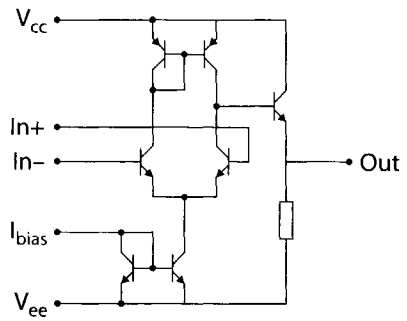


Figure 5.13 Circuit diagram of the comparator.

The above circuits are simple straightforward integrated solutions. Since the bandwidth of the electronic circuits is much higher than the thermal time constant of the TEC, straightforward high-frequency on-off control can be applied, resulting in simple logic circuits for driving.

In case the electronics need not be necessarily integrated on-chip, other thermal control solutions are available, either based on unchopped operational amplifiers [5.40], or even software [5.41]. These discrete implementations can be quite successful, and have proven to enable thermal stabilisation respectively to within 1 mK and 2.5 mK.

5.4 HEAT SPREADING

From the various discussions with representatives from industry, the most frequent discussion regarded the potential of integrated Peltier devices for thermal energy removal from localised hot spots in a substrate. In particular strongly dissipative components in a substrate with a low thermal conductivity, like GaAs-based power FETs [5.42][5.43] could benefit from the removal of hot spots. The typical thermal conductivity of GaAs can be approximated by $5.7{\cdot}10^{5}{\cdot}T^{-1.23}$ To provide an indication, $\lambda_{GaAs} \approx 51$ $Wm^{-1}K^{-1}$ at 25 °C, 38 $Wm^{-1}K^{-1}$ at 100 °C and 29 $Wm^{-1}K^{-1}$ at 200 °C. Thus, in case the power dissipation increases, the temperature will increase progressively, potentially leading to thermal runaway or even device failure.

Recall the expression for the thermal energy removal rate at the cold junction, Equation (2.15). Unlike in the conventional cooling regime, if thermal energy is removed from an object above ambient temperature, the thermal conductance term,

$K(T_a-T_c)$, changes sign (i.e. becomes negative.) Then, instead of working against the Peltier effect, the thermal conductivity suddenly works with the Peltier effect.

Only a simple calculation is required to reveal that on-chip integrated TECs are not suitable for hot spot removal from substrates. Consider the lateral TEC as analysed in this thesis. The Peltier cooling term (after optimization of geometry and electrical current) is equal to $5.6 \cdot 10^{-8} \cdot T_c$. Similarly, the conductance term equals $1.5 \cdot 10^{-6}(T_c-300)$. For these values, the Peltier cooling term and conductance term are already equal when T_c is 311.6 K (i.e. some 11 K above room temperature.) At higher temperatures, the conductance term will become dominant. At 200 °C, the conductance term is already an order of magnitude larger than the Peltier cooling term.

The conclusion from the above observation: for removal of hot spots in a substrate, a heat spreader in the shape of a film with a high thermal conductivity is far more effective than a thermoelectric cooler.

5.5 POLYSIGE FOR SENSING APPLICATIONS

Polycrystalline silicon germanium is not only suitable as thermoelectric material for Peltier devices, but can also be used very effectively in thermal sensing applications. In two applications in particular, the thermopile-based IR detector and the bolometer, polySiGe appears to be very potent.

In [5.28], the maximum sensitivity and detectivity of infrared detectors is provided. The maximum sensitivity [VKW^{-1}] is expressed as

$$S_{\max} = \frac{N\left(\alpha_p - \alpha_n\right)}{K_{rad} + K_{gas} + W\sum_m \lambda_m H_m \delta} \tag{5.5}$$

where K_{rad} and K_{gas} are the conductance by means of radiation and conduction/convection through air, N is the number of thermocouples and δ is equivalent to

$$\delta = \sqrt{\frac{K_{rad} + K_{gas}}{\sum_m \lambda_m H_m}} \tag{5.6}$$

This equation clearly reveals that $S_{\max} \propto \sqrt{\lambda_m}$. Similarly, the maximum detectivity of the detector, D* [$mHz^{1/2}W^{-1}$], is expressed as

$$D^*_{\max} = \frac{S_{\max}\sqrt{A}}{\sqrt{4k_B TR}} \tag{5.7}$$

So, also the specific detectivity $D^* \propto \sqrt{\lambda_m}$. Thus, when switching from polycrystalline silicon ($\lambda \approx 30$ $Wm^{-1}K^{-1}$) to polycrystalline silicon germanium ($\lambda \approx 5$ $Wm^{-1}K^{-1}$), the sensitivity and detectivity are roughly increased by a factor $\sqrt{6} \approx 2.5$.

Similarly, integrated bolometers may also benefit from the use of polySiGe rather than polySi [5.44]–[5.46]. The responsivity of a bolometer can be expressed as

$$\Re = \frac{|\alpha|\varepsilon V_{bias} R_B R_L}{K(R_B + R_S + R_L)^2 \sqrt{1+4\pi^2 f^2 \tau^2}} \tag{5.8}$$

where ε is the emissivity of the bolometer, R_B, R_L and R_S respectively are the resistances of the bolometer, load and electrical series connection between R_B and R_L. Furthermore, f is the actuation frequency and τ is the thermal time constant, equal to C/K, with C the thermal capacitance. The above equation reveals that, in a first-order approximation, $\Re \propto K^{-1}$. Thus, shifting from polySi to polySiGe the responsivity could be increased by up to a factor 6.

5.6 REFERENCES

[5.1] V. Székely, A. Nagy, S. Török, G. Hajas and M. Rencz, "Heat sinks with controlled thermal resistance: an experiment," *Proc. 5th Therminic Workshop*, 3–6 October 1999, Rome, Italy, pp. 249–251.

[5.2] C. Shafai and M.J. Brett, "A micro-integrated Peltier heat pump for localized on-chip temperature control," *Canadian J. of Physics*, Vol. 74, Suppl. 1, pp. S139–S142, 1996.

[5.3] D.D.L. Wijngaards, E. Cretu, S.H. Kong and R.F. Wolffenbuttel, "Modelling of integrated polySiGe Peltier elements," *Proc. Therminic 2000*, 24–27 September 2000, Budapest, Hungary, 4 pp.

[5.4] J.A. Chávez, J.A. Ortega, J. Salazar, A. Turó and M.J. Garcia, "SPICE model of thermoelectric elements including thermal effects," *Proc. 17th IEEE Instrumentation and Measurement Technology Conference*, 1–4 May 2000, Baltimore, Maryland, USA, pp. 1019–1023.

[5.5] P. Dziurdzia and A. Kos, "Electrothermal macromodel of active heat sink for cooling process simulation," *Proc. 5th Therminic Workshop*, 3–6 October 1999, Rome, Italy, pp. 76–81.

[5.6] V. Milanovic, M. Hopcroft, C. Zincke, M. Zaghloul and K.S.J. Pister, "Modeling of thermoelectric effects in planar micromachined structures using SPICE," *Proc. 6th Therminic Workshop*, 24–27 September 2000, Budapest, Hungary, 5 pp.

[5.7] R.E. Simons and R.C. Chu, "Application of thermoelectric cooling to electronic equipment: a review and analysis," *Proc. 16th Annual IEEE Semiconductor Thermal Measurement and Management Symposium*, 21–23 March 2000, San Jose, CA, USA, pp. 1–9.

[5.8] M.R. Holman and S.J. Rowland, "Design and development of a new cryosurgical instrument utilizing the Peltier thermoelectric effect," *J. of Medical Engineering & Technology*, Vol. 21, No. 3–4, pp. 106–110, 1997.

[5.9] D.-J. Yao, C.-J. Kim and G. Chen, "MEMS thermoelectric microcooler," *Proc. 20th International Conference on Thermoelectrics*, 8–11 June 2001, Beijing, China, pp. 401–404.

[5.10] C. Shafai and M.J. Brett, "On-chip thin film Peltier heat pumps," *Proc. Therminic '98*, 27–29 September 1998, Cannes, France, 5 pp.

[5.11] G. Zeng, A. Shakouri, C. La Bounty, G. Robinson, E. Croke, P. Abraham, X. Fan, H. Reese and J.E. Bowers, "SiGe micro-cooler," *Electronics Letters*, Vol. 35, No. 24, pp. 2146–2147, 1999.

[5.12] C. LaBounty, A. Shakouri and J.E. Bowers, "Design and characterization of thin film microcoolers," *J. of Appl. Phys.*, Vol. 89, No. 7, pp. 4059–4064, 2001.

[5.13] X. Fan, G. Zeng, C. LaBounty, D. Vashaee, J. Christofferson, A. Shakouri and J.E. Bowers, "Integrated cooling for Si-based microelectronics," *Proc. 20th International Conference on Thermoelectrics*, 8–11 June 2001, Beijing, China, pp. 405–408.

[5.14] P. Corrèges, E. Bugnard, C. Millerin, A. Masiero, J.P. Andrivet, A. Bloc, and Y. Dunant, "A simple, low-cost and fast Peltier thermoregulation set-up for electrophysiology," *Journal of Neuroscience Methods*, Vol. 83, Iss. 2, pp. 177–184, 1998.

[5.15] C.J. McKinney and M.W. Nader, "A Peltier thermal cycling unit for radiopharmaceutical synthesis," *Applied Radiation and Isotopes*, Vol. 54, Iss. 1, pp. 97–100, 2001.

[5.16] L. Elsgaard and L.W. Jørgensen, "A sandwich-designed temperature-gradient incubator for studies of microbial temperature responses," *Journal of Microbiological Methods*, Vol. 49, Iss. 1, pp. 19–29, 2002.

[5.17] Gordon Reid, Bogdan Amuzescu, Eberhard Zech, Maria-Luiza Flonta, "A system for applying rapid warm-ing or cooling stimuli to cells during patch clamp recording or ion imaging," *Journal of Neuroscience Methods*, Vol. 111, Iss. 1, pp. 1–8, 2001.

[5.18] J. Hodgson, "Gene sequencing's industrial revolution," *IEEE Spectrum*, Vol. 37, No. 11, pp. 36–42, 2000.

[5.19] E.H. Klaassen, R.J. Reay, and G.T.A. Kovacs, "Diode-based thermal RMS converter with on-chip circuitry fabricated using standard CMOS technology," *Proc. Transducers '95 / Eurosensors IX*, 25–29 June 1995, Stockholm, Sweden, pp. 43–50.

[5.20] R.J. Reay, E.H. Klaassen and G.T.A. Kovacs, "Thermally and electrically isolated single crystal silicon structures in CMOS technology," *IEEE Electron Device Lett.*, Vol. 15, No. 10, pp. 399–401, 1994.

[5.21] J. Puigcorbe, A. Vila, I. Gracia C. Cane and J.R. Morante, "Dielectric micro-hotplate for integrated sensors: an electro-thermo-mechanical analysis," *Digest of technical papers Transducers '01 / Eursensors XV*, 10–14 June 2001, Munich, Germany, pp. 312-315.

[5.22] S.K.H. Fung, Z. Tang, P.C.H. Chan, J.K.O. Sin and P.W. Cheung, "Thermal analysis and design of a micro-hotplate for integrated gas-sensor applications," *Sensors and Actuators*, Vol. A54, No. 1–3, pp. 482–487, 1996.

[5.23] P. Ancey, M. Gschwind and O. Vancauwenberghe, "New concept of integrated Peltier cooling device for the preventive detection of water condensation," *Sensors and Actuators*, Vol. B27, No. 1–3, pp. 303–307, 1995.

[5.24] O. Vancauwenberghe, J. Short, E. Giehler, P. Bildstein, P. Ancey and M. Gschwindt, "Microsensor for the preventive detection of water condensation: operating principle and interface electronics" *Sensors and Actuators*, Vol. A53, No. 1–3, pp. 304–308, 1996.

[5.25] A. Foucaran, B. Sorli, M. Garcia, F. Pascal-Delannoy, A. Giani and A. Boyer, "Porous silicon layer coupled with thermoelectric cooler: a humidity sensor," *Sensors and Actuators*, Vol. A79, No. 3, pp. 189–193, 2000.

[5.26] G. Festa and B. Neri, "Thermally regulated low-noise, wideband, I/V converter, using Peltier heat pumps," *IEEE. Trans. on Instrum. & Meas.*, Vol. 43, No. 6, 1994, pp. 900–905.

[5.27] P.P.L. Regtien, *Silicon dew point sensor for accurate humidity measurement systems*, Ph.D. Thesis, Delft University Press, Delft, The Netherlands, 1981.

[5.28] S.H. Kong, *Infrared micro-spectrometer based on a multi-slit grating*, Ph.D. Thesis, Eburon Publishers, Delft, The Netherlands, ISBN 90-5166-925-9.

[5.29] A.W. van Herwaarden, P.M. Sarro, J.W. Gardner and P. Bataillard, "Liquid and gas micro-calorimeters for (bio)chemical measurements," *Sensors and Actuators*, Vol. A43, No. 1–3, pp. 24–30, 1994.

[5.30] U. Dillner, E. Kessler, S. Poser, V. Baier and J. Müller, "Low power consumption thermal gas-flow sensor based on thermopiles of highly effective thermoelectric materials," *Sensors and Actuators*, Vol. A60, No. 1–3, pp. 1–4, 1997.

[5.31] M. Ashauer, H. Glosch, F. Hedrich, N. Hey, H. Sandmaier and W. Lang, "Thermal flow sensor for liquids and gases based on combinations of two principles," *Sensors and Actuators*, Vol. A73, No. 1–2, pp. 7–13, 1999.

[5.32] J.J. van Baar, R.J. Wiegerink, T.S.J. Lammers, G.J.M. Krijnen and M.C. Elwenspoek, "Micromachined structures for thermal measurements of fluid and flow parameters," *J. Micromech. Microeng.* Vol. 11, pp. 311–318, 2001.

[5.33] M. von Arx, O. Paul and H. Baltes, "Test structures to measure the heat capacity of CMOS layer sandwiches," *IEEE Trans. on Semicond. Manuf.*, Vol. 11, No. 2, 1998, pp. 217–224.

[5.34] MEMS Clearinghouse, http://www.memsnet.org.

[5.35] U.A. Dauderstädt, *A thermal accelerometer*, Delft University Press, Delft, The Netherlands, 1999, ISBN: 90-407-1885-7.

[5.36] *CRC Materials science and engineeringhandbook*, 3rd edition, CRC Press, Boca Raton, FL, USA, 2000, ISBN 0-849-32696-6.

[5.37] G.A. Slack and M.A. Hussain, "The maximum possible conversion efficiency of silicon-germanium thermoelectric generators," *J. Appl. Phys.*, Vol. 70, Iss. 5, pp. 2694–2718, 1991.

[5.38] G.C.M. Meijer and A.W. van Herwaarden, Thermal Sensors, IOP Publishing Ltd., London, UK, 1994.

[5.39] F.R. Riedijk, *Integrated smart sensors with digital bus interface*, Ph.D. Thesis, Delft University Press, Delft, The Netherlands, 1993.

[5.40] A.W. Sloman, P. Buggs, J. Molloy and D. Stewart, "A microcontroller-based driver to stabilize the temperature of an optical stage to within 1 mK in the range 4–38 °C, using a Peltier heat pump and a thermistor sensor," *Meas. Sci. Technol.*, Vol. 7, pp. 1653–1664, 1996.

[5.41] A. Amengual, "A feedback algorithm for programmed temperature control," *Meas. Sci. Technol.*, Vol. 3, pp. 1043–1046, 1992.

[5.42] J. Wilson, "Thermal issues in GaAs analog RF devices," *Electronics Cooling*, Vol. 8, No. 1, pp. 14–21, 2002.

[5.43] K. Decker, S. Ko and D. Rosato, "Thermal characterization of Gallium Arsenide FETs," *HDI Magazine*, Vol. 3, No. 12, 4 pp., 2000.

[5.44] S. Sedky, P. Fiorini, M. Caymax, A. Verbist, and C. Baert, "IR bolometers made of polycrystalline silicon germanium," *Sensors and Actuators*, Vol. A66, No. 1-3, pp. 193–199, 1998.

[5.45] S. Sedky, P. Fiorini, K. Baert, L .Hermans and R. Mertens, "Characterization and optimization of infrared poly SiGe bolometers," *IEEE Trans. on Electron Devices*, Vol. 46, No. 4, pp. 675–682, 1999.

[5.46] J-S. Shie and P.K. Weng, "Fabrication of micro-bolometer on silicon substrate by anisotropic etching technique," *Digest of technical papers Transducers '91*, 24–27 June 1991, San Francisco, CA, USA, pp. 627-630.

Chapter 6

CONCLUSIONS

This thesis discusses the lateral on-chip integrated Peltier element, which transports energy along the surface of a wafer (i.e. in-plane). These Peltier elements are generally very slim and long and need to be thermally isolated by means of micromachining. This ensures a high power-efficiency and a fast and accurate thermal response. However, due to the laws of downscaling, the thermal energy removal rate is decreased by a factor of 10^3, compared to the discrete Peltier device. The maximum temperature difference between the hot and cold junctions is not influenced by downscaling.

Fabrication

The lateral on-chip integrated TECs can be constructed using only a single layer of thermoelectric material, two dielectric layers and a single interconnect layer. This way a simple 7 mask fabrication process is obtained; Three masks are required to dope (P and B) and pattern the thermoelectric layer. Two masks are required for metallisation, to create the contact holes to the thermoelectric material, and to pattern the metallisation layer. The last two masks are required to thermally isolate the cold end of the Peltier device by a consecutive alkaline etch and dry etch, to respectively remove the substrate under the Peltier element and to create beams and cantilevers from the membrane resulting after the alkaline etch.

As thermoelectric material polycrystalline $Si_{0.7}Ge_{0.3}$ is chosen as this alloy has the highest figure-of-merit of any material that can be deposited using standard cleanroom facilities. Using APCVD a 300 nm to 600 nm poly$Si_{0.7}Ge_{0.3}$ layer is deposited at 700 °C, on a polySi nucleation layer. Recrystallisation took place at a temperature of 1050 °C. To protect the polySiGe structures, these are sandwiched in between a nitride and oxide layer. The total Si_3N_4/ polySiGe/SiO_2 stack can be made as thin as 650 nm (250 nm /300 nm /100 nm).

The largest technological difficulties occurred near the end of the fabrication process, which caused the yield to drop. In particular during the alkaline etch many wafers were lost, as KOH seeped through fractured membranes or even entirely fractures wafers, thereby removing the Al interconnect. Of the structures that survived processing, a significant number was destroyed during dicing. The violent agitation

of the wafer (by both the saw and the high-pressure water cooling system) causes beams and cantilevers to simply rupture. In particular the use of cantilever structures should be avoided as these had a yield of less than 20%.

Fabrication compatibility

In an operational amplifier, the input stage dominates the signal-to-noise performance of the system. Similarly, to improve the detection limit of a sensor system, in particular the input stage (i.e. the transducer itself) needs to be enhanced. This will become increasingly difficult without the use of non-standard materials and processing—which are poorly compatible with IC manufacturing.

At the Delft University of Technology, a significant effort is put into the development of smart transducers, i.e., transducers that are enhanced by on-chip electronics. I believe this concept is only feasible (both technically and economically) for transducers that can be constructed within standard IC processes. In all other cases, to obtain both an optimised transducer and optimised electronics, physical separation of both components is obligatory. Even though such a clear division goes against the smart sensor philosophy, only such a hybrid integration approach allows both parts to be independently optimised.

Non-idealities

Compared to the conventional TEC, the on-chip integrated TEC has a number of additional non-idealities arising from downscaling. In particular the contact resistance (that increases with the square of the feature size) and parasitic thermal conduction (due to the need of mechanically supporting and electrically insulating layers around the TEC) are to blame for the reduction in performance.

Consider a standard polySiGe thermocouple that is sandwiched in between two dielectric layers. In such a configuration, the parasitic thermal conduction through the dielectric layers accounts for 10%–30% of the total thermal conduction. The effect of electrical contact resistance is less significant, at a few percent of the total electrical resistance. Combined, the maximum temperature difference of an on-chip integrated polySiGe TEC will be as much as 25% less than the maximum temperature difference across a conventional TEC with the same material values. The effect of the non-idealities becomes more pronounced as the figure-of-merit of the thermoelectric material increases. This way, as much as 50%–60% of the total thermal conduction can come from the supporting dielectric membrane.

Most deteriorating to performance is the Al interconnect thar runs thermally in parallel with the TEC. As the thermal conductivity of Al is up to 45 times as high as the thermal conductivity of polySiGe, it was observed that as much as 70% of the total thermal energy flows to ambient through the Al interconnect. Again, this effect becomes more pronounced as the figure-of-merit increases. It is therefore imperative that metallic interconnect in parallel with the TEC is reduced to a bare minimum or preferably is avoided all together.

Another consequence of the increased magnitude of the non-idealities is that the maximum temperature difference no longer depends on the figure-of-merit only. Instead, also the non-idealities as well as the geometry of the device need to be considered in order to optimise device performance.

Modelling results

Both 1-D analytical models and 1-D lumped-element models have been developed. These serve to determine the maximum temperature difference, temperature profile and thermal response of the integrated TEC. Furthermore, 2-D finite element analyses have been performed to verify the analytical models and to determine the thermal resistance of the heat sink, i.e., the substrate. The results of the analytical models and the FEA model match to within 10%. The analytical solution to the maximum temperature difference varies less than 5% from actual measurement values.

Normally the optimum electrical driving current is expressed as a function of the cold junction temperature, which can only be solved by iteration. In this thesis for the first time the optimum electrical driving current is expressed as a function of the hot junction temperature, which provides a direct solution:

$$i_{opt} = \frac{KR\sqrt{KR(KR+2(\alpha_n-\alpha_p)^2 T_a)}}{R(\alpha_n-\alpha_p)}$$

Using FEA, the influence on the thermal resistance of the silicon substrate was investigated, in relation to 1) the geometry of the substrate and 2) a dielectric layer in between the hot junction and the substrate. The simulations revealed a front-side alkaline etch is preferred over a backside etch, as this reduces the thermal resistance by up to two thirds. Dominant, however, is the thin dielectric layer on top the substrate. For example, a 1 μm Si_3N_4 layer is sufficient to increase the thermal resistance between the hot junction and ambient by over 100%. Still, FEA also revealed that for the current power density levels of the polySiGe TEC, the temperature gradient in the substrate is only about 0.2% of the temperature difference between the hot and cold junctions.

Material characterisation

PolySiGe has been characterised extensively. The results are summarised in Table 6.1. These material values are the best thin-film polySiGe values reported to this date. Under ideal circumstances (zero load, no losses) these values are good for cooling to 6.9 K below ambient temperature. To put the material values up to par with the properties of conventional sintered polySiGe, further improvement of in particular the power factor is required.

Applications

In Chapter 1 three fundamental applications were identified for the TEC, cooling, thermal stabilisation near ambient temperature, and heat spreading. The maximum

Table 6.1 Material properties of APCVD deposited $polySi_{0.7}Ge_{0.3}$.

Property	n-polySiGe	p-polySiGe
Seebeck coefficient (α) [μVK^{-1}]	-174	129
Electrical resistivity (ρ) [$\mu\Omega m$]	29.2	28.9
Thermal conductivity (λ) [$Wm^{-1}K^{-1}$]	5.1	4.7
Power factor (ψ) [$Wm^{-1}K^{-2}$]	$1.0 \cdot 10^{-3}$	$5.8 \cdot 10^{-4}$
Figure-of-merit (z) [$10^{-3}\ K^{-1}$]	0.203	0.123

temperature difference between the hot and cold junctions was found to be 2.1 K at a pressure of 1.7 Pa. This low temperature difference is caused by the excessive use of Al wires in parallel with the TEC. Removing these wires increases the temperature difference to 5.3 K, which is close to the analytically predicted value of 5.5 K. It takes only 9 ms to stabilise within 1% from the final value.

For thermal stabilisation near ambient temperature it is essential to have an active cooling system to ramp down the temperature. A TEC is highly suitable as it is the only available solid-state cooling solution near room temperature. In Chapter 5 it is shown that the micro-thermostat with the largest thermal operating range uses the TEC for cooling, heating and sensing alike. To read out the temperture difference across the TEC, a special balanced voltage chopper was designed.

For the third application, heat spreading—in particular spreading of on-chip hot spots in substrates with a low thermal conductivity like GaAs—the on-chip integrated TEC is shown to be of limited use. A simple heat spreader in the shape of a thin film with a high thermal conductivity is far more effective.

Potential improvements

The largest improvement in performance will be obtained by switching to a material with a higher figure-of-merit, like thin-film bismuthides and antimonides. As these are generally fabrication incompatible, such materials can only be deposited at the very last step in the process flow, where patterning has to be implemented by less accurate methods like lift-off, and the material is left vulnerable to the environment. Thus, switching materials will prove to be very difficult.

Without considering alternative materials, room for performance improvement has to be sought primarily in technology. The most interesting modification is application of a dual polySiGe layer process, which only requires two extra masks compared to the current process. As a result, the polySiGe-to-dielectric layer ratio is improved (reducing parasitic losses), the material stack is made thicker (providing additional mechanical strength) and stresses in the different layers are cancelled out.

Appendix A

CONCISE THEORY OF THERMOELECTRICITY

The performance of a thermoelectric material is optimised when the figure-of-merit is. Thus, a high Seebeck coefficient is required, while the thermal conductivity and electrical resistivity should be low. To explain in detail how these three parameters can be optimised, a rigorous quantum-physical treatment is required. This goes far beyond the intent of this thesis. The reader who would like to go in-depth on the quantum physics behind thermoelectrics is referred to a number of books on this subject [3.32],[A.6]–[A.10]. This section only provides a high-level explanation of the thermophysical material parameters and how they can be influenced. In the following sections the material discussed is assumed to be an n-type semiconductor, unless stated otherwise.

A.1 THE SEEBECK COEFFICIENT

The first of the three material parameters that needs to addressed is the Seebeck coefficient. In previous chapters is was already pointed out that the Seebeck voltage develops across the ends of a material which is subjected to a temperature gradient between these ends. In [A.1] a simple explanation for this phenomena is provided: "*The charge carriers at the hot end of the sample have, on average, more energy than the charge carriers at the cold end. So, they are moving faster at the hot end than the cold end ... [so] ... they do diffuse to the cold end. But after only a few extra carriers have collected on the cold end, they set up a voltage which prevents further carriers from building up. At this point the hot end is deficient by a few carriers. The Seebeck coefficient represents the electrical potential required to balance the thermally driven diffusion.*" This explanation is somewhat simplified. Middelhoek *et al.* [A.2] point out that in fact a great number of effects contribute to the Seebeck effect and it is not exactly understood what the contribution of each effect is.

References [A.3] and [A.4] shed a more fundamental light on the origin of the Seebeck effect, in particular to what happens when two dissimilar materials are put into direct contact. Figure A.1a and A.1b show what happens at the interface of a metal when put into contact respectively with an n-type and p-type semiconductor.

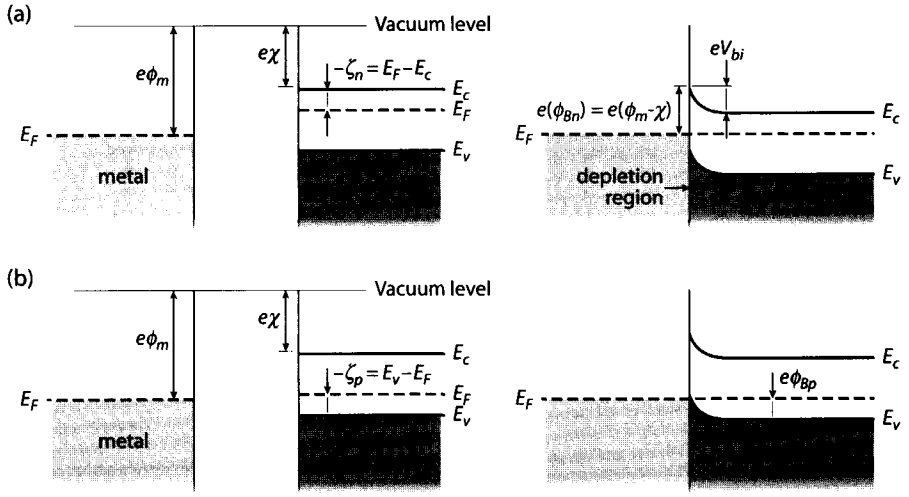


Figure A.1 Schematic representations of (a) a metal/n-type semiconductor and (b) a metal/p-type semiconductor interface.

At 0 K all the electron states in the metal up to the Fermi energy level, E_F, are occupied. At a finite temperature $T > 0$ *K*, there is a partial filling of states over a range of energy within $2k_BT$ around E_F, with k_B the Boltzmann's constant. For the semiconductor the situation is different, as there is a (much larger) energy gap $E_g = E_c - E_v$ between the conduction band (above E_c) and the valence band (below E_v). The main difference between an n-type and a p-type semiconductor is the position of the Fermi level. In the case of n-doped semiconductor material, E_F is situated close to the conduction band, while it is located close to the valence band in case of a p-doped semiconductor. What is mutual for both materials is that, when in contact, the Fermi level must be continuous from the metal to the semiconductor (in equilibrium).

Consider the n-type semiconductor junction. Close to the metal-semiconductor interface electrons will flow from the n-type semiconductor to the lower energy states in the metal, which results in the build-up of a depletion region. An electron that tries to move from the metal to the semiconductor experiences a barrier height ϕ_{Bn} equal to the difference between the work function of the metal ϕ_m and the electron affinity χ of the n-type semiconductor, i.e.,

$$\phi_{Bn} = \phi_m - \chi \tag{A.1}$$

For the larger part it is this barrier ϕ_{Bn} that causes the electrical contact resistance of a metal-semiconductor junctions. Ignoring the barrier, the electron experiences a built-in potential barrier V_{bi} (similar to the that of a pn-junction) which is defined as

$$V_{bi} = \phi_{Bn} - \zeta_n \tag{A.2}$$

where

$$\zeta_n = E_c - E_F \tag{A.3}$$

In carrying a current through the structure, the electron will travel at a certain kinetic energy $\varepsilon \gg e\phi_{Bn}$, where it will rise into the conduction band of the n-type material. To rise into this band, the electron needs to absorb an amount of potential energy equal to $-\zeta_n$. This absorption of energy is the origin of the Peltier cooling effect [A.3].

From the above observation that energy is absorbed or released as the electron passes from one material to another, often the erroneous conclusion is drawn that the Seebeck and Peltier effects occur due to the difference in contact potential at different temperatures. For example when a thermocouple of materials A and B is made, the Seebeck potential would be

$$V_{AB} = V_{bi}\big|_{T_1} - V_{bi}\big|_{T_2} \quad \text{(erroneous)}. \tag{A.4}$$

However, this can never explain why an electrical current starts to flow when the thermocouple loop is closed. Each individual space charge region is in thermal equilibrium, so there is a zero flow of electrons across this region. Even though the space charge region will change as the temperature of a junction changes, an equilibrium is reached that reduces the electron flow back to zero. Instead it is the charge build-up of electrons at the opposite ends of the semiconductor due to thermal diffusion that causes the current flow, i.e.,

$$V_{AB} = V_A\big|_{T_1-T_2} - V_B\big|_{T_1-T_2} \tag{A.5}$$

A.1.1 Seebeck coefficient in non-degenerate semiconductors

As stated, the Seebeck effect results from a number of different (quantum-)physical effects. The magnitude of the effect is attributed to three phenomena in particular, electron scattering, the level of Fermi potential, and the phonon drag [A.2][3.32]. Together these lead to the expression

$$\alpha = \mp\frac{k_B}{e}\left[\frac{5}{2} + s - \eta\right] + \varphi \tag{A.6}$$

A minus sign is used n-type materials and the positive sign for p-type materials. Furthermore, e is the electron charge, s is the scattering parameter, η is the reduced Fermi potential and φ is contribution of the phonon drag. The term 5/2 arises from mathematical manipulation of the contribution by the density of states and the average electron velocity in the material, which is shown most clearly in [A.5].

A.1.1.1 Electron and hole contribution

Let us consider the Seebeck coefficient of an n-type material. The reduced Fermi potential of such a material is expressed as

$$\eta_n = \frac{\zeta_n}{k_B T} = \frac{E_F - E_c}{k_B T} \tag{A.7}$$

Alternatively, this potential can be expressed as [3.24]

$$-\eta_n = \ln\left(\frac{N_c}{n_0}\right) \tag{A.8}$$

where

$$N_c = 2\left(\frac{2\pi m_n^* k_B T}{h^2}\right)^{3/2} \tag{A.9}$$

is the so-called effective density of states function in the conduction band. Assuming complete ionisation, n_0 can be expressed as a function of the intrinsic carrier concentration n_i, and the density of acceptor and donor impurity atoms, N_a and N_d:

$$n_0 = \frac{(N_d - N_a)}{2} + \sqrt{\left(\frac{N_d - N_a}{2}\right) + n_i^2} \tag{A.10}$$

For the n-type semiconductor, $N_d \gg \{n_i, N_a\}$, so $n_0 \approx N_d$. Consequently substituting n_0 with N_d in Equation (A.8), and inserting the solution into Equation (A.6), the Seebeck effect may be expressed as

$$\alpha_n = -\frac{k_B}{e}\left[\frac{5}{2} + \ln\left(\frac{N_c}{N_d}\right) + s_n\right] + \varphi_n \tag{A.11}$$

Similarly, for a p-type material, Equation (A.6) becomes

$$\alpha_p = \frac{k_B}{e}\left[\frac{5}{2} + \ln\left(\frac{N_v}{N_a}\right) + s_p\right] + \varphi_p \tag{A.12}$$

where N_v is the effective density of states function in the valence band:

$$N_v = 2\left(\frac{2\pi m_p^* k_B T}{h^2}\right)^{3/2} \tag{A.13}$$

Equations (A.11) and (A.12) reveal the Seebeck coefficient will roughly decrease with $\log(N_x)$ where N_x is the impurity concentration.

A.1.1.2 Electron scattering mechanisms

The second effect that determines the Seebeck coefficient is the scattering of electrons. There is a considerable number of different scattering effects, that interact and operate simultaneously. The contribution of each of these mechanisms varies strongly with temperature, impurity concentration, and even with the fabrication technique. Many of the scattering mechanisms are so complex, that only approximations of the effect can be provided, and the actual influence needs to be determined experimentally. Therefore, in this subsection, the scattering mechanisms discussed in [3.32], [A.5] and [A.6]–[A.10] are merely summarised and, if possible, an estimate of the magnitude of the effect is provided.

In general, three sources of charge carrier scattering processes can be identified:

- Scattering by lattice vibrations (phonons), either acoustically or optically.
- Scattering due to lattice defects. These include neutral impurities, ionised impurities, dislocations and grain boundaries.
- Scattering by other carriers. These can be binary or collective in nature, respectively through electron-electron or electron-hole collisions, and through oscillation in the carrier density, creating plasmons.

Below, the various scattering mechanisms are discussed in some more detail.

Acoustic phonon scattering. In pure and perfect crystals, the carriers are scattered by the vibrations of the crystal lattice, known as phonons. For long wavelengths, acoustic (low-frequency, harmonic) modes cause adjacent atoms to move almost in phase with each other while the optical (high-frequency, anharmonic) modes cause adjacent atoms to move almost exactly out of phase with each other.

Acoustic modes can upset the equilibrium situation in two ways. By locally changing the lattice parameter, the energy-band structure is slightly altered. This leads to a so-called deformation potential (i.e., a small change in the energy of a band edge) that scatters the charge carriers. Alternatively, if the atoms in the crystal are partially ionised, a displacement of these atoms creates piezoelectric potentials that scatter the charge carriers. It is usually assumed that deformation potential scattering is much stronger than piezoelectric scattering.

For acoustic scattering, the scattering parameter s is equal to -1/2. Generally, except at very high impurity concentrations, acoustic phonon scattering of charge carriers is the dominant scattering mechanism in semiconductors.

Optical phonon scattering. Optical vibrations can cause scattering in more or less the same two ways as acoustic phonons. In crystals such as ZnO and AgCl, the atoms carry a charge so the optical mode vibrations cause a polarisation of the lattice that will scatter electrons. In nonpolar crystals of materials like Si and Ge (where both atoms in the unit cell are identical) the charge of the two atoms is identical. Therefore, these optical modes produce no polarisation of the lattice. Still, the nonpolar lattice distortion produced by optical modes scatters the electrons. Generally, when optical phonon scattering dominates, $s = +1/2$.

Intervalley and intravalley scattering. For a multivalley semiconductor, there are two types of scattering, intervalley scattering (involving large changes in wave vector) and intravalley scattering (involving small changes in wave vector). In intervalley scattering the wave vector of the electron is in a different valley after scattering, while it remains in the same valley for intravalley scattering.

At low temperatures intervalley scattering should be negligible, as there are few phonons with the required energy. Also, emission of a phonon is rare because few carriers have enough energy to create a phonon with the required energy. At higher temperatures, when there is a substantial number of phonons with a large wave vector, intervalley scattering becomes as frequent to intravalley scattering.

Ionised impurity scattering. When electrons have been excited from the donor atoms into the conduction band or have gone to acceptor atoms, the donor carries a net positive charge. Similarly acceptors will acquire a net negative charge when they take electrons from the valence band or from donors. These charge centres will scatter electrons and reduce the mobility.

Thermoelectric semiconductor materials are usually very heavily doped. Consequently, the scattering of charge carriers by ionised impurities should be taken into account with such materials. For ionised impurity scattering, $s = +3/2$.

Neutral impurity scattering. Close to the centre of a neutral impurity atom, the potential starts to change. This potential difference will scatter the charge carriers. This type of scattering is important at low temperatures only when the charge carriers are in the bound levels of donor or acceptor atoms. This type of scattering is not considered important in thermoelectric devices.

Scattering by dislocations. Dislocations scatter charge carriers by two mechanisms. One is the microscopic deformation of the crystal about the dislocation. This mechanism is only important at low temperatures in pure crystals with high dislocation densities. Second, a dislocation may act as an acceptor. In n-type material the dislocation will be negatively charged and will be surrounded by a region of equal positive charge. This charged region is of the order of a micrometer in diameter.

The charged dislocations have two effects on the motion of charge carriers. First, if the mean free path is small compared to the spacing between dislocations, the dislocations produce an increase in the resistivity since the current tends to avoid the space charge region about the dislocations. Second, if the mean free path is large compared to the dislocation spacing, the scattering of charge carriers by the space charge region about the dislocation will reduce the mean free path.

Alloy scattering. In an alloy of multiple semiconducting materials, the charge carriers can be scattered by the fluctuations in the composition of the material. Since the energy of the band edges is a function of composition, the band edge energy will not be constant throughout the crystal. These variations are similar to the fluctuations produced by lattice waves that result in a deformation potential. Glicksman [A.11] has studied alloy scattering in Ge-Si alloys and found a temperature dependency in the order of $T^{-0.8}$.

Surface scattering. Electrons can scatter off the boundaries of the confining potential. If the confining potential is smooth, there is no effect but in reality it a crystal boundary atomically rough, so electrons do scatter. Consider the Si/SiO_2 interface in a MOS structure. As the oxide is somewhat rough, this produces fluctuations in the confining potential along the channel. This effect will increase as the gate voltage increases, which confines carriers more closely to the surface. For high gate biases, surface roughness scattering can even be the dominant mechanism at the Si/SiO_2 interface. At low inversion layer densities, ionised impurity scattering will dominate over surface scattering, but at high inversion layer densities, interface scattering becomes dominant.

Electron-electron and electron-hole scattering. Finally, charge carriers can also scatter each other, though [3.32] suggests it is doubtful that carrier-carrier scattering is significant for concentrations that are of interest to thermoelectrics. Electron-electron scattering does not affect the mobility directly as the total momentum of the electrons remains unchanged by the collision. However, such collisions will redistribute the electron momentum so that other scattering mechanisms can remove the momentum more effectively.

For example—only considering ionised impurity scattering—high-energy electrons have a much longer mean free path than low-energy electrons. Therefore, the high-energy electrons will acquire a relatively larger momentum from the electric field. Electron-electron collisions can redistribute this momentum to the lower energy electrons (which are more susceptible to ionised impurity scattering,) so the momentum will be removed more readily by ionised impurity scattering.

Only for the acoustic phonon scattering, optical phonon scattering and ionised impurity scattering it is possible to directly express the scattering parameter, respectively as $s = -1/2$, $s = +1/2$ and $s = +3/2$. Gaur *et al.* [A.12] have derived the Seebeck coefficient for SiGe as function of the reduced Fermi level η, for each of the three scattering parameters. The result is shown in Figure A.2.

It has to be noted that Equation (A.6) is derived using classical statistics, which has certain implications. Consider the phonon drag can be disregarded ($\varphi = 0$) and acoustic phonon scattering is the dominant scattering mechanism ($s = -1/2$.) Then, Equation (A.6) becomes

$$\alpha_{ac} = \mp \frac{k_B}{e}(2 - \eta) \tag{A.14}$$

As a result of these classical statistics, when $\eta \to 2$, the Seebeck coefficient becomes zero, indicated by the dashed curve in Figure A.2. This is not correct, as highly doped semiconductors and metals still exhibit a (small) Seebeck coefficient. Therefore, a more appropriate means of expressing the Seebeck coefficient is through Fermi-Dirac statistics, which—disregarding the phonon drag—yields

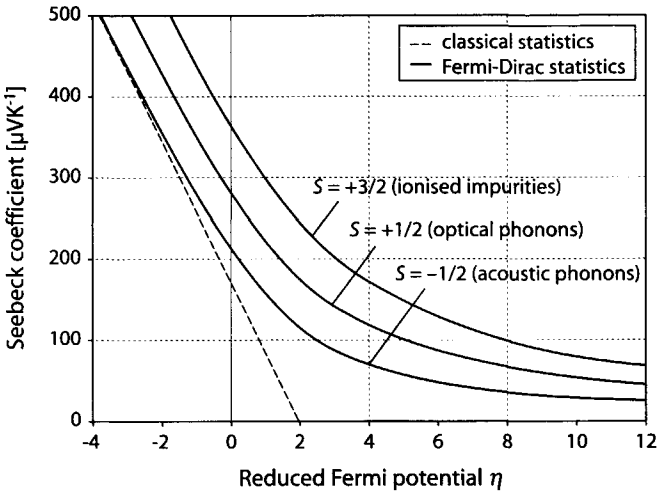


Figure A.2 Seebeck coefficient for silicon-germanium as a function of the reduced Fermi level η, plotted for different scattering parameters. Source: [A.12].

$$\alpha = \mp\frac{k_B}{e}[\delta - \eta] = \mp\frac{k_B}{e}\left[\frac{s+\frac{5}{2}}{s+\frac{3}{2}}\frac{F_{r+3/2}(\eta)}{F_{r+1/2}(\eta)} - \eta\right] \tag{A.15}$$

where $F_m(\eta)$ is the Fermi-Dirac integral, defined as

$$F_m(\eta) = e^{\eta}\int_0^{\infty}\xi^m f_o(\xi)\,d\xi \tag{A.16}$$

with $\xi = E/k_BT$ is the reduced energy and f_0 is the Fermi distribution,

$$f_0(E) = \left[\exp\left(\frac{E-E_F}{k_BT}\right)\right]^{-1} \tag{A.17}$$

Unfortunately the Fermi-Dirac integrals are quite elaborate to solve. Fortunately, for the most common integrals, tables with numerical solutions to the integrals are available, e.g. in [A.3] as well as some of the references referred to in [3.32].

A.1.1.3 Phonon drag

When a temperature gradient is present across a material, phonons will flow from the hot to the cold end, whereby they can transfer momentum to the charge carriers. Even though it is usually assumed that the charge carriers and phonons are scattered in random directions after a collision, the average heat flux will still be directed towards the cold end. Therefore the average momentum of the electrons from the col-

lisions will be in the direction of the flow of the phonons. This phonon drag will create an increase of charge carriers at one end of the material.

Consider a fraction x_d of collisions with charge carriers that involves phonons. For low charge carrier concentrations, the phonon drag φ can then be expressed as [3.32]:

$$\varphi = \mp \frac{x_d v^2 \tau_d}{\mu T} \tag{A.18}$$

where τ_d is the relaxation time for the loss of momentum from the phonon system, μ is the mobility of the charge carriers, T is the absolute temperature and v is the average speed of sound in the material.

Equation (A.18) reveals that the phonon drag depends strongly on temperature. As $\tau_d \propto T^{-5}$ and $\mu \propto T^{-3/2}$, the phonon drag component of the Seebeck coefficient roughly varies as $T^{-9/2}$. Also, the phonon drag effect decreases rapidly as the carrier concentration increases and becomes negligible when the material becomes degenerate. Due to the generally high level of doping concentration in thermoelectric materials it is valid to neglect the phonon drag component of the Seebeck coefficient in calculating the figure-of-merit for these materials.

A.1.2 Seebeck coefficient in degenerate semiconductors

The explanation of the Seebeck effect in Section A.1.1 is valid for non-degenerate semiconductors only, when the Fermi level lies within the forbidden band. As the impurity concentration increases, there is a point where the density of states N_c or N_v is exceeded. At this point the Fermi level either lies well above the conduction band edge or well below the valence band edge. (That is, $\eta / k_B T \gg 0$ for the n-type semiconductor.) In practise this implies that the semiconductor becomes degenerate and starts to behave like a metal. then the Seebeck coefficient can be expressed as

$$\alpha = \mp \left[\frac{\pi^2}{3} \frac{k_B}{e} \frac{s + \frac{3}{2}}{\eta} \right] + \varphi \tag{A.19}$$

As the charge carrier concentration increases, the phonon drag will decrease. To some extent this is caused by the increasing amount of charge carriers that cause the phonons to scatter. More importantly, momentum is transferred back from the phonons to the electrons, if these are very numerous. In the case of an n-type material, this saturation effect modifies Equation (A.18) to

$$\varphi = \mp \left(\frac{\mu T}{x_d v^2 \tau_d} + \frac{e}{k_B} \frac{3 n_0}{N_\varphi} \frac{x_d v^2 \tau_d}{\mu T} \right)^{-1} \tag{A.20}$$

where N_φ is the number of phonon modes that interact with the charge carriers.

A.2 ELECTRICAL CONDUCTIVITY

The electrical conductivity σ (or the reciprocal parameter, the electrical resistivity ρ) is the easiest to express parameter in the figure-of-merit, i.e.,

$$\sigma = \rho^{-1} = e(\mu_n n_0 + \mu_p p_0) \tag{A.21}$$

where μ is the mobility of the charge carrier and n and p are the concentrations of respectively the electrons and holes. As the electrical conductivity involves charge carriers only, the same scattering parameters as contribute to the Seebeck coefficient are expected to influence the electrical resistivity. Unfortunately, most books on thermoelectrics fail to break free from the quantum physical mathematics describing the scattering effects. Some simple approximation of the influence by lattice and ionised impurity scattering are provided in [3.24] and [A.6].

In general, the mobility, μ [$m^2V^{-1}s^{-1}$], of a charge carrier depends on the mean free time between collisions, τ_c, i.e.,

$$\mu = e\frac{\tau_c}{m^*} \tag{A.22}$$

Due to acoustic lattice scattering, for a nondegenerate single valley semiconductor, $\mu \propto T^{-3/2}$, while for a degenerate, multivalley semiconductor, $\mu \propto T^{-1/2}$ [A.6]. The second mechanism is ionised impurity scattering, where $\mu \propto T^{+3/2}/(N_a+N_d)$. Figure A.3 shows the mobility as function of the impurity concentration, for a number of semiconductor materials.

What the above is supposed to illustrate is that—due to the various interactions by the lattice, impurities and other charge carriers—the electrical conductivity exhibits a dependency on the doping concentration (or temperature) that is far from linear. Furthermore, Equation (A.22) indicates that the mobility μ and the effective mass m^* of the charge carrier are interdependent. The complicating matter of this situation is that this interdepency is governed by the charge scattering mechanisms. For example, for acoustic scattering, $\mu \propto (m^*)^{-5/2}$, while $\mu \propto (m^*)^{-1/2}$ in the case of ionised impurity scattering.

A.3 THERMAL CONDUCTIVITY

The thermal conductivity λ is the only parameter in the calculation of the figure-of-merit that is not necessarily dominated by the flow of charge carriers, as lattice vibrations (phonons) can also carry thermal energy across a material, i.e.,

$$\lambda = \lambda_L + \lambda_c \tag{A.23}$$

where λ_L is the lattice contribution and λ_c the charge carrier contribution. The latter can be broken down further, as charge carriers can be generated from impurities,

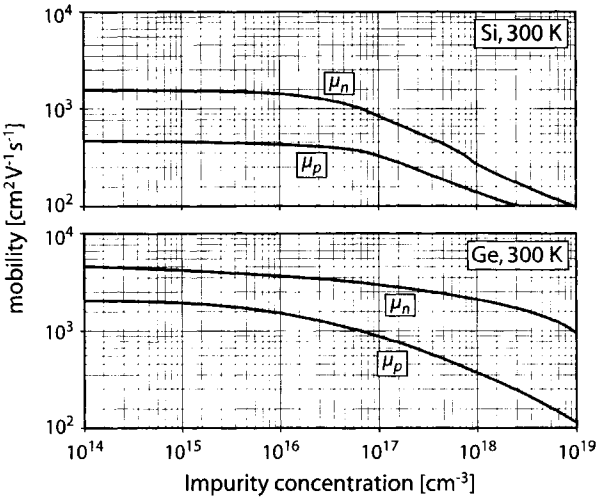


Figure A.3 The mobility of electrons and holes in silicon, germanium and gallium arsenide, as a function of the impurity concentration.

providing electrons or holes (polar contribution), or through thermal generation of electron-hole pairs (referred to as the bipolar contribution):

$$\lambda_c = \lambda_{pol} + \lambda_{bip} \tag{A.24}$$

For the temperature and doping range of interest for the thermoelectric cooler presented in this thesis (i.e., T ≈ 300 K and $N \approx 10^{20}$ cm^{-3}), the contribution by thermal excitation of charge carriers will be much smaller than the contribution due to impurities, so λ_{bip} can be safely ignored.

A.3.1 Electron thermal conduction

As stated, the electrons contribute to both the electrical and thermal conductivity. The polar (electron) contribution to the thermal conductivity can be expressed as

$$\lambda_{pol} = L\sigma T \tag{A.25}$$

where, L is the so-called Lorentz number and σ is the electrical conductivity. In itself, the Lorentz number is expressed as

$$L = \mathcal{L}\left(\frac{k_B}{e}\right)^2 \tag{A.26}$$

where the Lorentz factor $\mathcal{L}$ of a non-degenerate semiconductor is equal to

$$\mathcal{L} = \frac{5}{2} + s \tag{A.27}$$

in he case of classical statistics and

$$\mathcal{L} = \frac{(s+\frac{7}{2})F_{s+\frac{5}{2}}(\eta)}{(s+\frac{3}{2})F_{s+\frac{1}{2}}(\eta)} - \left[\frac{(s+\frac{5}{2})F_{s+\frac{3}{2}}(\eta)}{(s+\frac{3}{2})F_{s+\frac{1}{2}}(\eta)}\right]^2 \tag{A.28}$$

in the case Fermi-Dirac statistics are applied. Finally, in the case of a degenerate semiconductor, the Lorentz factor becomes

$$\mathcal{L} = \frac{\pi^2}{3} \tag{A.29}$$

Figure A.4 plots the Lorentz factor as a function of the reduced Fermi level, for the same three scattering parameters as in Figure A.2 ($s = -1/2$, $s = 1/2$ and $s = 3/2$.) Equation (A.25) is also known as the Wiedemann-Franz(-Lorentz) law, which directly relates the electrical conductivity and the electron contribution to the thermal conductivity of a material (as $\lambda_e/\sigma = LT$; See [A.13] for a simple derivation of this relationship.) Thus, the dependency of the polar thermal conductivity on impurity concentration follows directly from Equations (A.21) and (A.25).

In the case both the electron and hole contribution have to be considered (as these have different mobilities), Equation (A.25) should be expressed as

$$\lambda_{pol} = \frac{k_B^2 T}{e}\left((\tfrac{5}{2}+s_n)\mu_n n_0 + (\tfrac{5}{2}+s_p)\mu_p p_0\right) + \frac{k_B^2 T}{e}\left[\frac{\mu_n \mu_p n_0 p_0}{\mu_n n_0 + \mu_p p_0}\left(\frac{E_g(0)}{k_B T} + 5 + s_n + s_p\right)\right] \tag{A.30}$$

where $E_g(0)$ is the bandgap energy at 0 K. Figure A.3 shows that the mobility of the electrons and holes usually differs over an order of magnitude from moderately to heavily doped semiconductors, in which case the hole contribution can be neglected. However, for materials like bismuth telluride, the electron and hole contributions are of the same order of magnitude, so Equation (A.30) expresses the bipolar contribution more appropriately than Equation (A.25).

A.3.2 Lattice thermal conduction

Like with the Seebeck coefficient and electrical conductivity, scattering mechanisms can have a significant impact on the thermal conductivity. For example, intrinsic monocrystalline silicon has a thermal conductivity in the order of 145 $Wm^{-1}K^{-1}$, which must be attributed entirely to the lattice conduction, as there are hardly any charge carrier at room temperature. On the other hand, copper has a thermal con-

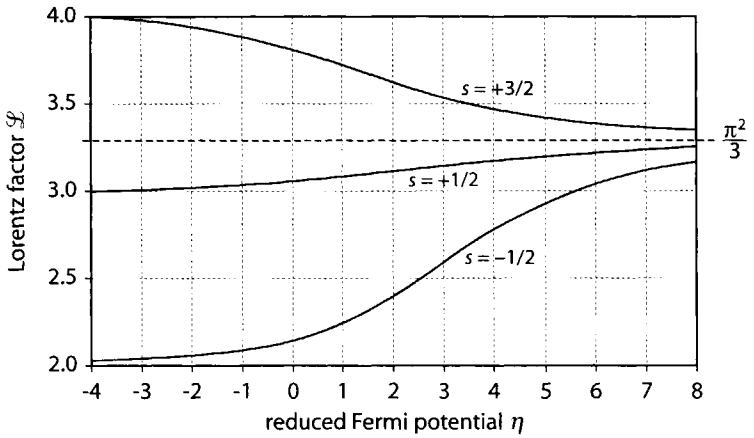


Figure A.4 The Lorentz factor as function of the reduced Fermi energy, for different values of the scattering parameter *s*. (After [2.2]).

ductivity in the order of 400 $Wm^{-1}K^{-1}$, of which only a few percent is lattice conduction. This can be explained from the high concentration of electrons that can scatter phonons very effectively [A.10].

Commonly, the thermal conductivity by the lattice is expressed as

$$\lambda_L = \frac{1}{3} C_v v_s l_t \tag{A.31}$$

where C_v is the specific heat per unit volume of the material, v_s is the average sound (phonon) velocity and l_t is the mean free path of the phonons. This mean free path decreases as the scattering increases.

Roughly speaking, the intensity of the lattice vibrations is proportional to the temperature. As a result, the mean free path of the phonons and, thus, the lattice thermal conduction will be inversely proportional to temperature. At very low temperatures, the thermal conductivity is proportional to T^3, which is the variation with temperature of the specific heat. The influence of the charge carriers on the lattice thermal conduction is only indirect, as electrons and holes can act as scattering sources for the phonons. Below, the most important scattering mechanisms of phonons discussed in references [2.2], [3.32], [A.3], [A.6] and [A.10] are summarised.

A.3.3 Phonon scattering mechanisms

In general, any mechanism that scatters phonons more effectively than charge carriers is likely to enhance the performance of a thermoelectric material. Going through literature, three groups of scattering sources can be identified: (1) scattering by other

phonons (2) scattering by charge carriers and (3) scattering by imperfections in the crystal lattice. Each is summarised below.

Phonon-phonon scattering. At high temperatures (near the Debye temperature and above) phonons are scattered predominantly by other phonons. The most significant phonon-phonon interaction is the so-called three-phonon process, where either two phonons combine to form a new phonon or one phonon breaks up into two new phonons. In general, these three-phonon interactions can be described as

$$\mathbf{q}_1 + \mathbf{q}_2 = \mathbf{q}_3 + \mathbf{G} \tag{A.32}$$

Where $\mathbf{q}_x$ is the phonon wave vector and $\mathbf{G}$ is a reciprocal lattice vector. In the case $\mathbf{G} = 0$, the phonon momentum is conserved, which is generally referred to as a Normal- or N-process. In the situation that $\mathbf{G} \neq 0$, the momentum is not conserved, which is referred to as an Umklapp- or U-process. Only the U-processes can limit the mean-free-path of a phonon. Therefore U-processes need to be considered in order to explain thermal resistance due to phonon-phonon interactions. In a first approximation, for high temperatures phonon scattering (primarily by the other phonons) the lattice thermal conductivity $\lambda_L \propto T^{-1}$. Instead, at very low temperatures, the thermal conductivity is proportional to T^3, caused by the variation with temperature of the specific heat (also see Equation (A.31).)

Scattering by electrons. In particular in degenerate semiconductors and metals charge carriers not only act as particles that aid the heat flow along the temperature gradient, but also act as scattering centres for phonons, obstructing the heat flow. In general the lattice thermal conductivity decreases with increasing carrier concentration. As such, the lattice thermal conductivity in metals will be small compared to the electronic contribution due to a considerable scattering of phonons by charge carriers. In semiconductors, the concentration of free charge carriers can be adjusted by suitable doping, so the ratio of electronic and lattice thermal conductivities will vary. Nevertheless, according to [A.6] scattering by electrons is negligible in semiconductors that are used for thermoelectric applications.

Scattering due to lattice imperfections (in general). Another important phonon scattering mechanism arises from the finite size of a crystal and various types of imperfections within. These include substitutional impurities, interstitials and dislocations. A phonon may be scattered by any or more of these imperfections. The effects of these defects may be considerable in thermoelectric materials, as impurities and polycrystallinity are deliberately introduced. The normal phonon modes can change depending on the nature of the impurity. In addition, new modes can appear, even in the frequency ranges forbidden to the normal modes of the host crystal.

In [A.10] the lattice defects are divided into different groups according to their dimensionality, i.e., point defects, line imperfections, surfaces of imperfections and volume disorder. If the diameter of an imperfection is smaller than a phonon wavelength, the imperfection is considered to be a point defect. These commonly include substitutional impurities, interstitial atoms and vacancies. Dislocations are typical

examples of line imperfections, whereas grain boundaries and stacking faults are classified as surfaces of imperfections. Finally, the nature of imperfections in substitutional alloys and glasses can be referred to as volume disorder.

Scattering by point defects. The scattering of phonons by point defects is caused by the differences in physical properties, as an impurity atom differs from the host atoms (in its mass, its size, etc.) Point defect scattering is very effective for short-wavelength (optical) phonons, because of an ω^4 dependency (where ω is the phonon frequency.)

Scattering due to alloy disorder. Alloy disorder should be seen as a special case of introduction of point defects. As such, according to [A.6] alloy scattering is no different from the scattering by impurity atoms. According to [A.3] in alloys of two isomorphous crystals (like SiGe) the long-range periodicity of the lattice is maintained but, owing to the difference between the sizes of atoms in adjacent corresponding lattice sites, there are considerable short-range distortions. The wavelength associated with an electron in a semiconductor is of the order of some tens of interatomic distances so that these are hardly scattered by disturbances in the short-range periodicity of the lattice. On the other hand, the wavelength of phonons is of the same order of magnitude as the interatomic distance. Thus, phonons will be scattered significantly if the short-range periodicity is upset.

The scattering of phonons by alloying is of significant importance in thermoelectricity. In many cases an alloy of different constituents will have a lattice thermal conductivity substantially lower than that of either pure constituent, while the mobility of the charge carriers remains largely unaffected. For example, Steele and Rosi (See [A.6]), have investigated the influence of the lattice conductivity as a function of the Si-Ge composition. Figure A.5 displays their measurement result. This preferential scattering is one of the main mechanisms behind the fairly good thermoelectric properties of polycrystalline silicon-germanium and other compound semiconductors.

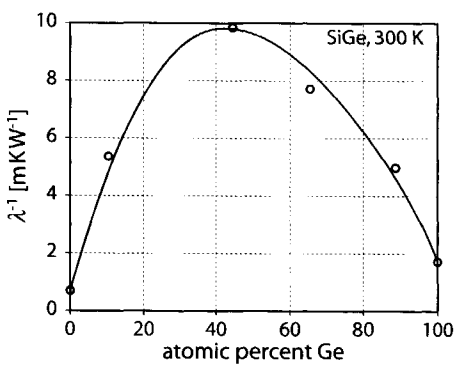


Figure A.5 Thermal resistivity at 300 K, as a function of the Si–Ge composition. (From: [A.6].)

Scattering by dislocations. The strain-field near a dislocation can cause significant phonon scattering. Still, dislocation scattering of phonons is not expected to play a significant role in thermoelectrics as the maximum dislocation density is in the order of 10^{12}, which appears to be at least an order of magnitude too small [A.6].

(Grain) boundary scattering. In a crystalline solid, phonons are scattered at the boundaries between crystallites of different orientation. The phonon mean free path l_t cannot exceed the diameter of the crystal sample, when it is scattered at its boundary. Although most prominent at low temperatures, boundary scattering can still be observed in certain materials at ordinary temperatures.

A high degree of disorder effectively scatters short-wavelength (mostly optical) phonons with the lattice conduction being primarily due to long-wavelength (acoustic) phonons. In amorphous solids such as glasses, these high-frequency phonons cannot even propagate for more than a few interatomic distances.

Grain boundary scattering occurs frequently in polycrystalline samples, where the maximum free path length of the phonons is expected to be limited by the grain size. Thus, preparation of fine-grained materials with the required degree of disorder can effectively reduce the thermal conductivity. Enclosed in this consideration is the assumption that the electrical conductivity is not influenced significantly by the grain boundaries. Varies studies have suggested appreciable improvements in the figure-of-merit of fine-grained SiGe alloys, although more recent studies disagree with the earlier results [A.14].

A.4 FIGURE-OF-MERIT AND DOPING CONCENTRATION

Theoretically it is possible to calculate the optimum figure-of-merit ($\alpha^2\sigma/\lambda$) and power factor ($\alpha^2\sigma$) directly from the reduced Fermi potential. For the power factor, the solution has been plotted in Figure A.6.

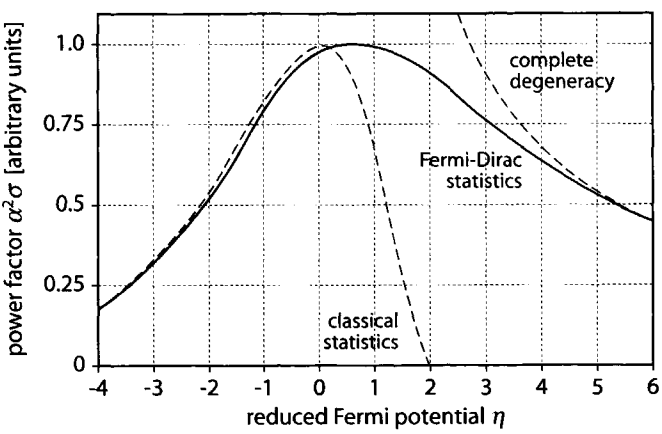


Figure A.6 The power factor as function of the reduced Fermi potential, for $s = -1/2$. [A.3]

According to Fermi-Dirac statistics, the optimum value of η is equal to -1/2. As this implies that the semiconductor has not yet degenerated, this value is quite close to the optimum predicted from the classical statistics (η = -1.) However, an exact calculation based on the optimum Fermi potential is of limited use in practise. At the optimum doping concentration, near the onset of degeneracy, it is very difficult to obtain a reliable value for the scattering parameter. At even higher impurity concentrations, the Fermi level moves into the conduction or valence band, at which point it becomes nearly impossible to calculate the exact thermoelectric property values from the Fermi potential. The only accurate determination of the thermoelectric properties remains through measurements.

As a final item, based on the observations and equations in this Appendix, it is possible to provide an estimate of the figure-of-merit as a function of the impurity concentration. Previously, Equations (A.11) and (A.12) revealed that the Seebeck coefficient will roughly decrease with $\log(N_x)$, where N_x is the impurity concentration. From Equation (A.21) it follows that the electrical conductivity $\sigma \propto N_x$. The third parameter, the thermal conductivity λ, consists of two components, the lattice thermal conductivity λ_L and the electron thermal conductivity λ_{pol}. The lattice thermal conductivity will only slightly increase with impurity concentration, as the phonon scattering will increase. The Wiedemann-Franz-Lorentz law revealed that the ratio of the electrical and electron thermal conductivity is constant, so $\lambda_{pol} \propto N_x$. All five parameters, α, σ, λ, λ_L and λ_{pol} are plotted in Figure A.7. When combining these parameters into the figure of merit $z = \alpha^2\sigma/(\lambda_L + \lambda_{pol})$, the typical bell-shaped curve as plotted in Figure A.7 appears. This curve in commonly used to show that the optimum thermoelectric material will be a heavily doped semiconductor or a semimetal.

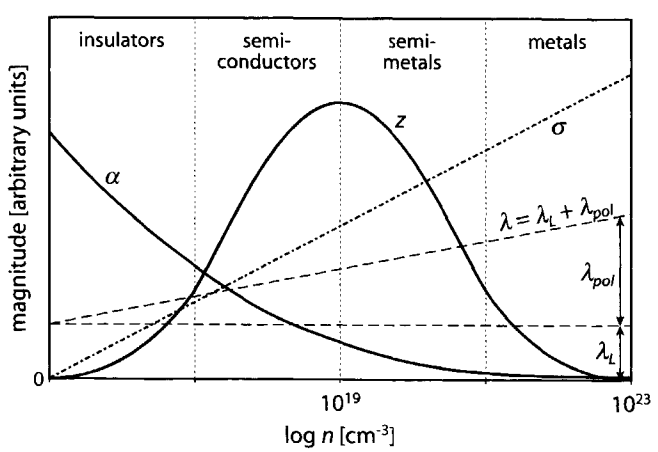


Figure A.7 The Seebeck coefficient α, electrical conductivity σ, thermal conductivity λ and figure-of-merit $\alpha^2\sigma/\lambda$, as a function of impurity concentration.

A.5 REFERENCES

[A.1] C.B. Vining, ITS Short Course on Thermoelectricity, 8 November 1993, Yokohama, Japan. (http://www.zts.com/ztreports/index.html)

[A.2] S. Middelhoek and S.A. Audet, *Silicon sensors*, Delft University Press, Delft, The Netherlands, 1994.

[A.3] H.J. Goldsmid, *Applications of thermoelectricity*, Methuen & Co, London, UK, 1960.

[A.4] J. Jäckle, "Über die Ursache der Thermospannung," July 1998, 15pp. (http://www.uni-konstanz.de/physik/Jaeckle/papers/thermopower/index.html)

[A.5] D.K.C. MacDonald, *Thermoelectricity: an introduction to the principles*, John Wiley & Sons, New York, NY, USA, 1962.

[A.6] R.R. Heikes and R.W. Ure jr., *Thermoelectricity: science and engineering*, Interscience Publishers, New York, NY, USA, 1961.

[A.7] M. Lundstrom, *Fundamentals of carrier transport*, second edition, Cambridge university press, Cambrige, UK, 2000, ISBN: 0-521-63134-3.

[A.8] D.M. Rowe and C.M. Bhandari, *Modern thermoelectrics*, Holt, Rinehart and Winston, London, UK, 1983, ISBN: 0-03-910433-8.

[A.9] A.F. Ioffe, *Semiconductor thermoelements and thermoelectric cooling*, Infosearch, London, UK, 1957.

[A.10] C.M. Bhandari and D.M. Rowe, *Thermal conduction in semiconductors*, John Wiley & Sons, New York, NY, USA, 1988, ISBN: 0-470-21049-4.

[A.11] M. Glicksman, "Mobility of electrons in germanium–silicon alloys," *Physical Review*, Vol. 111, No. 1, pp. 125–128, 1958.

[A.12] N.K.S. Gaur, C.M. Bhandari and G.S. Verma, "Temperature dependence of density-of-states effective mass and the electronic and phonon contributions to thermal resistance of doped Si-Ge alloys at high temperatures," *Physical Review*, Vol. 144, No. 2, pp. 628–641, 1966.

[A.13] N. Garcia and A. Damask, *Physics for computer science students*, Springer- Verlag, New York, NY, USA, 1991, ISBN: 0-387-97656-6.

[A.14] G.A. Slack and M.A. Hussain, "The maximum possible conversion efficiency of silicon-germanium thermoelectric generators," *J. Appl. Phys.*, Vol. 70, Iss. 5, pp. 2694–2718, 1991.

Appendix B

OVERVIEW OF THERMOELECTRIC MATERIALS

The current state of research on thin-film thermoelectric materials was discussed in Chapter 3. However, thin-film research covers only a small portion of the total research on thermoelectric materials. This appendix provides an overview of the other work performed. Section B.1 presents an overview of the established thermoelectric materials. Section B.2 discusses the current research directions towards improvement of existing materials as well as the development of new materials.

B.1 ESTABLISHED THERMOELECTRIC MATERIALS

In order to develop a good thermoelectric device, the key factor that will determine the maximum performance is the thermoelectric material itself. An application can never get better than the product of its material parameters. Depending on the application that will be either be the power factor $\alpha^2\sigma$ or the figure-of-merit, $\alpha^2\sigma/\lambda$. This section provides an overview of the various thermoelectric materials that are available, as well as the most prominent approaches towards the development of new materials. In the span of a few pages it is impossible to be complete. Instead, a high-level overview is provided, with reference to some more elaborate texts and articles.

B.1.1 Fabrication techniques for conventional materials

Before discussing the newer thermoelectric materials like the thin-film materials required for on-chip integration, the most important established (bulk) materials and the corresponding fabrication techniques are briefly reviewed.

The majority of bulk thermoelectric materials are either prepared from a melt or through powder metallurgy. The latter category includes sintering, hot-pressing, pulverised and intermixed elements sintering (PIES) and mechanical alloying. The basics of each of these technologies is outlined below. For more details the reader is referred to [2.2] and [3.32].

B.1.1.1 Preparation from a melt

In general, the best thermoelectric materials are those prepared from a melt, as a better level of homogeneity can be obtained. The basic process of preparing a material from a melt consists of two steps, synthesis and crystal growth. Synthesis refers to the melting and mixing of the various constituents of the thermoelectric material. To prevent contamination, synthesis has to be done highly purified elements, usually under a protective atmosphere while using a quartz tube in which the (solid) constituents are collected. The tube is sealed in the protective atmosphere, after which the tube is heated until all constituents melt. The resulting melt is carefully stirred until a homogeneous mixture is obtained.

After the melt has been homogenised, the actual crystal growth is performed. In [2.2] various techniques are discussed, that can be divided into two categories of growth, either from a stoichiometric or non-stoichiometric melt. In the first case, the final crystal will have (nearly) the same composition as the melt started with. In the latter case, the composition of the final crystal may vary significantly from that of the original melt. Figure B.1a illustrates the latter situation, for the bismuth–antimony system. To obtain a solid of composition $Bi_{1-y}Sb_y$, the original melt has to contain Bi and Sb in a $(1-x):x$ composition (with $x < y$). As the temperature of the liquid solution is slowly lowered to T_{x-y}, crystals of the required composition $Bi_{1-y}Sb_y$ start to appear.

One potential problem of the system in Figure B.1a is so-called constitutional supercooling, which occurs if the temperature is lowered too quickly or—equivalently—if the temperature gradient across localised melt in a solid ingot is too large, freezing will occur in advance of the interface, so the solid will contain multiple compositions. Constitutional supercooling is of particular importance in systems where there is a large transitional region in which both liquid and solid solutions are present, as in Figure B.1a. The systems of Figure B.1b and Figure B.1c are much less susceptible, as the intermediate region is much smaller. The difference between the compositions x and y at any given T is very small, limiting the range of compositions that can occur.

Examples of crystallisation from a stoichiometric melt are the Bridgman method, zone melting and gradient freezing. In the Bridgman method, the melt is lowered from a high-temperature zone to a low-temperature zone. The actual crystal growth occurs in the narrow region between the two zones, across which the temperature gradient between the hot- and cold zones is located. Using the Bridgman method, samples of for example PbTe, PbSe, Bi_2Te_3, Sb_2Te_3, $(Bi_2Te_3)_{25}(Sb_2Te_3)_{72}(Sb_2Se_3)_3$ as well as other Bi-Te-Sb compounds have been created.

With a zone melt, a pre-synthesised solid bar is locally heated above the melting point. By slowly moving the liquid zone across the bar, crystallisation in the required composition will occur. With a gradient freeze, no mechanical action is required. Instead, a temperature gradient across the whole width of the melt is applied. By slowly lowering the temperature, but keeping the gradient constant, the lower-temperature

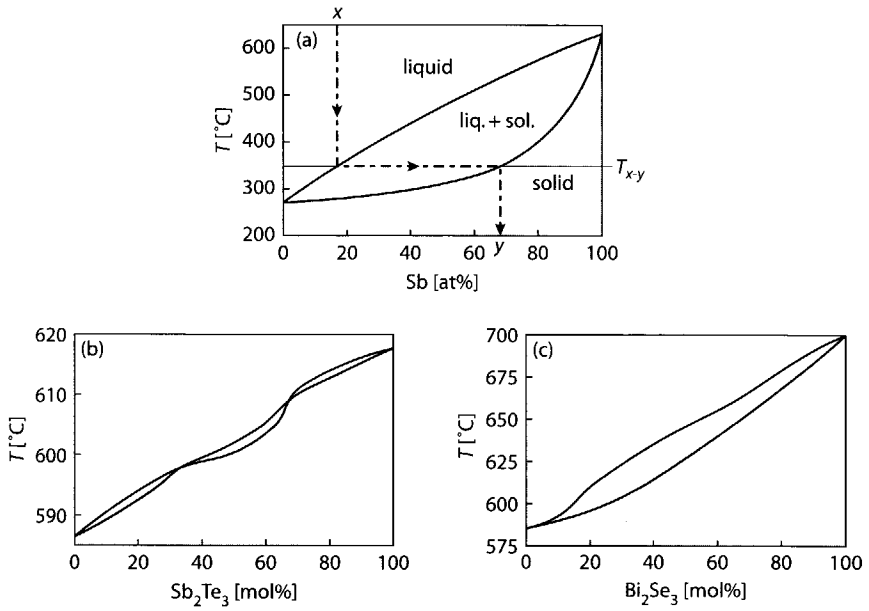


Figure B.1 (a) Solid solution phase diagram for BiSb, as a function of the atomic percentage of Sb, after [B.1]. In (b) and (c) the pseudobinary phase diagrams of the Bi_2Te_3–Sb_2Te_3 and Bi_2Te_3-Bi_2Se_3 systems are shown.

end of the melt will first start to crystallise, until the higher-end temperature reaches the temperature at which solidification takes place. At that point the entire bar has been solidified into the required composition. Examples of zone melting include Bi_2Te_3-Sb_2Te_3 quasi-binary compounds as well as $Bi_{1.6}Sb_{0.3}Te_3$ alloys.

One particularly well-known crystallisation technique from stoichiometric melts is the Czochralski method, where a large crystal is pulled from a liquid, starting from only a small seed crystal. This is the dominant technique in preparing high-quality silicon wafers. With respect to thermoelectrics, the Czochralski method has been used to create crystals of PbTe, Bi_2Te_3, Sb_2Te_3 as well as Bi_2Te_3-Bi_2Se_3-Sb_2Te_3 compounds.

Examples of crystal growth from non-stoichiometric melts are gradient freezing, zone levelling (a specific variety of zone melting), the travelling solution method (another kind of zone melting technique), and liquid phase epitaxy. In the latter case an material is grown epitaxially on an oriented seed layer, whereby the structure of the seed layer is reproduced during growth.

Preparation of silicon-germanium from melts

Silicon–germanium compounds are difficult to prepare from a melt. First, the liquidus point of the preferred composition ($Si_{0.8}Ge_{0.2}$) lies at 1350 °C, which rules out

the use of quartz ampoules, which soften at temperatures above 1250 °C. Moreover, there is a large gap of 70 °C between the liquidus and solidus points, combined with very low diffusion rates for both components in the solid alloy. The phase diagram of the silicon-germanium system is shown in Figure B.2. Unless an excessively long time is used for crystallisation, heavy segregation occurs (that separates the silicon- and germanium-rich crystals.)

Although difficult, alloying of polySiGe from melts is performed nonetheless, for example using an RF furnace [B.2]. An open graphite crucible holds both (10–20 nm) silicon and germanium particles, as well as the required dopants (P or GaP for n-type material and B for p-type material.) After evacuation of the furnace, a slightly overpressurised (~1.1 atm) Ar atmosphere is established. The temperature of the crucible is raised to 1370–1400 °C in 2–3 h. The liquid is stirred using Eddy currents whereby a homogeneous solution can be is obtained in under 50 min. The melt is then cast in a water-cooled copper mold, yielding rods with a diameter of 12.5 mm and a length of 100 mm.

Alternatively, zone levelling [B.3]-[B.5] as well as liquid phase epitaxy (LPE) can be applied to create high-quality polySiGe samples. The latter technique yields layers with a thickness in the range of 10–100 μm [B.6]. Liquid phase epitaxy has a two-fold advantage over other conventional techniques. First, polySiGe layers can be grown at a temperature well below the liquidus temperature, i.e at 900 °C in stead of 1350 °C. Second, by LPE the phosphorus solubility can be improved, which in turn improves the figure-of-merit of $Si_{0.8}Ge_{0.2}$ [B.7].

B.1.1.2 Powder metallurgy

In general powder metallurgy refers to the process whereby a powder mix of multiple constituents is compacted and annealed. In [2.2], four basic categories of powder metallurgy are distinguished, pressureless sintering, hot-pressing, pulverised and intermixed elements sintering (PIES) and mechanical alloying.

Even though a thermoelectric material created through powder metallurgy will generally have a lower figure-of-merit, there are a number of distinct advantages. First, materials created from powder metallurgy are more robust than melt-grown materials. Second, less precautions are required to prevent contamination as well as to ensure homogeneity. Third, a sintered material has the potential to improve the figure-of-merit by increasing phonon scattering at the resulting grain boundaries, in particular for polySiGe [B.8].

Pressureless sintering

The process of pressureless sintering is best illustrated in [3.32]. After the components are molten together, the cast is pulverised by milling and the powder is sieved to obtain grains of a particular size (around 100–250 μm; smaller particles become too susceptible to atmospheric contamination.) Next, the powder is compacted (at a pressure of about $5 \cdot 10^8$ Pa) to more-or-less the right dimensions, after which the resulting pellet is annealed / sintered (up to an hour at 400 °C in an inert atmosphere.)

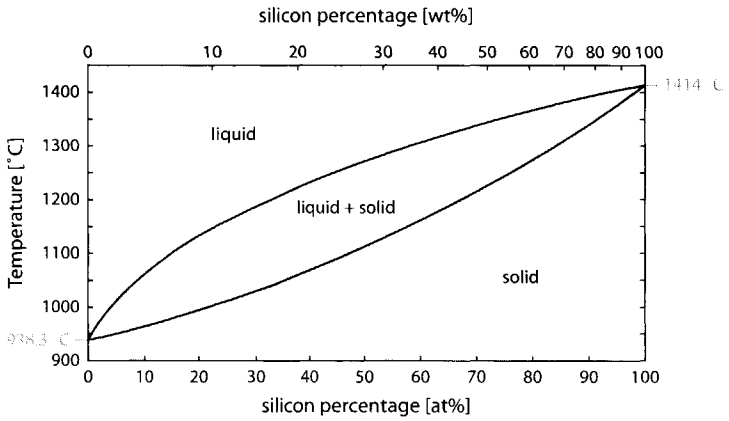


Figure B.2 Binary phase diagram of silicon-germanium.

During sintering, the grain size will increase, while the volume and pore size decreases. Small grains are consumed so the total number of grains will gradually decrease.

Hot pressing

Hot pressing is almost the same as pressureless sintering and is only applied when the aforementioned technique is impractical. If densification through pressing is insufficient at room temperature, the temperature can be raised to improve the densification rate. Addition of a small amount of liquified components helps increase both the rate of transport and rearrangement of the constituents. Nevertheless, hot pressing is less economical than pressureless sintering and is avoided whenever possible as such.

PIES

The largest difference between the more conventional cold-press sintering and the PIES method lies in the very beginning of processing. With cold pressing, the various elements and dopants are molten together, which is the primary stage for material alloying. This is followed by milling of the solidified solution, at which point a powder is created that consists of grains that all have the crystal structure of the solid solution. In contrast, with PIES, all elements and dopants are still intermixed but not molten together. A powder results that is a mixture of individual elements and dopants. As milling is the first step in proper alloying of the material, much finer grains are required (from 10 μm downto the submicron range). These steps are followed by cold pressing and sintering. In the case of the PIES method, sintering is the second step in alloying of the thermoelectric material.

As powder metallurgy generally is a much more economic method of material preparation, this method has received a significant amount of attention to replace material preparation from melts. Nevertheless, maybe with the exception of the PIES

preparation technique, the resulting thermoelectric material performance is always less than the performance of a melt-grown material. In the case of Bi_2Te_3 and related alloys (Sb_2Te_3 and Bi_2Se_3), the material performance is strongly direction-dependent. As powder metallurgy randomly orients all grains, a strong degradation in material performance is witnessed. Moreover, the electrical parameters of these alloys are degraded by oxidation of the powdered alloy, which is more noticeable as the grain size decreases. Still, good results have been obtained using hot-pressing, with a z of $2.74 \cdot 10^{-3}$ K^{-1} for n-type and a z of $1.9 \cdot 10^{-3}$ K^{-1} for p-type Bi_2Te_3 [2.2].

Mechanical alloying

Although primarily intended for the fabrication of high-strength alloys, mechanical alloying is successfully used for powder metallurgy.

Mechanical alloying is closest to the PIES method in how mixing of the elements is obtained yet differs significantly with respect to thermal treatment of the materials. No sintering or any other high-temperature processing steps are required. The high-energy collisions of the steel balls with the thermoelectric powder mixture are sufficient to achieve alloying (referred to as cold welding.) Due to the high-energy collisions during milling a significant amount of energy is generated, potentially raising the temperature of the alloy up to 350 K above ambient temperature. However, this is still much less than would be required for recrystallisation.

Like with the PIES method, mechanical alloying 1) removes the problem of inhomogeneity in the alloy due to segregation during solidification of a melt and 2) the volatility of dopants is much less of an issue. Moreover, during milling, nanometer-size particles can be introduced in the thermoelectric alloy, to increase phonon scattering without significantly affecting electron scattering [B.9]. The main concern regarding mechanical alloying is contamination of the thermoelectric material with the elements of which the milling balls are made, e.g., W or Fe.

B.1.2 Overview of established materials

It is impossible (and meaningless) to provide a comprehensive overview of established materials, as this only duplicates the matter found in e.g. [2.2][3.32] or [A.6]. Instead the emphasis is placed on the thermoelectric performance of these materials, as a reference for thin-film polySiGe.

B.1.2.1 Bismuthides and tellurides (Bi_2Te_3, Sb_2Te_3, Bi_2Se_3)

By far the most common materials used and studied in thermoelectric applications are the bismuthides (e.g., Bi_2Te_3, Bi_2Sb_3, etc.) and tellurides (Sb_2Te_3, etc.) Of these, materials, bismuth telluride (Bi_2Te_3) has been studied the most by far. Fabrication of these conventional bulk-type bismuthides and tellurides can be done by any of the techniques described in Section B.1.1. Bi_2Te_3 has a hexagonal layer structure in which each layer has atoms of a single type. The arrangement can be described as

—Te-Bi-Te-Bi-Te—Te-Bi-Te-Bi-Te—

where the bond between the two tellurium layers is much weaker than the ionic-covalent bonds between the Bi and Te layers. As a result of this weaker bonds—consisting of both Van der Waals bonds as well as some covalent bonds—bismuth telluride has the pleasant characteristic of being very easy to cleave, which greatly facilitates processing of this material. (This applies to Sb_2Te_3 and Bi_2Se_3 as well.)

When grown from a stoichiometric melt or by zone refining, the Bi_2Te_3 crystals are always p-type. Furthermore, an excess of Bi, or addition of Pb, Cd or Sn will also yield a p-type material. When grown from a melt that has an excess tellurium concentration, or when I, Br, Li, Al, CuBr, Cu_2S, or AgI are used, an n-type material is created. Sb_2Te_3 has the same (rhombohedral) crystal structure as Bi_2Te_3 and normally is n-type. A thorough review of the properties of Bi_2Te_3 and Sb_2Te_3 (and their solid solutions) is found in [2.2].

The best figures-of-merit reported for pure Bi_2Te_3 in the 1960's (e.g., as in [B.10]) are about $2.2 \cdot 10^{-3}$ K^{-1} for p-type material an $2.6 \cdot 10^{-3}$ K^{-1} for n-type material. Since the 1960's, these values have only improved marginally, e.g., as reported in [B.9][B.11]. The values from these references are shown in Table B.1. (For Sb_2Te_3, the values listed are those obtained along the cleavage plane.)

In Figure B.6 on page 201, the typical figure-of-merit of bismuthides and tellurides is plotted as a function of temperature. As can also be seen from Table B.1, Sb_2Te_3 (which is always p-type) and Bi_2Se_3 have a figure-of-merit roughly five times less than Bi_2Te_3. For this reason these two materials have not received nearly as much attention as Bi_2Te_3.

Table B.1 Thermoelectric material data for bismuthides and tellurides [B.11]

Material		Seebeck coefficient [$\mu V K^{-1}$]	Electrical resistivity [$\mu\Omega m$]	Thermal conductivity [$Wm^{-1}K^{-1}$]	Figure-of-merit [$10^{-3} K^{-1}$]
Bi_2Te_3	n	-240	10.0	2.02	2.89
	p	162	5.5	2.06	2.32
Sb_2Te_3* (Sb-rich)		53.9	1.36	6.35	0.373
Sb_2Te_3* (Te-rich)		83.4	1.89	5.62	0.654

B.1.2.2 The Bi_2Te_3–Sb_2Te_3–Bi_2Se_3 system—$(Bi_{1-x}Sb_x)_2(Te_{1-y}Se_y)_3$

In contrast to using individual Sb_2Te_3 or Bi_2Se_3 alloys, a solid solution mixed with Bi_2Te_3, can substantially reduce the lattice thermal conductivity. This effect has been reported by multiple authors. In Figure B.3, the thermal conductivity is plotted as a function of the material composition.

The difference in thermal conductivity about a factor of three is primarily caused by the difference in Lorentz number used to calculate the polar contribution. Rosi *et al.* (See [2.2]) used degenerate samples, where λ_{pol} can be described in terms of the

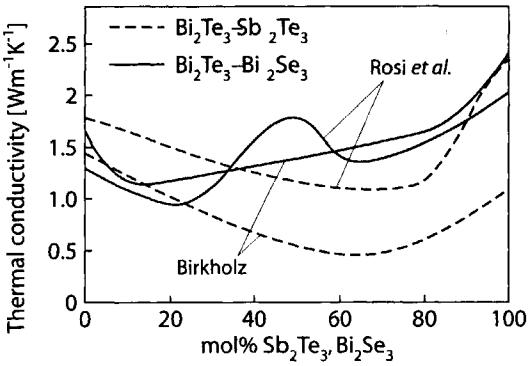


Figure B.3 Thermal conductivity as a function of alloying composition for Bi_2Te_3 systems. From [2.2].

Lorentz number of Equation (A.29), i.e. $\lambda_{pol} = (\pi^2/3)(k_B/e)^2\sigma T$. Birkholz used the nondegenate approximation, so $\lambda_{pol} = 2(k_B/e)^2\sigma T$.

For the $(Bi_{1-x}Sb_x)_3Te_3$ system, the figure-of-merit peaks around $x = 0.75$, where z becomes in the order of $3.3{\cdot}10^{-3}$ K^{-1} [B.12]. In general, when $x > 0.5$, $(Bi_{1-x}Sb_x)_3Te_3$ becomes p-type. Similarly, $Bi_2(Te_{1-y}Se_y)_3$ is an n-type material at its optimal composition. The optimal $Bi_2(Te_{1-y}Se_y)_3$, composition is obtained for $y = 0.15$, for which z reaches a value of $3.0{\cdot}10^{-3}$ K^{-1} [3.32]. In Figure B.6 the figure-of-merit as a function of temperature is shown for various $(Bi,Sb)_2(Se,Te)_3$ alloys.

The appeal of these alloys, like with bismuth- and antimony telluride, is that the figure-of-merit peaks near room temperature, making these alloys highly suitable for near- and sub-room temperature applications.

B.1.2.3 Lead telluride and lead selenide alloys (PbTe, PbSe, $Pb_{1-x}Sn_xTe$)

Lead telluride (PbTe) and similar alloys are intermetallic compounds that have a cubic NaCl lattice. In contrast to bismuthides and tellurides, these $A^{IV}B^{VI}$ materials are most effective in the 600 – 900 K region [B.13].

PbTe and $Pb_{1-x}Sn_xTe$ can be fabricated from both fine-grained polycrystalline material as well as single crystals. Due to the NaCl lattice, there is no preferred material direction, so polycrystalline and monocrystalline samples should perform equally. Both sintering and growth techniques are used to fabricate PbTe-based compounds (by vapour phase epitaxy for small crystals, from melts for larger crystals.) Polycrystalline material is obtained from a direct synthesis of oxygen-free, > 99.90% pure elements.

Undoped PbTe has a doping concentration of around 10^{18} cm^{-3}. This can be increased to about 10^{19} cm^{-3} by including an excess of Pb or Te (in the order of a few hundreth of a percent.) Still, 10^{19} cm^{-3} is about an order of magnitude too low for thermoelectric applications. To increase the doping concentration beyond this level,

either the Pb or Te atom must be substituted. Materials such as O_2, Na, Au and Tl act as p-type dopants, while materials such as Zn, Cd, In, Br_2, Cl_2 and Bi act as donors.

In Figure B.6, the figure-of-merit for a number of different cast PbTe compositions is shown. At room temperature, z is in the order of $1 \cdot 10^{-3}$ K^{-1} only, but becomes superior to bismuthides and tellurides around 450 – 550 K. Generally, the optimum figure-of-merit can be tuned in the range from 400 K to 600 K, by changing the doping profile. Generally, a higher doping concentration reduces the figure-of-merit for PbTe, but shifts the relative optimum to a higher temperature.

B.1.2.4 Polycrystalline silicon germanium (polySiGe)

Silicon germanium is one of the best examples of the benefit of how phonon scattering can improve the figure-of-merit of a material. Consider Figure A.5; Neither silicon nor germanium are good thermoelectric materials. They respectively have a thermal conductivity of 113 and 63 $Wm^{-1}K^{-1}$, which is far too high for thermoelectrics. However, when both materials are alloyed, the thermal conductivity decreases to around 5 $Wm^{-1}K^{-1}$ [B.14]. This scattering due to alloying reduces the thermal conductivity primarily by scattering the high-frequency phonons [3.32]. A λ_L of 5 $Wm^{-1}K^{-1}$ corresponds to a phonon mean free path of about 3 interatomic distances only [A.3]. This is confirmed from Slack and Hussain [B.15] who estimate that the minimal lattice thermal conductivity of poly$Si_{0.7}Ge_{0.3}$ is 0.9 $Wm^{-1}K^{-1}$ (when the phonon mean free path is reduced to about 1 interatomic distance). When this minimum lattice thermal conductivity is obtained, the dimensionless figure-of-merit is expected to go up from 0.3 (at 500 K) and 1.1 (at 1200 K) to 1.0 (at 500 K) and 1.8 (at 1200 K), respectively.

Conventional SiGe / polySiGe fabrication techniques

As was already discussed in Sections B.1.1.1, preparation from a melt is quite complicated, as there exists a large separation between the liquidus and solidus (See Figure B.2). As a result, severe alloy segregation may occur. Two other techniques, zone levelling and hot pressing, have been applied with more success [2.2]. Nevertheless as zone levelling also creates SiGe from a melt, care has to be taken to prevent constitutional supercooling. This can only be avoided by keeping the growth rate below 1 mm/h.

A significant advantage of hot-pressing is that process is relatively insensitive to variations in the preparation conditions [2.2]. A 4:1 atomic ratio of Si:Ge is molten in a sealed and evacuated crucible. For p-type material 0.08 wt% B is added, while 0.55 wt% P is added for n-type material. After the melt is cast, the solidified solution is pulverized / milled to obtain particles in the order of 2–10 μm. Smaller particles are favourable for both homogeneity as well as mechanical strength. Finally, pressure sintering is performed at a pressure of 180 MPa and a temperature of 1513 K [2.2].

Hot-pressed and PIES-based materials have a thermoelectric performance comparable to that of melt-grown materials [B.16]–[B.18]. Moreover, as is reported in [B.19][B.20], the dimensionless figure-of-merit is relative insensitive to composition.

So, even when a material is not perfectly homogenised, a good performance is obtained. It has been demonstrated that the fabrication of poly$Si_{0.8}Ge_{0.2}$ by means of mechanical alloying can be done in as little as 4–6 h, e.g. in [B.21].

Grain size

One interesting discussion, which hasn't been decided yet, is what the optimum grain size is to maximise the phonon scattering. The experiments of Savvides and Goldsmid [B.22][B.23] on fine-grained polySiGe revealed that the lattice conductivity of undoped $Si_{0.7}Ge_{0.3}$ at 300 K falls from 8.2 $Wm^{-1}K^{-1}$ to 4.3 $Wm^{-1}K^{-1}$ when the mean grain size is reduced to 2 μm. Such a reduction is even more remarkable in the sense that the electron and hole mobility was hardly influenced.

Slack and Hussain also discussed the changes in the thermal and electrical conductivities of polySiGe alloys that result from boundary scattering due to small grain sized specimen [B.15]. Their conclusions were that a small increase in z due to the associated reduction in λ_L is obtained as the grain size is reduced to 2 μm. They estimate a further reduction of the grain size below this value will drastically reduces z, as charge carrier scattering will decrease the electrical conductivity more than the thermal conductivity. However, this estimation is contradictory to the findings of Bhandari and Rowe [B.24], which indicate that the lattice thermal conductivity may be further reduced by a grain size smaller than 2 μm, as is revealed from Figure B.4.

When the SiGe crystals are large (i.e. when the number of grain boundaries is small) a reduction of the thermal conductivity may be obtained by inclusion of extremely small particles (~ 40–100 μm) inside the grains that further scatter the phonons without influencing the electrical properties. Some small improvements using this technique have already been reported [2.2].

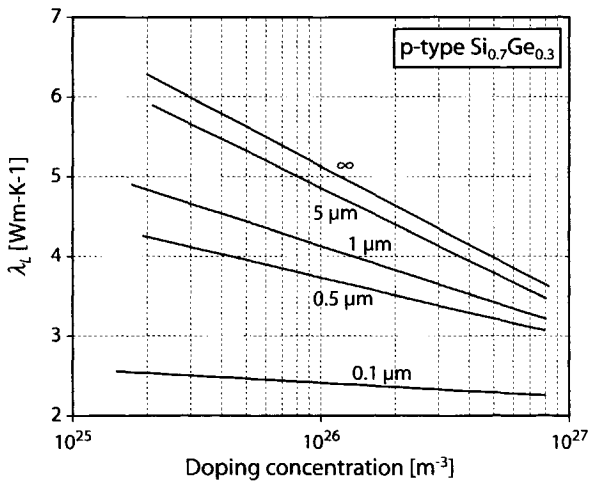


Figure B.4 Lattice thermal conductivity as a function charge carrier concentration, for different grain size values, for p-type $Si_{0.7}Ge_{0.3}$ at 300 K. (Source: [B.24])

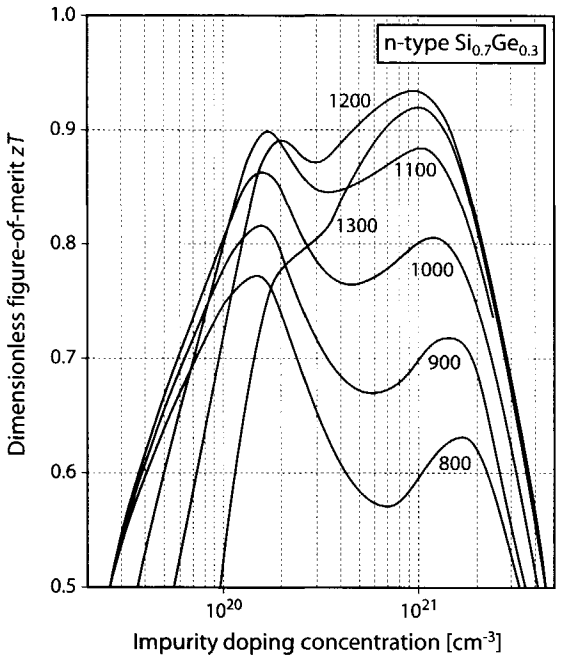


Figure B.5 Figure-of-merit versus doping concentration for n-type $Si_{0.7}Ge_{0.3}$. (After [B.15]).

SiGe and polySiGe carrier concentration

There seems to be a general agreement as to what the optimum doping concentration is for (poly)SiGe. In Chapter 12 of [2.2], the optimum carrier concentration for standard $Si_{0.8}Ge_{0.2}$ is said to be around $(1.4\text{-}1.8)\cdot10^{20}$ cm^{-3}, and $2.2\cdot10^{20}$ cm^{-3} for n-type $Si_{0.8}Ge_{0.2}$. Again in [2.2], but his time in Chapter 28, a wider doping concentration range is mentioned, that is, $(1\text{–}3)\cdot10^{20}$ cm^{-3} for n-type SiGe and $(2\text{–}4)\cdot10^{20}$ for p-type SiGe. The Seebeck effect should not change significantly with the doping concentration n as, as it varies with $\log(n)$.

The optimum doping concentration remains difficult to predict in practise, as Figure B.5 illustrates: This set of data obtained by Slack and Hussain [B.15] reveals that the figure-of-merit varies strongly with anneal temperature. As this anneal temperature is increased, the optimum doping concentration actually shifts from 10^{20} cm^{-3} towards 10^{21} cm^{-3}. This shift is attributed to the occurrence of new phases in the material, but is still under investigation [2.2]. Clearly, this temperature dependency must be kept in mind during (thin film) device fabrication.

In general, GaP is used when an n-type material is desired (where P acts as donor atom), while SiB_4 is used when a p-type material is required (where B acts as acceptor atom.) For example, when adding 0.63 at% GaP (with a Ga:P ratio of 1:3) during

mechanical alloying, Cook *et al.* [Rowe, 12.12] managed to obtain n-type $Si_{0.8}Ge_{0.2}$ that has an average figure-of-merit of 0.93 10^{-3} K^{-1} over the temperature range from 300 °C to 1000 °C.

Typical values for n-type and p-type SiGe are reported in Chapter 28 of [Rowe]. The electrical resistivity varies from 11 μΩm at 300 K to 33 μΩm at 1300 K. Similarly, the typical Seebeck coefficient varies from 120 μVK^{-1} at 300 K to 230 μVK^{-1} at 1300 K, while the thermal conductivity decreases from 4.8 $Wm^{-1}K^{-1}$ at 300 K to 4.1 $Wm^{-1}K^{-1}$ at 900 K at which it rises again to 4.8 $Wm^{-1}K^{-1}$ at 1300 K.

A number of good overviews as well as in-depth reviews of conventionally prepared SiGe alloys are provided in [B.14],[B.15],[B.25]–[B.28],[B.29] and [B.30]. The fabrication and properties of thin-film polySiGe is discussed extensively in Chapter 5.

B.1.2.5 Germanium telluride (GeTe) and TAGS $(AgSbTe_2)_{1-x}(GeTe)_x$

Germanium telluride is prepared by melting together the right amounts of Ge and Te and slowly freezing the melt. Alternatively, GeTe may be fabricated by hot pressing (at the cost of a slight decrease in performance.) Intrinsic GeTe is always p-type and has a carrier concentration of $0.9 \cdot 10^{21}$ cm^{-3}. Figure B.6 shows the figure-of-merit of GeTe versus temperature. The carrier concentration of GeTe is too high to provide a very good figure-of-merit. As such, Bi is dissolved in GeTe to obtain an optimal doping concentration. It seems Bi^{+3} substitutes Ge^{+2}, thereby removing excess holes and changing the concentration of vacancies. The figure-of-merit of 95% GeTe with 5% Bi_2Te_3 is plotted in Figure B.6 and reveals a practical doubling of the figure-of-merit compared to pure GeTe, in the range of 700 – 800 K.

Even better results can be obtained from a TAGS compound. The acronym TAGS stands for a Te–Ag–Ge–Sb alloy, essentially between the compounds $AgSbTe_2$ and GeTe. These alloys are closely related to PbTe as they have the same sodium chloride lattice, when the GeTe concentration is below 80%. At this concentration a transformation to a rhombohedral structure occurs, which is also the point of highest figure-of-merit. It is believed that lattice strains due to this transformation effectively reduce the lattice thermal conductivity [2.2].

Local minima in the lattice thermal conductivity are observed at both 80% and 85% GeTe, at which compositions the figure-of-merit also peaks [2.2]. The composition with 80% GeTe, referred to as TAGS-80, has the superior figure-of-merit. However, the largest mechanical strength / and thermal stability is obtained when 85% GeTe is used (TAGS-85). As TAGS is primarily used in RTG's (radioisotope thermoelectric generators) and other high-temperature applications, materials stability is of primary importance, so TAGS-85 is generally preferred. There still appears to a room for improvement in the figure-of-merit by using a 1.5:1 ratio of Sb to Ag, instead of a 1:1 ratio.

TAGS compounds are normally produced by casting, where—If casting is done properly—no further treatment is required. TAGS compounds can also be produced

by means of hot-pressing. This does not influence the thermoelectric properties, but provides somewhat better mechanical properties.

One problem that remains with TAGS compounds is that they are very difficult to dope, so it is hard to obtain an extrinsic material. Moreover, TAGS-compounds are always p-type, so an n-type leg of another material is required, mostly PbTe.

B.1.2.6 Bismuth antimonide (BiSb)

In the very low temperature range (< 200 K), the figure-of-merit of Bi_2Te_3 and related room-temperature materials rapidly decreases. In this temperature range BiSb is preferred as in the thermal conductivity is significantly lower, i.e. near 80 K, λ_L of $Bi_{0.88}Sb_{0.12}$ is 3.1 $Wm^{-1}K^{-1}$, compared to 11 $Wm^{-1}K^{-1}$ for bismuth. Even though the Seebeck coefficient will drop with temperature, at 80 K values of -180 μVK^{-1} have been reported for $Bi_{0.88}Sb_{0.12}$. Consequently, undoped $Bi_{0.88}Sb_{0.12}$ has a maximum figure-of-merit around $5.2 \cdot 10^{-3}$ K^{-1} at 80 K, which drops off rapidly for higher temperatures ($2 \cdot 10^{-3}$ K^{-1} around 200 K and $1 \cdot 10^{-3}$ K^{-1} around 300 K.)

B.1.2.7 Zinc antimonide (ZnSb)

Although not as well known as other thermoelectric materials, zinc antimonide is used as p-type material in some thermoelectric generators [A.6]. Like PbTe, it should be regarded as an intermetallic compound. The major problem with this material is that it apparently is very difficult to reproducibly obtain high-quality and homogeneous samples.

Sb atoms form a pair, with the distance between the atoms equal to an Sb-Sb covalent bond. As such, the best way to describe the material with respect to it's valances is as $(Zn_2)^{+2}(Sb\text{-}Sb)^{-4}$. Addition of Sb, Sn and Ag yields a p-type material, while In gives an n-type material. Addition of a excess amount of Sb causes the formation of Zn vacancies with twice as many holes. Tin will substitute for antimony and form and $(Sb\text{-}Sn)^{-5}$ anion plus one hole. Ag^{+} will substitute the Zn^{+2} ion, creating a hole. In^{+3} will also substitute the Zn^{+2} ion, giving a free electron. In Figure B.6 the figure-of-merit of pressed/sintered $Zn_{0.346}Sb_{0.643}Sn_{0.01}Ag_{0.001}$ is shown. For this composition, the highest figure-of-merit is obtained near a temperature of 450 K.

B.1.2.8 Manganese telluride (MnTe)

Manganese telluride (MnTe) has been proposed as a high-temperature (800–1300 K) p-type thermoelectric material. Sodium is used to provide the necessary free holes for electrical conduction. It appears difficult to predict the behaviour of MnTe based on quantum theory, in particular as the material data obtained depends strongly on the method of material preparation [A.6]. Unfortunately, as both Na and Li are insufficiently soluble, the optimum doping level of MnTe is not obtained.

B.1.2.9 III-V compounds (B, Al, Ga, In)(P, As, Sb)

Group III-V compounds are formed from the elements in groups IIIA (B, Al, Ga and In) and VA (P, As and Sb) and usually form a zinc-blende lattice. The compounds that received initial attention were InAs, InSb and $InAs_{1-x}P_x$ [A.6]. For the latter

compound, the average figure-of-merit is some 15% higher than of a pure InAs compound. First this is due to the increase of the energy gap by addition of InP to InAs. Second, this is due to the creation of scattering centres that decrease the thermal conductivity. In Figure B.6 the figure-of-merit can be is plotted against temperature, in relation

The energy gap for III-V compounds is rather small, so the carrier concentration soon exceeds the concentration for optimal thermoelectric performance as the temperature increases. Furthermore, the charge carriers in these compounds have small effective masses, so the optimum carrier concentration is in the order of 10^{18} rather tan 10^{19} to 10^{20} for most thermoelectric materials.

B.1.2.10 Silicides

The best known material seen as a silicide is silicon-germanium, discussed in Section B.1.2.4. Still, various other silicides have been considered. In particular, $MnSi_{1.75}$ and $Mg_2(Si,Ge)$ appear capable of a dimensionless figure-of-merit $zT_{max} > 1$ [2.2]. $Mg_2Si_{0.6}Ge_{0.4}$, for example, has a value zT of 1.68 at 300 K. Moreover, Ru_2Si_3 is said to be capable of outperforming SiGe, by a factor of 2 for p-type material and a factor 1.5 for n-type material [B.31], although this still has to be proven.

Besides these high-temperature materials, $FeSi_2$ is considered a low-temperature thermoelectric semiconductor. It appears in both a high-temperature conducting phase (α-$FeSi_2$) and a low-temperature semiconducting phase (β-$FeSi_2$). The latter component is of interest to thermoelectrics and may be prepared, by means of vapour-phase epitaxy, e-beam evaporation or powder metallurgy. The doping concentration can be controlled by doping with Co or Mn for n-type materials and Al for p-type materials, respectively yielding $Fe_{1-x}(Co,Mn)_xSi_2$ and $FeSi_{2-x}Al_x$. In Figure B.6, the figure-of-merit over temperature as provided in [B.32] is included.

The composition of manganese silicide ranges from $MnSi_{1.71}$ to $MnSi_{1.77}$, while most references concern $MnSi_{1.75}$. The complexity of the crystal structure (in particular due to the existence of large superlattices like $Mn_{11}Si_{19}$ and $Mn_{15}Si_{26}$) and strong anisotropy of the thermoelectric material properties make it hard to predict the thermoelectric performance. Empirically, it has been established that hot-pressing yields materials with poor performance. This is due to the unpredictable changes in the lattice structure as well as carrier concentration. Nevertheless, the properties of doped $MnSi_{1.75}$ closely approach those of SiGe.

As a final silicide, semimetallic n-type CoSi has already been used in RTG's. At about 400 K, the maximum figure-of-merit is obtained, which is close to 0.36 K^{-1}. To further improve this value, it is not useful to reduce the lattice thermal conductivity as λ_L is only about 20% of the total thermal conductivity. Instead efforts are made to reduce the overlap of the valence and conduction bands, e.g. by introducing Fe. At the same time, an effort is made to increase the selective scattering of carrier, e.g. by $Mn_4Al_3Si_5$.

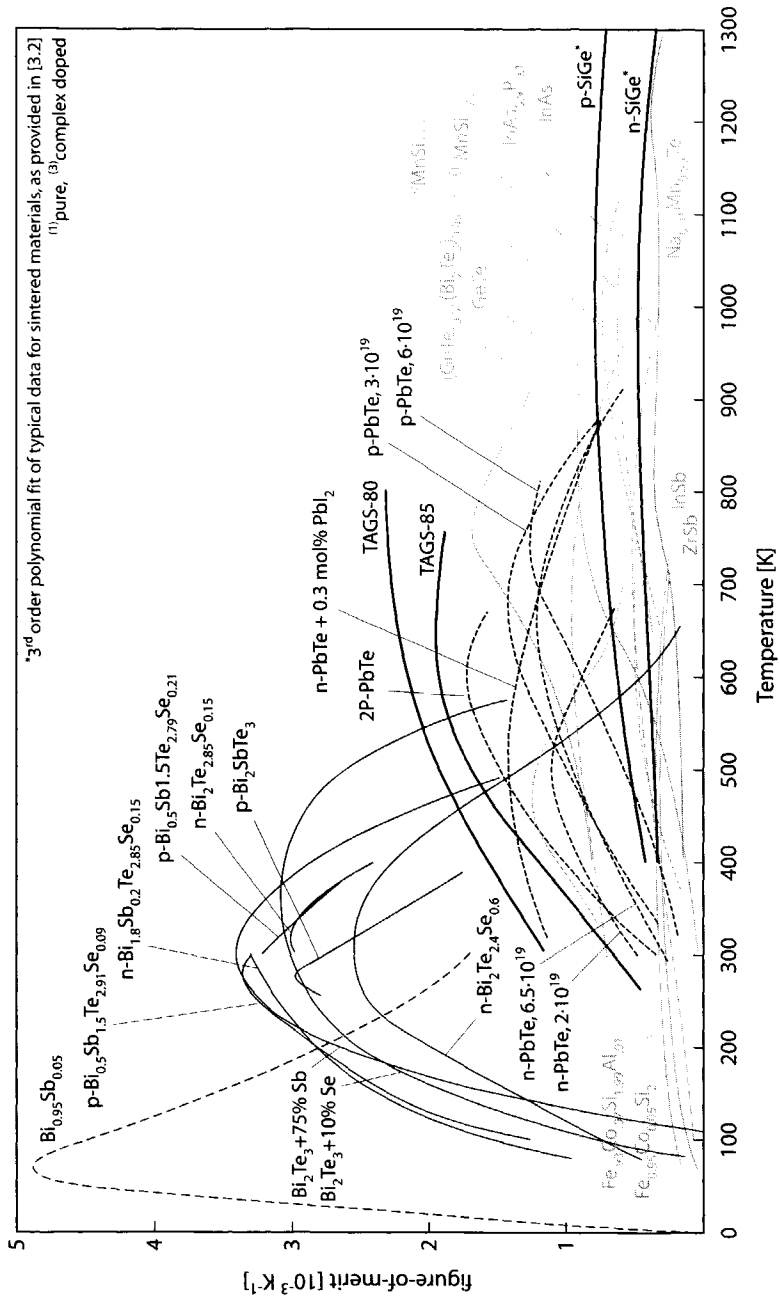


Figure B.6 Qualitative overview of the figure-of-merit for various conventional thermoelectric materials, plotted against temperature.

B.2 APPROACHES TOWARDS NEW MATERIALS

The basic concept of manufacturing a good thermoelectric material is to take a semiconducting compound of heavy elements from the lower right region of the periodic table and alloying these elements to reduce the lattice thermal conductivity. It appears as though these more conventional thermoelectric materials (tellurides, antimonides, silicon-germanium, TAGS, etc.) are now quite close to the limit of how far the thermoelectric material properties of existing compounds can be improved.

Examples supporting this observation were already given in Section B.1.2.1. In 30 years, the figure of merit of bulk n-type Bi_2Te_3 only improved from $2.6 \cdot 10^{-3}$ K^{-1} to $2.9 \cdot 10^{-3}$ K^{-1}. For p-type Bi_2Te_3 the improvement was even less. Alloy scattering was successfully introduced in systems like $(Bi_{1-x}Sb_x)_2(Te_{1-y}Se_y)_3$ and provided some room for improvement (currently still only up to $3.3 \cdot 10^{-3}$ K^{-1} at room temperature.)

According to Goldsmid and Nolas [B.33] there should still be some room for improvement: "*the lattice thermal conductivity in bismuth telluride and its alloys is quite low for a crystalline material but an order of magnitude greater than one would expect for a glass. There is, thus, some prospect for improving the figure of merit by reducing the lattice thermal conductivity.*" Nevertheless, the near-magical "$zT = 1$" barrier remains virtually impossible to break by means of conventional materials. This frustration has lead to an elaborate search for new thermoelectric materials [B.33][B.34]. The most extensive overview to date of new research directions is provided in [3.32], where four chapter are devoted to the subject of new materials.

One of the approaches has recently smashed the "$zT = 1$" barrier by means of low-dimensional materials. Low-dimensional materials are outlined in Section B.2.2. Before that the phonon-glass electron-crystal concept is introduced.

B.2.1 Phonon-Glass Electron-Crystal (PGEC)

One of the approaches to find new materials with an improved figure-of-merit (or with plenty of room of improving it) is the phonon-glass electron-crystal, or PGEC for short [2.2]. As the name directly suggests, the material should behave like a glass for phonons, i.e., the lattice thermal conductivity should be minimised. (It was shown by Kittel that glasses possess phonon mean free paths of near-atomic dimensions, which is necessary if phonons are to be scattered effectively [B.35].) At the same time, the electrons should experience the material as a regular crystalline material, so that the electrical conductivity and Seebeck coefficient remain unaffected.

The likelyhood of a material that possesses PGEC-like properties is largest for crystals that have open positions in the lattice, so-called cages. In these cages, an atom can be loosely trapped so that it 'rattles' about its position. This rattling is thought to be an effective scatterer of phonons.

A low lattice thermal conductivity cannot just be contributed to the rattling of atoms. Also a complex crystal structure, with a very large number of atoms in a unit cell causes the lattice thermal conductivity to decrease. Slack argued [B.36] and re-

cently could see confirmed [B.37] that $\lambda_l \propto n^{-2/3}$, where n is the number of atoms in a unit cell. For example, the Ge_{46} crystal has a thermal conductivity of about one tenth of Ge with a diamond lattice. The thermal conductivity value of $Sr_8Ga_{16}Ge_{30}$—where $Ga_{16}Ge_{30}$ has the same structure as Ge_{46}, but has 8 Sr atoms 'trapped'— has a thermal conductivity over 100 times less than crystalline Ge. This clearly illustrates the effect of increasing the acoustic phonon scattering by means of 'rattling' atoms.

In practise, four groups of materials are considered with respect to PGEC-like properties. These are skutterudites [B.38], clathrates [B.39], half-Heusler compounds [B.33][B.40] and chalcogenides.

Skutterudites

The name is derived from the mineral $CoAs_3$, first discovered in Skutterud, Norway. Typical skutterudites are $CoSb_3$, $IrSb_3$. These crystals have a large unit cell with voids that can hold interstitial atoms. For example, without interstitial atoms, λ_l of $CoSb_3$ is around 10 $Wm^{-1}K^{-1}$. When all voids are filled, for example in $LaFe_3CoSb_{12}$ and $CeFe_3CoSb_{12}$, the thermal conductivity is nearly halved. Most remarkably, when the voids are only partially filled, as with $La_{0.9}Co_4Sn_3Sb_9$, the lattice thermal conductivity may be reduced even more, by as much as an order of magnitude. Combined with a good power factor, zT values larger than unity have already been obtained at temperatures around 600–700 K [B.41].

Clathrates

For over a century, water that is turned into ice is known to form clathrate compounds with numerous small molecules and compounds. Like with skutterudites, voids are formed in the crystal that can hold these smaller atoms and molecules.

Clathrates can be divided into two categories, based on the size of the unit cell. Type I clathrates have a cubic cell of 46 atoms from group IV, with 8 voids. Type II clathrates also have cubic lattice but with 136 atoms with 24 voids. As was already stated $Sr_8Ga_{16}Ge_{30}$ is a very effect type I clathrate that has a lattice thermal conductivity in the order of 1 $Wm^{-1}K^{-1}$ [B.41]. At room temperature Ge-based clathrates have a reported zT value of 0.34, which is still less than for bismuth-telluride based materials. However, as the band gap of Ge clathrates is larger, zT increases with temperature. Recently it was discovered that zT of $Sr_8Ga_{16}Ge_{30}$ can be increased threefold, when subjected to a pressure of 7 GPa [B.42]. This improvement is attributed to a significant rise in the power factor. Based on these findings, it has been concluded that clathrates have the best potential to become the next-generation room-temperature thermoelectric materials.

Half-Heusler alloys

Half-Heusler alloys share the same kind of large unit cells found in skutterudites and clathrates. Unfortunately, even though Half-Heusler alloys appear to have a power factor comparable to that of bismuth telluride, the lattice thermal conductivity is still

significant. ZrNiSn, for example, has a λ_L of about 10 $Wm^{-1}K^{-1}$. Alloying, for example by forming $Zr_{0.5}Hf_{0.5}NiSn$, can reduce the lattice thermal conductivity [B.43].

Most of the lattice thermal conductivity appears to be due to acoustic phonons that seem to have a mean free path greater than the electron mean free path. Therefore, boundary scattering is expected to be very effective in reducing the lattice thermal conductivity.

Chalcogenides

The skutterudites, clathrates and half-Heusler compounds are based on extremely large unit cells, and in no way resemble any of the crystal structures of conventional thermoelectric materials. This is different for chalcogenides. These are based on Bi_2Te_3 and similar compounds.

As stated in Section B.1.2.1, the two neighbouring Te atoms in the Bi_2Te_3 lattice are loosely bound, which creates room for interstitial layers in between the Te layers. Examples are $CsBi_4Te_6$, $HfTe_3$, $ZrTe_3$ and Tl_2SnTe_5. The driving force behind the performance improvement is the same as with the PGEC concept, where phonons scattering is promoted over electron scattering. Performance is not yet up to par with conventional bismuth telluride, but $CsBi_4Te_6$ has been reported with a zT value of 0.8 at 300 K, while Tl_2SnTe_5 has a zT value of 0.85 at 400 K and $Bi_{2-x}Sb_xTe_6$ has a zT value of 0.9 at 325 K.

Discussion

In most review articles on the PGEC approach, two issues are encountered. The first concerns which approach to follow in optimising the existing materials as well as searching for new materials. Most of the studies deal with reducing the thermal conductivity without significantly affecting the electrical power factor, $\alpha^2\sigma$. It can be rightfully argued that this approach will lead to considerable problems, as most of the thermoelectric semiconductors are heavily doped, so a significant contribution to thermal conductivity arises from the charge carriers rather than the lattice. As the charge carrier contribution to the total thermal conductivity is linked to the electrical conductivity by means of the Wiedemann-Franz-Lorentz law, it becomes very hard to reduce λ without reducing the power factor as well.

Second, even though the PGEC-based materials offer plenty of room for improvement in thermoelectrics, it will be the economic issues that will prove most difficult to tackle in the future. Most of the PGEC-based materials use exotic elements like Cs, Nf and Tl. Many of these elements are very toxic, risking a ban from health agencies. Of a more practical nature, many of these exotic elements come in a limited supply, raising the question whether production of potent PGEC-based materials can be scaled up at all (i.e., if sufficient supplies can be acquired) and whether upscaling yields an economic product (i.e., if supplies can be acquired at a reasonable cost.)

B.2.2 Low-dimensional materials

Research related to improved PGEC materials is still waiting for the big breakthrough. Currently, the largest advances are made in low-dimensional thermoelectrics research. The concept is simple. Reducing at least one of the geometrical dimensions to about the lattice constant, the band structure is altered. This change can be used in two ways, either to improve the power factor or to reduce the lattice thermal conductivity. This depends on the type of operation, as will be explained.

The potential improvement of the thermoelectric properties by reducing the dimensions of a thermoelectric material was first reported in 1993, by Hicks and Dresselhaus [B.44]. Most of these structures as based on superlattice crystals, that can be used to confine the electron movement to two dimensions. Nanowires and -whiskers can be used to further confine the flow of electrons to one dimension.

In essence such a crystal consists of several tens or even hundreds of repetitively alternated monolayers of different materials/compounds. The basic construction of a superlattice is illustrated in Figure B.7. Two modes of operation can be distinguished, lateral and transversal, in similarity with the definition of directions used in this thesis. For clarity, these two directions are indicated in Figure B.7.

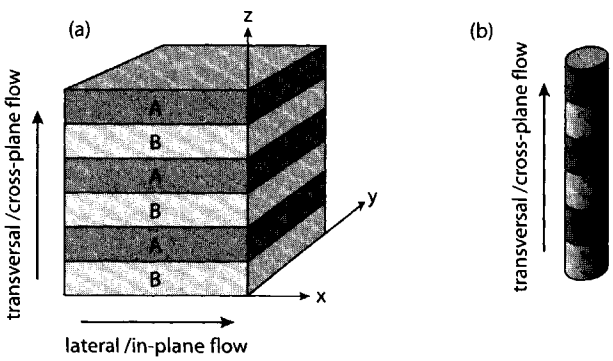


Figure B.7 Schematic representation of (a) a superlattice structure and (b) a nanowire.

Different micro-/nanofabrication technologies are available for the fabrication of superlattices and nanowires. Molecular beam epitaxy (MBE) has been used to create Si/Ge and PBTe/$Pb_{1-x}Eu_xTe$ superlattices. Chemical vapour deposition (CVD) and pulsed laser ablation (PLA) have been used to create Bi_2Te_3/Sb_2Te_3 superlattices. To create nanowires, two methods are encountered. The first uses lithography to etch pores into a substrate, mostly anodic alumina. These holes are then filled with the required materials, either by pressure injection from a liquid source, vacuum evaporation or electrochemical deposition. By tuning the etch, a surface fill factor of around 50% can be obtained. More recently, a hybrid pulsed laser ablation/chemical vapour deposition (PLA-CVD) technique were applied to create $Si/Si_{0.82}Ge_{0.12}$ and

GaAs/GaP nanowires [B.45][B.46]. The main advantage of the latter technique is that freestanding nanowires can be created rather than nanowires that are embedded in a substrate with low thermal conductivity.

It is my believe that in the years to come only 2-D quantum structures will remain interesting for industry. Application of nanowires for thermoelectrics will prove too difficult, due to the fragility of the wires as well as the difficulties in contacting these wires. For this reason, only 2-D quantum devices are outlined below.

Quantum well superlattice structures

The basic philosophy behind the quantum well principle (i.e., the lateral/in-plane mode) is to contain the electrons in a conducting layer, by sandwiching this layer between to insulating barriers. Quantum confinement of electrons to 2 dimensions as in quantum wells or even to 1 dimension as in nanowires can effectively increase the electrical conductivity of a material. Nolas [3.32] expresses the ratio of 2-D to 3-D electrical conductivity as

$$\frac{\sigma_{2D}}{\sigma_{3D}} \cong \frac{a \cdot h}{2} \sqrt{2 m_z k_B T} \tag{B.1}$$

where m_z is the effective mass in the z-direction and a is the height of the quantum well, also in the z-direction. It can be shown a typically has to be in the order of 10 nm or less for $\sigma_{2D}/\sigma_{3D} > 1$, so that the 2-D electrical conductivity improves. Similarly, the diameter of nanowires has to be in the order of 5 nm or less.

Abstreiter *et al.* [B.47] identify the formation of a 2-D electron gas within the Si layer of a $Si/Si_{0.5}Ge_{0.5}$ superlattice. Confinement is obtained due to a 0.15 eV conduction band offset caused by strain effects. Harman *et al.* [B.48] achieved similar results with $PbTe/Pb_{1-x}Eu_xTe$ superlattices. The 2–5 nm PbTe wells form the dominant in-plane conductive paths, as the 40-60 nm $Pb_{0.93}Eu_{0.07}Te$ barrier layers have an energy about 0.2 eV higher than the wells. When the well height approached 2 nm, an increase of a factor 2 is observed in the Seebeck coefficient. When the well height was increased to 4 nm, an increase was no longer observed. This clearly indicates how critical the well height is.

Non-quantum well superlattice structures

In the above lateral mode, electrons are transported along the wells. This approach is primarily focused on improvement of the electrical parameters. Instead, in the transversal (cross-plane) mode, the electrons will flow across all layers, so each layer should have good electrical properties. In this mode of operation, a reduction in the lattice thermal conductivity is pursued.

Both Bi_2Te_3/Sb_2Te_3 [B.49] and $Si/Si_{1-x}Ge_x$ [B.50] superlattices have been studied. With both compositions, the expected reduction in the lattice thermal conductivity was observed. For the Bi_2Te_3/Sb_2Te_3 superlattice, thermal conductivity values in the

range of 0.11 $Wm^{-1}K^{-1}$ to 0.2 $Wm^{-1}K^{-1}$ were obtained. For the Si/Ge superlattices, values in the range of 3.4 $Wm^{-1}K^{-1}$ to 4.5 $Wm^{-1}K^{-1}$ were observed.

Two groups in particular have gone beyond mere characterisation of superlattice structures and performed cooling experiments with their supperlattice structures. The University of California [B.51]–[B.54] has studied InGaAs/InGaAsP, $Si/Si_{0.75}Ge_{0.25}$ and $Si/Si_{0.89}Ge_{0.10}C_{0.01}$ superlattices, for which they respectively observed a ΔT of 0.65 K, 2.5 K and 4.2 K (at room temperature). These rather small temperature differences are mainly caused by the bond wire placed thermally in parallel with the superlattice structure. As was already emphasized in Chapter 5, metal wiring is certain to ruin cooling performance unless used cautiously.

At the Research Triangle Institute Bi_2Te_3/Sb_2Te_3 superlattice coolers are studied. This group holds the current 'world record' regarding the dimensionless figure-of-merit: a zT value of 2.38 (!) was obtained at 300 K for their p-type material. Even for n-type Bi_2Te_3/Sb_2Te_3, a zT of 1.43 was obtained at 300 K. In cooling mode, their structures obtained a ΔT of 32.2 K, when the substrate was kept at 298 K [B.55]. The only remark about their results is the lack of clarity on how these values were obtained. Nevertheless, these result remain impressive.

B.3 REFERENCES

[B.1] M. Hansen, *Der Aufbau er Zweistofflegierungen*, Springer Verlag, Berlin, Germany, 1936.

[B.2] A. Borshchevsky, J.-P. Fleurial, J.W. Vandersande and C. Wood, "Preliminary results on zone-leveling of multidoped SiGe thermoelectric alloys," *Proc. 7th Symposium on Space Nuclear Power Systems*, 7–11 January, 1990, Albuquerque, NM, USA, p. 220.

[B.3] J.P. Dismukes and L. Ekstrom, "Homogeneous solidification of Ge-Si alloys," *Trans. Met. Soc. AIME*, Vol. 233, pp. 672, 1965.

[B.4] A. Barz, U. Kerat, P. Dold and K.W. Benz, "Growth of $Ge_{1-x}Si_x$ (x = 0 to 0.5) single crystals by a zone melting method," *Cryst. Res. Technol.*, Vol. 33, No. 6, pp. 849–855.

[B.5] B.R. Pamplin, Ed., *Crystal growth*, second edition, Pergamon Press, Oxford, UK, 1980.

[B.6] A. Borshchevsky and J.-P. Fleurial, "Growth of heavily doped Si-Ge from metallic solutions," *J. Cryst. Growth*, Vol. 128, Iss. 1–4, pp. 331–337, 1993.

[B.7] J.-P. Fleurial, A. Borschevsky and J. Vandersande, "Optimization of the thermoelectric properties of hot-pressed n-type SiGe materials by multiple doping and microstructure control," *Proc. 10th International Conference on Thermoelectrics*, 10–12 September 1991, Cardiff, UK, p. 156.

[B.8] D.M. Rowe and C.M. Bhandari, "???", *Appl. Energy*, Vol. 6, pp. 347–351, 1980.

[B.9] D.P. White and P.G. Klemens, "Thermal conductivity of thermoelectric $Si_{0.8}Ge_{0.2}$ alloys," *J. Appl. Phys.*, Vol. 71, No. 9, pp. 4258–4263, 1992.

[B.10] H.J. Goldsmid, A.R. Sheard and D.A. Wright, "The performance of bismuth telluride thermojunctions," *Br. J. Appl. Phys.*, Vol. 9, No. 9, pp. 365–370, 1958.

[B.11] J.-P. Fleurial, L. Gaillard, R. Triboulet, H. Scherrer and S. Scherrer, "Thermal properties of high quality single crystals of bismuth telluride—part I: Experimental characterization," *J. Phys. Chem. Solids*, Vol. 49, Iss. 10, pp. 1237–1247, 1988.

[B.12] F.D. Rosi, B. Abeles and R.V. Jensen, "Materials for thermoelectric refrigeration," *J. Phys. Chem. Solids*, Vol. 10, Iss. 2–3, pp. 191–200, 1959.

[B.13] D.M. Rowe and C.M. Bhandari, *Modern thermoelectrics*, Holt, Rinehart and Winston, London, UK, 1983.

[B.14] J.P. Dismukes, L. Ekstrom, E.F. Steigmeier, L. Kudman and D.S. Beers, "Thermal and electrical properties of heavily doped Ge-Si alloys up to 1300 K," *J. Appl. Phys.*, Vol. 35, p. 2899, 1964.

[B.15] G.A. Slack and M.A. Hussain, "The maximum possible conversion efficiency of silicon-germanium thermoelectric generators," *J. Appl. Phys.*, Vol. 70, Iss. 5, pp. 2694–2718, 1991.

[B.16] B.A. Cook, B.J. Beaudry, J.L. Harringa and W.J. Barnett, "Mechanical alloying as an alternative method of producing n-type $Si_{80}Ge_{20}$ thermoelectric materials," *Proc. 9th International Conference on Thermoelectrics*, 19–21 March 1990, Pasadena, CA, USA, p. 234.

[B.17] B.A. Cook, B.J. Beaudry, J.L. Harringa and W.J. Barnett, "Oxygen effects in mechanical alloyed $Si_{80}Ge_{20}$ doped with GaP and P," *Modern Perspectives on Thermoelectrics and Related Materials*, MRS Symp. Proc. 234, D.D. Allred, C.B. Vining and G.A. Slack, eds., Material Research Society, Pittsburgh, 1991, p. 111.

[B.18] S.H. Han, B.A. Cook and K.A. Gschneidner, Jr., "Effect of the doping process on the electrical activity of phosphorous in $Si_{80}Ge_{20}$," *Proc. 11th International Conference on Thermoelectrics*, 7–9 October, 1992, Arlington, Texas, USA, p. 52.

[B.19] H.J. Goldsmid, *Thermoelectric Refrigeration*, Plenum Press, New York, NY, USA, 1964.

[B.20] W.M. Yim and F.D. Rosi, "Compound tellurides and their alloys for Peltier cooling—a review," *Solid-State Electronics*, Vol. 15, Iss. 10, pp. 1121–1134, 1972.

[B.21] R.M. Davis and C.C. Koch, "Mechanical alloying of brittle components: silicon and germanium," *Scr. Metall.*, Vol. 21, Iss. 3, pp. 305–310, 1987.

[B.22] N. Savvides and H.J. Goldsmid, "Boundary scattering of phonons in fine-grained hot-pressed Ge-Si alloys. I. The dependence of lattice thermal conductivity on grain size and porosity," *J.Phys. C*, Vol. 13, No. 25, pp. 4657–4670, 1980.

[B.23] N. Savvides and H.J. Goldsmid, "Boundary scattering of phonons in fine-grained hot-pressed Ge-Si alloys. II. Theory," *J.Phys. C*, Vol. 13, No. 25, pp. 4671–4678, 1980.

[B.24] C.M. Bhandari and D.M. Rowe, "Boundary scattering of phonons," *Journal of Physics C: Solid State Physics*, Vol. 11, 1978, pp. 1787–1794.

[B.25] F.D. Rosi, "Thermoelectricity and thermoelectric power generation," *Solid-State Electronics*, Vol. 11, Iss. 9, pp. 833–848, 1968.

[B.26] C.M. Bhandari and D.M. Rowe, "Silicon-germanium alloys as high-temperature thermoelectric materials," *Contemporary Physics*, Vol. 21, No. 3, pp. 219–242, 1980.

[B.27] D.M. Rowe, Recent advances in silicon-germanium alloy technology and a assessment of the problems of building the modules for a radioisotope thermoelectric generator," *J. Power Sources*, Vol. 19, Iss. 4, pp. 247–259, 1987.

[B.28] C.B. Vining and J.-P. Fleurial, "Silicon germanium—an overview of recent developments," *Proc. 10th International Conference on Thermoelectrics*, 10–12 September 1991, Cardiff, UK, pp. 1.

[B.29] C.B. Vining, "A model for the high-temperature transport properties of heavily doped n-type silicon-germanium alloys," *J. Appl. Phys.*, Vol. 69, No. 1, pp. 331–341, 1991.

[B.30] C.B. Vining, "A model for the high-temperature transport properties of heavily doped p-type silicon-germanium alloys," *Modern Perspectives on Thermoe,lectrics and Related Materials*, MRS Symp. Proc. 234, D.D. Allred, C.B. Vining and G.A. Slack, eds., Material Research Society, Pittsburgh, 1991, p. 95.

[B.31] C.B. Vining, "Extrapolated thermoelectric figure-of-merit of ruthenium silicide," *Proc. 9th Symposium on Space Nuclear Power systems*, AIP Conf. Proc. 246, M.S. El-Genk and M.D. Hoover, eds., AIP, New York, NY, USA, 1992, p. 338.

[B.32] U. Stöhrer, "High temperature measurements on thermoelectric materials," *Proc. 11th International Conference on Thermoelectrics, Proc. 11th International Conference on Thermoelectrics*, 7–9 October, 1992, Arlington, Texas, USA, p. 191.

[B.33] H.J. Goldsmid and G.S. Nolas, "A review of the new thermoelectric materials," *Proc. 20th International Conference on Thermoelectrics*, 8–11 June 2001, Beijing, China, pp. 1–6.

[B.34] G.S. Nolas, G.A. Slack, J.L. Cohn and S.B. Schujman, "The next generation of thermoelectric materials," *Proc. 17th International Conference on Thermoelectrics*, 29 August–2 September1998, Nagoya, Japan, pp. 294–297.

[B.35] C. Kittel, "Interpretation of the thermal conductivity of glasses," *Phys. Rev.*, Vol. 75, Iss. 6, pp. 972–974, 1949.

[B.36] G.A. Slack, "The thermal conductivity of nonmetallic crystals," in *Solid State Physics*, H. Ehrenreich *et al.*, Eds., Academic Press, New York, NY, USA, 1979.

[B.37] J. Dong, O.F. Sankey and C.W. Myles, "Theoretical study of the lattice thermal conductivity in Ge framework semiconductors," *Phys. rev. lett.*, Vol. 86, Iss. 11, pp. 2361–2364, 2001.

[B.38] G.S. Nolas, J.L. Cohn and G.A. Slack, "Effect of partial void filling on the lattice thermal conductivity of skutterudites," *Phys. Rev. B*, Vol. 58, Iss. 1, pp.164–170, 1998.

[B.39] G.S. Nolas, T.J.R. Weakley and J.L Cohn, "Structural, chemical and transport properties of a new clathrate compound: $Cd_8Zn_4Sn_{42}$," *Chem. Mater*, Vol. 11, No. 9, pp. 2470–2473, 1999.

[B.40] Y. Xia, S. Bhattacharya, V. Ponnambalem, A.L. Pope, S.J. Poon and T.M. Tritt, "Thermoelectric properties of semimetallic (Zr, Hf)CoSb half-Heusler phases," *J. Appl. Phys.*, Vol. 88, Iss. 4, pp. 1952–1955, 2000.

[B.41] T.M. Tritt, ed., *Semiconductors and semimetals, Volume 69: Recent trends in thermoelectric materials research, Part one*, Academic Press, San Diego, CA, USA, 2001, ISBN 012752178X.

[B.42] J.F. Meng, N.V. Chandra Shekar, J.V. Badding and G.S. Nolas, "Threefold enhancement of the thermoelectric figure of merit for pressure tuned $Sr_8Ga_{16}Ge_{30}$," *J. Appl. Phys.* Vol. 89, Iss. 3, pp. 1730–1733, 2001.

[B.43] C. Uher, J. Yang, S. Hu, D.T. Morelli and G.P. Meisner, "Transport properties of pure and doped *M*NiSn (*M* = Zr, Hf)," *Phys. Rev. B*, Vol. 59, Iss. 13, pp. 8615–8621, 1999.

[B.44] L.D. Hicks and M.S. Dresselhaus, "Effect of quantum-well structures on the thermoelectric figure of merit," *Phys. Rev. B.*, Vo.l 47, Iss. 19, pp. 12727–12731, 1993.

[B.45] Y. Wu, R. Fan, and P. Yang, "Block-by-block growth of single-crystalline Si/SiGe superlattice," *Nano Lett.*, Vol. NL0156888, 4pp, 2001.

[B.46] M.S. Gudiksen, L.J. Lauhon, J. Wang, D.C. Smith and C.M. Lieber, "Growth of nanowire superlattice structures for nanoscale photonics and electronics," Nature, Vol. 415, pp. 617-620, 2002.

[B.47] G. Abstreiter, H. Brugger, T. Wolf, H. Jorke and H.J. Herzog, "Strain-induced two-dimensional electron gas in selectively doped Si/Si_xGe_{1-x} superlattices" *Phys. Rev. Lett.*, Vol. 54, Iss. 22, p. 2441-2444, 1985.

[B.48] T.C. Harman, D.L. Spears and M.J. Manfra, "High thermoelectric figures of merit in PbTe quantum wells," *J. Electron. Mat.*, Vol. 25, No. 7, pp. 1121–1127, 1996.

[B.49] R. Venkatasubramanian T. Colpitts, B. O'Quinn, S. Liu, M. El-Masry and M. Lamvik, "Low-temperature organometallic epitaxy and its application to superlattice structures in thermoelectrics," *Appl. Phys. Lett.*, Vol. 75, Iss. 8, pp. 1104-1106, 1999.

[B.50] S.-M. Lee, D.G. Cahill and R. Venkatsubramanian, "Thermal conductivity of Si–Ge superlattices ," *Appl. Phys. Lett.*, Vol. 70, Iss. 22, p. 2957-2959, 1997.

[B.51] A. Shakouri, P. Abraham, C. LaBounty and J.E. Bowers, "III-V based heterostructure integrated thermionic coolers," *Proc. 17th international conference on thermoelectrics, Proc. 17th International Conference on Thermoelectrics,* 29 August–2 September1998, Nagoya, Japan, pp. 218–220.

[B.52] A. Shakouri, C. LaBounty, J. Piprek, P. Abraham and J.E. Bowers, "Thermionic emission cooling in single barrier heterostructures," *Appl. Phys. Lett.*, Vol. 74, No.1, pp. 89–89, 1999.

[B.53] X. Fan, G. Zeng, C. LaBounty, D. Vashaee, J. Christofferson, A. Shakouri and J.E. Bowers, "Integrated cooling for Si-based microelectronics," *Proc. 20th International Conference on Thermoelectrics*, DATE 2001, PLACE, pp. 405–408.

[B.54] X. Fan, G. Zeng, C. LaBounty, J.E. Bowers, E. Croke, C.C. Ahn, S. Huxtable, A. Majumdar and A. Shakouri, "SiGeC/Si superlattice microcoolers," *Appl. Phys. Lett.*, Vol. 78, No. 11, pp. 1580–1582, 2001.

[B.55] R. Venkatasubramanian, E. Siivola, T. Colpitts and B. O'Quinn, "Thin-film thermoelectric devices with high room-temperature figures of merit," *Nature*, Vol. 413, pp. 597-602, 2001.

SUMMARY

The Peltier effect is the driving force behind the coolers described in this thesis. Due to this effect, at the junction of two materials through which an electrical current is running, heat is absorbed or released in that junction, depending on the direction of the electrical current. The Peltier effect is unique in more ways than one. First and foremost, it is the only solid-state cooling effect available. No external (off-chip) liquid reservoirs or moving components (prone to mechanical failure) are needed. Moreover, like the other two thermoelectric effects, the Peltier effect has the unique characteristic that it is reversible.

The conventional discrete Peltier device is primarily used for cooling and thermal stabilisation of objects (e.g. soda cans) and systems (e.g. humidity detectors) and—less frequently reported but no less important—for rapid thermal control in medical applications like DNA sequencing. Miniaturisation of Peltier elements to chip-scale sizes is particularly useful for the latter two applications, thermal stabilisation and rapid thermal cycling, as this allows the operations to become less space-consuming, more power-efficient, more accurate and faster.

Chapter 1 treats the more general aspects of the miniaturisation and application of a thermoelectric cooler (also referred to as Peltier device or Peltier cooler.) Miniaturisation to chip-scale sizes leads to a a huge decrease in the volume:area ratio of a device. This leads to a large change in design, fabrication and behaviour of a thin-film Peltier device. For example, the Peltier cooling power decreases by a factor of 1000 when going from a discrete element to a thin-film element.

Chapter 2 provides extensive models to describe the performance and behaviour of integrated Peltier coolers. These models include the influence of parasitic effects, that are not (or less dominantly) present in the conventional elements, i.e., losses by parasitic thermal conduction, power dissipation by electrical contact resistance, limited heat sinking capabilities and increased radiant heat exchange. In total, these parasitics can cause the performance of the TEC to decrease by as much as 30%. As the

substrate is the only heat sink for the integrated TEC, its resistance is investigated against the etch technique used as well as the against the thickness of dielectric layers in between the hot junction and ambient. Finally it is proven that if these parasitics are considered early in the design, the performance of the integrated Peltier element is easily optimised, by carefully choosing the width ratio between the n-type and p-type arms of the Peltier element.

In *Chapter 3* the thermoelectric performance of polycrystalline silicon germanium (polySiGe) is evaluated. The performance is based on three material parameters, the Seebeck coefficient, the electrical resistivity and the thermal conductivity. These three parameters together determine the theoretical cooling limit of the Peltier device. Additionally, the (parasitic) electrical contact resistance is measured in order to predict the magnitude of this effect on device performance.

Chapter 4 treats the device fabrication. First, some more general considerations are introduced like the fabrication techniques employed (thin film deposition, micromachining, etc.) and the motivation for co-integration of TEC and electronics in the same chip. The Peltier device itself can be fabricated using a seven-mask process, For co-integration two additional masks from the electronics process are required. Furthermore, the performance of the present devices can be further improved by the addition of two more masks, simultaneously increasing performance and mechanical strength. As a final part of this chapter, a generic process is described for the deposition of nearly any thin-film metal on-chip, without sacrificing fabrication compatibility. Such a process is essential for, for example, temperature-independent metal films, e.g. for accurate voltage-controlled Joule heating in micro-thermostats or for reference resistors.

Chapter 5 discusses the applications in which the on-chip integrated lateral Peltier device is of interest. Three main applications are discussed. The first and most obvious is cooling (near and below ambient temperature). The second is thermal stabilisation and the third is heat spreading. Finally, the potential for polySiGe to replace polySi in IR detectors is addressed.

SAMENVATTING

Het Peltier effect is de drijvende kracht achter de koelers beschreven in dit proefschift. Door dit effect wordt er op het grensvlak van twee materialen waardoor een elektrische stroom loopt warmte geabsorbeerd of afgestaan, al naar gelang de stroomrichting. Het Peltier effect is uniek in meer dan een manier. Bovenal is dit het enige beschikbare vaste-stof koelingseffect. Er zijn geen externe vloeistofreservoirs of bewegende onderdelen (gevoelig voor mechanische slijtage) nodig. Bovendien heeft het Peltier effect, net als de andere twee thermoelektrische effecten, de unieke eigenschap dat het effect omkeerbaar is.

Het conventionele diskrete Peltier element wordt vooral gebruikt voor de koeling en thermische stabilisatie van objecten (zoals blikjes frisdrank) en systemen (zoals vochtigheidssensoren.), en—hoewel minder vaak beschreven maar daarom niet minder belangrijk—voor snelle thermische aansturing in medische toepassingen zoals in DNA onderzoek. Miniaturisatie van Peltier elementen tot chip-niveau is vooral van nut voor de laatste twee toepassingen, d.w.z., voor thermische stabilisatie en snelle thermische cycli, aangezien integratie het mogelijk maakt om de bewerkingen compacter, efficiënter, nauwkeuriger en sneller te maken.

Hoofdstuk 1 behandelt de meer generieke aspecten van de minaturisatie en toepassing van een thermoelektrische koeler (ook wel Peltier element genoemd.) Miniaturisatie tot op chip-niveau leidt tot een enorme vermindering van de volume:oppervlakte verhouding. Dit leidt tot een grote verandering in ontwerp, fabricage en gedrag van een dunne-film Peltier element. Zo verminderd het koelvermogen een factor 1000 wanneer men van een discreet element naar een dunn-film element gaat.

Hoofdstuk 2 bevat uitgebreide modellen om de prestatie en het gedrag van geïntegreerde Peltier koelers te beschrijven. Deze modellen bevatten de invloed van parasitaire effecten die niet (of in mindere mate) aanwezig zijn in de conventionele elementen, d.w.z., verliezen door parasitaire thermische geleiding, vermogensdissi-

patie door elektrische contactweerstand, beperkte warmte-afvoer via het substraat en toegenomen stralingsuitwisseling. In totaal veroorzaken deze parasitaire effecten een vermindering van de prestaties met wel 30%. Aangezien het substraat de enige warmte-afvoer is voor de geïntegreerde thermoelektrische koeler (TEK), is de thermiche weerstand hiervan onderzocht, in relatie tot de ets techniek alswel de dikte van de diëlektrische laag tussen de warme las en de omgeving. Tenslotte wordt het bewezen dat, indien de parasitaire effecten vroeg in het ontwerp worden meegenomen, men de presaties van de TEK eenvoudig kan optimaliseren, door het zorgvuldig kiezen van de juiste breedte-verhouding tussen de n-type en p-type armen van het element.

In *Hoofdstuk 3* worden de thermoelectrische prestaties van polykristallijn silicium germanium (polySiGe) geëvalueerd. De prestatie is gebaseerd op drie parameters, de Seebeck coëfficiënt, de elektrische weerstand en de thermische geleiding. Deze drie parameters tezamen bepalen de theoretische koel-limiet van het Peltier element. Aanvullend is de (parasitaire) elektrische contactweerstand gemeten, om zodoende een voorspelling te kunnen doen van de grootte van dit effect op de prestaties.

Hoofdstuk 4 behandelt de fabricage van de elementen. Allereerst volgen wat meer algemene beschouwingen zoals welke fabricage-technieken worden gebruikt (voor dunne-film depositie, microbewerkingen, etc.) en de motivatie voor co-integratie van de TEK met elektronica in dezelfde chip. Het Peltier element zelf kan worden gefabriceerd middels een proces met zeven maskers. Voor co-integratie zijn twee aanvullende maskers van het electronisch proces nodig. Verder kunnen de prestaties van het element worden verbeterd too twee extra maskers, die gelijktijdig zorgen voor minder verliezen en meer mechanische sterkte. Als een laatste onderdeel van dit hoofdstuk wordt een generiek proces beschreven voor de depositie van welhaast elk dunne-film metaal op een chip, zonder fabricage-compatibiliteit op te moeten offeren. Zo een proces is essentieel voor, onder andere, temperatuur-onafhankelijke stookweerstanden in micro-thermostaten or voor referentie-weerstanden.

Hoodstuk 5 bespreekt de toepassingen waarbij een geïntegreerd lateraal Peltier element van belang kan zijn. Drie hoofd-toepassingen worden besproken. De eerste en meest voor de hand liggende is koeling (rondom en beneden omgevingstemperatuur). De tweede is thermische stabilisatie en de derde is warmte spreiding. Tot slot wordt de potentie van polySiGe ter vervanging van polySi in infrarood detectoren aangehaald.

LIST OF PUBLICATIONS

Journal publications

1. D.D.L. Wijngaards, M. Bartek and R.F. Wolffenbuttel, "IC process compatible metal film postprocessing module," *Sensors and Actuators*, Vol. A68, 1998, pp. 419–428.
2. D.D.L. Wijngaards, S.H. Kong, M. Bartek and R.F. Wolffenbuttel, "Design and fabrication of on-chip integrated polySiGe and polySi Peltier devices," *Sensors and Actuators*, Vol. A85, 2000, pp. 316–324.
3. S.H. Kong, D.D.L. Wijngaards and R.F. Wolffenbuttel, "Infrared microspectrometer based on a diffraction grating," *Sensors and Actuators*, Vol. A92, 2001, pp. 88–95.
4. D.D.L. Wijngaards and R.F. Wolffenbuttel, "Opportunities for microtechnology in metrology," *IEEE Instrumentation & Measurement Magazine*, Vol. 4, No. 3, 2001, pp. 24–29.
5. D.D.L. Wijngaards and R.F. Wolffenbuttel, "IC fabrication-compatible processing for instrumentation and measurement applications," *IEEE Trans. on Instrumentation & Measurement*, Vol. 50, No. 6, 2000, pp. 1475–1484.
6. D.D.L. Wijngaards and R.F. Wolffenbuttel, "Study on temperature stability improvement of on-chip reference elements using integrated Peltier coolers," *IEEE Trans. on Instrumentation & Measurement*, Vol. 52, No. 2, 2003, pp. 478-482.
7. D.D.L.Wijngaards, G. de Graaf and R.F. Wolffenbuttel, "Single-chip microthermostat applying both active heating and active cooling," submitted to *Sensors and Actuators A*.

Editorships

1. R.F. Wolffenbuttel, D.D.L. Wijngaards and C.J. van Mullem; Editorship: *Proc. Symposium on Microtechnology in Metrology and Metrology in Microsystems*, 31 August–1 September 2000, Delft, The Netherlands, ISBN: 90-9014166-9.

Conference publications

1. D.D.L. Wijngaards, M. Bartek and R.F. Wolffenbuttel, "IC process compatible thin metal film postprocessing module," *Proc. ProRISC Workshop*, 27–28 November 1997, Mierlo, The Netherlands, pp. 591–596.
2. D.L. Wijngaards, M. Bartek and R.F. Wolffenbuttel, "IC process compatible metal film postprocessing module," *Proc. Eurosensors XI*, 21–24 September 1997, Warsaw, Poland, pp. 587–590.
3. D.D.L. Wijngaards, M. Bartek and R.F. Wolffenbuttel, "IC post-processing module for deposition of thin metal films," *Proc. of the Dutch National Sensor Conference*, 2–3 March 1998, Enschede, The Netherlands, pp. 241–246.
4. D.D.L. Wijngaards, S.H. Kong, M. Bartek and R.F. Woffenbuttel, "Thermal stabilisation of integrated silicon platforms using polysilicon and SiGe Peltier elements," *Proc. Transducers '99*, 7–10 June 1999, Sendai, Japan, pp. 310–313.
5. S.H. Kong, D.D.L. Wijngaards, M. Bartek and R.F. Woffenbuttel, "Study of active on-chip cooling using integrated Peltier elements," *Proc. Eurosensors XIII*, 12–15 September 1999, The Hague, The Netherlands, pp. 319–322. (C,O)
6. D.D.L Wijngaards, E. Cretu, S.H. Kong and R.F. Wolffenbuttel, "Modelling of integrated Peltier elements," *Technical Proc. MSM 2000*, 27–29 March 2000, San Diego, CA, USA, pp. 652–655.
7. S.H. Kong, D.D.L. Wijngaards, M. Bartek and R.F. Wolffenbuttel, "Infrared Micro-spectrometer based on a diffraction grating," *Proc. Eurosensors XIV*, 27–30 August 2000, Copenhagen, Denmark, pp. 321–324.
8. S.H. Kong, D.D.L. Wijngaards and R.F. Wolffenbuttel, "Fabrication of an integrated silicon infrared micro-spectrometer," *Proc. MME 2000*, 1–3 October 2000, Uppsala, Sweden, pp. 65–68.
9. D.D.L. Wijngaards, S.H. Kong and R.F. Wolffenbuttel, "Fabrication and application of on-chip integrated polySiGe and polySi Peltier elements," *Proc. MME 2000*, 1–3 October 2000, Uppsala, Sweden, pp. 12–15.

10. D.D.L. Wijngaards, E. Cretu, S.H. Kong and R.F. Wolffenbuttel, "Modelling of integrated polySiGe Peltier elements," *Proc. Therminic 2000*, 24–27 September 2000, Budapest, Hungary, pp. 275a–275d.

11. D.D.L. Wijngaards and R.F. Wolffenbuttel, "Post-processing modules in IC-compatible microsystem fabrication," *Proc. Symposium on Microtechnology in Metrology and Metrology in Microsystems*, 31 August–1 September 2000, Delft, The Netherlands, pp. 41–58.

12. S.H. Kong, D.D.L. Wijngaards and R.F. Wolffenbuttel, "IC compatible integrated Si infrared microspectrometer," *Proc. NanoSpace 2001*, 13–16 March 2001, Galveston, TX, USA.

13. D.D.L. Wijngaards, G. de Graaf, S.H. Kong and R.F. Wolffenbuttel, "Study on the behaviour of various Si substrate geometries for use as passive heat sinks," *Proc. of the Sensor Technology Conference 2001*, 14–15 May 2001, Enschede, The Netherlands, pp. 47–52.

14. S.H. Kong, D.D.L. Wijngaards G. de Graaf and R.F. Wolffenbuttel, "Fabrication and optical measurements of multi-slit grating based infrared microspectrometer," *Proc. Transducers '01 / Eurosensors XV*, 10–14 June 2001, Munich, Germany, pp. 548–551.

15. D.D.L. Wijngaards, S.H. Kong, P.M. Sarro and R.F. Wolffenbuttel, "Thermophysical characterisation of $PolySi_{0.7}Ge_{0.3}$ for use in thermoelectric devices," *Proc. Transducers '01 / Eurosensors XV*, 10–14 June 2001, Munich, Germany, pp. 1010–1013.

16. D.D.L. Wijngaards, G. de Graaf, S.H. Kong and R.F. Wolffenbuttel, "Temperature profile comparison of various substrate geometries for passive heat sinks," *Proc. MME 2001*, 16–18 September 2001, Cork, Ireland, pp. 143–146.

17. G. de Graaf, D.D.L. Wijngaards and R.F. Wolffenbuttel, "Evaluation of an oscilloscope training course," *Proc. Symposium "Virtual and real tools for education in measurement,"* 17–18 September 2001, Enschede, The Netherlands, pp. 67–74.

18. D.D.L. Wijngaards and R.F. Wolffenbuttel, "Characterisation and modelling of on-chip integrated Peltier elements for highly localised thermal stabilisation and cooling," *Proc. SemiTherm XVIII*, 12–14 March 2002, San Jose, CA, USA, pp. 105–112 .

19. D.D.L. Wijngaards and R.F. Wolffenbuttel, "Temperature stability improvement of on-chip reference elements using integrated Peltier coolers," *Conference Digest CPEM 2002*, 16–21 June 2002, Ottawa, Canada, pp. 522–523.

20. D.D.L. Wijngaards, G. de Graaf and R.F. Wolffenbuttel, "Single-chip microthermostat for stabilization at ambient temperature applying both active

heating and cooling," *Proc. Eurosensors XVI*, 15–18 September 2002, Prague, Czech Republic, 4pp.

21. D.D.L. Wijngaards, G. de Graaf and R.F. Wolffenbuttel, "Design of a single-chip micro-thermostat employing both active heating and cooling," *Proc. MME 2002*, 6–8 October 2002, Shinaia, Rumania, 4pp.

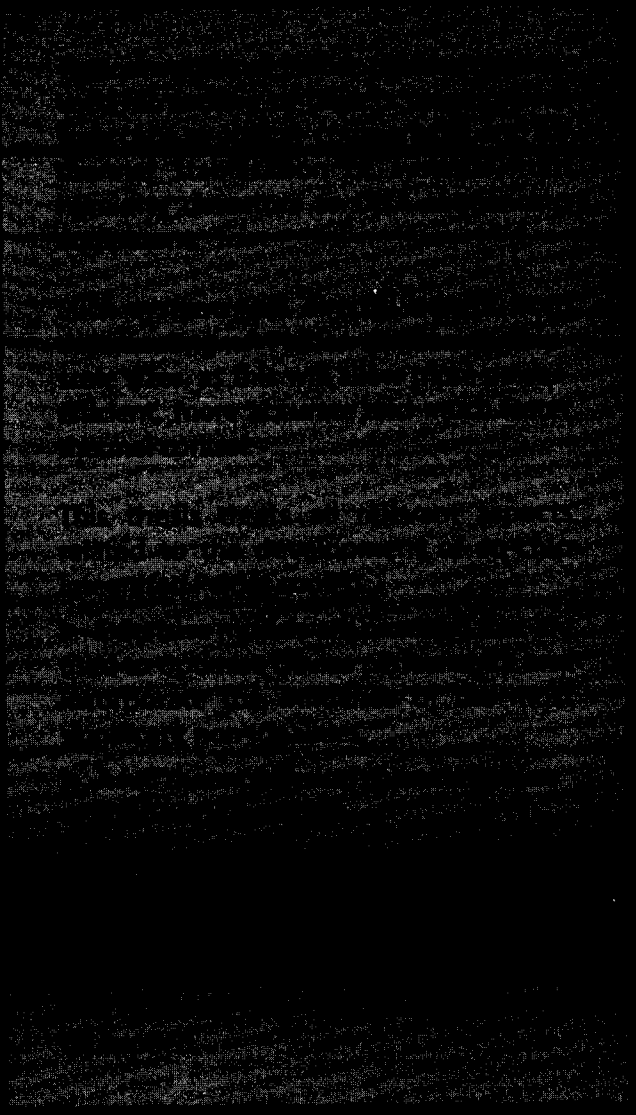
22. D.D.L. Wijngaards and R.F. Wolffenbuttel, "Thermo-electric performance of poly-$Si_{0.7}Ge_{0.3}$," *MRS Proc. Vol. 770, Optoelectronics and Group IV-based Materials*, 21–25 April 2003, San Francisco, CA, USA, accepted for publication.

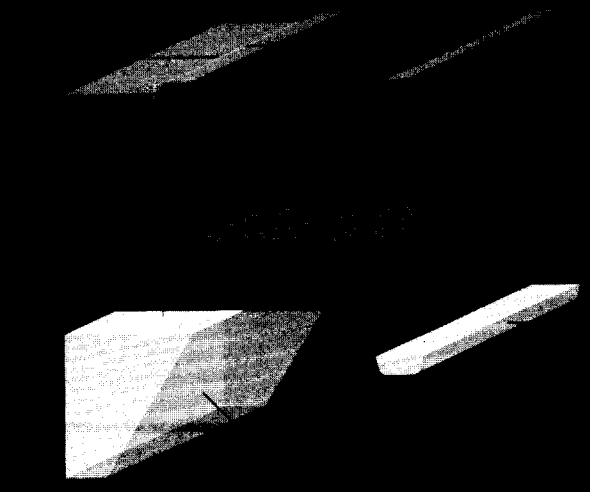
ABOUT THE AUTHOR

Davey Wijngaards was born in Bergen op Zoom, The Netherlands, on 6 February 1975. In 1993 he passed his pre-university education (*cum laude*) at the Gertrudis Lyceum in Roosendaal, and became an Electrical Engineering student at the Delft University of Technology. In March 1998 he received his M.Sc. degree in Electrical Engineering (*cum laude*). The graduation assignment included research on high-accuracy integrated resistive voltage dividers for which a generic thin-film metal deposition technique was developed.

From March 1998 until October 2002 he was a Ph.D. candidate at the Department of MicroElectronics, within the Delft University of Technology. There he performed research on on-chip integrated Peltier devices, as part of the STW project "Microinstrumentation platform in silicon." Since January 2003 he is employed at Océ-Technologies B.V. in Venlo, The Netherlands.

At the time of printing this thesis he has authored / co-authored a total of 27 publications, to be divided in 7 jounal publications and 22 conference publications. Furthermore he has chaired two symposia and edited two proceedings.





ISBN 90-9017216-5