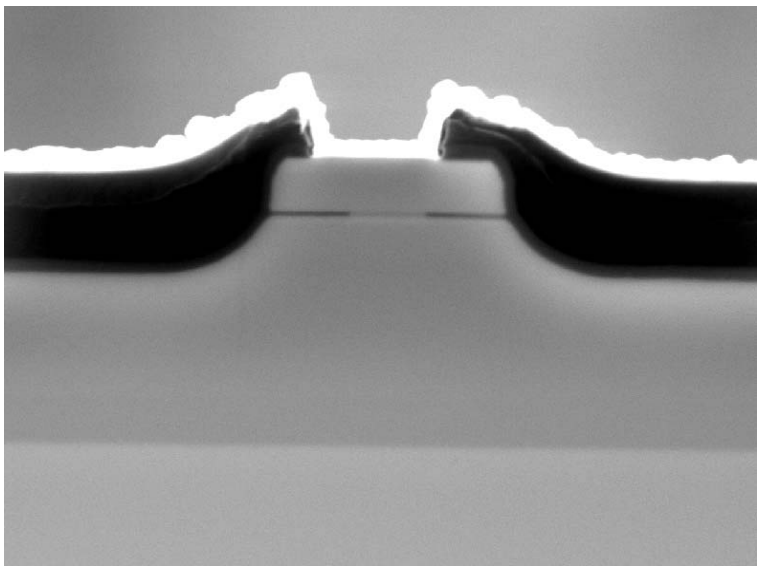
Universiteit Gent
Faculteit Toegepaste Wetenschappen
Vakgroep Informatietechnologie

Optische bundelexpansiestructuren voor fotonische geïntegreerde schakelingen

Spot-size converters for photonic integrated circuits

Kurt De Mesel



Proefschrift tot het bekomen van de graad van
Doctor in de Toegepaste Wetenschappen:
Elektrotechniek
Academiejaar 2002-2003

Promotor:

Prof. dr. ir. R. Baets	INTEC, Universiteit Gent, België

Examencommissie:

Prof. dr. P. Kiekens (voorzitter)	UGT, Universiteit Gent, België
Prof. dr. ir. M. Smit (leescommissie)	TU Eindhoven, Nederland
Prof. dr. ir. I. Moerman (leescommissie)	INTEC, Universiteit Gent, België
Dr. ir. D. Van Thourhout (leescommissie)	Lucent Technologies, Holmdel, USA
Dr. ir. J. Vanfleteren (leescommissie)	TFCG, Universiteit Gent, België
Prof. dr. ir. P. Van Daele	INTEC, Universiteit Gent, België
Prof. dr. ir. G. Morthier	INTEC, Universiteit Gent, België
Prof. dr. ir. P. Lagasse	INTEC, Universiteit Gent, België

Universiteit Gent
Faculteit Toegepaste Wetenschappen

Vakgroep Informatietechnologie (INTEC)
Sint-Pietersnieuwstraat 41,
B-9000 Gent
België

Tel.: +32-9-264.33.24
Fax.: +32-9-264.35.93

Dankwoord

Het dankwoord bij de aanvang van een eindwerk is traditioneel het laatste stukje tekst dat op papier wordt gezet. De eindfase is meestal ook een heel spannende periode waarbij de deadline plots veel sneller nadert dan verwacht. Vandaar dat ik het hier bij een kort, maar daarom niet minder gemeend, woord van dank houd.
Het onderzoekswerk zou uiteraard onmogelijk geweest zijn zonder de technologische faciliteiten van de Vakgroep Informatietechnologie. Hiervoor dank ik dan ook de vakgroepvoorzitter Prof. Paul Lagasse, Prof. Ingrid Moerman en Prof. Peter Van Daele. Prof. Roel Baets dank ik voor het in goede banen leiden van dit doctoraat.
Voor het programmeren van de meetsofware en het ter hulp snellen bij dienst weigerende meetinstrumenten bedank ik Geert Diet. Hij wist zich de laatste week verder in dit dankwoord naar voor te werken door de kuren van *Matlab*, bij het maken van de laatste figuren, nog gauw te verhelpen. Mathias Vanwolleghem verdient een woord van dank voor het uitmeten van de laagverliesstructuren. I would also like to thank Marko Galarza for the pleasant cooperation during the last years. Hendrik Sergeant dank ik voor het werk in de meetkamer en ook Kristien De Meulder mogen we zeker niet vergeten voor het draaiende houden van het netwerk, printers en PC's.
Prof. Bart Dhoedt bedank ik voor het in extremis in orde brengen van enkele administratieve zaken, en ook voor zijn hulp bij de start van dit werk.
Voor de epitaxiaalgroei van het halfgeleidermateriaal en het verdere verwerken ervan tot opto-elektronische componenten ben ik Carl Sys, Steven Verstuyft, Peter Geerinck, Liesbet Van Landschoot en Steven Van Put een woord van dank verschuldigd.
Voor de logistieke steun dank ik het administratief en technisch personeel van de vakgroep.
Verder ben ik ook de vele programmeurs van LaTeXdankbaar voor het verschaffen van een programma dat werkelijk doet wat men het opdraagt. Stefaan Vanhastel dank ik voor de vele LaTeX-tips en -technieken en mijn zus Ilse voor het uitlenen van haar computer. De traditionele 'last-but-not-least'-dank gaat uiteraard uit naar Mieke voor haar steun en het begrip dat ze opbracht voor de vele avonden die ze de laatste maanden alleen moest doorbrengen.

Gent, juni 2002
Kurt De Mesel

Contents

Nederlandstalige samenvatting

English text

Nederlandstalige

Samenvatting

1. Inleiding

1.1 Optische telecommunicatie: een kort overzicht

Momenteel worden massaal optische vezelsystemen geïnstalleerd om aan de steeds toenemende vraag naar bandbreedte te kunnen voldoen. Een lange geschiedenis van cruciale uitvindingen en technische innovaties hebben de optische communicatie uiteindelijk gemaakt tot wat ze nu is. Veel van haar potentieel moet echter nog verder ontdekt worden.

Reeds in de jaren '50 waren telecommunicatie ingenieurs op zoek naar meer bandbreedte als aanvulling van de reeds druk bezette radio- en microgolfkanalen. Het potentieel van hogere frequenties werd daarbij onderzocht om te kunnen blijven voldoen aan de steeds toenemende belasting door telefonie en televisie, waarvan het gebruik verwacht werd te zullen blijven toenemen. Daarbij werd voorgesteld de optische frequenties te benutten, maar het was pas met de uitvinding van de laser in 1960 dat men zich echt op dit domein begon te concentreren. Glasvezel leek het aangewezen transportmedium, maar vezels uit die tijd vertoonden echter veel te hoge verliezen ($\approx$1dB/m) om bruikbaar te zijn voor telecommunicatietoepassingen. Jarenlang onderzoek leidde uiteindelijk tot een drastische reductie van de vezelverliezen tot minder dan 20dB/km (bij 633nm). Ondertussen werden ook halfgeleiderlasers ontwikkeld die steeds langere golflengtes emitteerden. Dit leidde tot nog lagere propagatieverliezen, zoals ook blijkt uit figuur 1.1, die de evolutie van de vezelverliezen in functie van de golflengte schetst.

De eerste veldexperimenten voor telefoonverkeer werden in 1977 uitgevoerd met 850nm laserlicht in punt-tot-punt verbindingen over multimodale (MM) vezels, die boven monomodale (single mode, SM) vezels verkozen werden om de problemen met het inkoppelen van het laserlicht in de vezel te omzeilen. De vele geleide modes propageren echter elk met een lichtjes verschillende snelheid door de vezel (modale dispersie), waardoor de lichtpuls uitspreidt. Hierdoor zijn in dergelijke systemen dure 'repeaters' vereist, die het signaal weer een goede vorm geven en tegelijkertijd versterken om de attenuatie (2dB/km) te compenseren.
Algauw kwamen de eerste 1.3μm InGaAsP/InP lasers beschikbaar. Doordat de verliezen bij deze golflengte slechts 0.5dB/km bedragen, ontstond de interesse om een transatlantische kabel te ontwerpen en wel met SM vezel om de afstand tussen de 'repeaters' te kunnen opdrijven. Deze eerste kabel werd operationeel in 1988,

met een capaciteit die het dubbele was van alle transatlantische koperen kabels samen. Met de ontwikkeling van de 1.55μm lasers, de golflengte waarbij de vezelattenuatie zijn minimum van 0.2-0.3dB/km bereikt, kon de 'repeater'-afstand nog verder vergroot worden tot 100km.

Ondertussen werden ook steeds meer optische verbindingen geïnstalleerd op het land en niettegenstaande de enorm toegenomen bandbreedte, was de capaciteit veel vroeger dan verwacht volledig ingenomen door de opkomst van nieuwe applicaties, met het Internet in het bijzonder. De meest voor de hand liggende oplossing is meer vezels te installeren (Space Division Multiplexing, SDM), maar dit is uiteraard een dure optie. Er kan ook naar hogere modulatiesnelheden overgestapt worden door tijdsdomeinmultiplexering (Time Domain Multiplexing, TDM), waarbij verschillende signalen tot één hoge-snelheidssignaal worden gecombineerd. De maximale op deze manier bereikte modulatiesnelheid bedraagt tot nog toe echter 'slechts' 320Gbit/s, wat maar een fractie voorstelt van de 25THz bandbreedte die een optische vezel te bieden heeft. Golflengtemultiplexering (Wavelength Division Multiplexing, WDM) laat echter wél toe de volledige beschikbare bandbreedte te benutten. Figuur 1.3 toont hoe in een WDM-systeem verschillende golflengtes met een TDM-signaal gemoduleerd worden, om dan met behulp van een multiplexer gecombineerd te worden en over de vezel verzonden te worden. Aan de ontvangerzijde worden de verschillende golflengtes gedemultiplexeerd en opnieuw omgezet tot een elektrisch signaal. De doorbraak van WDM kon er enkel komen door de uitvinding van de 'erbium doped fibre amplifier' (EDFA), die toelaat het volledige signaal te versterken zonder het eerst te demultiplexeren (zogenaamde 'transparantie').
Alcatel demonstreerde in november 2001 een 3.65Tbit/s WDM systeem, dat 365 golflengtes, elk gemoduleerd aan 10Gbit/s, combineerde en over een afstand van 6850km transporteerde.

De eerste generatie WDM systemen waren punt-tot-punt verbindingen. Ondertussen zijn ze echter uitgegroeid tot heuse optische *netwerken*, waarbij de vezel de koperleidingen steeds verder verdringt en steeds dichter bij de gebruiker komt. Langzaamaan wordt ook de routering, het schakelen en het verwerken van de signalen verschoven van het elektrische naar het optische domein. Uiteindelijk zullen de trage elektronische knooppunten volledig vervangen worden door snelle en volledig optische componenten, die de tijdrovende en energieverslindende omzetting van elektrisch naar optisch en terug naar elektrisch signaal vermijden. Figuur 1.4 schetst deze evolutie. Het moge duidelijk zijn dat hier nog heel veel onderzoek en ontwikkeling moet verricht worden vooraleer deze verschuiving naar het optische domein volledig kan doorgevoerd worden. Essentiële componenten zijn hierbij 'add-drop multiplexers' en 'cross connects' die toelaten kanalen toe te voegen of af te splitsen en het signaal doorheen het netwerk routeren louter uitgaande van de golflengte van het signaal, en in een verdere toekomst komen daar ook nog de optische 'routers' bij.

1.2 Fotonische geïntegreerde schakelingen

Door de toenemende optische functionaliteit in netwerkknooppunten, die het gevolg is van het verschuiven van functies van het elektrische naar het optische domein, is er een nood ontstaan voor geavanceerde optische componenten of fotonische geïntegreerde schakelingen (Photonic Integrated Circuits, PICs). Het ontwikkelen van een PIC vormt een serieuze uitdaging die eruit bestaat een brede waaier van verschillende opto-elektronische basiscomponenten samen te brengen tot één module, via hybride of monolithische integratie. Vooral het samen integreren van actieve en passieve componenten is een niet-triviale opgave door de verschillende materialeneigenschappen en lagenstructuren die door beide types vereist zijn.

1.2.1 Actieve en passieve componenten

Passieve componenten zijn eenvoudige rechte of gebogen golfgeleiders of een samenvoeging ervan. Ze zijn opgebouwd uit materialen die transparant zijn bij de gebruikte golflengte. De energie van de verboden zone van de halfgeleidmaterialen is met andere woorden groter dan de foton energie. Actieve componenten bevatten dan weer meestal een materiaal met een verboden zone die gelijk is aan of kleiner is dan de energie van het foton en ze hebben ook steeds een extra ingang. Die ingang kan een stroom, spanning, een ander optisch signaal of een temperatuurcontrole zijn.

1.2.2 Materiaalsystemen

Bij geïntegreerde electronica worden nagenoeg alle componenten uit silicium gefabriceerd. Bij geïntegreerde optica is er echter niet zo'n 'ideaal' materiaalsysteem voorhanden. Elk van de verschillende materiaalsystemen heeft zijn eigen voordelen en dus ook typische toepassingen.

Halfgeleiders Momenteel is het op InP gebaseerde materiaalsysteem het best geschikte en meest gebruikte materiaalsysteem voor telecommunicatietoepassingen bij 1.3 en 1.55μm. Dit systeem bestaat uit InP, en het ternaire InGaAs en quaternaire InGaAsP, die via epitaxiaalgroeitechnieken roosteraangepast op het InP kunnen gedeponeerd worden. Door de samenstelling van het quaternaire materiaal te wijzigen, kan de verboden zone gevarieerd worden tussen 0.92 en 1.65μm. Dit maakt het mogelijk actieve en passieve functies te realiseren binnen het golflengtegebied dat voor optische telecommunicatie van toepassing is. Een heel gamma van zowel actieve als passieve componenten werd dan ook reeds gedemonstreerd. Een ander belangrijk halfgeleider materiaalsysteem is GaAs/AlGaAs, maar dan voor korte golflengtetoepassingen, zoals verticale caviteit oppervlakte emitterende lasers (VCSELs) binnen het golflengtegebied van 630nm tot 870nm. Er wordt momenteel echter veel onderzoek verricht voor het uitbreiden van het toepassingsgebied van GaAs-gebaseerde componenten naar langere golflengtes. Dit materiaalsysteem heeft namelijk een aantal belangrijke voordelen te bieden in vergelijking

met InP/InGaAsP. Zo zijn de substraten goedkoper en beschikbaar in grotere formaten, en vertonen de GaAs-gebaseerde lasers een veel beter temperatuurgedrag. Hier wordt in hoofdstuk 6 dieper op in gegaan. Het onderzoek heeft geleid tot de ontwikkeling van GaInNAs, een nieuw quaternair materiaal dat roosteraangepast kan groeien op GaAs substraten, en emissiegolflengtes tot 1.3μm en hoger kan realiseren. GaInNAs-gebaseerde VCSELs en zijdelings emitterende laserdiodes met een excellent temperatuurgedrag werden reeds gedemonstreerd. Een ander interessant kenmerk van het GaAs-systeem is de mogelijkheid om AlGaAs-lagen met een hoog aluminium-gehalte selectief te oxideren. De laag wordt dan omgezet in een hoog-kwalitatief oxide dat voor allerhande interessante toepassingen kan aangewend worden. Zo wordt het bijvoorbeeld aangewend om stroomaperturen te creëren die tegelijkertijd een optische golfgeleider definiëren. Een gedetailleerde beschrijving van het selectief nat oxidatieproces wordt gegeven in hoofdstuk 5.

Polymeren Polymeren hebben sterk aan belang gewonnen, vooral door de ontwikkeling van polymeer optische vezel (soms ook plastiek optische vezel genoemd of 'plastic optical fibre', POF), maar ook voor het fabriceren van planaire golfgeleidercircuits. Positieve aspecten van polymeergolfgeleiders zijn vooral hun goede koppelingseigenschappen naar monomode optische vezel. Actieve toepassingen beperken zich echter tot componenten die steunen op het relatief trage thermo-optische effect.

Silica Silica (SiO_xN_y) vormen momenteel het meest gebruikte materiaalsysteem voor het realiseren van puur passieve geïntegreerde schakelingen. De redenen hiervoor zijn de heel lage propagatieverliezen (0.02dB/cm bij 1550nm) en optische modeprofielen die heel goed aangepast zijn aan monomodale optische vezels. Er werden ook pogingen ondernomen om actieve functionaliteiten te realiseren rond 1530-1550nm door silica te doperen met erbium atomen en optisch te pompen met 980nm of 1480nm laserlicht. Hoge winstfactoren kunnen echter niet gerealiseerd worden.

$LiNbO_3$ Het ferro-electrische materiaal $LiNbO_3$ is gekend om zijn sterke acousto-optische en niet-lineaire electro-optische eigenschappen, wat het vooral geschikt maakt voor toepassingen zoals externe modulatoren (reeds beschikbaar bij verschillende producenten).

Glasvezel Glasvezel is niet langer enkel het passieve transportmedium, maar verschaft ondertussen een resem aan nieuwe functionaliteiten. De 'Erbium-doped fibre amplifiers' (EDFA) en Raman versterkers zijn hierbij zeker de belangrijkste nieuwe toepassingen. Het zijn breedbandige versterkers die voor de doorbraak van WDM-systemen hebben gezorgd, zoals hierboven reeds uitgelegd.

1.2.3 Hybride versus monolithische integratie

Voor het samenbrengen van de diverse subcomponenten tot een complexe geïntegreerde optische module zijn reeds veel hybride en monolithische integratietechnieken voorgesteld, maar een algemene doorbraak wordt nog steeds verhinderd door hun hoge kostprijs, die vooral veroorzaakt wordt door de problematische koppeling van het optische signaal van chip naar vezel en omgekeerd.
Bij *hybride* integratieschema's worden de verschillende subcomponenten in het voor hun functionaliteit best geschikte materiaalsysteem vervaardigd en vervolgens op een drager geassembleerd. Het voordeel hiervan is het feit dat elke component optimaal gerealiseerd kan worden en dat de componenten geselecteerd kunnen worden vooraleer ze te integreren. Het samenbrengen van de componenten op de drager en het nauwkeurig en stabiel fixeren blijkt dan weer een zware opdracht. De aligneringstoleranties zijn extreem klein als gevolg van de compacte optische modes. Sub-micrometer toleranties op de montage zijn meestal van toepassing wegens de sterk toenemende verliezen naarmate de misalignering verder toeneemt. Een grondige discussie van deze problematiek en mogelijke oplossingen worden in sectie 1.3 gegeven.
Door de problematische koppeling is meestal een actieve aligneringstechniek vereist. Hierbij wordt licht van de ene component in de andere gekoppeld tijdens het monteren. De relatieve positie van beide componenten wordt hierbij gewijzigd tot het ingekoppeld vermogen een maximum bereikt. Eenmaal dit maximum bereikt, worden beide componenten gefixeerd.
Veel beter zou het zijn over een passieve aligneringstechniek te beschikken, waarbij de verschillende componenten rechttoe rechtaan op de drager kunnen gemonteerd worden, enkel en alleen door aligneringstekens op drager en component te aligneren. Dit vereist uiteraard een tolerantere koppeling, typisch in de orde van 1-2μm. Een belangrijke passieve aligneringstechniek is zogenaamde 'silicon bench' (figuur 1.5).
Hybride integratietechnieken kunnen tijdelijk een antwoord bieden aan de integratievraag, maar uiteindelijk is het voor een economisch verantwoorde massaproductie van optische componenten nodig over een *monolithische* integratietechniek te beschikken die op volledige plakken halfgeleidermateriaal kunnen uitgevoerd worden. De verschillende componenten worden hierbij gedefinieerd via fotolithografische technieken en worden met elkaar verbonden door middel golfgeleiders die in dezelfde lithografische stap worden gedefinieerd. De intra-chip koppelingsproblematiek is hiermee van de baan, maar de problematiek van een efficiente en tolerante koppeling naar vezel blijft. Een van de grootste problemen bij monolithische integratietechnieken is echter het samen integreren van actieve en passieve componenten. Speciale groeitechnieken, zoals selectieve groei en gemaskeerde groei werden hiertoe reeds ontwikkeld.

1.3 Het koppelingsprobleem

1.3.1 Vezel-chip koppeling

De koppelingsproblematiek die optreeds bij vezel-chip koppeling werd reeds kort aangehaald. Het is een moeilijk te verhelpen probleem waar men sinds het begin van optische communicatie mee worstelt. De koppelingsverliezen kunnen gemakkelijk oplopen tot 10dB en de aligneringstoleranties, die een maat zijn voor de verliestoename in functie van een misalignering in de laterale, transversale of longitudinale richting, zijn enorm streng ($<1\mu$m) en leiden tot hoge verpakkingskosten, die kunnen oplopen tot 80% van de totale modulekost.
De oorsprong van de koppelingsproblematiek is te herleiden tot de kleine modeafmetingen en de sterk verschillende profielen van vezel- en chipmode, zowel qua grootte als qua vorm. De chipmode is typisch 1.5x2-3μm, terwijl de vezelmode circulair symmetrisch is met een diameter van 8-9μm.

1.3.2 Koppelingsefficiëntie en aligneringstoleranties

In de engelstalige tekst wordt de koppeling tussen chip- en vezelveld theoretisch behandeld om een beter inzicht te krijgen in de verschillende factoren die de koppelingsefficiëntie beïnvloeden. We vermelden hier de belangrijkste besluiten en verwijzen naar de engelstalige tekst voor meer details.
De algemene koppelingsconfiguratie is voorgesteld in figuur 1.6. Het veld dat aan het chipfacet geëmitteerd wordt, diffracteert in de luchtspleet tussen beide componenten vooraleer het op het vezelfacet invalt (figuur 1.7). De koppelingsefficiëntie wordt dan gegeven door (1.2). Om tot analytische uitdrukkingen voor de koppelingsefficiëntie te komen, worden de optische velden door Gaussische profielen van de vorm (1.3) benaderd, waarvoor het diffracterende veld eveneens analytisch gekend is (1.5). Evaluatie van de overlapintegraal voor een perfecte alignering van beide bundels leidt dan uiteindelijk tot (1.7), die de maximaal haalbare koppelingsefficiëntie weergeeft. De koppelingsverliezen bij perfecte alignering van beide bundels blijken sterk toe te nemen naarmate de modediameters sterker van elkaar verschillen. Figuur 1.8 geeft dit verloop grafisch weer.
Het verloop van de koppelingsverliezen in functie van de breedte van de luchtspleet tussen beide componenten is geschetst in figuur 1.9(a). Bij breder wordende luchtspleet blijkt het verlies toe te nemen, hoewel men zou kunnen verwachten dat het door diffractie uitdeinende chip-veld beter met de afmetingen van de vezelmode overeen zou moeten stemmen en de verliezen bijgevolg zouden moeten dalen. De daling van de curve wordt echter verklaard door de fasekromming die de diffracterende bundel vertoont en waardoor de overlapintegraal gereduceerd wordt.
In (1.8) en figuur 1.9(b) wordt de invloed van een laterale of transversale misalignering geschetst. We komen hier in de volgende sectie op terug.

1.3.3 Micro-optische technieken

In figuur 1.10 worden schematische voorstellingen gegeven van de verschillende technieken die aangewend worden om de koppelingsefficiëntie te verhogen en/of aligneringstoleranties te verruimen. Een eerste mogelijkheid is het gebruik van micro-optische lenzen (figuur 1.10(b)), die in het ideale geval echter asferisch dienen te zijn om de asymmetrie van de chipmode op te vangen. Bovendien blijkt een verhoging van de koppelingsefficiëntie gerealiseerd te worden ten koste van de aligneringstoleranties (vergelijk ook de curves in figuur 1.9(b)). Een andere techniek is het gebruik van gelensde vezels (figuur 1.10(c)). Ook hier gelden echter dezelfde opmerkingen als bij het gebruik van micro-lenzen. Uit (1.7) en (1.8) kan echter afgeleid worden dat de beste oplossing erin bestaat om aan de chip in- en uitgang(en) bundelexpansiestructuren te integereren, die de kleine asymmetrische chipmode omvormen tot een groot vezelaangepast profiel. De overeenstemmende mode diameters leiden dan tot een hoge koppelingsefficiëntie, doch niet langer ten koste van de aligneringstoleranties (zie de 'fibre-fibre'-curves in figuur 1.9(a) en figuur 1.9(b)). De grote nadelen van bundelexpansiestructuren zijn dan weer de complicaties van het epitaxiaal groeiproces en/of verdere bewerkingsproces die ze meestal met zich meebrengen. Het doel van deze thesis is dan ook het ontwerpen van nieuwe concepten voor dergelijke componenten, die echter het fabricageschema van de chip niet bemoeilijken.

1.4 Overzicht van het proefschrift

Na dit eerste inleidende hoofdstuk, waarin beschreven wordt hoe de evolutie naar volledig optische netwerken op gang gekomen is en hoe zo de nood aan complexe PICs ontstaat, gaan we in de rest van dit werk op zoek naar geïntegreerde oplossingen om de vezel-chip koppelingsproblematiek op te lossen. Meer bepaald wordt de nadruk gelegd op het bekomen van een eenvoudig fabricageschema, wat bij de meeste gangbare oplossingen niet het geval is.
In hoofdstuk 2 wordt een overzicht van de grote verscheidenheid aan bundelexpansiestructuren gegeven. Er wordt een classificatieschema voor deze componenten voorgesteld, dat verder verduidelijkt wordt door talrijke voorbeelden met vermelding van hun performantie. Van alle componenten vormen de 'adiabatische' de belangrijkste categorie. Ze vormen de kleine chipmode geleidelijk om tot een vezel-aangepast profiel, en dit met een verwaarloosbaar vermogenverlies. Wegens het belang van deze componenten wordt in hoofdstuk 3 het vaak geciteerde adiabaticiteitscriterium van Love rigoureus afgeleid door stukken bestaande theorie te combineren. We illustreren het criterium met enkele ontwerpsvoorbeelden en maken ook een kritische opmerking over de geldigheid ervan.
Hoofdstuk 4 beschrijft het ontwerp, fabricage en karakterisering van InP-gebaseerde componenten. Zowel structuren voor het bereiken van een efficiënte koppeling naar vezel als intra-chip bundelomvormers worden er besproken.
Een van de troeven van het GaAs/AlGaAs materiaal systeem is de mogelijkheid om lagen met een hoog aluminium-gehalte selectief te oxideren. Hoofdstuk 5 be-

schrijft het oxidatieproces en bespreekt de verschillende oxidatieparameters, om dan af te sluiten met enkele toepassingen. Hoofdstuk 6 behandelt het door ons voorgestelde concept van een geöxideerde laser met geïntegreerde bundelexpansiestructuur. Bij dit ontwerp worden de eigenschappen van het oxide en de hoge selectiviteit van het oxidatieproces ten volle benut om er de mode-evolutie in de component mee te bepalen. Het ontwerp, de fabricage en de karakterisering van dergelijke korte golflengte (980nm) componenten komen aan bod. Tot slot wordt ook het werk dat in de context van de projecten OPTIVAN en GIFT (beide financieel gesteund door de Europese Unie) werd uitgevoerd kort besproken. Hierbij wordt ingegaan op het karakteriseren van GaInNAs-gebaseerde componenten en het ontwerpen van 1.3μm geoxideerde bundelexpansiestructuren.
Hoofdstuk 8 sluit af met een korte samenvatting van het werk en schetst de toekomstperspectieven.

1.5 Publicaties in de context van dit werk

- K. De Mesel, R. Baets, C. Sys, S. Verstuyft, I. Moerman and P. Van Daele, "First demonstration of a 980nm oxide confined laser with integrated spot size converter", El. Lett., Vol. 36, No. 12, pp. 1028-1029, 2000.

- G. Flamand, K. De Mesel, I. Moerman, B. Dhoedt, W. Hunziker, A. Kalmar, R. Baets, P. Van Daele and W. Leeb, "InP-Based PIC for an optical phased array antenna at 1.06μm", IEEE Photonics Technology Letters, Vol. 12, No. 7, pp. 876-878, 2000.

- M. Galarza, K. De Mesel, D. Fuentes, R. Baets and M. López-Amo, "Modeling of InGaAsP-InP 1.55μm lasers with integrated mode expanders using fiber-matched leaky waveguides", Applied Physics B, Lasers and Optics, B 73, pp. 585-588, 2001.

- M. Galarza, K. De Mesel, S. Verstuyft, C. Aramburu, I. Moerman, P. Van Daele, R. Baets and M. López-Amo,"1.55μm InP-InGaAsP Fabry-Perot lasers with integrated spot size converters using Antiresonant Reflecting Optical Waveguides", IEEE Photonics Technology Letters, Vol. 14, No. 8, pp. 1043-1045, 2002.

- D. Taillaert, W. Bogaerts, P. Bienstman, T. Krauss, P. Van Daele, I. Moerman, S. Verstuyft, K. De Mesel and R. Baets, An out-of-plane grating coupler for efficient butt-coupling between compact planar waveguides and single-mode fibres, IEEE Journal of Quantum Electronics, Vol. 38, No. 7, pp. 949-955, 2002.

- M. Galarza, K. De Mesel, S. Verstuyft, C. Aramburu, I. Moerman, P. Van Daele, R. Baets and M. López-Amo,"A new spot-size converter concept using fiber-matched antiresonant reflecting optical waveguides", wordt gepubliceerd in IEEE Journal of Lightwave Technology in januari 2002.

- K. De Mesel, C. Sys, K. Caekebeke, S. Verstuyft, B. Dhoedt, I. Moerman, P. Van Daele and R. Baets, "Wet lateral oxidation of AlGaAs for passive and active guided wave components", Proceedings Annual Symposium IEEE/LEOS Benelux Chapter, pp.109-112, Gent, Belgium, November 1998.

- K. De Mesel, I. Moerman, R. Baets, B. Dhoedt, P. Van Daele and J. Stulemijer, "Spot size converters for low cost PICs", Proceedings 9th European Conference on Integrated Optics, pp. 253-258, Turin, Italy, 1999.

- K. De Mesel, C. Sys, S. Verstuyft, I. Moerman, P. Van Daele and R. Baets, "First demonstration of a 980nm oxide confined laser with integrated spot size converter", Proceedings Integrated Photonics Research, pp.126-128, Québec, Canada, 2000.

- K. De Mesel, C. Sys, S. Verstuyft, I. Moerman, P. Van Daele and R. Baets, "Demonstration of a novel type of tapered laser based on lateral selective wet oxidation", Proceedings Annual Symposium IEEE Lasers and Electro-Optic Society Benelux Chapter, pp.159-162, Delft, The Netherlands, October 2000.

- M. Galarza, K. De Mesel, R. Baets and M. López-Amo, "Design of InGaAsP-InP tapered ridge mode transformer using an underlying ARROW coupling waveguide", Proceedings Annual Symposium IEEE/LEOS Benelux Chapter, pp.79-82, Delft, The Netherlands, October 2000.

- M. Galarza, K. De Mesel, R. Baets and M. López-Amo, "Modelling of InGaAsP-InP 1.55μm lasers with integrated mode expanders using fibre-matched Antiresonant Reflecting Optical Waveguide", Proceedings 10th European Conference on Integrated Optics, pp. 421-424, Paderborn, Germany, 4-6 April 2001.

- M. Galarza, K. De Mesel, T. Van Caenegem, S. Verstuyft, C. Aramburu, I. Moerman, P. Van Daele, R. Baets and M. López-Amo, "InGaAsP-InP 1.55μm lasers with integrated mode expanders using fiber-matched leaky waveguides", Proceedings Lasers and Electro-Optic Society 2001 Annual Meeting, pp. 798-799, San Diego, California, USA, 12-15 November 2001.

- M. Galarza, K. De Mesel, S. Verstuyft, I. Moerman, P. Van Daele, R. Baets and M. López-Amo, "InGaAsP-InP 1.55μm lasers with integrated mode expanders using fibre-matched Antiresonant Reflecting Optical Waveguides", European Semiconductor Laser Workshop Gent, Belgium, 28-29 September 2001.

- G. Flamand, K. De Mesel, J. Derluyn, I. Moerman, B. Dhoedt, R. Baets, P. Demeester and P. Van Daele, "InP-based PIC for an optical phased array antenna at 1.06μm", Annual IEEE Lasers and Electro-Optic Society Benelux Chapter Workshop on Integrated Optics and Passive Micro-optics, p. 19, Brugge, Belgium, June 1998.

- G. Flamand, K. De Mesel, I. Christiaens, I. Moerman, B. Dhoedt, W. Hunziker, A. Kalmar, R. Baets, P. Van Daele and W. Leeb, "InP-Based PIC for an Optical Phased Array Antenna at 1.06μm", Proceedings 12th International Conference on Indium Phosphide and Related Materials, pp. 25-28, May 2000, Williamsburg, United States.
- K. De Mesel, S. Verstuyft, I. Moerman, P. Van Daele and R. Baets, "Oxide confined laser diodes with an integrated spot size converter", ingediend voor publicatie in IEEE Journal of Selected Topics in Quantum Electronics in September 2002.
- M. Galarza, K. De Mesel, S. Verstuyft, D. Fuentes, C. Aramburu, M. López-Amo, I. Moerman, P. Van Daele and R. Baets, "Mode-Expanded 1.55 μm InP-InGaAsP Fabry-Prot Lasers Using ARROW Waveguides for Efficient Fiber Coupling", ingediend voor publicatie in IEEE Journal of Selected Topics in Quantum Electronics in September 2002.

1.6 Octrooi aanvragen

Marko Galarza, Kurt De Mesel, Candido Aramburu and Roel Baets, 'Semiconductor optical component utilizing leaky structures to match the mode of rib waveguides to that of the fiber'

2. Overzicht van bundelexpansiestructuren

2.1 Inleiding

Geïntegreerde bundelexpansiestructuren kunnen op veel manieren gerealiseerd worden. Bijgevolg werden reeds talloze ontwerpen en verschillende principes voor modetransformatie gepubliceerd. Daarom is het interessant om bij het begin van dit werk even stil te staan bij die diversiteit aan concepten en er ook enige structuur in proberen te scheppen. We doen dit door een classificatieschema te presenteren, dat opgebouwd is vanuit een optische invalshoek. Hiermee wordt bedoeld dat er een onderscheid gemaakt wordt tussen de verschillende manieren waarop de modetransformatie gerealiseerd wordt enerzijds, en de wijze waarop de geometrie verandert anderzijds.

2.2 Classificatie

De twee dimensies –geometrie en optisch werkingsprincipe– die onderscheiden worden bij het opstellen van het classificatieschema, bestaan elk uit twee onderscheiden gevallen en een hybride overgangsvorm. Dit leidt uiteindelijk tot de negen categorieën zoals voorgesteld in tabel en verder toegelicht in de volgende paragrafen.

	Monomodaal 'Adiabatisch' **(A)**	Hybride **(A+I)**	Multimodaal 'Interferentie' **(I)**
Lateraal **(L)**	A/L	A+I/L	I/L
Hybride **(L+T)**	A/L+T	A+I/L+T	I/L+T
Transversaal **(T)**	A/T	A+I/T	I/T

Tabel 2.1: Optische bundelomvormers: een classificatieschema.

2.2.1 Geometrie

De dwarsdoorsnede van een bundelomvormer verandert, ofwel geleidelijk ofwel abrupt, in de lengterichting. De specifieke manier waarop die evolutie gebeurt, leidt tot een eerste onderverdeling in de klassen 'lateraal' (L), 'transversaal' (T) of hybride lateraal/transversaal (L+T). Een eenduidige definitie van deze klassen is de volgende: een modetransformator is van het 'laterale' ('transversale') type wanneer er geometrische veranderingen zichtbaar zijn in de 'transversale' ('laterale') projectie van de component (tenzij, uiteraard, deze veranderingen geen impact hebben op het optisch veld) en niet in de 'laterale' ('transversale') projectie. Bij de hybride klasse zijn er in beide projecties veranderingen zichtbaar. Het is belangrijk op te merken dat de termen 'lateraal' en 'transversaal' enkel duiden op geometrische veranderingen en niet op veranderingen van het optisch veld. Sommige zuiver laterale bundelexpansiestructuren leiden immers tot een belangrijke transversale modeconversie, zoals duidelijk zal worden uit de voorbeelden.

2.2.2 Optisch werkingsprincipe

Er zijn twee onderscheiden principes om een optische mode om te vormen tot een meer vezel-aangepast profiel. Een eerste categorie, met de zogenaamde 'adiabatische' componenten, realiseert een geleidelijke expansie van de fundamentele locale mode. De andere groep exciteert meerdere modes die met elkaar interfereren en zo het gewenste veld aan de uitgang opleveren.

Monomodale of adiabatische bundelomvormers (klasse A) De overgrote meerderheid van de bundeltransformatoren steunen op een adiabatische modeomzetting, waarbij de kleine en sterk asymmetrische chipmode geleidelijk omgevormd wordt tot een groot en meer cirkelvormig profiel. De term 'adiabatisch' duidt op het feit dat het volledige vermogen in één bepaalde locale mode geconcentreerd blijft tijdens de transformatie. Om de transformatieverliezen, die ontstaan door koppeling van vermogen naar hogere orde geleide modes of stralende modes, beperkt te houden, zijn relatief lange componenten met een traag variërende dwasdoorsnede vereist.
Zoals reeds aangegeven werd, maken adiabatische componenten een geleidelijke overgang tussen een sterk geleidende golfgeleider en een doorsnede die een uitgestrekte mode geleidt. Dit kan op twee manieren verwezenlijkt worden. Ofwel worden de afmetingen van de kern geleidelijk vergroot om zo te eindigen in een grote en vaak multimodale uitgang, waarvan wegens de adiabaticiteit van het ontwerp enkel de fundamentele mode geëxciteerd wordt. Een tweede optie is over te gaan op een zwak geleidende structuur met een kleine kern, waarbij de mode zich sterk in de omgevende mantellagen uitstrekt. In beide gevallen wordt een grote mode bekomen, in het ene geval een sterk geleide en in het andere geval een zwak geleide. Dit wordt geïllustreerd in figuur 2.1(b), waar de evolutie van de effectieve modebreedte van de fundamentele mode van de gelaagde structuur uit figuur 2.1(a) geschetst wordt in functie van de dikte van de kern (bij 1.55μm). Met de effectieve modebreedte wordt hier $w = d_k + d_{sub} + d_{mantel}$ bedoeld, waarbij d_k de kerndikte

is en d_{sub} respectievelijk d_{mantel} de penetratiediepte van het veld in het substraat, respectievelijk de mantel. We merken een minimum mode-afmeting op rond een kerndikte van 0.5μm. Voor deze dikte is het licht maximaal geconcentreerd in de kern, wat dan ook betekent dat dit het ideale ontwerp voor een lasergolfgeleider is. Het is intuïtief duidelijk dat dikkere kernen grotere modes zullen geleiden. Dit is ook duidelijk in figuur 2.1(b), die aangeeft dat de modebreedte evenredig met de kerndikte varieert. Anderzijds nemen de mode-afmetingen ook heel sterk toe voor dunner wordende kernen. De modebreedte evolueert nu omgekeerd evenredig met de kerndikte. Een nauwkeurige controle over de kerndikte is uiteraard vereist indien men voor dit laatste werkingsgebied kiest om een geëxpandeerde mode te bekomen.
Adiabatische componenten worden in hoofdstuk 3 gedetailleerd bestudeerd. Er wordt onder meer een criterium opgesteld om het ideale verloop van de mode-omvormer te bepalen.

Multimodale of interferentie bundelomvormers (klasse I) Een tweede klasse bestaat uit de multimodale of interferentie bundelomvormers. Bij deze componenten worden meerdere modes geëxciteerd die elk met hun eigen propagatieconstante verder doorheen de structuur propageren. Het veld op een willekeurige positie wordt dan bepaald door de interferentie van deze velden. Het totale veld varieert dus met de propagatie-afstand. De positie van het uitgangsfacet wordt dan zó gekozen dat het totale veld een maximale koppeling naar de vezel oplevert. Over het algemeen hebben deze componenten het voordeel veel korter te zijn dan de adiabatische types.

Hybride adiabatische/interferentie bundelomvormers (klasse A+I) Sommige componenten kunnen niet bij de klasses A of I ondergebracht worden omdat ze op een combinatie van beide principes gebaseerd zijn. Het gaat hierbij dan om een adiabatische omvorming in de laterale dimensie, gecombineerd met een interferentie-omzetting in de transversale richting, of omgekeerd.

2.2.3 Voorbeelden

In de engelstalige tekst worden heel wat voorbeelden van de verschillende klassen gegeven. Het zou ons echter te ver leiden dit hier te herhalen. We beperken ons dan ook tot het samenvatten van enkele algemene opmerkingen.
Er is een groot verschil tussen de complexiteit van het fabriceren van zuiver laterale en zuiver transversale adiabatische bundelomvormers (A/L resp. A/T). De A/L-componenten worden gefabriceerd door een combinatie van planaire groei- en etsprocessen. Het creëren van een *tranversaal* dunner of dikker wordende structuur vereist echter complexe groei en/of etstechnieken. Als bijzonderste worden de selectieve groei (figuur 2.2.3) en de schaduwmasker groeitechniek (figuur 2.6) besproken. Bij de niet-planaire etstechnieken worden de herhaalde natte etstechniek (figuur 2.7), etsen met ionenbundel en schaduwmasker (figuur 2.8), diffusiegelimiteerd etsen of ook nog dip-etsen vermeld. Het moge duidelijk zijn dat bun-

deltransformatoren die enkel het gebruik van conventionele, planaire groei-en etstechnieken vereisen, de voorkeur genieten.

2.3 Simulatiesoftware voor bundelexpansiestructuren

Bij het ontwerpen van opto-elektronische componenten is het van primordiaal belang over goede software te beschikken om optische modes te berekenen en de evolutie van het optisch veld doorheen een medium met willekeurige evolutie van de brekingsindex te simuleren. We bespreken hier kort de verschillende pakketten die in de vakgroep aanwezig zijn en die intensief gebruikt werden bij het ontwerpen van de bundelexpansiestructuren. Het gaat meer bepaald om een modebepaler, een mode-expansie algorithme en een bundelpropagatie methode (BPM). Dit zijn meteen ook de meest krachtige methoden die momenteel beschikbaar zijn voor het doorrekenen van optische structuren.

2.3.1 Modebepalers

Modebepalers zijn essentieel voor het bepalen van de propagatieconstanten van geleide modi in longitudinaal invariante structuren. In de context van dit werk werd de tweedimensionale modebepaler *Fimmwave* gebruikt, dat steunt op een transversale resonantie methode ('film mode matching'). Deze methode is in het bijzonder geschikt voor dwarsdoorsnedes die opgebouwd zijn uit rechthoekige gebieden met een constante brekingsindex, zoals het geval is voor halfgeleider optische golfgeleiders. De kennis van de propagatieconstante is van belang voor:

- het berekenen van het modeprofiel. Dit leidt tot inzicht in het gedrag van een optische component, bijvoorbeeld door de evolutie van de fundamentele locale mode in een bundelomzetter na te gaan (voor meer uitleg over het local modeconcept verwijzen we naar hoofdstuk 3).
- het berekenen van de opsluitfactor van een optisch veld in een bepaald gebied, bijvoorbeeld in de actieve laag van een laser.
- het bepalen van het verre veldpatroon. De kennis van het verre veld is belangrijk voor toepassingen zoals optische opnamesystemen of bij vezel-chip koppeling.
- het bepalen van de koppelingsefficiëntie tussen twee golfgeleiders door de overlapintegraal tussen beide velden te berekenen.

Het pakket laat ook toe 'projecten' te bouwen, waarmee gemakkelijk achterhaald kan worden hoe een aantal parameters, zoals propagatieconstanten of opsluitfactoren, veranderen wanneer de golfgeleiderdoorsnede wordt gewijzigd. Een dergelijke scanning is bijvoorbeeld nodig om de evolutie van de propagatieconstanten te bepalen, om dan hiermee de ideale vorm van het adiabatisch ontwerp vast te leggen (zie hoofdstuk 3).

Een ander type project laat toe een samengestelde component te scannen, bijvoorbeeld om na te gaan hoe de koppeling tussen twee componenten varieert in functie van de misalignering tussen beide. Figuur 2.16 geeft een voorbeeld van de berekende evolutie van de koppelingsefficiëntie tussen een laser en vezel als functie van de transversale misalignering. Hieruit kunnen dan gemakkelijk de -1dB aligneringstoleranties afgeleid worden.

2.3.2 Mode expansie methode

Een eerste algorithme om de evolutie van een willekeurige mode doorheen een arbitraire golfgeleider te berekenen is het mode-expansie algorithme. Hierbij wordt de golfgeleider benaderd door een aaneenschakeling van longitudinaal invariante golfgeleidersecties, zoals voorgesteld in figuur 2.17. Een willekeurig veld in zo'n sectie kan geschreven worden als een oneindige som van geleide, stralende en evanescente modes. Het op een sectie invallende veld wordt dan ook uitgedrukt als een lineaire combinatie van dergelijke locale modes, waarbij het spectrum van stralende modes uiteraard tot een discrete bijdrage beperkt wordt (de discretisatie wordt ingevoerd door het aanbrengen van 'harde' wanden rondom de golfgeleider). De propagatie van elke mode doorheen een sectie is triviaal. Bij de overgang naar het volgende element koppelen de modes naar de modes van de volgende sectie. Ook met het gereflecteerde vermogen wordt rekening gehouden, zodat het mode-expansie algorithme een bidirectionele methode wordt. Door deze procedure te herhalen van de eerste sectie tot de laatste, wordt de volledige component doorgerekend.
Het mode-expansie programma dat hier op de vakgroep beschikbaar is en voor dit werk intensief gebruikt werd, is *Fimmprop-3D* en werkt met *Fimmwave* als onderliggende modebepaler. Dit betekent dat een volledig vectoriële propagatie doorheen structuren met hoge indexcontrasten kan berekend worden. Bovendien is het een algorithme dat ook grote-hoek propagatie aankan, wat bij een BPM niet het geval is.

2.3.3 Bundel Propagatie Methode (BPM)

Bij BPM technieken wordt, uitgegaande van een veldverdeling op positie $z=z_k$, het veld op $z=z_{k+1}=z+\Delta z$ berekend. Door dit te herhalen van het begin tot het einde van de component, wordt het veldverloop doorheen de volledige structuur berekend. De belangrijkste beperkingen van de methode zitten in de benaderingen die doorgevoerd worden voor de berekeningen. Zo zijn BPM-programma's enkel geschikt voor lage indexcontrasten en propagatie van velden die rond de lengte-as geconcentreerd blijven. Hoewel er ondertussen al talrijke uitbreidingen uitgewerkt werden, blijft het een meer restrictief schema dan het mode-expansie algorithme.

2.3.4 Ontwerp van bundelexpansiestructuren: *BPM_CAD* versus *Fimmprop-3D*

Fimmprop-3D is zonder enige twijfel het meest geschikte simulatieprogramma voor het ontwerpen van bundelexpansiestructuren. Een heel belangrijk nadeel van het 3D-BPM-programma, vooral wanneer men in gedachten houdt dat het ontwerp van een dergelijke structuur een iteratief proces is, is het feit dat het doorrekenen van een bundelomvormer gemakkelijk meer dan 20 uur in beslag neemt. Bovendien is een geschikte keuze van het rooster voor de discretisering van de structuur en de propagatiestap Δz niet voor de hand liggend en is het niet meteen duidelijk of vreemde simulatieresultaten te wijten zijn aan deze keuze of aan de structuur zelf. Het instellen van de verschillende simulatieparameters in *Fimmprop-3D* is veel transparanter, mede dankzij de 'calculation diagnostic tool'. De mogelijkheid om de velden te visualiseren in een willekeurige doorsnede (zowel dwarsdoorsnedes als langsdoorsnedes) leidt tot een goed fysisch inzicht in de werking van de component. Het gedrag van de structuur wordt volledig bepaald door de locale modes en de koppeling ertussen. Eenmaal de koppelingsmatrices opgesteld zijn, kan een heel gamma aan informatie bekomen worden zonder verder zware berekeningen uit te voeren. Zo impliceert bijvoorbeeld –en dit is heel belangrijk bij het ontwerp van bundelomvormers– het 'uitrekken' van een structuur niet dat de structuur volledig opnieuw moet uitgerekend worden zoals dit wel het geval is bij een BPM. Aan de hand van de reeds berekende koppelingsmatrices is het nieuwe resultaat in een tiental seconden berekend.

3. Adiabaticiteitscriterium

3.1 Inleiding

Een eenvoudig te gebruiken criterium voor het ontwerpen van adiabatische componenten werd afgeleid door Love en medewerkers. Er werd al talloze keren naar dit criterium gerefereerd met de vermelding dat het intuïtief duidelijk in te zien is. In werkelijkheid vereist dit echter een heel sterke intuïtie en daarom vonden we het aangewezen om dit criterium hier rigoureus af te leiden aan de hand van de theorie van de gekoppelde locale modes.
Omdat het hoofdstuk grotendeels bestaat uit een wiskundige afleiding, verwijzen we naar de engelstalige tekst voor details, en houden het hier bij een schets van de gevolgde redenering. Op het einde van het hoofdstuk worden ook nog een paar voorbeelden gegeven van de praktische toepassing van het criterium. We maken ook nog een kritische opmerking over de exactheid ervan.

3.2 Locale modes

In een golfgeleider met longitudinaal (in de z-richting) invariant profiel propageert elektromagnetische straling onder de vorm van modes zoals beschreven door (3.1) en gekarakteriseerd door hun propagatieconstante β. Een bundelomvormer, van de andere kant, wordt getypeerd door een veranderende dwarsdoorsnede. Van eigenmodes is niet langer sprake. De bundelexpansiestructuur kan echter benaderd worden door een aaneenschakeling van z-invariante secties, zoals geschetst in figuur 3.2. In elke sectie zijn eigenmodes van de vorm (3.1) opnieuw oplossingen van de Maxwell vergelijkingen. Ze worden 'locale' modes genoemd, omdat ze de oplossingen zijn van een oneindig lange golfgeleider met als dwarsdoorsnede de locale dwarsdoorsnede. Doordat de eigenmodes een complete set vormen, kan het exacte veld op een willekeurige positie in de bundelomvormer exact geschreven worden als een superpositie van de locale modes, zoals voorgesteld in (3.4) en (3.5).
Over een sectie neemt de fase van mode j toe met het product van de sectielengte en de propagatieconstante van die mode op die plaats. In het geval van een perfect adiabatische mode-omvormer (die uiteraard niet bestaat aangezien enkel een z-invariante structuur met behoud van vermogen verder propageert) kan het verloop van het optisch veld geschreven worden zoals in (3.7). In woorden uitgedrukt

betekent dit dat de beschouwde locale mode haar vermogen behoudt (constante amplitude), haar vorm verandert (via de z-afhankelijkheid van het elektrisch en magnetisch veld) en in fase toeneemt naarmate ze verder doorheen de structuur propageert. In deze eerste orde benadering wordt uiteraard nog geen rekening gehouden met de interactie met andere modes.

3.3 Locale modekoppeling

In een z-afhankelijke structuur wordt continu vermogen uitgewisseld tussen de verschillende locale modes. Dit gedrag wordt in rekening gebracht door de koppeling tussen de locale modes nader te beschouwen. We vermelden hier enkel in grote lijnen de gevolgde werkwijze.
De koppeling die ontstaat tussen de verschillende locale modes kan ingezien worden door de overgang tussen twee secties te beschouwen (zie figuur 3.3). Een invallende locale mode is niet langer een locale mode van de volgende sectie. Ze kan echter wel geschreven worden als een lineaire combinatie van de locale modes van die volgende sectie. Op die manier wordt vermogen overgedragen van de invallende mode op de locale modes van de volgende sectie. In sectie 3.3 wordt uitgaande van dit idee een stelsel gekoppelde differentiaalvergelijkingen opgesteld (3.21)-(3.22) die de verandering van de amplitudes van de verschillende modes beschrijven. Hierbij spelen de koppelingscoëfficiënten C_{jk}, die de graad van koppeling tussen mode j en k beschrijven, een centrale rol. De koppelingscoëfficiënten worden gegeven door (3.24).

3.4 Adiabaticiteitscriterium

De opgestelde gekoppelde modevergelijkingen dienen nu als basis om een criterium op te stellen dat een voldoende trage variatie garandeert om een adiabatisch gedrag te garanderen. Aan de ingang van de component wordt enkel de fundamentele mode geëxciteerd. Uit de gekoppelde modevergelijkingen kan dan een uitdrukking opgesteld worden voor het vermogen in een willekeurige andere mode j (die vermogen uit de fundamentele mode oppikt). Die uitdrukking wordt gegeven door (3.35). Het vermogen in de j-de orde mode blijkt nu verwaarloosbaar laag te blijven als de factor F(z), die in de uitdrukking voorkomt, voor elke positie z verwaarloosbaar klein blijft. Die factor wordt gegeven door de verhouding van de locale koppelingslengte en het gemiddelde verschil tussen de propagatieconstantes van de fundamentele mode en mode j. De redenering die nu door Love gevolgd wordt, is op te merken dat de koppelingscoëfficiënt sterk afhankelijk is van veranderingen van de dwarsdoorsnede (zie 3.24) en dat het verschil in propagatieconstanten omgekeerd evenredig is met de zwevinglengte ('beat' lengte) tussen beide modes. Hiermee komt hij tot de volgende formulering van het criterium voor trage variatie (en dus adiabatische mode-evolutie):

Een locale mode propageert met een verwaarloosbaar vermogenverlies naar een andere locale mode indien de veranderingen van de golfgeleiderdoorsnede gebeuren op een lengteschaal die groot is vergeleken met de locale zwevingslengte tussen beide modes.

3.5 Practisch ontwerpscriterium

Uit het criterium voor trage variatie kan een praktisch ontwerpscriterium voor adiabatische bundelexpansiestructuren, die de fundamentele locale mode omvormen, afgeleid worden. We beschouwen hiervoor een lateraal versmallende structuur zoals voorgesteld in figuur 3.4. Het criterium wordt in dit geval vertaald tot de eis dat de relatieve veranderingen van de breedte over een lengteschaal die correspondeert met de locale zwevingslengte, verwaarloosbaar klein is. Hierbij dient het langste zwevingsverschijnsel in acht genomen te worden, of, met andere woorden, zweving tussen de fundamentele mode en de mode met dichtstbijgelegen propagatieconstante. Verdere uitwerking leidt uiteindelijk tot (3.37), of, alternatief, tot (3.38) voor de locale wighoek $\Omega(z)$ zoals gedefinieerd in figuur 3.4. Uit deze laatste twee vergelijkingen blijkt dat een snellere versmalling toegelaten is in het gebied waar de propagatieconstanten het verst uit elkaar gelegen zijn. Er dient wel opgemerkt te worden dat het criterium ons enkel iets leert over de *vorm* en niet over de lengte van de component, aangezien de lengte afhangt van de interpretatie van het '$\ll$'-teken in 3.38 (of dus van het vermogenverlies dat toegestaan wordt).

Voorbeeld 1: Ontwerp van een adiabatische multimodale modetransformator Bij dit eerste voorbeeld wordt het criterium toegepast om de 'ideale' vorm te bepalen van de bundelomvormer die een adiabatische overgang realiseert naar een brede multimodale golfgeleider. Door de geleide modes te benaderen door een sinusvormig verloop (zie figuur 3.5), waarvoor gemakkelijk een analytische uitdrukking voor de propagatieconstanten kan opgesteld worden, komen we uiteindelijk tot het besluit dat het optimale breedteverloop parabolisch van vorm is (zie (3.42) en figuur 3.6(b)). De lengte van de component hangt af van de beoogde graad van adiabaticiteit (het toegestane verlies voor de fundamentele mode), en dient verder bepaald te worden door middel van een BPM of een mode expansie algorithme.

Opmerking Men zou zich kunnen afvragen of het berekende breedteverloop echt wel overeenkomt met het optimale ontwerp dat uit de eis $F(z) \ll 1$ (zie (3.35)) volgt, aangezien het praktische ontwerpcriterium uit een toch wel vrije interpretatie van deze eis volgt. In de engelstalige tekst wordt de eis $F(z) \ll 1$ rigoureus uitgewerkt en komen we tot het besluit dat het parabolisch verloop dat hierboven met het praktische criterium werd afgeleid inderdaad wel het ideal verloop voorstelt. Of dit ook zo is voor andere structuren (zoals voor het volgende voorbeeld) is hiermee echter niet bewezen.

Voorbeeld 2: Ontwerp van een adiabatische bundelexpansiestructuur
In dit hoofdstuk berekenen we het verloop van een meer complexe structuur, die voorgesteld is in figuur 3.8. Bij deze monomodale component wordt vermogen overgekoppeld naar het continuüm van stralingsmodes. De zwevingslengte die nu in aanmerking dient genomen te worden is de zweving tussen de fundamentele mode en de stralende mode met dichtstbijgelegen (en dus hoogste) effectieve index. Een analytische behandeling is nu niet meer mogelijk en er dient uitgegaan te worden van het gesimuleerde verloop van de betreffende propagatieconstanten (zie figuur 3.9). Via het iteratief schema (3.48)-(3.49) kan dan opnieuw het verloop van de versmalling van de bovenste rib berekend worden. Figuur 3.10 toont het berekende verloop van de breedte, het verloop van het vermogen in de fundamentele mode en een boven- en zijaanzicht op het expanderende veld (berekend met *Fimmprop-3D*). Ter vergelijking worden in figuur 3.11 dezelfde grafieken getoond voor een lineair verloop van de versmalling. Het is duidelijk dat het adiabatisch ontwerp reeds van in het begin van de component een modetransformatie realiseert, terwijl de lineaire component in het begin te traag en op het einde te snel verandert.

4. InP-gebaseerde bundelexpansiestructuren

4.1 Inleiding

Zowel in de onzoekswereld als bij industriële fabricage van optische componenten zijn twee golfgeleider implementatieschema's gangbaar, die elk hun specifieke voor- en nadelen vertonen. Hierbij gaat het om zogenaamde 'overgroeide' structuren enerzijds en golfgeleiders van het 'rib-type' anderzijds. In de context van dit werk werden bundeltransformatoren van beide types bestudeerd.
Componenten van het rib-type vereisen slechts één enkele epitaxiale groeistap, gevolg door één of meerdere etsstappen. Figuren 4.1(a) en 4.1(b) tonen een typische dwarsdoorsnede van een passieve, respectievelijk een actieve component.
Bij overgroeide structuren wordt voor passieve golfgeleiders na het etsen nog een extra groeistap uitgevoerd (figuur 4.1(c)). Voor actieve componenten zijn twee extra groeistappen vereist om een invers gepolariseerde junctie te creëren naast de actieve sectie, die de geïnjecteerde stroom door de actieve sectie dwingt (figuur 4.1(d)).
Dank zij haar typisch lagere interne verliezen en een betere stroomopsluiting wordt voor actieve componenten meestal de overgroei-technologie verkozen. De lagere elektrische opsluiting voor rib-componenten wordt veroorzaakt doordat het etsen meestal boven de actieve laag gestopt wordt om oppervlakrecombinatie aan de geëtste golfgeleiderwanden te vermijden, echter met een zijdelings uitspreidende stroom tot gevolg. De hogere interne efficiëntie van overgroeide structuren is toe te schrijven aan de lagere verstrooiingsverliezen door het lagere indexcontrast aan de randen van de actieve golfgeleider.
Ribgolfgeleiders zijn wegens het hogere laterale contrast dan weer beter geschikt voor passieve golfgeleiders, omdat dit kortere bochtstralen toelaat.
Beide types golfgeleiders vereisen uiteraard een verschillende aanpak om een bundelexpansiestructuur te realiseren. Terwijl overgroeide structuren meestal evolueren naar een zwak geleidende golfgeleider waarvan het veld tot diep in de omgevende mantellagen dringt, zullen rib-type modetransformatoren het optisch veld dwingen te verschuiven naar een onderliggende vezelaangepaste rib, zoals geschetst in figuur 2.9(a). De dikke laag halfgeleidermateriaal die vereist is voor de vezelaangepaste golfgeleider moet hierbij bestaan uit een quaternair InGaAsP met een

brekingsindex die gelegen is tussen de index van het InP-substraat en die van de bovenliggende kern. Dit vereist een heel strenge controle over de samenstelling van het roosteraangepaste quaternaire materiaal. Zelfs wanneer het groeiproces voldoende nauwkeurig kan gecalibreerd worden om de gewenste materiaalcompositie te bereiken, dan nog is het niet evident de groeiparameters constant te houden tijdens een groei die verschillende uren duurt. In sectie 4.3 hebben we een oplossing voor dit probleem uitgewerkt. We bespreken er het principe van 'Anti-Resonant Reflecting Optical Waveguides' (ARROW) en hoe dit concept kan toegepast worden in modetransformatoren.
Het ontwerp, de fabricage en de karakterisering van rib-type bundelexpansiestructuren worden behandeld in sectie 4.2.

Tot slot wordt ook het gebruik van intra-chip tapers voor het realiseren van laag-verlies golfgeleiders besproken in sectie 4.4. Het adiabatisch karakter van deze componenten zorgt er immers voor dat enkel de fundamentele mode van een multimodale golfgeleider, die met lage verstrooiingsverliezen propageert, wordt geëxciteerd.

4.2 Overgroeide bundelexpansiestructuren

4.2.1 Situering

De overgroeide bundelexpansiestructuur die in deze sectie wordt besproken maakt deel uit van een PIC die de optische signalen van een 'phased array' antenne combineert tot één sterk uitgangssignaal. Dergelijke PIC kan ingeschakeld worden in een vrije ruimte communicatiesysteem voor inter-satelliet communicatie. Dit werk past dan ook in het kader van een project voor de Europese ruimtevaartorganisatie ESA. Het principe van het systeem wordt verder uitgelegd in de engelstalige tekst. Een belangrijke vereiste was dat de module niet meer dan 10dB tussenschakelverlies mocht vertonen. Het belang van bundeltransformatoren aan de in- en uitgangen van de chip is dan ook meteen duidelijk, aangezien een rechttoe-rechtaan koppeling resulteert in 7.5dB koppelingsverlies per facet.

4.2.2 Algemene beschouwingen

De golfglengte van het vrije ruimte communicatiesysteem is vastgelegd op 1.06μm. Dit is de golflengte van de Nd:YAG laser, die verkozen wordt als zender wegens zijn hoge vermogen en smalle lijnbreedte. Een vergelijkende studie van de verschillende materiaalsystemen leidde uiteidelijk tot de keuze van het InP/InGaAsP materiaalsysteem om de PIC te realiseren. 1.06μm is echter een eerder ongebruikelijke golflengte voor dit halfgeleidersysteem. Bovendien is het bruikbare gebied van quaternair InGaAsP sterk beperkt tot samenstellingen met een verboden zone die correspondeert met een golflengte λ_g van 0.98-1.03μm, zoals geschetst in figuur 4.3. Voor $\lambda_g < 0.98\mu$m is het indexcontrast met InP te laag om efficiënte golfgeleiding te bekomen, terwijl voor $\lambda_g > 1.03\mu$m de absorptie sterk toeneemt.

De epitaxiaalgroei van dergelijke materialen bleek een uitdaging op zich, wegens de vereiste hoge compositionele nauwkeurigheid (zie tabel 4.1).

4.2.3 Ontwerp

Figuur 4.4 schetst een bundelexpansiestructuur die verenigbaar is met de lagenstructuur van de PIC (zie tabel 4.2). De overgroeide structuur bestaat uit een kern waarvan het bovenste deel langzaam wordt versmald tot nul. Tegelijkertijd wordt het dunne onderste deel verbreed en wordt op die manier een structuur bekomen die zwak geleidend is in transversale richting, maar sterk geleidend in laterale richting. De mode spreidt uit in de InP-mantellaag en het InP-substraat. Op die manier wordt een tweedimensionale expansie gerealiseerd door een louter laterale versmalling, wat het gebruik van niet-planaire groei- of etstechnieken vermijdt.

De uitgangsdwarsdoorsnede werdt ontworpen door de koppeling naar een 1.06μm monomodale polarisatie-behoudende vezel te optimaliseren. Uiteindelijk werd besloten een 100nm dikke, 6μm brede kern te implementeren met een 2.75μm dikke bovenste mantellaag. De berekende koppelingsefficiëntie voor deze structuur bedraagt 68% (1.75dB verlies), wat vergeleken dient te worden met een efficiëntie van 17.8% (7.5dB) bij een koppeling zonder modetransformator.

Om de totale lengte van de component te reduceren, werd geopteerd voor een stuksgewijs lineaire evolutie van de dwarsdoorsnede. Dit leidde uiteindelijk tot een 200μm lang ontwerp, waarvan de adiabaticiteit met behulp van tijdrovende 3D-BPM simulaties werd geverifieerd. Het longitudinaal ontwerp en de evolutie het optisch veld zijn geschetst in figuur 4.5.

4.2.4 Fabricage

Figuur 4.6 geeft de evolutie weer van de dwarsdoorsnede aan het begin, halfweg en aan het einde van de bundelomvormer. Aan de hand van deze figuur bespreken we nu kort het fabricageproces.

Groei van de wafer We vertrekken van een lagenstructuur die via planaire epitaxiaalgroei (MOCVD, MetalOrganic Chemical Vapour Deposition) werd gerealiseerd. De lagenstructuur bestaat uit een onderste InP mantellaag, de InGaAsP kernlaag en daar bovenop een 100nm dik laagje InP, dat de kwaliteit van de latere overgroeistap moet bevorderen (zie verder).

Etsen van de versmallende rib In een volgende stap wordt met een combinatie van een droog (RIE, Reactive Ion Etching) en een nat etsproces de versmallende punt geëtst, 600nm diep in de kernlaag. Een eerste droge etsstap etst nagenoeg de volledige diepte en resulteert in verticale golfgeleiderwanden. Omdat bij een droge ets de diepte niet goed gecontroleerd kan worden, wordt daarna nog een korte

natte selectieve ets uitgevoerd die automatisch stopt (of toch sterk vertraagt) op het 30nm dunne etsstoplaagje dat hiervoor in de kern werd aangebracht. Dit laagje is zo dun dat het de werking van de PIC niet beïnvloedt.

Overgroeien met de bovenste mantellaag Vooraleer de bovenmantel over de geëtste structuur te laten groeien , wordt het monster kort gedipt in H_2SO_4, wat de hergroei gunstig beïnvloedt. Het grensgebied met de mantellaag die over de structuur groeit is enkel van goede kwaliteit als de onderliggende laag uit InP bestaat in plaats van InGaAsP. Daarom werd in de eerste groeistap een laagje InP voorzien bovenop de kernlaag. Uiteindelijk werd slechts een 1.4μm dikke bovenmantel aangebracht (in plaats van 2.75μm), om problemen met de volgende etsstap te vermijden. Uiteraard heeft dit zijn gevolgen voor de haalbare koppelingsefficiëntie, die nu nog 58% (2.3dB verlies) bedraagt.
De natte ets bleek ook een gunstig effect te hebben op de kwaliteit van de bovenmantellaag. Hergroei op een zuiver drooggeëtste rib gaven aanleiding tot zogenaamde 'rabbit ears' (figuur 4.7), die verdere etsstappen sterk bemoeilijken. Verschillende etsmengsels werden uitgeprobeerd en de beste resultaten werden behaald met een zwavelzuurmengsel.

Etsen van de brede golfgeleider Ten slotte wordt in een tweede etsstap (RIE, 2.2μm diep) de brede golfgeleider geëtst die de geëxpandeerde uitgangsmode definieert. Het masker dat de brede golfgeleiders definieert, dient hierbij uiterst nauwkeurig met de reeds aanwezige punten gealigneerd te worden om fouten, zoals geschetst in figuur 4.9, te vermijden. Om de impact van kleine aligneringsfouten op de uiteindelijk structuur te reduceren, werden in het masker speciale voorzieningen getroffen (zie figuur 4.10).

4.2.5 Karakterisering

Twee belangrijke parameters om de performantie van een bundelomvormer in te schatten, zijn het koppelingsverlies naar vezel en de transformatieverliezen. De transformatieverliezen werden opgemeten door met de Hakki-Paoli methode (zie appendix) de verliezen van een gewone golfgeleider op te meten en dezelfde meting te herhalen op een golfgeleider met een bundelomvormer aan beide uiteinden. Door beide resultaten met elkaar te vergelijken kunnen dan de transformatieverliezen geëxtraheerd worden. We vonden op die manier een waarde van 1.15dB. Dit verlies kan nagenoeg volledig toegeschreven worden aan de beperkingen van de fotolithografie, waardoor smalle golfgeleiderpunten niet continu doorlopen en op verschillende plaatsen onderbroken worden. Het vezel-chip verlies werd opgemeten in ETH Zürich, aangezien een meetopstelling voor dergelijke metingen op dat moment nog niet voorhanden was in onze vakgroep. Er werd een koppelingsverlies van 2.8dB opgemeten, iets meer dus dan het theoretisch voorspelde 2.3dB. De oorzaak hiervan wordt duidelijk wanneer het nabije veld van de uitgangsmode van de modeomvormer geïnspecteerd wordt (figuur 4.11). Hieruit blijkt dat de uitgangsmode 6μm breed en 3.5μm hoog is, terwijl een hoogte van 4.5μm verwacht werd. Uit berekening met (1.7) blijkt dit inderdaad tot het extra verlies van 0.5dB

te leiden. Deze afwijking is te wijten aan een iets te dikke kernlaag aan de uitgang of aan een lichtjes afwijkende materiaalsamenstelling van de kern. Desalniettemin kunnen we toch besluiten dat het integreren van dergelijke bundelomvormers uiteindelijk leidt tot een reductie van het vezel-naar-vezel verlies met 7dB.

4.2.6 Verbeterd ontwerp

Om het propagatieverlies verder te reduceren, werd besloten om een tweede maskerset aan te maken die het mogelijk maakt om smallere ribben te definiëren. Hiervoor wordt de zogenaamde 'knife-etch' techniek aangewend. De punt wordt nu gedefinieerd door middel van twee fotolithografische stappen, zoals geschetst in figuur 4.12. In de engelstalige tekst wordt in detail beschreven hoe de maskers aangepast werden om tot een robuust ontwerp te komen dat afwijkingen in de uiteindelijke structuur, zoals geschetst in figuur 4.13, elimineert.

Alhoewel met deze techniek heel smalle punten gerealiseerd werden, zijn we er wegens problemen met de hergroei niet in geslaagd om de performantie van de componenten te verbeteren. De kwaliteit van de hergroei blijkt immers kritisch afhankelijk te zijn van de parameters van het RIE-proces, die echter moeilijk te reproduceren zijn, vooral in een systeem dat voor het etsen van verschillende materiaalsystemen dienst doet.

4.3 Bundelexpansiestructuren van het rib-type met tranversale 'ARROW' opsluiting

4.3.1 Inleiding

In deze sectie bespreken we het ontwerp van een 1.55μm laser met een monolithisch geïntegreerde bundelomvormer. De component is gebaseerd op de structuur die in figuur 4.18 wordt voorgesteld. Deze component bestaat uit een laserstructuur, die bovenop een brede vezelaangepaste rib is geplaatst. In de lasersectie is de laserrib relatief breed, en is het optisch veld sterk in de de laserkern opgesloten. Naarmate de laserrib in de bundelomvormer geleidelijk aan versmald wordt, zakt die sterk opgesloten mode uit naar de brede onderliggende structuur. Dit is een heel interessant modetransformator-concept, omdat één planaire epitaxiale groeistap en conventionele etstechnieken volstaan voor de fabricage ervan.
De koppelingsefficiëntie naar een monomodale vezel wordt bepaald door de afmetingen van de onderste golfgeleider. De groei van dikke lagen quaternair InGaAsP materiaal is dus vereist. Bovendien moet de brekingsindex van die laag tussen de index van de laserkern en die van het InP gelegen zijn, omdat ze enerzijds dienst doet als mantel voor de lasermode en anderzijds ook de kern van een golfgeleider met InP-cladding vormt. De epitaxie van dergelijk dikke lagen met lage Ga- en As-fracties is echter niet evident. Het ware interessant, mocht de dikke laag kunnen vervangen worden door het gemakkelijk te groeien InP. Op het eerste gezicht

is dit echter niet van toepassing in deze structuur wegens het vereiste contrast met het substraat. Indien echter, zoals wij hier voorstellen, een hoogreflectieve Bragg-achtige structuur onder die dikke kern wordt aangebracht, dan kan wel een transversale mode-opsluiting bekomen worden. Natuurlijk wordt zo geen geleide mode bekomen, maar een lekkende mode die vermogen naar het substraat toe verliest. Een geschikt ontwerp van die reflecterende structuur kan er echter voor zorgen dat de propagatieverliezen van een dergelijke mode minder dan 0.1dB/cm bedragen.
Dit opsluitingsprincipe bleek al te zijn toegepast in de jaren '80 in het silica-op-silicium materiaalsysteem, zij het om totaal andere redenen (zie engelstalige tekst). Dergelijke structuren zijn er gekend onder het acroniem 'ARROW', wat staat voor 'AntiResonant Reflecting Optical Waveguides'. De implementatie van ARROW's in bundelexpansiestructuren leidde uiteindelijk tot een octrooiaanvraag.

Het werd al gauw duidelijk dat ARROWs een heel gamma aan toepassingen verschaft, gaande van rib-type bundelomvormers (die hier behandeld worden) en overgroeide modetransformatoren, tot het reduceren van stralingsverliezen in bochten. Door de het uitgebreide spectrum aan toepassingen werd het ARROW-concept dan ook het onderwerp van een doctoraatsthesis op zich, door Marko Galarza.

4.3.2 Principe en ontwerp van 'ARROWs'

Implementatie van het ARROW-concept in de hierboven besproken bundelomvormer leidt tot een structuur zoals voorgesteld in figuur 4.20. Uiteraard is de ARROW-opsluiting enkel van belang op het einde van de bundelomvormer, waar de mode al voldoende is uitgezakt en heeft ze geen invloed op de lasermode. De lagenstructuur die bijgevolg beschouwd dient te worden voor het ontwerp van de ARROW is de structuur zonder de laserrib, zoals getoond in figuur 4.21. Het ontwerp zelf wordt gedetailleerd uiteengezet in de engelstalige tekst. We beperken ons hier tot de essentie ervan. Door de twee lagen quaternair materiaal, met een hogere brekingsindex dan het omgevende InP, in de structuur aan te brengen, ontstaan discontinuïteiten in het transversale brekingsindexverloop die reflecties creëren. De laagdiktes worden nu zó gekozen, dat de verschillende reflecties, die de InP kern opnieuw binnendringen, constructief interfereren. De mantellaag wordt met andere woorden gevormd door een serieschakeling van anti-resonante Fabry-Pérot caviteiten. De kern zelf kan beschouwd worden als een resonante caviteit, waaruit meteen volgt dat de dikte van de InP-laag in de mantel half zo dik dient te zijn als de kern zelf. De andere laagdiktes volgen uit de voorwaarde 4.2 voor anti-resonantie.

We merken ten slotte nog kort op dat, hoewel de verliezen voor de fundamentele mode heel laag kunnen gehouden worden, de hogere orde modes aanzienlijke verlieswaarden vertonen. De structuur blijkt zich ook nagenoeg polarisatie-onafhankelijk te gedragen.

4.3.3 Ontwerp van de bundelexpansiestructuur

Door middel van overlapberekeningen werd de uitgangsdwarsdoorsnede van de component geoptimaliseerd voor koppeling naar een standaard lange golflengte monomodale vezel. Hierbij werd er wel op gelet dat de totale dikte van de te groeien lagenstructuur niet meer dan 10-11μm mag bedragen om problemen met de epitaxiaalgroei te vermijden. De uiteindelijke structuur realiseert een koppelingsefficiëntie van 52% (2.8dB verlies), een verbetering met 6.6dB.

Er werd gekozen om een stuksgewijs lineaire benadering van het ideale adiabatische verloop te implementeren. Figuur 4.23 toont een bovenaanzicht op het verloop van de laserrib, samen met de evolutie van het optisch veld in enkele representatieve dwarsdoorsnedes.

4.3.4 Resultaten

Uiteindelijk werden twee implementaties van de componenten gefabriceerd. Het verschil tussen beide versies is de etsdiepte van de laserrib. De eerste optie etst tot boven de actieve laag, wat eventuele oppervlakrecombinatie aan de geëtse wanden vermijdt, maar wel het risico oplevert dat de expanderende mode gedeeltelijk geabsorbeerd wordt in het ongepompte actieve materiaal naast de laserrib. Bij de tweede optie wordt dwars door de quantumputten geëtst. Op die manier is er wel kans op een verhoogde oppervlakrecombinatie, maar wordt absorptie in de modetransformator uitgesloten. Voor een gedetailleerde bespreking van de opgemeten verschillen tussen beide verwijzen we naar de doctoraatsthesis van Marko Galarza. We wensen hier enkel de demonstreren dat het door ons voorgestelde principe in de praktijk ook werkt.

Figuur 4.24 toont enkele opgemeten optisch vermogen-versus-stroom curves (PI-curves) van diepgeëtste mode-geëxpandeerde lasers en ook van een referentielaser. De hogere drempelstroom is toe te schrijven aan het feit dat het actieve volume van de component aanzienlijk groter is (750μm laser met daarbij nog de lengte van de actieve modetransformator) dan voor de 750μm lange referentielaser. De differentiële efficiëntie blijkt zelfs iets hoger te zijn dan voor de referentielasers. Dit is toe te schrijven aan de verlaagde reflectie voor de optische mode aan het uitgangsfacet van de bundeltransformator. Figuur 4.25 toont hoe het verre veld stralingspatroon, in goede overeenstemming met de simulaties, efficiënt wordt gereduceerd van 30^ox40^o (lateraal x transversaal) tot 10^ox27^o.

4.3.5 Suboptimaal 'ARROW'-ontwerp

Er werd reeds opgemerkt dat de InP-laag in de ARROW mantel half zo dik dient te zijn als de bovenliggende InP kernlaag. Die kernlaag dient nu juist dik te zijn, om een goede koppeling naar vezel te bekomen. Dit resulteert in een ARROW-ontwerp met een dik lagenpakket, wat een speciaal voor de groei van dikke lagen

geoptimaliseerd groeiproces vereist. In de vakgroep is het groeiproces echter geoptimaliseerd voor dunnere lagen. Bovendien wordt over het algemeen wordt een kortere groeitijd sterk op prijs gesteld.
We kunnen ons echter de vraag stellen of het werkelijk nodig is een *optimale* ARROW te implementeren, vooral gelet op het feit dat de ARROW-opsluiting slechts over een korte afstand, namelijk op het einde van de bundelomvormer, dienst doet.
Door in de ARROW-mantel van het ideale ontwerp af te wijken en dunnere lagen te gebruiken, nemen de lekverliezen uiteraard toe. Uiteindelijk kon, zoals in de engelstalige tekst beschreven wordt, de totale dikte van de ARROW-mantel gereduceerd worden van 3.15μm tot 1.3μm en werd hierdoor slechts 0.08dB extra verlies geïntroduceerd. Meteen wordt zo het belangrijkste nadeel van de ARROW-opsluiting verholpen.

4.3.6 Laterale 'ARROW' opsluiting

Het ARROW-principe kan ook aangewend worden om een laterale optische opsluiting te creëren, zoals geïllustreerd wordt in figuur 4.26. Het principe kan ruwweg ingezien worden door de structuur in de effectieve index-benadering te beschouwen. De laterale ribben worden getypeerd door een andere effectieve index dan de gebieden er vlak naast. Op die manier ontstaat een laterale ARROW. Het exacte werkingsprincipe is echter complexer en een gedetailleerde studie van het modegedrag is vereist bij het ontwerpen van een dergelijke structuur. Voor meer details verwijzen we opnieuw naar het werk van Marko Galarza.
De belangrijkste beperking schuilt in het feit dat de dikte van de InP-kern beperkt is, aangezien de ribben hun invloed verliezen wanneer die kern te dik wordt. De structuur die in figuur 4.26 voorgesteld is realiseert uiteindelijk een koppelingsefficiëntie van 43% naar een standaard monomodale vezel (3.7dB verlies), wat een verbetering van 4.7dB inhoudt.

Het is heel belangrijk hierbij op te merken dat het fabricageschema van deze component uiterst eenvoudig is. Het is meer bepaald het meest eenvoudige schema voor een laser met een monolithisch geïntegreerde bundelomvormer dat ooit gepubliceerd werd. Er is slechts één enkele planaire groeistap en één conventionele etsstap vereist. Aangezien de laterale ribben in dezelfde etsstap als de laserrib geëtst worden, is het fabricageproces niet gecompliceerder dan voor een conventionele laser.

Figuur 4.28 toont enkele opgemeten PI-curves. De overeenkomst tussen gesimuleerde en opgemeten verre velden is duidelijk uit figuur 4.29.

4.4 Multimodale laagverlies golfgeleiders

4.4.1 Inleiding

Golfgeleiders in PICs worden meestal gerealiseerd door droge etsing omwille van de verticale wanden die hiermee kunnen gerealiseerd worden, onafhankelijk van de oriëntatie van de rib ten opzichte van de kristalrichting. Dergelijke golfgeleiders vertonen echter redelijk hoge propagatieverliezen (typisch 2dB/cm) die nagenoeg volledig toe te schrijven zijn aan de verstrooiing van het veld aan wandruwheden (figuur 4.30). Dit verstrooiingsproces werd theoretisch bestudeerd door Marcuse en een van de logische besluiten was dat de verstrooiing evenredig is met de veldintensiteit ter hoogte van de ruwheden.

We stellen hier voor om brede, multimodale golfgeleiders te gebruiken om de verliezen te reduceren. In die multimodale golfgeleider wordt enkel de fundamentele mode geëxciteerd, waarvan de intensiteit aan de rand veel lager is dan voor de geleider mode in een monomodale golfgeleider. Bijgevolg zullen ook de verstrooiingsverliezen veel lager zijn. De monomodale excitatie, die ook vereist is om ongewenste interferentie-effecten te vermijden, wordt verzekerd door adiabatische modeomvormers die een overgang maken vanuit een monomodale golfgeleider. Figuur 4.31 schetst het principe van de structuur.

4.4.2 Ontwerp

Er werd besloten te werken met diepgeëtste golfgeleiders (zie figuur 4.32) die typisch hogere verstrooiingsverliezen vertonen, zodat het verliesverlagende effect van ons principe dan ook duidelijker moet waargenomen worden. De grote vraag bij het ontwerp is uiteraard hoe breed die brede golfgeleiders precies moeten gekozen worden. Om hier enigszins een idee over te verkrijgen, werd de evolutie van de effectieve index van de fundamentele mode in functie van de golfgeleiderbreedte uitgezet in figuur 4.33. Deze index evolueert voor brede golfgeleiders naar de index van het kernmateriaal (cfr. vlakke golf wanneer de kern oneindig dik is). Het vlakke verloop van de effectieve index voor brede golfgeleiders is een indicatie voor de sterke opsluiting. Een andere aanwijzing voor de vereiste breedte is gegeven in figuur 4.34, die de evolutie van het vermogen in een 50nm (typische ruwheidsamplitude) dunne randlaag weergeeft in functie van de golfgeleiderbreedte. Een absoluut idee van welke breedte voldoende moet geacht worden is er met deze figuren uiteraard nog niet, maar het moge duidelijk zijn dat de verwachte evolutie van de golfgeleiderverliezen gelijkvormig zal zijn met de curve in figuur 4.34. Uiteindelijk werd geopteerd om 4, 6 en 10μm brede structuren (en 1.9μm brede monomodale referentiegolfgeleiders) op het testmasker te implementeren, om de evolutie van de verliezen duidelijk te kunnen observeren.

Er werden verschillende modetransformatoren op het masker geïmplementeerd. Er werd gekozen voor lineaire en exponentiële componenten. De verantwoording voor het gebruik van exponentiële, en niet voor parabolische evoluties, zoals in

Breedte	**Configuratie met exponentiële modetransf.**	**Configuratie met lineaire modetransf.**
1.9 μm	5 dB/cm	
4 μm	0.6 dB/cm	0.8 dB/cm
6 μm	0.5 dB/cm	0.6 dB/cm
10 μm	0.15 dB/cm	0.2 dB/cm

Tabel 4.2: Measured average loss of monomodal reference waveguides and broad low-loss waveguides.

hoofdstuk 3 uit het adiabaticiteitscriterium werd afgeleid, is in de engelstalige tekst terug te vinden. Het gedrag van de bundelomvormers werd aan de hand van 2D-BPM simulaties gesimuleerd.

4.4.3 Karakterisering en resultaten

De verliezen van de golfgeleiders en de taper werden bepaald via Hakki-Paoli metingen. In de engelstalige tekst wordt in detail ingegaan op het meetprincipe en de meetresultaten. We beperken ons hier tot de essentie. Tabel vat de opgemeten verliezen voor verschillende breedtes van de brede golfgeleider samen (zie ook figuur 4.38). De 5dB/cm voor de monomodale golfgeleiders is relatief hoog, maar onderlijnt de indrukwekkende reductie van de propagatieverliezen (tot 0.15dB/cm voor 10μm brede structuren!) nog sterker.
Alle modetransformatorverliezen blijken heel laag te zijn (<0.05dB), zoals voorgesteld in figuren 4.39 en 4.40.

5. Selectieve natte oxidatie

5.1 Situering

Reeds van in het begin van de studie van het GaAs/AlGaAs materiaalsysteem is geweten dat AlGaAs-lagen met een hoog aluminiumgehalte onstabiel zijn en oxideren door blootstelling aan de atmosfeer. Deze oxidatie leidde zelfs tot volledige vernietiging van de componenten in het geval van AlAs. De degradatie voltrekt zich typisch over een tijdsspanne van jaren.
Bij een poging om het oxidatieproces te versnellen door het in een stoomomgeving te plaatsen, kwam men tot de bevinding dat zich onder deze omstandigheden een stabiel en optisch hoog-kwalitatief oxide vormt, dat dus totaal verschilt van het oxide dat zich bij lagere temperaturen en over de jaren heen bij atmostferische omstandigheden vormt. Tabel 5.1 vergelijkt beide oxidevormen.

5.2 De oxidatie-opstelling

Figuur 5.2 schetst een typische oxidatie-opstelling en figuur 5.3 toont de opstelling die in de vakgroep gebouwd werd. Stikstof wordt doorheen een 'bubbler' gestuurd die water bevat op 80-95^{o}C. De verzadigde waterdamp wordt dan naar de oven geleid die typisch op een temperatuur van 370-450^{o}C gehouden wordt. In de engelstalige tekst wordt verder uitleg gegeven over het opbouwen van de opstelling.
In wat volgt beschouwen we, tenzij anders aangegeven, *laterale* oxidatie (figuur 5.1), waarbij een rib geëtst wordt die de laag met hoog aluminium-gehalte blootlegt. Het oxide vormt zich vanaf de geëtste wand en groeit verder naar binnen.

5.3 De oxidatiereactie

De engelstalige tekst beschrijft gedetailleerd de reactiemechanismen die leiden tot de vorming van het Al_2O_3.

5.4 Eigenschappen van het oxide

De natte oxidatie van $Al_xGa_{1-x}As$ leidt tot een stabiel oxide, in tegenstelling tot AlAs, waar een heel sterke inkrimping (20%) wordt waargenomen en bovenliggende lagen hierdoor soms zelfs volledig kunnen loskomen. De aanwezigheid van Ga creëert een stabiele matrix die de krimp van het oxide sterk beperkt. Het verschil tussen beide gevallen is duidelijk zichtbaar in figuur 5.6.
De eigenschappen van het oxide zijn drastisch verschillend van die van het halfgeleidermateriaal waaruit het is ontstaan. Zo is het een goede elektrische isolator en wordt het getypeerd door een lage brekingsindex (zie figuur 5.7), wat interessante mogelijkheden schept voor toepassingen in opto-elektronische componenten (zie verder).

5.5 Kinetica en oxidatieparameters

5.5.1 Tijdsafhankelijkheid

De reactie is een typisch diffusie/reactie proces. Het doet zich voor als een diffusiegelimiteerd proces wanneer de reactie vlugger verloopt dan de aanvoer van reagentia doorheen het reeds gevormde oxide. Dit is typisch het geval bij langere oxidatietijden (dus grotere oxidedieptes) en ook bij hogere temperaturen. Een reactiegelimiteerd proces wordt dan weer waargenomen wanneer de reactie trager verloopt dan de aanvoer van reagentia, meer bepaald voor korte oxidatietijden en lage temperaturen. Figuur 5.8 illustreert dit.

5.5.2 Temperatuurafhankelijkheid

De oxidatiesnelheid is sterk temperatuurafhankelijk. Een exponentiële afhankelijkheid wordt hierbij waargenomen (zie figuur 5.9).

5.5.3 Afhankelijkheid van alminiumgehalte

Ook het aluminiumgehalte heeft een sterke invloed op oxidatiesnelheid (figuur 5.10). Deze eigenschap leidt er toe dat bij laterale oxidatie enkel de lagen met een hoog aluminiumgehalte oxideren. Deze hoge selectiviteit is een heel interessante eigenschap met het oog op toepassingen.

5.5.4 Afhankelijkheid van de laagdikte

Voor heel dunne lagen (<80nm) neemt de oxidatiesnelheid af met de dikte van de laag en stopt nagenoeg volledig voor diktes van 10-20nm. Figuur 5.11 toont een electronenmicroscoop-foto van een structuur met een aantal dunne lagen van verschillende dikte. De hieruit opgemeten oxidatiesnelheid in functie van de laagdikte wordt in figuur 5.12 weergegeven.

5.5.5 Afhankelijkheid van doperingstype

Ook het type van dopering blijkt zijn invloed te hebben op de oxidatiesnelheid: p-type materialen oxideren sneller dan n-type halfgeleiders.

5.6 Toepassingen

Aanvankelijk werd het oxide gebruikt om door verticale oxidatie stroomaperturen te definiëren voor winst-geleide lasers (figuur 5.13(a)). Kort daarna werd ook de lage brekingsindex van het oxide benut om de stroomapertuur tegelijkertijd te laten dienst doen als optische golfgeleider (figuur 5.13(b)). Later werden via laterale oxidatie begraven stroomaperturen gerealiseerd, die de stroomspreiding sterk reduceren (5.14).
Nog andere toepassingen, gaande van passieve golfgeleiders, het realiseren van hoge dubbelbreking, hoog-reflectieve DBR spiegels voor VCSELs, heel kleine stroomaperturen in VCSELs,... worden in verder behandeld in de engelstalige tekst.

6. GaAs-gebaseerde bundelexpansiestructuren

6.1 Inleiding

Zoals in het inleidende hoofdstuk werd vermeld, situeert het toepassingsgebied van GaAs-gebaseerde componenten zich in het korte golflengtegebied tussen 700nm en 1.1μm. Belangrijke golflengtes hierbij zijn 850-870nm, de emissiegolflengte van (strained) GaAs en typisch gebruikt voor korte afstand communicatie over multimode vezels, en 980nm, voor componenten met InGaAs actieve lagen en typisch gebruikt voor hoog-vermogen pomplasers voor EDFAs. Bij deze laatste toepassing heeft het gebruik van bundelexpansiestructuren naast het realiseren van een hoogefficiënte koppeling naar vezel ook nog het voordeel het hoge uitgangsvermogen over een groter facetoppervlak uit te spreiden. Hierdoor kunnen hogere uitgangsvermogens bereikt worden omdat men zo langer onder de COD-vermogendichtheid blijft (Catastrophic Optical Damage level, $\approx$1-10MW/cm^2).

Bij GaAs-gebaseerde componenten is het echter enkel mogelijk rib-type componenten te fabriceren, omdat het overgroeien op AlGaAs wegens de aanwezigheid van het sterk reactieve aluminium niet mogelijk is. Een interessant ontwerp is de structuur waarop de ARROW-modeomvormer gebaseerd is (zie hoofdstuk 4), omdat een tweedimensionale mode-expansie bereikt wordt door een zuiver laterale versmalling van de bovenste rib. Bovendien vereist de fabricage slechts een planaire groei en twee conventionele etsstappen.

Dit hoofdstuk vat samen hoe we het fabricageschema verder vereenvoudigd hebben door de mogelijkheden die het oxidatieproces bieden ten volle te benutten. We bespreken hoe een welgekozen plaatsing van de oxidelaag een stroomapertuur creëert die tegelijkertijd het mode-gedrag van de laser met geïntegreerde bundelomvormer bepaalt. Op deze manier kunnen we het fabricageschema verder reduceren tot een planaire epitaxiale groei en één enkele *niet-kritische* etsstap. Bovendien dienen er enkel brede maskerpatronen (>8μm) fotolithografisch gedefinieerd te worden.

6.2 Principe

Het selectieve oxidatieproces heeft heel wat te bieden voor het vereenvoudigen van fabricageschema's en het verbeteren van de performantie van opto-elektronische componenten. Dit werd in hoofdstuk 5 uitvoerig aangehaald. Hier gaan we een stap verder en gebruiken we het oxide om het mode-gedrag van de component te dirigeren.

Figuur 6.1 toont het conventionele ontwerp, samen met ons voorstel dat gebruik maakt van selectieve oxidatie. De conventionele component wordt gedefinieerd door twee etsstappen waarbij een nauwkeurige alignering van beide maskerplaten vereist is. Bij onze geoxideerde component volstaat echter één enkele etsstap die, zoals verder zal worden aangetoond, niet kritisch is. Bovendien zijn in deze component de optische en de stroomapertuur auromatisch gealigneerd. De lasergolfgeleider wordt gedefinieerd door het laterale indexcontrast dat door het oxide gecreëerd wordt.

Bij het ontwerp dient er op gelet te worden dat de transversale mode in de niet-geoxideerde centrale lagenstructuur ('slab'-mode) zich sterk in de actieve laserkern concentreert, terwijl die voor de geoxideerde lagenstructuur zo diep mogelijk in de onderliggende dikke kern moet gesitueerd zijn. De drastische impact die de overgang van niet-geoxideerd AlGaAs naar oxide veroorzaakt wordt getoond in figuur 6.2. Figuur 6.3 toont een bovenaanzicht op de component en de evolutie van de fundamentele locale mode. In de lasersectie, waar het oxidevenster het breedst is (2μm), is de mode sterk geconcentreerd in de actieve laserkern. In de bundelomvormer wordt de brede rib versmald. Dit brengt de oxidestroken dichter bij elkaar, tot ze elkaar uiteindelijk raken. Op die manier wordt de optische mode geleidelijk aan van de laserkern in de vezelaangepaste onderliggende golfgeleider gedwongen. Hierbij dient opgemerkt te worden dat op deze manier een volledige versmalling tot nulbreedte wordt gerealiseerd zonder dat dit de definitie van smalle lijnen vereist: bij een versmalling van het masker van 10 naar 8μm, versmalt de breedte van het centrale oxidevenster (=lasergolfgeleider) van 2 naar 0μm.

6.3 Ontwerp

6.3.1 Ontwerp van de dwarsdoorsnede

Het ontwerp van deze eerste generatie korte golflengte (980nm) componenten is er vooral op gericht het voorgestelde principe te demonstreren en niet om meteen de ultieme performantie na te streven. De aandacht ging dus specifiek naar het realiseren van een tolerant ontwerp.

Bij dit ontwerp komt het er op aan een zo sterk mogelijke verandering van de transversale 1-D mode te bekomen door over te gaan van een niet-geoxideerde naar een geoxideerde lagenstructuur, zoals hierboven reeds vermeld. Factoren die een sterke opsluiting in de laserkern bevorderen in de niet-geoxideerde slab, werken een

expansie van de geoxideerde slabmode in de onderliggende dikke kernlaag tegen, en omgekeerd. Een gedetailleerde beschrijving van het ontwerpsproces wordt in de engelstalige tekst uitgewerkt. We beperken ons hier opnieuw tot de essentie.
De ontwerpsprocedure is een iteratief proces, waarbij eerst met de modebepaler de opsluiting van de mode in de laserkern bepaald wordt, de koppelingsefficiëntie naar vezel berekend wordt en het kritische breedtebereik, waar de eigenlijke modetransformatie plaatsvindt, nagegaan wordt. Eenmaal een geschikt locale modegedrag gerealiseerd is, wordt de evolutie van het veld doorheen de structuur gesimuleerd met behulp van het *Fimmprop-3D* mode-expansie algorithme. De resultaten hiervan kunnen dan weer wijzigingen aan de lagenstructuur suggereren, wat ons weer aan het begin van de ontwerpscyclus brengt.
Er werd uitgegaan van een 100nm dikke oxidelaag. Dikkere lagen leveren mogelijks problemen op met geïnduceerde spanningen door het krimpende oxide. Dunnere lagen worden vermeden om niet in het laagdikte-afhankelijke oxidatieregime te belanden. Het aluminium-gehalte in de te oxideren laag werd gereduceerd tot 90% om een goede controle over de oxidediepte te realiseren (langere oxidatietijden).

De uitgangsdwarsdoorsnede werd geoptimaliseerd voor koppeling naar een 980nm monomodale vezel. Figuur 6.5 toont de evolutie van de koppelingsefficiëntie in functie van de breedte van de brede golfgeleider, en dit voor verschillende diktes van de dikke kernlaag. Figuur 6.6 schetst de evolutie van de koppelingsefficiëntie in functie van de etsdiepte. Uit beide grafieken werd besloten een 8μm brede uitgangsgolfgeleider te implementeren, met een 4μm dikke kern, waarbij 3μm diep geëtst wordt. We merken hierbij op dat de exacte etsdiepte totaal niet kritisch is, zoals duidelijk is uit de figuur. We merken ook op dat een dikkere kern weliswaar een hogere koppelingsefficiëntie oplevert, maar aangezien we hier enkel het concept van de component willen demonstreren en dikkere lagen speciale aandacht voor de groei vereisen, werd voor een dunnere laag geopteerd. Uiteindelijk levert deze structuur 75.5% koppeling naar vezel (1.22dB verlies). De uiteindelijke lagenstructuur is in tabel 6.1 samengevat.

Alle simulaties werden voor de eenvoud uitgevoerd voor een drooggeëtste structuur zoals geschetst in figuur 6.4. In werkelijkheid zullen natgeëtste structuren, zoals geschetst in figuur 6.3, gefabriceerd worden. Het verloop van de koppelingsefficiëntie van de natgeëtste structuur, in functie van de laterale en transversale misalignering is gegeven in figuur 6.7, zowel voor laser-vezel als voor bundelomvormer-vezel koppeling. Het koppelingsverlies voor deze natgeëtste structuur is slecht weinig hoger (1.4dB of 72.5% efficiëntie). Voor de laser-vezel koppeling vinden we 6.25dB koppelingsverlies, of 23.7% efficiëntie. De -1dB aligneringstoleranties nemen slechts marginaal toe.

6.3.2 Longitudinaal ontwerp

Er werd gekozen voor een robuust longitudinaal ontwerp, meer bepaald een stuksgewijs lineair verloop dat voor verschillende lagenstrukturen en eveneens voor lan-

gere golflengtes (zie verder) dienst kan doen . De eerste sectie kan snel versmald worden. Figuur 6.8 toont het verloop van het vermogen in de fundamentele mode voor verschillende hellingen van het tweede deel. Figuur 6.9 toont de evolutie van het veld (bovenaanzicht en zijaanzicht) in een 475μm lange component. Het uiteindelijke longitudinale verloop wordt geschetst in figuur 6.10. Bij het maken van dit ontwerp werd ook een sensitiviteitsanalyse uitgevoerd, waarbij er op gelet werd dat de component nog correct functioneert bij afwijkingen op de laagdiktes en/of brekingsindex(contrasten). Voor meer details omtrent het ontwerp verwijzen we naar de engelstalige tekst.

6.3.3 Maskerontwerp

Figuur 6.11 geeft een overzicht van de ontworpen maskerset, die verschillende groepen lasers met bundelexpansiestructuren definieert. Het verschil tussen die vier groepen is de helling van het tweede, traag versmallende deel van de bundelomvormer. Daarnaast zijn ook referentielasers zijn voorzien.
Er werd gekozen voor een actieve uitvoering van de bundelomvormers, om eventuele absorptieverliezen in het niet-gepompte deel van de component te vermijden. Dit leidt uiteraard tot hogere drempelstromen, maar eveneens tot een robuuster ontwerp.

6.4 Fabricage

De fabricage bestaat uit een 3μm diepe natte etsstap in een H_2SO_4:H_2O_2:H_2O (1:1:8) oplossing, gevolgd door een natte selectieve oxidatie aan 410^{o}C. Door de onderets, eigen aan de natte etsstap, volstaat 1.5μm oxide om een 2μm brede lasergolfgeleider te definiëren. Hiervoor is een 45 minuten lange oxidatie vereist. Figuur 6.12 toont een electronenmicroscoop-foto, waarop duidelijk het oxide zichtbaar is. Na het oxideren volgt dan nog een planarisatie met polyimide, waar openingen in geëtst worden als contactopeningen. Daarna volgt nog de metallisatie, het electrochemisch aandikken van het metaal, het verdunnen van het substraat en het aanbrengen van het achtercontact.

6.5 Karakerisering

Verschillende metingen werden uitgevoerd om de verschillende aspecten van de component te karakteriseren. Hiervoor werden 800μm lange referentielasers gemonteerd en lasers met geïntegreerde bundelomvormers. Deze laatsten bevatten eveneens een 800μm lange lasersectie en bundelomvormers met verschillende lengtes.

Figuur 6.13 toont enkelerepresentatieve PI-curves. De hogere drempelstromen van de mode-expansielasers zijn te verklaren door de actieve uitvoering van de bundelomvormers. De externe efficiëntie van de kortste componenten is iets hoger

	Laser-Vezel	
	Simulatie	**Meting**
Kopp. Eff.	23.7% (-6.25dB)	23.2% (-6.34dB)
-1dB Align. Tol. (LxT)	2.6 x 2.55μm	2.35 x 2.8μm
	Bundelomvormer-Vezel	
	Simulatie	**Meting**
Kopp. Eff.	72.5% (-1.4dB)	71% (-1.49dB)
-1dB Align. Tol. (LxT)	3.1 x 2.7μm	3 x 3.2μm

Tabel 6.3: Vergelijking tussen de gesimuleerde en gemeten koppelingsefficiënties en -1dB toleranties van een laser-vezel en een bundelomvormer-laser koppeling.

dan voor de referentielasers (23%/facet tegenover 21%/facet voor de referentielasers). De lagere efficiëntie van de langste component is te verklaren door het feit dat de mode hier langer doorheen het slecht gepompte gebied propageert (bij heel smalle oxidevensters), waar de mode nog niet voldoende is uitgezakt.

De gemeten en gesimuleerde laterale en transversale verre velden worden in figuur 6.14 weergegeven. We merken een heel goede overeenkomst tussen meting en simulatie, waarbij het verre veld efficiënt wordt herleid van 21^ox23^o (LxT, volle breedte bij half maximum-waarden, FWHM-waarden) voor de referentielasers tot 7.5^ox13.5^o voor de bundelexpansiestructuren. Dit demonstreert meteen het succes van het voorgestelde mode-expansieprincipe.

Ook de stabiliteit van het verre veld werd gecontroleerd en is voorgesteld in figuur 6.15. We merken een perfecte stabiliteit van de bundel voor stromen tot minstens drie maal de drempelwaarde.

Er werd een geautomatiseerde opstelling gebouwd om de koppelingsefficiëntie naar vezel en de aligneringstoleranties op te meten. Hierbij werd de positie van de vezel afgescand in een vlak evenwijdig aan het facet. Figuur 6.16 toont het resultaat voor een referentielaser-vezel koppeling en voor een bundelomvormer-vezel koppeling (de afstand tussen twee contouren komt overeen met een koppelingsverschil van 5%). De resultaten van de koppelingsmetingen worden in tabel 6.3 samengevat. We zien een verrassend goede overeenkomst tussen gemeten en gesimuleerde waarden. De koppelingsefficiënties werden gecorrigeerd voor een Fresnel reflectieverlies aan het vezelfacet van 4%.

In figuur 6.17(a) wordt de evolutie van het koppelingsverlies in functie van de afstand tussen chip- en vezelfacet getoond. De oscillaties op de curves zijn te wijten aan resonanties die ontstaan in de luchtcaviteit tussen beide facetten. Het koppelingsverlies neemt sterker toe voor de laser-vezel configuratie. Dit is duidelijker zichtbaar door het verschil tussen laser-vezel en bundelomvormer-vezel curves te beschouwen, zoals in figuur 6.17(b) is weergegeven.

6.6 Reproduceerbaarheid van het fabricageproces

Tot slot werd ook de reproduceerbaarheid van het fabricageproces nagegaan door een tweede reeks componenten op een identieke manier te fabriceren. De opgemeten drempelstromen, externe efficiënties en FWHM-waarden worden in figuur 6.18 vergeleken en geven een excellente reproduceerbaarheid aan.

6.7 Lange-golflengte componenten op GaAs substraten

6.7.1 Situering

Zoals reeds enkele keren opgemerkt, is het traditionele toepassingsgebied van GaAs-gebaseerde componenten in het korte golflengtegebied te situeren. Onlangs werd echter een nieuw materiaal, GaInNAs, voorgesteld, dat roosteraangepast kan groeien op GaAs-substraten en het emissiebereik van actieve GaAs-gebaseerde componenten kan uitbreiden tot 1.3μm en verder. De vele voordelen van het GaAs-systeem ten opzichte van het InP-systeem en meer uitleg over het GaInNAs materiaal worden in de engelstalige tekst gegeven.
Om toegang te krijgen tot dit nieuwe en heel interessante materiaal, hebben we meegewerkt aan twee Europese projecten, OPTIVAN en GIFT, die de ontwikkeling ervan tot doel hadden. In het kader van die projecten stonden wij in voor de karakterisatie van elementaire structuren. De resultaten van die metingen werden dan naar de partners, die het groeiproces optimaliseerden, teruggestuurd voor een verdere bijsturing van hun proces.
Eveneens waren wij verantwoordelijk voor het ontwerpen van geoxideerde bundelomvormers voor golflengts tot 1.3μm.

6.7.2 Karakterisering van GaInNAs-gebaseerde componenten

Voor een gedetailleerde bespreking van de metingen wordt opnieuw naar de engelstalige tekst verwezen. We vermelden hier enkel dat we aan de hand PI- en IV-metingen en het opmeten van spontane emissiecurves de volgende parameters bepaald hebben: drempelstroom, externe quantumefficiëntie, karakteristieke temperatuur, inwendige efficiëntie, junctiespanning, serieweerstand, winstcurves bij verschillende stromen, differentiële winst,.. Enkele opgemeten winstcurves worden in figuur 6.22 gegeven. Het verloop van de maximale winst en de piekwinstgolflengte in functie van de stroom worden in figuur 6.23(a), respectievelijk figuur 6.23(b) voorgesteld.

6.7.3 Ontwerp van lange-golflengte bundelexpansiestructuren

Er werden verschillende ontwerpen van geoxideerde bundelexpansiestructuren gemaakt, die voldoen aan de specifieke wensen van de verschillende partners. Lo-

gischerwijze ziet elke partner liefst zijn specifieke laserstructuur geïmplementeerd omdat daar reeds het meest ervaring mee is opgedaan en het groeiproces er ook voor gekarakteriseerd is. Enkele voorbeelden worden in tabel 6.3-6.5 gegeven.
Jammer genoeg zijn we er niet in geslaagd werkende structuren te fabriceren uit de weinige plakken materiaal die tot nog toe werden geleverd. De problemen zijn echter niet gerelateerd aan het principe of het ontwerp van de bundelomvormers, maar wel aan de kwaliteit van het materiaal (dat nog volop in ontwikkeling is!).
Toch kunnen we met een postieve opmerking besluiten dat de componenten met uiteenlopende materiaalsamenstellingen en lagenstructuren goed blijken te werken in de simulaties, wat dan ook aangeeft dat het geoxideerde bundelomvormerconcept een breed toepassingsgebied kent.

7. Besluiten en perspectieven

In dit hoofdstuk worden de behaalde resultaten van dit eindwerk nog eens kort samengevat. Centraal stond het ontwikkelen van bundelexpansiestructuren voor het realiseren van een efficiënte vezel-chip koppeling. Twee nieuwe principes werden met succes gedemonstreerd: het gebruik van 'ARROW'-opsluiting in InP-gebaseerde componenten, wat tot een octrooi-aanvraag leidde, en toepassing van de selectieve natte oxidatie voor GaAs-gebaseerde mode-expansie structuren. Het hoofddoel hierbij was het bekomen van eenvoudige fabricageschema's.

7.1 Besluiten

In hoofdstuk 1 werd de evolutie naar volledig optische netwerken, waarbij de nood aan PICs ontstaat, besproken. De problematische koppeling naar vezel, die leidt tot een hoge kostprijs, staat een echte doorbraak van PICs echter in de weg. Geïntegreerde bundelexpansiestructuren werden aangewezen als de beste oplossing voor dit probleem. Deze oplossing vereist echter meestal een complex fabricageschema. Hoofdstuk 2 gaf een overzicht van de verschillende concepten voor dergelijke componenten. Het overgrote deel van deze structuren steunt op een 'adiabatisch' werkingsprincipe, dat verder theoretisch uitgediept werd in hoofdstuk 3. Hoofdstuk 4 vatte het werk op InP-gebaseerde componenten samen. Een overgroeide structuur, voor integratie in een PIC die zorgt voor het combineren van de optische signalen in een 'phased array'-antennesysteem, werd ontworpen bij de voor InP ongebruikelijke golflengte van 1.06μm. Koppelingsefficiënties van 2.8dB werden gedemonstreerd, wat een verbetering van de vezel-naar-vezel verliezen van 7dB oplevert.
Verder werd het gebruik van het 'ARROW'-concept voorgesteld voor de optische opsluiting van de vezel-aangepaste mode. Dit leidt tot een belangrijke vereenvoudiging van het fabricageschema en resulteerde in een octrooi-aanvraag. Door de 'ARROW'-opsluiting ook lateraal toe te passen, wordt het meest eenvoudige schema, ooit gerealiseerd, bekomen: het fabricageschema dat bestaat uit een enkele planaire epitaxiale groeistap en een enkele planaire ets, wordt niet beïnvloed door een modetransformator te integreren.
We demonstreerden ook het gebruik van intra-chip modetransformatoren, die de geleide mode van een monomodale golfgeleider adiabatisch omvormen tot de fundamentele mode van een brede, multimodale golfgeleider. De fundamentele mode

in de brede golfgeleider propageert met veel lagere verstrooiingsverliezen, zodat op deze manier laag-verlies interconnecties kunnen gerealiseerd worden. De verliezen van diep-geëtste structuren werden zo gereduceerd van 5dB/cm voor monomodale golfgeleiders, tot 0.15dB/cm voor 10μm brede structuren. De bundelexpansiestructuren vertoonden minder dan 0.05dB modetransformatieverlies.
In hoofdstuk 5 werd, als inleiding op hoofdstuk 6, het selectieve natte oxidatieproces bij GaAs-gebaseerde componenten besproken. Hoofdstuk 6 behandelt het door ons voorgestelde concept van de geoxideerde bundelexpansiestructuur. Hierbij worden de eigenschappen van het oxide (de lage brekingsindex en de elektrisch isolatie) ten volle benut om het modale gedrag in de component te dirigeren en een geëxpandeerde bundel te verkrijgen aan het uitgangsfacet. De fabricage behelst enkel een planaire epitaxiaalgroei en een niet-kritische conventionele ets. De gemeten koppelingsefficiëntie naar vezel (71.2%, of 1.49dB verlies) stemt, net als de alignatietoleranties (3.1x2.7μm, LxT) heel goed overeen met de theoretisch berekende waarden. De hoge efficiëntie van de laser is een aanwijzing voor de lage transformatieverliezen in de bundelomvormer. De reproduceerbaarheid van het fabricageschema werd geverifieerd. Verder werd ook het werk dat uitgevoerd werd in het kader van de Europese projecten OPTIVAN en GIFT, die de ontwikkeling van het nieuwe GaInNAs-materiaal beogen, samengevat. Zo werd de karakterisering van GaInNAs-gebaseerde componenten besproken, en werden verschillende ontwerpen van geoxideerde lasers met geïntegreerde bundelomvormers en GaInNAs actieve lagen bij 1.3μm behandeld.

7.2 Perspectieven

Het werk op de 'ARROW'-gebaseerde bundelomvormers van het rib-type heeft uiteindelijk ook geleid tot het introduceren van dit concept in andere toepassingen zoals overgroeide modetransformatoren en voor het reduceren van de stralingsverliezen in gebogen golfgeleiders. Dit wordt momenteel verder onderzocht in het doctoraatswerk van Marko Galarza.
Er werden verschillende geoxideerde bundelomvormers ontworpen voor GaInNAs-structuren bij 1.3μm. Deze structuren worden momenteel nog gefabriceerd en zullen in de nabije toekomst uitgemeten worden.
Het potentieel van de selectieve oxidatie wordt nog steeds verder ontgonnen. We zijn er van overtuigd dat het gebruik van oxide voor het definiëren van geoxideerde passieve golfgeleiders (analoog aan de definitie van de lasergolfgeleider in de geoxideerde laser met bundelomvormer) kan leiden tot een aanzienlijke reductie van de verstrooiingsverliezen, doordat ruwheden in de golfgeleiderwand uitgevlakt worden in het oxidefront. Dit concept, aangevuld met het gebruik van bundelomvormers om op de chip een overgang te maken tussen een actieve en passieve lagenstructuur, kan leiden tot de ontwikkeling van een GaAs-gebaseerde integratietechnologie voor telecommunicatietoepassingen, met alle voordelen van dien.

English Text

Chapter 1

Introduction

1.1 Optical communications: a short history

At the moment optical fibre systems are being installed massively all over the world to meet the ever increasing demand for bandwidth. Optical communications has known a long history of crucial inventions and technical innovations to bring the technology to what it is today and a lot more of its potential still has to be explored.

Already in the 1950's, telecommunications engineers were seeking for more transmission bandwidth since radio and microwave frequencies were already in heavy use at those days. They looked at higher frequencies to carry loads they expected to continue increasing with the growth of telephone and television traffic. Alec Reeves, a forward-looking engineer at **S**tandard **T**elecommunications **L**aboratories (STL), proposed to use the optical frequencies, but it was only in 1960, when Maiman demonstrated the first laser, that others started to concentrate on the optical domain. However, the ruby laser only operated in pulsed regime and serious work on optical communication had to wait some months for the continuous-wave helium-neon laser.

While air is highly transparent at optical wavelengths, researchers soon found that rain, haze, clouds and atmospheric turbulence limited the reliability of long-distance atmospheric laser links. Optical waveguides would be essential and optical fibres had attracted the attention because they were analogous in theory to dielectric waveguides used in microwave applications. Optical fibre bundles were already being used as endoscopes in the medical sector, but they were much too lossy for communications. Attenuation of a decibel per meter was fine for looking inside the body, but telecommunications operated over much longer distances and required losses no more than 10 to 20 decibel per kilometer.

This made almost everyone to abandon the path of optical fibre, except for Charles Kao at STL. He investigated the properties of optical fibre thoroughly and came to the conclusion that the high losses of early fibres were due to im-

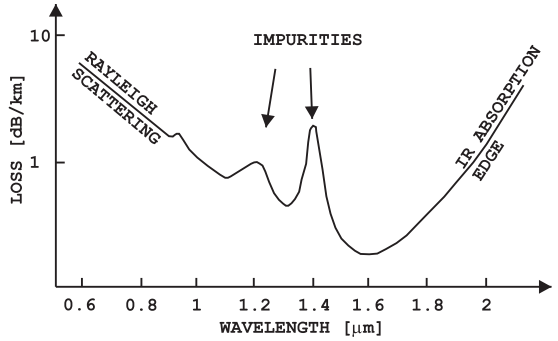


Figure 1.1: Single mode fibre loss

purities and not to the silica glass itself. In an article in 1966 he forecasted that fibre loss could be reduced below 20dB/km. In the same publication he also described a proposal for long-distance communications over single-mode fibres. This attracted the interest of the British Post Office that saw the possibilities of fibre-optic communication systems. Kao's work and the interest of the British Post Office triggered several laboratories around the world to try to reduce the fibre losses. It took four years before the goal of 20dB/km was finally reached at Corning Glass Works (at the 633nm helium-neon line). This was one of the most important of many developments that opened the door to fibre-optic communications.

In the same year, Bell Labs and the Ioffe Physical Institute made the first semiconductor laser diodes able to emit continuous wave at room temperature. This enabled the use of lasers to transmit optically encoded telephone conversations over fibre-optic cables. Over the next several years, fibre losses dropped dramatically, aided both by improved fabrication methods and by the shift to longer wavelengths where fibres have inherently lower attenuation (figure 1.1).

A first generation of telephone field trials in 1977 used graded-index multimode fibres in combination with 850nm AlGaAs-laser diodes in point-to-point links. Multi-mode fibre was preferred at that time because developers were bothered by the small single-mode fibre core size of only a few micrometers. In a multi-mode fibre the core is large (50 to 85 μm) and light propagates in the form of multiple modes, which travel all at a slightly different velocity. This modal dispersion causes spreading of the optical pulses. To overcome this signal degradation and the attenuation of approximately 2dB/km, regenerative amplifiers or *'repeaters'* were needed every few kilometers. Regenerative repeaters are complex opto-electronic devices that convert the weak optical signal to a weak electronic signal, which is amplified, retimed, reshaped and used to drive a semiconductor laser that launches the original, undistorted optical signal at full intensity on the next span of cable.

Shortly after these first transmission experiments new InGaAsP lasers which

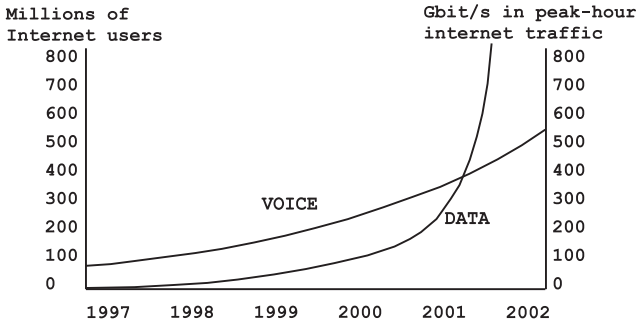


Figure 1.2: Evolution of the Internet traffic

emitted at 1.3μm became available and the low fibre loss at this wavelength (0.5dB/km) renewed the interest in single-mode fibre. This should be compared with copper coaxial cable, whose loss is 2.5dB/km at 1MHz and 50dB/km at 1GHz. AT&T, British Post Office and STL committed to developing a single-mode transatlantic fibre cable using the new 1.3μm window. Since light is guided in a single mode, modal dispersion is ruled out and repeaters were needed only every 40-50km to overcome chromatic and material dispersion and fibre attenuation. This new cable, which became operational in 1988, was able to handle double the traffic of all the trans-Atlantic copper cable predecessors combined. The cable contained two pairs of single-mode fibre (two fibres for each direction), each carrying signals at 280Mbit/s, the combined equivalent of 35000 phone calls. The limiting factor in the number of fibre pairs implemented in submarine cables is the repeaters. These are very expensive devices that make up a substantial part of the total cost of the cable system since they have to guarantee a reliable operation in harsh submarine conditions for at least 25 years. With the advent of the 1.55μm laser technology, new trans-Atlantic optic cable systems could increase repeater spans to more than 100km because of the lower fibre attenuation (only 0.2-0.3dB/km) at this wavelength.

At the same time also more and more terrestrial fibres were being installed as point-to-point backbone links between large cities and major traffic-generating regions. However, in spite of the enormous increase in bandwidth, the link capacity was exhausted much sooner than anticipated. This was due to a serious underestimation of the bandwidth usage as a result of an upsurge of new applications. The major catalyst for this evolution is of course the Internet, which generates an amount of traffic that doubles approximately every year (figure 1.2). Not only is the number of Internet-users growing rapidly, also the number of applications is increasing. E-mail (attachments!), data transport (ftp, grid computing [1],..), audio (mp3, internet radio), video,... are all booming applications. It is expected that in the future, when new access technologies like DSL (**D**igital **S**ubscriber **L**ine) and cable modems bring more bandwidth to the end-user, new applications will emerge that will push the bandwidth demand even higher.

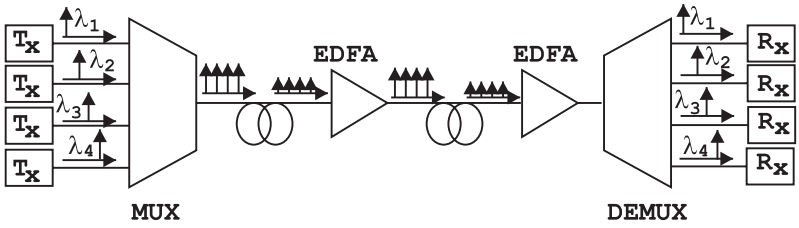


Figure 1.3: Typical configuration of a WDM point-to-point link.

There are different ways to obtain larger link capacities. The most straightforward but at the same time costly and labour-intensive solution is to install more fibre, also called **S**pace **D**ivision **M**ultiplexing (SDM). A second option is to migrate to higher transmission speeds, following the standardized SONET optical channel rates, from 2.5Gbit/s to 10Gbit/s to 40Gbit/s and even further to 160Gbit/s. These higher speeds are obtained by **T**ime **D**ivision **M**ultiplexing (TDM), where the bits of different tributary signals are combined into a single high-bitrate signal. Presently available systems allow multiplexing up to speeds of 40Gbit/s (10Gbit/s in commercial systems). At higher bit-rates, the signals have to be multiplexed in the optical domain. Optical time domain multiplexing (OTDM) is obviously more challenging, which explains why OTDM systems have so far not made it to real-life applications. The present record in the laboratory for OTDM stands at 320Gbit/s [2], which is still only a small fraction of the 25THz bandwidth that is available in the optical fibre. A better exploitation of the immense fibre bandwidth is obtained by the use of Wavelength Division Multiplexing (WDM) (figure 1.3). In WDM, the low bit-rate signals are first time division multiplexed to a certain level (2.5Gbit/s or 10Gbit/s in present commercial systems). Subsequently, each TDM signal is modulated onto an optical carrier wave with a different wavelength and fed into a single-mode fibre. At the receiver side the signal is demultiplexed again and each channel is converted back to an electrical signal. A WDM system thus provides many parallel data channels on a single optical fibre.

Although the principle was already known for a long time, the breakthrough for WDM-systems was only made after the invention of the **E**rbium **D**oped **F**ibre **A**mplifier (EDFA). This all-optical amplifier (meaning that there is no conversion to intermediate electrical signals) is capable of amplifying many signals simultaneously, regardless of their modulation scheme and rate (so-called *'transparancy'*). Amplification is obtained by mixing a high-powered 'pump laser' operating at a wavelength of 980nm or 1.48μm with the data traffic and feeding the combined signal into the Er-doped fibre. The pump-light excites the Er-ions, which decay rapidly to a lower metastable state. Photons entering the pumped region with a wavelength between 1530nm and 1565nm, often called the conventional band or 'C-band', stimulate the excited ions to fall back to their lower energy level, generating new photons with the same wavelength and phase.

Recent progress in the development of optical amplifiers has stretched the wavelength range even further with the Raman amplification process. Raman amplification uses the glass fibre as active medium by injecting a pump beam, which excites fibre molecules to higher vibrational states. In glass, with its extended structure of small molecular units bonded together with no long range order, there is an infinite number of local structural arrangements and these all affect the local energy of the structure and the vibrational energies. This leads to a broad continuum of vibrational energy states between which absorption and recombination can occur. Amplification of a signal occurs when a signal is transmitted with a frequency which falls within the Raman scattering spectrum of the pump laser. Raman amplifiers in principle allow amplification at any frequency of the 25THz broad window of the fibre and opens the door to a further extension of WDM-wavelengths into the long-wavelength band ('L-band' from 1565nm to 1625nm).

The **I**nternational **T**elecommunication **U**nion (ITU) standards body has standardized the optical spectrum and assigned specific wavelengths to the different optical channels. A regularly spaced frequency grid, anchored at 193.1THz (about 1.5525μm wavelength), with a channel spacing of 100 or 50GHz, was specified [3]. Present-day commercially available systems use the entire C-band. For instance, in 2001 a transatlantic cable has been put in service offering 4.8Tb/s by implementing 6 fibre pairs, each carrying 80 wavelengths at 10Gbit/s per channel. To keep up with capacity demands, the number of channels will probably have to be increased even further, both by reducing the channel spacing and by extending the frequency band (e.g. by using the L-band). This has recently been demonstrated in the laboratory by Alcatel. In November 2001 they set a new world record for single fibre capacity. They demonstrated the transmission of 365 wavelengths, each operating at 10Gbit/s. The aggregate 3.65Tbit/s was sent over 6850km, a distance comparable to most transatlantic cable links. The channel spacing was reduced to 22.2GHz, less than half of the 50GHz standard with an optical bandwidth (66nm) reaching into the L-band.

Figure 1.4 sketches the expected evolution of WDM-networks. The first generation of WDM-systems were used as point-to-point connections, but soon different optical links were interconnected and real optical *networks* came into being with mesh and ring topologies as its basic architectural building blocks. With ever increasing bandwidth capacity at the core of the network and higher-bandwidth applications originating from the user, a serious amount of pressure has been placed on the metropolitan and access networks. As a consequence, the need to enhance capability and performance is driving networks towards end-to-end optical connectivity. In other words, optical fibre is getting closer to our homes.
In the first generation networks all switching, routing and processing was done electronically. The 'slow' electronics formed the bottleneck in these systems, generating delays and reducing the network's quality of service. In the next generation networks they will be replaced by 'all-optical' components, which form the subject of current research in photonics. These devices will avoid the optic-electronic-optic translation and electronic signal processing. All-optical reconfigurable add/drop

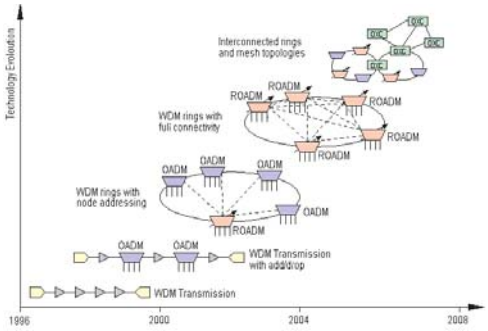


Figure 1.4: Evolution scenario of WDM-networks (from [4]).

multiplexers and cross-connects will guide signals through the network based on the wavelength of the specific channel. The use of wavelength converters will result in a more flexible and efficient use of available bandwidth throughout the network. In a more remote future optical routing might even become possible, although some high hurdles still have to be taken (e.g. resolving the problem of optical buffering).
Clearly, a serious amount of research is needed to move into this higher level of optical functionality.

1.2 Photonic Integrated Circuits

The increasing level of optical functionality in network nodes, resulting from shifting functions from the electrical to the optical domain, has created a need for complex optical devices or **P**hotonic **I**ntegrated **C**ircuits (PICs). The development of PICs is a challenging task which consists of assembling a variety of different basic opto-electronic components into a single module in a hybrid or a monolithic manner. PICs have so far not found a widespread use because of performance limitations. One major performance limitation is associated with the connection loss between weakly guiding circular single mode fibres and strongly guiding planar deposited waveguides. The problematic coupling, requiring sub-micron alignment accuracies, forces the use of time-consuming active alignment schemes that cause the packaging cost to amount up to about 80% of the total module cost. The coupling problem and common ways to circumvent it are discussed in detail in paragraph 1.3.

1.2.1 Active and passive devices

The sub-components that build up a PIC can be either 'passive' or 'active' structures. Passive devices are simple waveguiding structures like straight and bent waveguides or an arrangement of them as e.g. in phased array (de)multiplexers.

The band gap of the materials in passive devices is typically larger than the photon energy of the signal, making them transparent to the guided light. Active components on the other hand mostly incorporate a layer of material with a bandgap that equals the photon energy and they always involve an extra input terminal to the device. This extra input can be a current, a voltage, another optical signal or a temperature control. For example, in an optical amplifier or in a laser, which is basically an amplifier placed in a cavity, this additional input is a current. The optical signal is amplified through stimulated emission by injecting current through a layer of material with a band gap that equals the photon energy of the signal. Electro-optic modulators are controlled by an external voltage that changes the refractive index of the semiconductor. This way an extra phase shift is introduced for the guided mode. Such phase modulators can be arranged in a Mach-Zehnder configuration to form an optical switch. Another example of a voltage driven active device is an electro-absorber, which can be used as an external modulator for high bit-rate laser modulation (>10Gbit/s). By switching the voltage across the semiconductor the material changes from absorptive to transparent since the presence of a voltage also influences the imaginary part of the refractive index or the material absorption. An optically pumped laser is an example of an active device with an additional optical input and a thermal control is used in thermo-optic phase modulators, where a tuning of the device temperature results in a change of the refractive index of the material as for instance in polymeric and silica-on silicon phase modulators.

1.2.2 Material systems

Unlike integrated electronics, where almost all devices are fabricated in silicon, there exists no 'ideal' material for integrated optics. The most important material systems are semiconductors, polymers, silica, ferro-electric materials and glass optical fibre.

Semiconductors The most promising semiconductor material for opto-electronic telecommunication purposes is InP with the ternary and quaternary alloys InGaAs and InGaAsP or InAlGaAs, which can be epitaxially grown lattice-matched on an InP substrate. By changing the composition of the quaternary, the bandgap can be tuned anywhere between 0.92μm and 1.65μm. This offers the possibility to provide active and passive functions for any wavelength in the telecommunication window giving the material system a strong potential for monolithic integration. As the photon energy is close to the bandgap, electro-optical effects are relatively strong. InP-based passive devices are typically an order of magnitude smaller than silica or polymer waveguides because of the high index contrast that allows smaller bend radii. The losses, on the other hand, are higher (typ. 1dB/cm) but still low enough for optical devices. A wide variety of components and PICs has already been demonstrated. Some examples are DFB lasers, SOAs, (de)multiplexers, multi-wavelength lasers, switches, wavelength converters, modulators, detectors,.. Another extensively studied semiconductor system is GaAs/AlGaAs/In(Al)GaAs,

which is mainly used for the fabrication of edge emitting lasers and Vertical Cavity Surface Emitting Laser (VCSELs) operating between 700nm and 1100nm. They are used as lasers for compact disk players, bar-code scanners or transmitters to send moderate bit-rates over short distances (2-3 kilometer) of multimode fibre. A lot of research is currently being done to extend the domain of GaAs-based emitters towards longer wavelengths because of the advantages this material system has to offer in comparison with InP-based devices (substrates are larger and cheaper, better temperature characteristics,..). This effort has led to the development of GaInNAs, a new quaternary material that can be grown lattice matched on GaAs substrates, with a shift in emission wavelength to 1.3μm and beyond. 1.3μm edge emitters and VCSELs with superior temperature characteristics as compared to InP-based devices have been demonstrated. An interesting feature of the GaAs-system is the possibility to selectively oxidize AlGaAs layers with a high Al-content. The oxide is an electrical isolator and has a very low refractive index. These oxide properties can for instance be exploited to form current apertures which are at the same time optical apertures. A detailed discussion on the selective wet oxidation of AlGaAs-layers and its applications will be given in chapter 5.

Polymers Polymers have attracted considerable interest through the development of Plastic Optical Fibre (POF), but they are also attractive for planar waveguiding by their low polarization dependence and insertion loss (0.05 to 0.5dB/cm depending upon optical contrast). The latter is mainly due to the fact that the index contrasts are very similar to single mode glass fibre and hence the waveguides can be designed to closely match the fibre mode profile. The waveguides are fabricated by successive low-temperature spin-coating of a polymer buffer layer and a core layer with a higher refractive index on a silicon substrate. Next, the core is etched using conventional etching techniques like wet etching or reactive ion etching and finally an upper cladding layer is spun. Losses are typically about 0.01dB/cm at 840nm and 0.1dB/cm at 1550nm, making them suitable for passive optical components like (de)multiplexers or splitters. The most successful polymer-on-silicon applications are the thermo-optic switches. The coexistence of a large thermo-optic coefficient (dn/dT) and a low thermal conductivity makes that the required heat or switching power is very low. The drawbacks are the rather large switching times (2 ms) and the limits on large-scale integration due to the low contrast and hence the large bending radii. Another disadvantage is the fact that active devices like lasers or detectors cannot be implemented.

Silica SiO_xN_y-based waveguides with low losses (0.02dB/cm at 1550nm) and mode profiles that are well-matched to single mode optical fibre are currently the most used type of waveguide for passive devices. They are fabricated by a combination of flame hydrolysis deposition on silicon or silica (Si_3N_4) substrates and reactive ion etching. Thermo-optic switching is much less efficient with regard to power consumption than in polymer-based devices because of a seven times higher thermal conductivity and a low thermo-optic coefficient. Doping of the silicate

glasses with rare-earth erbium ions and optical pumping at 980nm or 1480nm provides limited gain around 1530-1550nm. High gain, however, can not be obtained.

$LiNbO_3$ The ferro-electric material $LiNbO_3$ is known for its strong acousto-optic and (non-linear) electro-optic properties. This makes the material system especially interesting for the fabrication of external modulators as already available from several vendors. Waveguides are formed by diffusing Ti into the crystal and doping with Er-ions provides a gain that is substantially higher than in silica. The drawbacks are the high material cost and the difficult processing.

Fibres Glass optical fibres have surpassed the stadium of being a mere passive transmission medium. The emerging availability of broadband fibre amplifiers –EDFAs and Raman amplifiers– were essential in the breakthrough of WDM networks as explained in paragraph 1.1. Also a wide variety of passive devices are already commercially available. Fibre Bragg gratings are used to compensate the chromatic dispersion on long-haul backbone links or to flatten the gain curve of EDFAs. WDM (de)multiplexers are fabricated from a concatenation of fused fibre couplers, if necessary supplemented with fibre bragg gratings to enhance the channel isolation up to -40dBm.

1.2.3 Hybrid versus monolithic integration

There have been several hybrid and monolithic approaches to the integration problem, but a widespread use of opto-electronic components is still restricted by their high cost, mainly caused by the problematic coupling to single mode fibres.
In *hybrid* integration schemes the different sub-components are fabricated separately and then assembled on a carrier. This implies that each sub-component can be fabricated in the material system that is best suited for its specific functionality and that devices can be selected prior to integration, improving the module yield. On the other hand the assembly on a common carrier is found to be a tough task requiring an accurate and stable fixation of the sub-components. The alignment tolerances are extremely tight as a consequence of the small optical mode sizes. Submicron tolerances are mandatory in many cases because of the steep decrease of the coupling efficiency as the components are further misaligned. This is particularly true when coupling from a strongly confined semiconductor laser to a single mode optical fibre. A profound discussion of this problem and an overview of common techniques to obtain a more efficient and/or more tolerant coupling will be given in section 1.3.
The problematic nature of the coupling necessitates the use of an active alignment technique in most cases. Active alignment means that the amount of light coupled from one device into another is monitored during the alignment of the two devices. This can be done with continuous measurement, or on a sequential basis. When the coupling is finally optimized, the assembly is fixed with e.g an epoxy, UV-curable glue, solder or laser welding. The advantages of this method are low excess loss, good loss uniformity and high yield, but the disadvantages are an extremely time consuming alignment procedure and a high cost for labour,

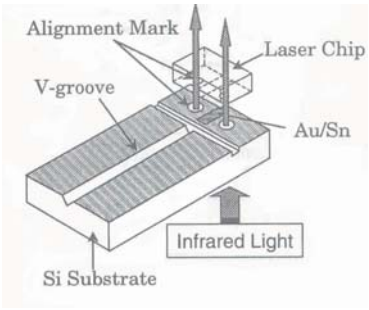


Figure 1.5: Passive alignment by means of a silicon bench.

equipment, precision-machined assembly parts and operating devices. In addition, the throughput is very low and scaling up to large-scale manufacturing is not straightforward.
Extensive research has therefore been dedicated to the development of passive alignment schemes where, as opposed to active techniques, there is no measurement of the coupling efficiency during the alignment. The components are assembled on a carrier and positioned by means of alignment marks or bumps on the carrier and mating features on the components. The accuracies obtained with most of the passive alignment schemes are in the 1-2μm range, with a potential for submicron capability with improved process control. One key enabling technology is the silicon optical bench system. Silicon as a substrate material is advantageous because of its thermal and mechanical properties and because micro-structures and alignment marks can be micro-machined by the use of photolithographic techniques. In addition, anisotropic etching of Si using e.g. potassium hydroxide (KOH) is used to form etched structures in the Si along crystalline planes to achieve very precise V-grooves to hold the optical fibre. An example is shown in figure 1.5. The passive alignment technique is simpler and less time consuming than the active alignment method, but it is hindered by a higher insertion loss, increased loss non-uniformity and a lower yield of packaged modules.

Although hybrid techniques can satisfy many near-term needs, they involve a great deal of piecework and assembly and therefore significantly increase costs and affect robustness. Before opto-electronic integrated circuits and systems can be produced in high volumes and at low cost it will be necessary to develop *monolithic* integration techniques which can be executed on full wafers and in multi-wafer lots. Monolithic integrated circuits are fabricated from a single piece of wafer by consecutive photolithographic exposures, etching steps and possibly additional growth runs. The passive waveguides connecting the sub-components are defined with the same photolithographic masks as used for the definition of the different components, thus eliminating the alignment problem and reducing the mounting and packaging cost. Of course, the critical coupling from fibre to chip remains. One of the main difficulties in developing a monolithic integration technology is to combine components with different layer structures on a single wafer. Especially the combination of active components incorporating layers with a bandgap that equals the photon energy and passive waveguides that are transparent to the guided light

is a challenging task. Several techniques, e.g. selective area growth, shadow mask growth, quantum well intermixing,.. have been developed for this purpose. For more details we refer to [5].

It should be noted that traditional pure monolithic processing is not always essential, and pseudo-monolithic integration can be acceptable. By 'pseudo-monolithic' we mean integration processes which involve monolithic processing of separate wafers and then the combining of two or more such wafers using a bonding technique followed in most cases by additional monolithic processing and testing of this new 'composite wafer' prior to dividing it into individual integrated circuit chips. The primary objective is to achieve the packing density, ruggedness and processing economics of traditional monolithic processing by doing as much processing as possible at the wafer level in batches, and to avoid handling individual circuits until the final stage of packaging.

1.3 The coupling problem

1.3.1 Fibre-chip coupling

We already briefly mentioned the severe problems encountered when coupling light from a PIC to a single mode glass fibre or vice versa. Already from the very beginning of the concept of telecommunication by means of light have the interfaces between optical component and optical fibre proven to be a critical issue. The coupling problem is difficult to tackle and results in very high optical losses with values up to 10dB being no exception. The alignment tolerances, which are a measure of the excess power loss as a function of displacement in the transverse, lateral or axial dimensions, are extremely tight and result in high module costs –packaging cost amounts up to 80% of the total cost– that hamper a further breakthrough of optical modules. Finding a solution to this problem is especially important for low-cost applications such as the emerging 10Gbit/s optical ethernet or Fiber-To-The-Home (FTTH) [6].
The origin of the problematic coupling can be found in the small mode sizes and the size and shape mismatch between the on-chip mode and the fibre mode. The chip mode is typically small and asymmetric because of the high refractive index contrast and the planar nature of these devices. Such a strongly confined mode ensures compact modules through the use of short bending radii. A small optical field profile is also desired in active devices where it is beneficial for the device performance to confine the optical power in an as small as possible active volume. Typical mode sizes of devices operating at a wavelength of 1.55μm are in the order of 1.5μm x 2-3μm, which is substantially different from the 8-9μm large circular mode of the weakly guiding fibre.

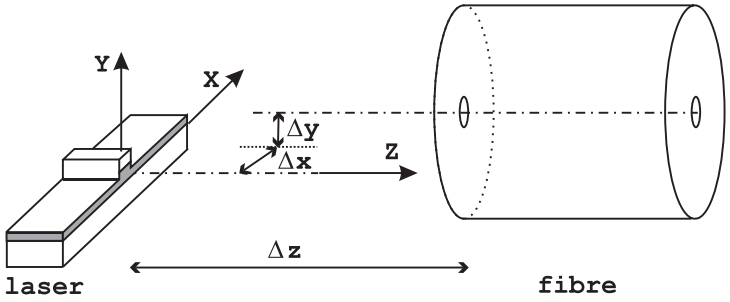


Figure 1.6: Coupling configuration.

1.3.2 Coupling efficiency and alignment tolerances

We now derive expressions for the coupling efficiency and the alignment tolerances between a PIC and an optical fibre in a configuration as sketched in figure 1.6. The mode emitted at the laser facet diffracts in the free-space gap as it propagates towards the fibre facet. The power coupling efficiency η between chip and fibre is determined by the projection of the incident field on the fibre mode[1]. This projection is expressed by the overlap integral[2] [7]

$$\eta = \frac{\mid \int_{S_f} \mathbf{E}_{ct}(x,y) \cdot \mathbf{E}^*_{ft}(x,y) dS \mid^2}{\int_{S_c} \mid \mathbf{E}_{ct}(x,y) \mid^2 dS \cdot \int_{S_f} \mid \mathbf{E}_{ft}(x,y) \mid^2 dS} \tag{1.1}$$

with $\mathbf{E}_{ct}$ and $\mathbf{E}_{ft}$ the transverse component of the chip and fibre electric fields respectively and S_c and S_f the planes of integration as shown in figure 1.6. (1.1) takes on the following more elegant form when the normalized chip and fibre modes $\widehat{\mathbf{e}}_c$ and $\widehat{\mathbf{e}}_f$ are considered:

$$\eta = \left| \int_{S_f} \widehat{\mathbf{e}}_{ct}(x,y) \cdot \widehat{\mathbf{e}}^*_{ft}(x,y) dS \right|^2 \tag{1.2}$$

The evaluation of the overlap integral is preferably accomplished in closed form, but this requires the exact modal field expressions. For this reason it is convenient to approximate both modes by Gaussian distributions ψ_c and ψ_f:

$$\widehat{\mathbf{e}}_{ct}(x,y,0) \approx \psi_c(x,y,0) = \frac{\sqrt{2/\pi}}{\sqrt{w_{x0} w_{y0}}} \, e^{-\left(\frac{x^2}{w_{x0}^2} + \frac{y^2}{w_{y0}^2}\right)} \tag{1.3}$$

$$\widehat{\mathbf{e}}_{ft}(x,y,\Delta z) \approx \psi_f(x,y,\Delta z) = \frac{\sqrt{2/\pi}}{w_0} \, e^{-(\frac{x^2+y^2}{w_0^2})} \tag{1.4}$$

with $2w_{x0}$ and $2w_{y0}$ the horizontal and vertical on-chip spot-sizes and $2w_0$ the spot-size of the circular fibre mode. For crude purposes a two-point fit at the 1/e

[1] Fresnel reflection loss will be ignored for the moment.

[2] This formulation of the overlap integral is exact in uniform media and forms an accurate approximation for weakly guiding waveguides like optical fibres.

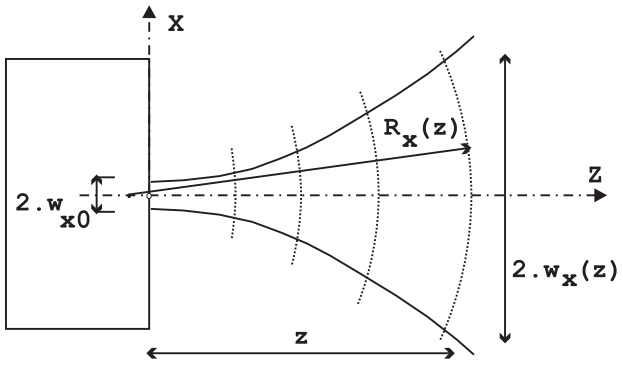


Figure 1.7: Diffraction of a Gaussian beam.

intensity points or at the half-power points may suffice, but a better choice is usually that Gaussian which has the maximum overlap integral with the actual mode profile. The approximation by Gaussian fields yields simple closed-form solutions for the coupling loss as a function of chip mode eccentricity and alignment errors, but the exact fibre mode is a Bessel function and planar waveguide modes tend more to an exponential form, especially when thin core layers are used. However, the approximation is very accurate in most cases and the qualitative conclusions drawn from the Gaussian approximation remain valid for the exact fields.

The field $\widehat{\mathbf{e}}_c(x, y, \Delta z)$ impinging on the optical fibre and appearing in the overlap integral (1.2) is, as mentioned before, the diffracted field of $\widehat{\mathbf{e}}_c(x, y, 0)$ emitted at the chip facet. The propagation, or diffraction, of a Gaussian beam through a uniform medium has a known solution and the field at a position z from the chip facet is given by [8]

$$\psi_c(x,y,z) = \frac{\sqrt{2/\pi}}{\sqrt{w_x(z)w_y(z)}}\, e^{-\frac{x^2}{w_x^2(z)} - \frac{y^2}{w_y^2(z)}}\; e^{-\frac{jk}{2}\left[\frac{kx^2}{R_x(z)} + \frac{y^2}{R_y(z)}\right]}\; e^{-jkz-\phi(z)} \tag{1.5}$$

with

$$\begin{aligned} k &= \frac{2\pi}{\lambda} \\ w_i(z) &= w_{i0}\sqrt{1 + \frac{z^2}{z_{i0}^2}} \\ R_i(z) &= z + \frac{z_{i0}^2}{z} \\ \phi(z) &= \arctan\left(\frac{z}{z_{i0}}\right) \\ z_{i0} &= \frac{kw_{i0}^2}{2} \end{aligned}$$

The field described in (1.5) is sketched in figure 1.7. The field emerges with a flat phase front from the chip facet[3] and retains its Gaussian shape as it expands while

[3]The phase front at a laser can be curved due to gain-guiding or anti-guiding effects.

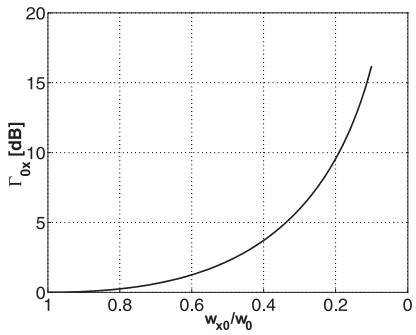


Figure 1.8: Coupling loss as a function of the beam waist ratio w_{x0}/w_0.

propagating further away from the facet and the phase front becomes curved with a radius of curvature R(z).

Now that all the fields appearing in the overlap integral (1.2) are known we can start the evaluation of the coupling efficiency[4]. Since both the fibre and the chip field can be written as products of a function of x and a function of y, the power coupling efficiency can be written likewise as $\eta = \eta_x\,\eta_y$. In case both modes are aligned ($\Delta x = \Delta y = 0$), we find for the x-z plane transmissivity [9]

$$\eta_x = \frac{2}{\sqrt{\left(\frac{w_0}{w_{x0}} + \frac{w_{x0}}{w_0}\right)^2 + \left(\frac{2\Delta z}{k w_0 w_{x0}}\right)^2}} \tag{1.6}$$

A similar expression gives the y-z plane transmissivity η_y. The greatest coupling efficiency is obviously achieved when $\Delta z \to 0$, or, in other words, when the waists of both modes coincide[5]. This minimal coupling loss, expressed in dB, is given by:

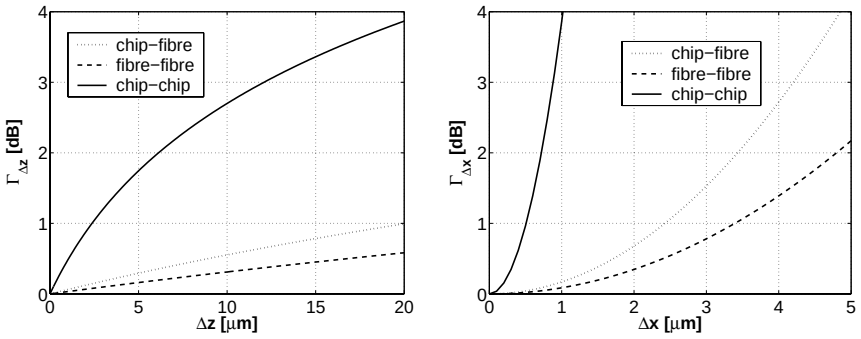
$$\begin{aligned} \Gamma_0\ [dB] &= \Gamma_{0x} + \Gamma_{0y} \\ &= 10\log\left(\frac{2w_0 w_{x0}}{w_0^2 + w_{x0}^2}\right) + 10\log\left(\frac{2w_0 w_{y0}}{w_0^2 + w_{y0}^2}\right) \end{aligned} \tag{1.7}$$

Γ_{0x} is plotted in figure 1.8 as a function of mode size mismatch w_{x0}/w_0. A maximum coupling efficiency of unity is reached for the case of perfectly matched modes ($w_{0x} = w_0$).

The positional tolerances can be estimated by taking into consideration the additional losses caused by a misalignment of the fibre. The additional loss $\Gamma_{\Delta z}$ due to a separation Δz between fibre and chip is calculated from (1.6) and is sketched

[4]The phase components jkz and $\phi(z)$ do not affect the coupling efficiency since they are independent of x and y.

[5]Note that this maximum coupling efficiency might not be achieved with a laser-fibre butt-coupling since the (virtual) waist is located inside the laser.



(a) (b)

Figure 1.9: Coupling loss for chip-fibre, fibre-fibre and chip-chip coupling, caused by (a) a longitudinal separation between the components and (b) a lateral misalignment.

in figure 1.9(a) for a fibre-chip, fibre-fibre and chip-chip coupling ($1.55\mu m$-devices). When a small chip mode is coupled to a fibre, one might expect the coupling efficiency to increase when the fibre is moved away from the chip because a more expanded beam reaches the fibre facet and a better overlap would be obtained. However, a steadily increasing coupling loss is found as both components are further separated. This effect is a consequence of the phase front curvature, which makes the phase in the plane of the fibre facet a rapidly oscillating function that reduces the overlap integral with the flat fibre mode. The sharper increase in coupling loss when smaller modes come into play can be explained by the stronger diffraction for smaller modes, causing a stronger curvature of the phase front and a stronger broadening of the propagating beam.

The additional loss due to a lateral misalignment Δx and Δy from the optimum coupling position in the x and y directions is given by:

$$\begin{aligned} \Gamma_{\Delta xy}\ [dB] &= \Gamma_{\Delta x} + \Gamma_{\Delta y} \\ &= 10\log(e) \cdot \left(\frac{\Delta x^2}{w_0^2 + w_{x0}^2} + \frac{\Delta y^2}{w_0^2 + w_{y0}^2} \right) \end{aligned} \tag{1.8}$$

(1.8) confirms the intuitive feeling that the tolerances are larger between two large modes than between two modes of small dimension. Figure 1.9(b) shows $\Gamma_{\Delta x}$ for a typical chip mode and fibre mode in a chip-fibre, fibre-fibre and chip-chip configuration. The corresponding plot for $\Gamma_{\Delta y}$ shows the same behavior, but with even stricter tolerances because of the stronger vertical confinement of the chip mode.

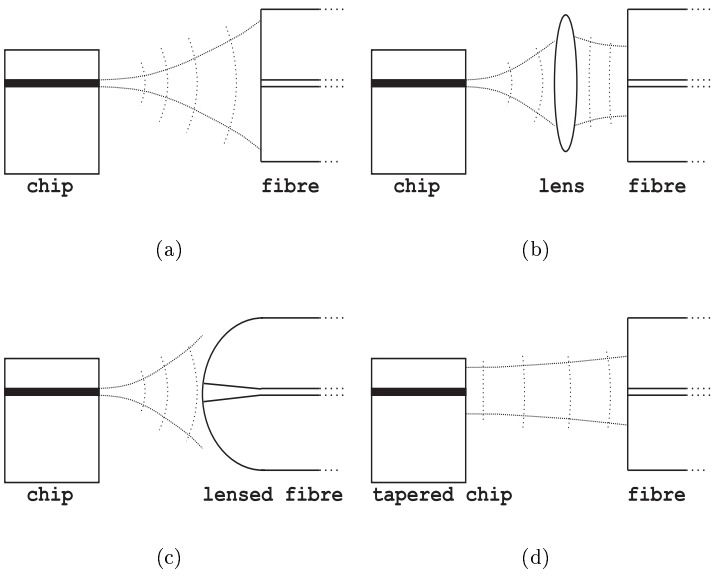


Figure 1.10: Schematic overview of coupling schemes. a) chip-fibre butt-coupling; b) use of a lens; c) tapered/lensed fibres; d) monolithic integrated spot-size converter.

1.3.3 Micro-optic solutions

Figure 1.10 gives schematic representations of several approaches to raise the coupling efficiency. As seen in the previous section, the mode size mismatch leads to a high coupling loss. A straightforward way to match the mode sizes and increase the coupling efficiency significantly, is through the introduction of a magnifying lens system that images the chip mode on the fibre facet. The elliptical nature of the chip mode suggests departures from simple hemispherical surfaces to aspheric lenses with different magnification factors M in lateral ($M_x = w_0/w_{0x}$) and transverse ($M_y = w_0/w_{0y}$) directions. Also combinations of lenses are sometimes used to overcome the asymmetry of the chip mode. However, because of the complexity and cost of such aspheric lenses, a hemispheric surface with a magnification factor $M = \sqrt{M_x.M_y}$ is mostly chosen.

But lenses do not only improve the maximum coupling efficiency considerably, they also make the arrangement increasingly more sensitive to alignment errors. For instance, when light is travelling from fibre to chip, two smaller modes have to be aligned at the chip facet, but the alignment error of the chip with respect to the lens remains the same Δx. When light is propagating in the opposite direction, from chip to fibre, two large modes have to be aligned but the alignment error Δx of the chip is magnified to $M_x\Delta_x$. It is easily calculated from (1.8) that in both cases the system becomes a factor $\frac{1}{2}\left(1 + w_0^2/w_{0x}^2\right)$ more sensitive to misalignments Δx if a lens with magnification $M_x = w_0/w_{0x}$ is introduced. For instance, if a

fibre mode with $w_0 = 5\mu m$ and a chip mode with $w_{0x} = 1\mu m$ are considered, the loss $\Gamma_{\Delta x}$ is multiplied by a factor of 13. This makes the alignment sensitivity as severe as for the coupling between two chip modes.
Since waist matching apparently brings about high coupling efficiency at the expense of reduced alignment tolerances, it may in some cases be preferable to design the lens system in order to trade coupling efficiency for increased offset tolerance. A procedure to find the optimum trade-off is presented in [10].

Another consideration is the mechanical stability of the coupling arrangement, especially to vibrations or to temperature fluctuations. For these reasons, the most common way of coupling light between chip and fibre is via a tapered or lensed fibre. Tapered fibres are mostly used for multimode fibres, where the fibre tip is etched by a controlled HF-dip [11]. Lensed fibres are generally fabricated by tapering the fibre down to a point and melting the end. Another technique is to form the lens with photoresist [12] or by micro-machining the fibre tip by means of laser ablation [13] or grinding the facet in a wedge-shape with a polishing tool [14]. In all cases, the fibre mode is focussed down to a more chip-adapted size and the coupling efficiency is again improved at the cost of the alignment tolerance.

From (1.7) and (1.8) it is clear that the best solution to the coupling problem is the integration of a **S**pot-**S**ize-**C**onverter (SSC) at the inputs and outputs of the chip. This SSC transforms the small on-chip mode to a field profile that resembles the fibre mode as closely as possible. The size and shape matching of both large modes assures again a high coupling efficiency, but this time not at the expense of the alignment tolerances. It should be noted that some authors mention a significantly increased alignment tolerance for tapered devices as one of its advantages. However, if we compare the chip-fibre and fibre-fibre coupling loss curves in figure 1.9, we see that the improvement is only marginal. The true merit of integrating SSCs is the fact that, as opposed to discrete lenses and lensed fibres, a high coupling efficiency is obtained without compromising the alignment tolerances. The main disadvantage of integrated SSCs are the complications introduced in the device fabrication. The search for a simple and robust fabrication scheme that allows to integrate SSCs on a PIC without compromising the other components forms the main subject of this thesis.

1.4 Conclusion and outline of the thesis

In this chapter we have explained that optical communications has evolved from simple single-wavelength point-to-point links to complex wavelength division multiplexed optical networks requiring complex photonic integrated circuits. Hybrid integration techniques provide near-future solutions, but monolithic integration will eventually be needed to obtain a low-cost mass-production of photonic modules. One of the main obstacles to a commercial breakthrough remains the lack of a highly efficient and alignment tolerant coupling between module and optical fibre. This problem can be solved by providing the modules with monolithic in-

tegrated spot-size converters, an approach that is being studied intensively, but which mostly involves serious complications for the wafer growth and processing. The main part of this thesis describes the design, fabrication and characterization of various types of spot-size converters, with particular attention to obtaining an uncomplicated fabrication scheme. Both InP-based and GaAs-based devices are treated. Chapter 2 discusses the coupling problem in detail and gives an overview of possible approaches to the design of spot-size converters. An important class of components is formed by the 'adiabatic' spot-size converters, which are devices with a slowly varying geometry that gradually transforms the small 'on-chip' mode into a fibre-adapted field profile without loss of power. Chapter 3 treats the theory of adiabatic mode transformers and concludes with a practical design criterion. Also some illustrative examples are elaborated. The design of several types of InP-based spot-size converters, both for an improved fibre-chip coupling and an on-chip mode transformation, is presented in chapter 4. An interesting feature of the GaAs/AlGaAs material system is the possibility to selectively oxidize AlGaAs-layers with a high Al-content, as explained in chapter 5. Chapter 6 then treats the development of a novel type of short-wavelength (980nm) laser with a monolithic integrated spot-size converter. By taking full advantage of the oxide properties, a simple fabrication scheme, involving a single planar growth and a single conventional etch step is obtained. Also designs for an extension to longer wavelengths through the incorporation of GaInNAs active layers is discussed. The very recently developed GaInNAs material broadens the wavelength range of GaAs-based active devices to 1.3μm and beyond. Chapter 6 further summarises the work that was done in the context of the European projects OPTIVAN and GIFT, including the characterisation of GaInNAs-based components and the design of oxide-confined tapered lasers operating at 1.3μm. Chapter 7 finally recapitulates the main achievements of this thesis and presents some perspectives of this work.

1.5 Publications in the context of this work

- K. De Mesel, R. Baets, C. Sys, S. Verstuyft, I. Moerman and P. Van Daele, "First demonstration of a 980nm oxide confined laser with integrated spot size converter", El. Lett., Vol. 36, No. 12, pp. 1028-1029, 2000.
- G. Flamand, K. De Mesel, I. Moerman, B. Dhoedt, W. Hunziker, A. Kalmar, R. Baets, P. Van Daele and W. Leeb, "InP-Based PIC for an optical phased array antenna at 1.06μm", IEEE Photonics Technology Letters, Vol. 12, No. 7, pp. 876-878, 2000.
- M. Galarza, K. De Mesel, D. Fuentes, R. Baets and M. López-Amo, "Modeling of InGaAsP-InP 1.55μm lasers with integrated mode expanders using fiber-matched leaky waveguides", Applied Physics B, Lasers and Optics, B 73, pp. 585-588, 2001.
- M. Galarza, K. De Mesel, S. Verstuyft, C. Aramburu, I. Moerman, P. Van Daele, R. Baets and M. López-Amo,"1.55μm InP-InGaAsP Fabry-Perot lasers

with integrated spot size converters using Antiresonant Reflecting Optical Waveguides", IEEE Photonics Technology Letters, Vol. 14, No. 8, pp. 1043-1045, 2002.

- D. Taillaert, W. Bogaerts, P. Bienstman, T. Krauss, P. Van Daele, I. Moerman, S. Verstuyft, K. De Mesel and R. Baets, An out-of-plane grating coupler for efficient butt-coupling between compact planar waveguides and single-mode fibres, IEEE Journal of Quantum Electronics, Vol. 38, No. 7, pp. 949-955, 2002.

- M. Galarza, K. De Mesel, S. Verstuyft, C. Aramburu, I. Moerman, P. Van Daele, R. Baets and M. López-Amo,"A new spot-size converter concept using fiber-matched antiresonant reflecting optical waveguides", to be published in IEEE Journal of Lightwave Technology in Januari 2003.

- K. De Mesel, C. Sys, K. Caekebeke, S. Verstuyft, B. Dhoedt, I. Moerman, P. Vandaele and R. Baets, "Wet lateral oxidation of AlGaAs for passive and active guided wave components", Proceedings Annual Symposium IEEE/LEOS Benelux Chapter, pp.109-112, Gent, Belgium, November 1998.

- K. De Mesel, I. Moerman, R. Baets, B. Dhoedt, P. Van Daele and J. Stulemijer, "Spot size converters for low cost PICs", Proceeding 9th European Conference on Integrated Optics, pp. 253-258, Turin, Italy, 1999.

- K. De Mesel, C. Sys, S. Verstuyft, I. Moerman, P. Van Daele and R. Baets, "First demonstration of a 980nm oxide confined laser with integrated spot size converter", Proceedings Integrated Photonics Research, pp.126-128, Québec, Canada, 2000.

- K. De Mesel, C. Sys, S. Verstuyft, I. Moerman, P. Van Daele and R. Baets, "Demonstration of a novel type of tapered laser based on lateral selective wet oxidation", Proceedings Annual Symposium IEEE Lasers and Electro-Optic Society Benelux Chapter, pp.159-162, Delft, The Netherlands, October 2000.

- M. Galarza, K. De Mesel, R. Baets and M. López-Amo, "Design of InGaAsP-InP tapered ridge mode transformer using an underlying ARROW coupling waveguide", Proceedings Annual Symposium IEEE/LEOS Benelux Chapter, pp.79-82, Delft, The Netherlands, October 2000.

- M. Galarza, K. De Mesel, R. Baets and M. López-Amo, "Modelling of InGaAsP-InP 1.55μm lasers with integrated mode expanders using fibre-matched Antiresonant Reflecting Optical Waveguide", Proceedings 10th European Conference on Integrated Optics, pp. 421-424, Paderborn, Germany, 4-6 April 2001.

- M. Galarza, K. De Mesel, T. Van Caenegem, S. Verstuyft, C. Aramburu, I. Moerman, P. Van Daele, R. Baets and M. López-Amo, "InGaAsP-InP 1.55μm lasers with integrated mode expanders using fiber-matched leaky waveguides", Proceedings Lasers and Electro-Optic Society 2001 Annual Meeting, pp. 798-799, San Diego, California, USA, 12-15 November 2001.

- M. Galarza, K. De Mesel, S. Verstuyft, I. Moerman, P. Van Daele, R. Baets and M. López-Amo, "InGaAsP-InP 1.55μm lasers with integrated mode expanders using fibre-matched Antiresonant Reflecting Optical Waveguides", European Semiconductor Laser Workshop Gent, Belgium, 28-29 September 2001.

- G. Flamand, K. De Mesel, J. Derluyn, I. Moerman, B. Dhoedt, R. Baets, P. Demeester and P. Van Daele, "InP-based PIC for an optical phased array antenna at 1.06μm", Annual IEEE Lasers and Electro-Optic Society Benelux Chapter Workshop on Integrated Optics and Passive Micro-optics, p. 19, Brugge, Belgium, June 1998.

- G. Flamand, K. De Mesel, I. Christiaens, I. Moerman, B. Dhoedt, W. Hunziker, A. Kalmar, R. Baets, P. Van Daele and W. Leeb, "InP-Based PIC for an Optical Phased Array Antenna at 1.06μm", Proceedings 12th International Conference on Indium Phosphide and Related Materials, pp. 25-28, May 2000, Williamsburg, United States.

- K. De Mesel, S. Verstuyft, I. Moerman, P. Van Daele and R. Baets, "Oxide confined laser diodes with an integrated spot size converter", submitted to IEEE Journal of Selected Topics in Quantum Electronics in September 2002.

- M. Galarza, K. De Mesel, S. Verstuyft, D. Fuentes, C. Aramburu, M. López-Amo, I. Moerman, P. Van Daele and R. Baets, "Mode-Expanded 1.55 μm InP-InGaAsP Fabry-Prot Lasers Using ARROW Waveguides for Efficient Fiber Coupling", submitted to IEEE Journal of Selected Topics in Quantum Electronics in September 2002.

1.6 Patent applications

Marko Galarza, Kurt De Mesel, Candido Aramburu and Roel Baets, 'Semiconductor optical component utilizing leaky structures to match the mode of rib waveguides to that of the fiber', Patent Application, 2002.

References

[1] "http://www.gridcomputing.com."

[2] G. Raybon, B. Mikkelsen, R. Essiambre, A. Stentz, T. Nielsen, D. Peckham, L. Hsu, L. Grüner-Nielsen, K. Dreyer, and J. Johnson, "320 Gbit/s single-channel pseudo-linear transmission over 200km of nonzero-dispersion fiber," in *Proceedings Conference on Optical FiberCommunications 2000 - Postdeadline Papers*, (Baltimore, USA), March 2000. paper PD29.

[3] "International Telecommunication Union - Telecommunication Standardization Sector (ITU-T), Recommendation G.692: 'Optical interfaces for multichannel systems with optical amplifiers', Series G: Transmission systems and media, digital systems and networks (Transmission media characteristics - Chracteristics of optical components and sub-systems)," October 1998.

[4] "http://www.alcatel.com/telecom/ssd/keytech/evolut/overview.htm."

[5] T. VanCaenegem, *Development of monolithic integration techniques by means of selective area growth and MOVPE for the realisation of photonic integrated ciruits for telecom applications.* PhD thesis, Ghent University, 2001.

[6] "Gigabit ethernet empowers broadband," *Lightwave Europe*, vol. 1, p. 12, May 2002.

[7] A. Snyder and J. Love, *Optical waveguide theory.* No. 6, London, England: Chapman and Hall Ltd., 1983.

[8] A. Yariv, *Introduction to optical electronics.* New York: Holt, Rienhart and Winston, 2nd ed., 1976.

[9] W. Joyce and B. DeLoach, "Alignment of gaussian beams," *Appl. Opt.*, vol. 23, no. 23, pp. 4187–4196, 1984.

[10] W. Joyce and B. DeLoach, "Alignment-tolerant optical-fiber tips for laser transmitters," *IEEE J. Lightwave Technol.*, vol. 3, no. 4, pp. 755–757, 1985.

[11] T. Ozeki and B. Kawasaki, "Efficient power coupling using taper-ended multimode optical fibres," *Electron. Lett.*, vol. 12, no. 23, pp. 607–608, 1976.

[12] L. Cohen and M. Schneider, "Microlenses for coupling junction lasers to optical fibers," *Appl. Opt.*, vol. 13, no. 1, pp. 89–94, 1974.

[13] K. Naessens, *PhD in preparation.* PhD thesis, INTEC, Ghent University.

[14] R. Modavis and T. Webb, "Anamorphic microlens for laser diode to single-mode fiber coupling," *IEEE Phot. Technol. Lett.*, vol. 7, no. 7, pp. 798–800, 1995.

Chapter 2

Spot-size converters: an overview

2.1 Introduction

Integrated spot-size converters can be conceived in many different ways. As a consequence, a lot of designs and several mode conversion techniques have been reported during the last years. It is therefore instructive to produce some order out of this large variety of designs by constructing a general classification scheme. The classification elaborated in this chapter was introduced in [1] and is built from an optical point of view by making a distinction between the ways the device geometry is changed and between the ways the optical field is transformed. For a classification from a technological point of view we refer to [2].

2.2 Classification

The two dimensions –geometry and optical operation principle– that are distinguished in the construction of the classification scheme consist each of two cases plus an intermediate hybrid case. This leads to the nine classes as presented in table 2.2 and further clarified in the next paragraphs.

	Single Mode 'Adiabatic' **(A)**	Hybrid **(A+I)**	Multi Mode 'Interference' **(I)**
Lateral **(L)**	A/L	A+I/L	I/L
Hybrid **(L+T)**	A/L+T	A+I/L+T	I/L+T
Transverse **(T)**	A/T	A+I/T	I/T

Table 2.1: Taper classification: schematic overview.

2.2.1 Geometry

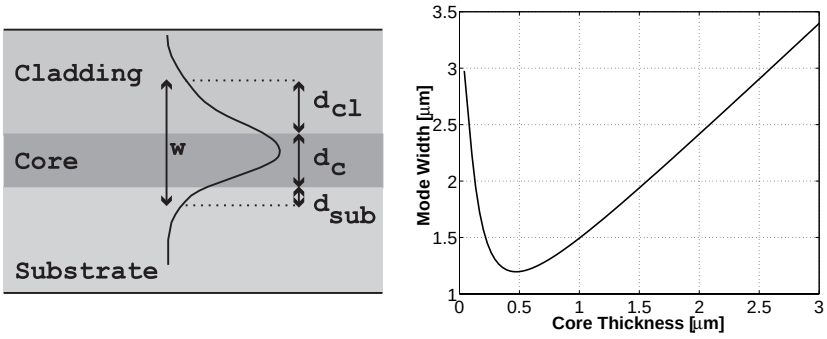
The cross-section of a spot-size converter changes, either gradually or abruptly, along its length. The specific geometrical evolution leads to a first categorisation into the classes 'lateral' (L), 'transverse' (T) or hybrid lateral/transverse (L+T). An unambiguous definition of these subclasses is the following: a spot-size converter is classified as 'lateral' ('transverse') if geometrical changes are visible in the transverse (lateral) projection of the device (unless, of course, these changes have no impact on the optical field) and not in the lateral (transverse) projection. In the hybrid 'L+T' subclass changes are visible in both projections. It is important to emphasise that the terms 'lateral' and 'transverse' only relate to geometrical changes and not to the changes in the optical field. Many purely lateral spot-size converters do actually induce an important transverse spot-size conversion, as will be made clear in the examples in section 2.2.3.

2.2.2 Optical operation principle

There are two distinct principles to transform a mode into a more fibre-adapted field profile. A first category, containing the so-called 'adiabatic' mode transformers, implies a gradual expansion of the fundamental local mode, while the other group involves the excitation of multiple modes that interfere with each other to produce the desired field at the output.

Single mode or adiabatic spot-size converters (class A) The vast majority of the spot-size converters are based on an adiabatic mode evolution, whereby the small and highly asymmetric fundamental mode is gradually transformed to a large and more circular profile. The term 'adiabatic' indicates that all power is kept in one (local) mode during propagation. To avoid excessive loss[1], caused by coupling to higher order guided modes and/or radiation modes, relatively long devices with a slowly varying geometry –hence 'tapered' devices or 'tapers'– and often complicated shapes are needed. This class forms the first column of table 2.2 and will further be referred to as 'adiabatic spot-size transformers' or 'single mode spot-size transformers'. We label them as 'class A' for short.
Adiabatic devices make a gradual transition between a strongly confining waveguide and a cross-section that carries a large mode. This can be accomplished in two ways. A first option is to increase the core dimensions, such that a broad and/or thick (multi-mode) waveguide is obtained. The adiabaticity of the design then ensures that only the fundamental output mode is excited. A second option is to make the output core sufficiently narrow and/or thin to obtain a more loosely confined optical field. In both cases, a large mode is obtained. This can be nicely illustrated by considering the simple slab waveguide of figure 2.1(a), consisting of a core layer that is sandwiched between a substrate and an upper cladding. The effective width w of the guided mode is expressed as $w = d_c + d_{sub} + d_{cl}$, where

[1] i.e. the difference in fundamental-mode power between the beginning and the end of the taper.



(a) (b)

Figure 2.1: a) slab waveguide and b) evolution of the mode width as a function of the core thickness for an InP-based waveguide ($n_{cl} = n_{sub} = n_{InP} = 3.18$, $n_c = 3.25$).

d_c is the core thickness and d_{sub} and d_{cl} are the penetration depths of the exponentially decaying field in the substrate and the cladding respectively[2]. Figure 2.1(b) shows the evolution of the mode width as a function of core thickness for a typical 1.55μm InP-based waveguide ($n_{cl} = n_{sub} = n_{InP} = 3.18$, $n_c = 3.25$). We note the occurrence of a minimum mode width for a core thickness of about 0.5μm. At that point, the light is maximally confined in the core layer, making this design the obvious choice for e.g. a laser structure. It is clear by intuition that thicker cores will guide larger fundamental modes. This is also seen in figure 2.1(b) where the mode width varies directly as the core thickness. On the other hand, also sufficiently thin waveguides guide larger modes because of the reduced guiding effect and mode tails penetrating deeper into the cladding and substrate. The mode width is now inversely proportional to the core thickness, implying that this thickness has to be controlled very accurately if thin cores are used to define the taper output mode.

The behaviour and design of adiabatic devices are studied in detail in chapter 3, where a criterion is constructed to delineate those taper shapes, which ensure an approximately adiabatic propagation, and those shapes, which exhibit significant loss. Such a delineation is not rigorous, but provides a useful guide to the design of taper shapes, which minimise excess loss and simultaneously are short enough for practical application.

[2]The fields have decreased by a factor e^{-1} at a distance d_{sub} and d_{cl} in the cladding and the substrate respectively. The values d_{sub} and d_{cl} are also related to the Goos-Hänchen shift appearing in the ray description of slab waveguides.

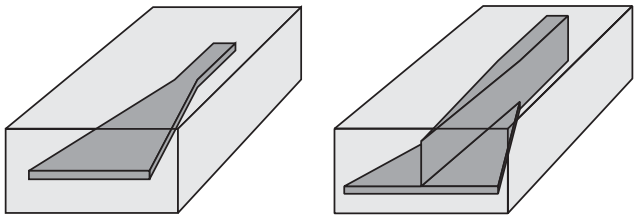
Multi-mode or interference spot-size converters (class I) A second class of devices are the 'multi-mode spot-size converters' or 'interference spot-size converters'. Spot-size transformers of this class are not designed to keep all power in a slowly changing fundamental mode. Instead, several modes are excited, each propagating with its own propagation constant so that the total field changes as one travels towards the output facet. The position of the output facet is chosen such that the interference of these multiple modes yields a maximum coupling to fibre. This category generally has the advantage of being much shorter than adiabatic spot-size transformers [3][4][5][6].

Hybrid adiabatic/interference spot-size converters (class A+I) Some devices don't fit in one of the classes A or I, because they are based on a combination of both principles. We call them the hybrid adiabatic/interference spot-size converters, type A+I (column II in table 2.2).

2.2.3 Illustrative examples

We will now illustrate the diversity of spot-size converter concepts by discussing some representative examples of each class. All devices have, for the ease of drawing, a linear profile and the darker regions correspond with a higher refractive index.

Class A/L Lateral spot-size converters have the advantage of being easy to fabricate, since only standard processing techniques (planar growth, conventional etching) are needed for the definition of the devices. A simple adiabatic lateral spot-size converter is shown in figure 2.2(a) [7]. The optical field expands as it propagates from the small end of the device to the large end, keeping all power in the fundamental mode. Of course, only a lateral mode expansion is obtained. The tapered core is overgrown with a cladding material to obtain a more symmetrical mode profile (buried waveguide). A similar device, but with a core that is sufficiently narrowed towards the output facet has also been demonstrated at 1320nm, with 6dB coupling loss (an improvement of 3dB) to a standard long-wavelength



(a) (b)

Figure 2.2: Adiabatic spot-size converters: (a) lateral; (b) lateral, double core.

fiber and $\pm 2.5\mu$m alignment tolerance (-1dB) in all directions [8].
Another, and very popular, adiabatic lateral taper design is the overlapping waveguide structure shown in figure 2.2(b) [9]. This is a very attractive approach because both a lateral and transverse mode expansion are obtained by a pure lateral tapering of the core layer. The structure is again overgrown with a cladding material. By slowly narrowing the upper section, the mode is gradually transferred to the lower part, which is weakly guiding in the transverse direction. The lower section is broadened at the same time, leading to a broad output field as well. The drawback of this design, however, is the need for extremely narrow taper tips ($\approx$ 100nm), which may cause problems with the photolithography of the device and require special techniques as e.g. direct e-beam writing. Very low coupling losses of 1.1dB to single mode fibre have been demonstrated at 1.55μm for a tapered waveguide [9]. Also tapered lasers with coupling efficiencies of 1.85dB and -1dB alignment tolerances of ± 2.35x2.15μm (LxT) were reported [10].

The examples in figure 2.2 require a second growth of an InP-cladding after the etching of the lateral taper to create a well-expanded output mode. Regrowth has become an established technique in the InP/InGaAsP material system, but is problematic in the GaAs/AlGaAs system. Especially layers with a high aluminium content are difficult to overgrow due to oxide formation [11]. Other types of spot-size converter designs, without regrown cladding layers, are therefore required in the GaAs/AlGaAs system. An interesting device is the laterally tapered laser structure shown in figure 2.3 and first proposed in [12]. It consists of a high confinement waveguide on top a large ridge waveguide, which is optimised for coupling to fibre. The optical mode is concentrated in the upper waveguide when the upper rib is broad (i.e. in the laser section). A gradual narrowing of this rib forces the tightly confined mode to expand into the underlying mesa. The fabrication of this device requires only one planar epitaxial growth and two etch steps (and hence also two masks which have to be carefully aligned) to define both mesas. The definition of the narrowing section can be made less demanding as for the design in figure 2.2(b) by taking special care in the design of the layer structure. The short wavelength (980nm) tapered laser reported in [12] incorporated a 10μm wide and 8μm thick fibre-matched waveguide, resulting in coupling losses of only 0.9dB for coupling to a standard long-wavelength single mode fibre (design wavelength=1.3μm). The farfield full width at half maximum (FWHM) divergence angles were as low as 5.6^ox7.4^o.

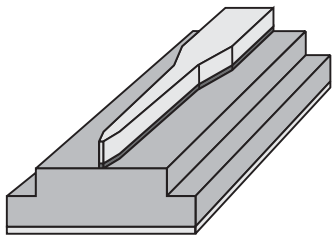
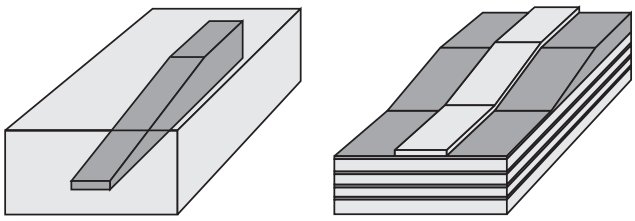


Figure 2.3: An adiabatic lateral spot-size converter without regrowth.

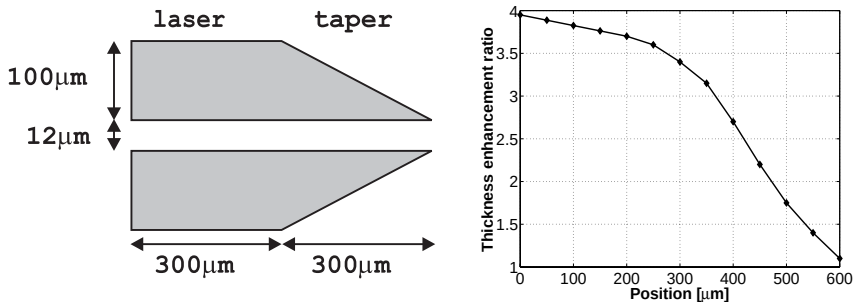
The same design has also been demonstrated in the InP/InGaAsP material system, where the elimination of a second growth step helps to reduce the fabrication complexity [13]. Coupling losses of 3.7dB to standard single mode fibre were measured at 1.55μm with a reduction of the farfield FWHM from 34x39^oC (LxT) to 9.5x22^{o}C.

Class A/T A transverse adiabatic taper is shown in figure 2.4(a) where the thickness of the core layer of a buried waveguide is reduced to end in a weakly guiding structure carrying a transversally expanded mode [14][15][16]. Figure 2.4(b) shows a tapered strip-loaded waveguide where a rib of constant width and height covers a layer stack with a transversely tapered core to define a guided mode that expands into the diluted underlaying cladding towards the output facet [17].
Standard growth techniques like MOCVD, MBE or CBE, however, deposit planar layers of semiconductor material, making the implementation of transverse tapers not straightforward. Modified growth techniques and complicated etching processes are therefore required for the fabrication of this type of device. We will now shortly discuss some of these techniques to illustrate both the variety of fabrication techniques and the increased complexity as compared to conventional planar growth and etching.
A first important non-planar growth technique is Selective Area Growth (SAG), where pairs of dielectric (SiO_2, Si_3N_4) stripes are defined on the substrate prior to the growth [18]. The nucleation properties for the growth of species on the dielectric mask and the semiconductor surface are completely different, resulting in an enhanced growth rate for the material deposited between or near the mask stripes as compared to material deposited in a region far from any mask. This growth rate enhancement increases with increasing width of the mask and is caused by lateral vapor phase diffusion. A gradual narrowing of the mask stripes, as shown in figure 2.5(a), thus results in a decreasing layer thickness as required in the devices shown in figure 2.2.3. Takiguchi et al., for instance, used this technology for the fabrication of 1.3μm tapered lasers [14]. The tapering of the mask width



(a) (b)

Figure 2.4: Transverse adiabatic spot-size converters: (a) buried structure; (b) strip-loaded design.



(a) (b)

Figure 2.5: a) Mask design and b) corresponding thickness enhancement of InP for a laser diode with integrated spot-size converter (after [14]).

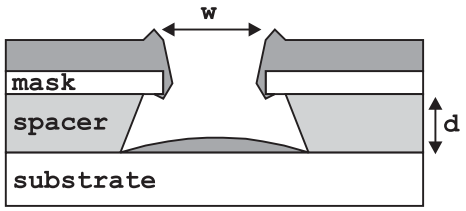


Figure 2.6: Principle of shadow masked growth.

from 100μm near the active region to zero near the facet, over a length of 300μm, resulted in a thickness reduction by a factor of four and improved the coupling by 5dB.

Another modified growth technique is Shadow Masked Growth (SMG), which uses a mask that is mechanically held at a certain distance above the substrate, as shown in figure 2.6 [19]. During epitaxial growth, deposition on the substrate will take place through the window in the shadow mask, resulting in a reduced growth rate with respect to the growth rate on a planar substrate. By changing the distance d between the mask and the substrate or the width w of the window, it is possible to control the thickness reduction in the middle of the window. The higher d and the lower w, the higher the growth rate reduction. With a constant mask-substrate spacing, it is possible to control the thickness of a layer over the substrate by changing the channel width. In [16] coupling efficiencies of 3.3dB, reducing the coupling loss by 6dB, and -1dB alignment tolerances of $\pm$2.1x1.7μm (LxT) were obtained at 1.55μm for a thickness reduction by a factor of three over a length of 200μm.

Transverse tapers can also be fabricated by combining planar growth with special non-planar etching techniques. Perhaps the most straightforward technique is dip-etching, where the sample is dipped with a well-chosen speed in an etching

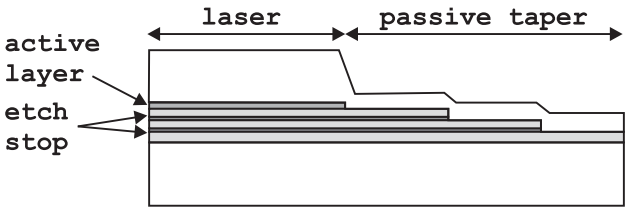


Figure 2.7: Longitudial evolution of a transverse taper fabricated by consecutive selective wet etching steps [23].

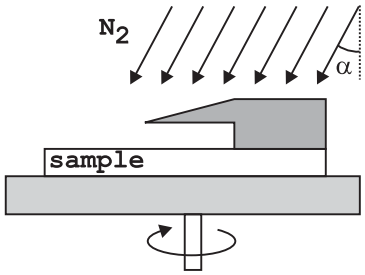


Figure 2.8: Principle of ion beam shadow etching.

solution [20]. This technique is simple, but leads to relatively long tapers and it is not a wafer-scale technique.
With diffusion-limited etching the wafer surface is covered by a mask pattern, similar to the ones used for SAG [21][22]. The samples are etched in a diffusion-limited etchant, e.g. bromine-methanol for InP-based devices. In this situation, where mass transport is the limiting factor for dissolution of the semiconductor, the etch rate is enhanced for narrower mask openings and lateral mask changes are translated in etch depths variations.
Consecutive selective wet etching steps are sometimes used to obtain a staircase-approximation of a continuous transverse adiabatic taper [23][24]. This is illustrated in figure 2.7, showing a laser integrated with a passive transverse spot-size converter with thin InP etch-stop layers inserted in the core for a controlled sequential removal of the core material. This very time consuming work requires up to five photolithographic exposures and etching steps if the transformation loss is to be reduced below 1dB [23].

Yet another technique is ion beam shadow etching, where a wedge-shaped silicon shadow mask is fixed above the sample on top of a spacer (figure 2.8) [25]. The arrangement is mounted on a rotating table that is tilted by an angle α with respect to the N_2 ion beam. With this set-up one region of the sample is constantly exposed to the beam, while another region is completely shaded by the mask. Between these two regions, however, there is a transitional region in which the total exposure time gradually decreases from maximum to zero and where the etch depth will vary accordingly.

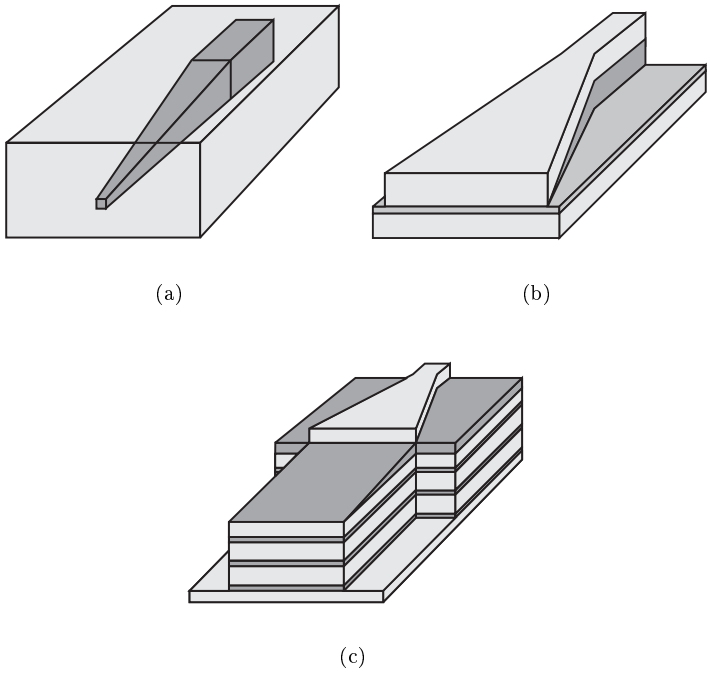


Figure 2.9: Hybrid lateral/transverse adiabatic spot-size converters.

Class A/L+T A two-dimensional mode expansion will in many cases involve the integration of a hybrid lateral/transverse spot-size converter. Figure 2.9(a) shows an adiabatic hybrid design with a core that is narrowed in both the lateral and transverse dimensions so that a two-dimensional weakly guiding structure is obtained at the output facet. Coupling losses of 4dB to standard $1.55\mu m$ single mode fibre have been reported for this type of device [26]. Several other designs involving combinations of down- and up-tapering, be it sequentially or simultaneously, have been reported. The device in figure 2.9(b), for example, has a core that becomes thinner as the waveguide is broadened [27]. This component was fabricated using consecutive selective wet etching steps (see section 2.2.3) and should be compared with the more elegant approach of figure 2.2(b) that only requires two etching steps to evolve towards a thin and broad output waveguide. In addition, the measured coupling loss of the hybrid lateral/transverse design is significantly higher and amounts up to 4.3dB. Another example of a hybrid adiabatic spot-size converter is shown figure 2.9(c). The transition from the high confinement waveguide to the fibre-matched waveguide is made by concatenating an adiabatic lateral taper in the high confinement waveguide with an adiabatic transverse taper. This forces the mode to expand transversally into the thin guiding layers of the un-

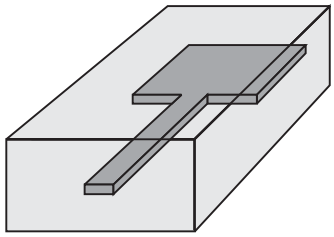


Figure 2.10: A lateral interference spot-size converter.

derlying diluted structure [28]. Coupling losses (including mode transformation losses) of 3.5dB have been obtained with this structure.

Class I/L A lateral interference spot-size converter can be as simple as the device shown in figure 2.10. It is in fact just a narrow straight waveguide with a well-defined length, which is butt-coupled to the on-chip waveguide. Due to the discontinuity at the junction, the guided mode incident from the on-chip waveguide will excite the guided mode and radiation modes of the narrow waveguide. As these modes propagate further towards the output facet, they interfere with each other and constitute a composite modal field pattern. The waveguide is terminated at a point where this composite field, comprising a linear combination of the guided and radiation modes, generates a better overlap with the fibre mode than the fundamental guided mode alone. Simulations showed that a device as short as 60μm could improve the coupling loss to a 1.3μm graded index fibre to less than 3dB. This should be compared with the performance of a linear adiabatic device, which still exhibits a coupling loss of more than 4dB for a length of 300μm.
It is interesting to note that the adiabatic lateral spot-size converter presented in figure 2.3 can also be designed as a lateral interference component [13]. The device supports two supermodes: one mainly centered in the upper core and one in the lower. The laser mode entering an adiabatic design will change its shape gradually and the narrowing is chosen slow enough to avoid coupling to the higher order supermode (and the radiation modes). When designed as an interference device, the first section has a steeper angle and power is being transferred to the second supermode. Both modes become closely matched in the second, slowly changing, section of the taper, where mode beating is observed in a way similar as in e.g. a directional coupler. If this second section is made long enough, then the power of the composite field is seen to shift back and forth between the upper and lower waveguide. The length of this section in the spot-size converter, however, is chosen equal to the beat length and a steep section ends the device to prevent coupling back to the upper waveguide.

Class I/T The principle of a transverse interference spot-size converter is shown for a slab waveguide in figure 2.11, as proposed in [3]. Just like in the device in figure 2.10 there is a discontinuous transition to the spot-size converter section. This multi-mode section supports three guided supermodes. Symmetrical alignment of

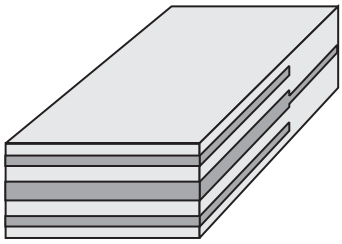


Figure 2.11: A transverse interference spot-size converter.

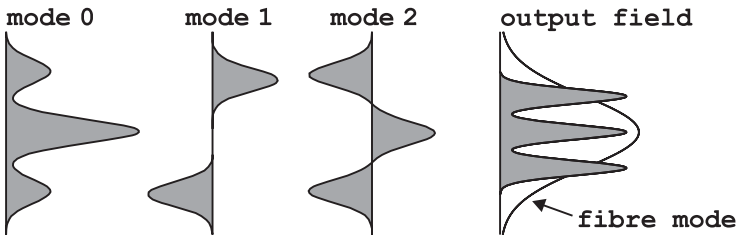


Figure 2.12: Supermodes supported in the multi-mode structure and the fibre mode overlapped with the output field (= mode 0 - mode 2)

the on-chip waveguide core with the multi mode section as sketched in figure 2.11 ensures that only the symmetric modes (mode 0 and mode 2) are excited. The interference pattern after a propagation distance equal to the beat length between both modes yields a higher mode field distribution and hence a better overlap to fibre (figure 2.12). Coupling efficiencies up to 80% (1dB loss) have been simulated.

Class I/L+T The non-periodic segmented waveguide described in [6] and shown in figure 2.13 is an example of a hybrid lateral/transverse interference spot-size converter, since many modes are excited at the consecutive discontinuous transitions and changes are visible in both the lateral and transverse projections of the device. The component was fabricated in the SiO_2/SiON material system

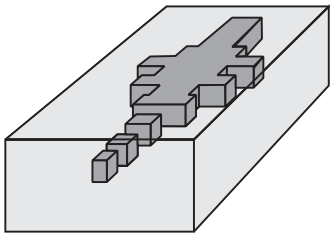


Figure 2.13: A hybrid lateral/transverse interference spot-size converter.

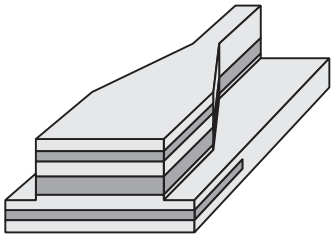


Figure 2.14: A hybrid adiabatic/interference, hybrid lateral/transverse spot-size converter.

and designed by using an evolutionary optimisation procedure, which relies on a collective learning process within a population of potential solutions, comparable to natural selection. It can be fabricated by simple planar growth and conventional etching techniques. The coupling loss to standard single mode fibre, which is already low for waveguides fabricated in this material system, was reduced from 3.5dB to 1.5dB at 1.55μm.

Class A+I No examples of hybrid adiabatic/interference spot-size converters were found in literature. It is even not straightforward to think of a reasonable design of such a type of component for a pure lateral or pure transverse device. A hybrid lateral/transverse component, however, can easily be conceived as a concatenation of e.g. a lateral adiabatic and a transverse interference device as shown in figure 2.14.

2.3 Spot-size converter simulation tools

When designing optical waveguide devices it is of primary importance to dispose of good simulation software to calculate optical modes in a specific device cross-section and to simulate the evolution of an optical field through a medium with an arbitrary refractive index distribution. In this section, we will shortly discuss the principle of the simulation tools that are available at our department and that were intensively used for the design of the spot-size converters reported in this thesis. These programs, a mode solver, a mode expansion algorithm and a Beam Propagation Method (BPM), are the most powerful tools for the simulation of photonic devices available today. A comprehensive overview of the numerous simulation methods can be found in [29].

2.3.1 Mode solvers

Mode solvers are essentially used to find the propagation constants of guided modes in longitudinally invariant waveguides. Early solvers were 1D slab mode solvers, mainly because of the success of the effective index method, which allows to reduce the 2D cross-section to an equivalent 1D geometry and which gives fairly accurate results (see further). A few years ago, with the advent of powerful computers and

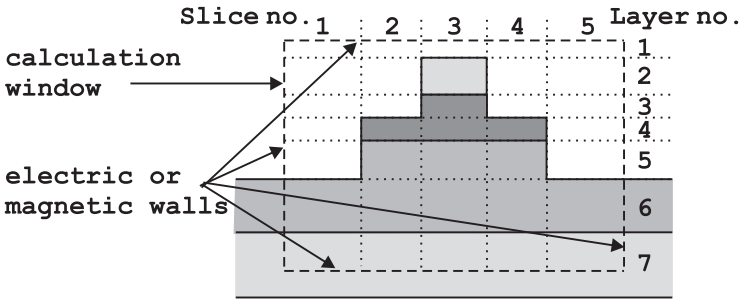


Figure 2.15: Waveguide cross-section modelled by a number of guide 'slices'.

more sophisticated algorithms, 2D mode solvers that can analyse a component within a reasonable time, have become available.
In the context of this work we used *Fimmwave* [30], a 2D mode solver based on a transverse resonance method (film mode matching) [29]. This method is particularly interesting for device geometries consisting of rectangular areas with a constant refractive index, like most of the semiconductor structures used for optical waveguiding. In this technique, the device cross-section is seen as a set of slab waveguides or 'slices' as illustrated in figure 2.15. The field in each slice is written as a superposition of the slab modes of that particular slice[3]. These slab modes are the exact solutions of the wave equations in that region. At the interface between two slices, the continuity of the tangential components of the electric and magnetic fields is imposed. From this matching procedure and by applying the appropriate boundary conditions ('electric' or 'magnetic' walls) at the edges of the calculation window, an indirect eigenvalue problem for the propagation constant, which can be readily solved, is formulated. The film mode matching method can be employed for full-vectorial, semi-vectorial or scalar calculations.
The knowledge of the propagation factor is for instance essential to:

- generate and display the spatial distribution of the optical field. This gives a lot of insight in the behaviour of a particular device (e.g. to visualise the evolution of the fundamental local mode of a spot-size converter[4]);
- calculate the confinement factor of the optical field in a specific region, e.g. in the active region of a laser diode;
- calculate the far field radiation pattern of a mode from its near-field distribution. The knowledge of the far field properties may be very important for

[3] Any slab field can be written exactly as a superposition of an infinite series of slab modes. A mode solver algorithm will of course restrict this series to a finite number of modes to approximate the exact field profile. The number of modes to be taken into account depends on the geometry and the desired accuracy. The effective index method forms the lowest order approximation by only taking the fundamental slab mode of each slice into account.
[4] Local modes of an arbitrary waveguide at a certain position z along the propagation axis are the eigenmodes of a straight waveguide with the same cross-section. The local mode concept will be explained in detail in chapter 3.

applications like optical recording or laser-to-fibre coupling. *Fimmwave* first calculates the near field profile just outside the facet and then decomposes this field into its plane-wave spectrum by performing 2D Fourier transforms. The propagation of these plane waves is trivial, since they are the eigenmodes of homogeneous space. Finally, the 2D far-field profile is calculated, taking the mixing of the field components into account;

- determine the coupling efficiency between two waveguides by calculating the overlap integral between two optical fields.

The software package also includes a module to build so-called 'projects', which allow to see how a variety of mode parameters, like propagation constants or confinement factors, change as the shape of a waveguide is altered. The scanner works by being given two versions of the waveguide, one representing the structure at the beginning and the other at the end of the scan respectively. A third structure is then automatically constructed as a weighted average of the two waveguides. A scan giving the evolution of the propagation constants is for instance needed to determine the optical shape of an adiabatic spot-size converter (see e.g. section 3.5).
Another type of project allows to perform calculations on compound devices. A compound device is a structure made up of two or more sections of straight waveguides that are butt-coupled together. Each section is specified by the waveguide cross-section, its length and a specification of the joint with the previous section. This might include a lateral or transverse offset and a small ($< 10°$) tilt angle. The program will calculate the coupling between the n modes in the input waveguide and the m modes in the output waveguide, constructing a matrix (the S matrix) for the coupling coefficients[5]. Compound devices were used in this work to calculate the coupling efficiency between waveguides and optical fibres.

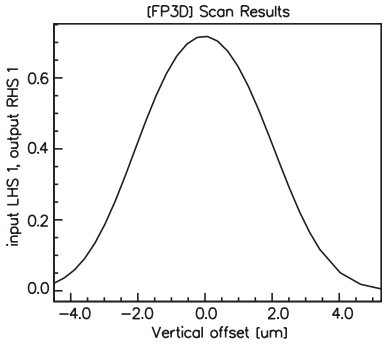


Figure 2.16: Coupling between a tapered waveguide and a fibre, calculated with the *Fimmwave* 'compound device scanner'.

[5]The calculation uses a vector overlap integral between modes of adjoining sections.

There is also a compound device scanner, which tells how the S-matrix of the compound device varies with one or more of the device parameters. The compound device scanner works similarly to the waveguide scanner: one compound device representing the starting point of the scan and one representing the end-point have to be entered. The program then proceeds by calculating a varying weighted average of the two. The compound device scanner is e.g. a very useful tool to determine lateral and transverse alignment tolerances of the coupling between a spot-size converter and an optical fibre (see e.g. figure 2.16).

2.3.2 Mode expansion method

A first method to simulate the propagation of an optical field through an arbitrary waveguide is the mode expansion algorithm. In this method, the waveguide is approximated by a concatenation of longitudinally invariant waveguide sections as shown in figure 2.17. Any field in a cross-section of an optical waveguide can be written exactly as an infinite sum of the guided, radiation and evanescent local modes of the waveguide. The mode matching tools calculate a finite number of modes and express the incident field as a superposition of these local modes. The radiation modes, which form a continuous spectrum, are replaced by a discrete set of modes because of the presence of the hard boundary conditions. This superposition will of course only be an approximation of the exact field and the accuracy (but also the total simulation time) will increase by increasing the total number of local modes in the expansion. The propagation of each mode through a section is trivial. At the end of a section these modes all couple to the modes of the next section. Also the power reflected back into the modes of the preceding section is taken into account, making the mode matching technique a bidirectional algorithm. The complete structure is calculated by repeating this procedure from the first section to the last.

The mode expansion tool that was used in this work is *Fimmprop-3D* [30], a program that is built around the *Fimmwave* mode solver engine. This means that the propagation can be fully vectorial and that high index contrasts can be handled. It is also, in contrast to BPM algorithms, a simulation tool that can handle large-angle propagation.

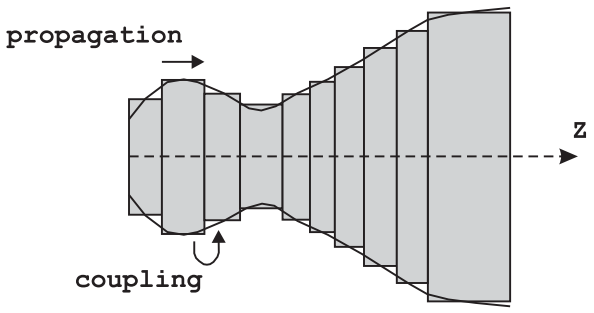


Figure 2.17: Stepwise approximation of a z-dependent waveguide structure.

Even though mode expansion methods are particularly suited for structures that have a small number of distinct cross-sections, such as MMI's (Multi Mode Interference couplers) and periodic structures, *Fimmprop-3D* also effectively deals with continuously varying structures like spot-size converters. It is found that taking into account a limited number of modes (10-20 modes are sufficient in many cases) already yields accurate results.

2.3.3 Beam Propagation Method (BPM)

BPM techniques were initially developed for two-dimensional and uni-directional paraxial propagation of scalar fields through media with (very) low refractive index differences. Meanwhile, less restrictive but more time-consuming schemes allowing three-dimensional, full-vectorial, bi-directional, wide(r)-angle propagation have been elaborated.
Assuming a wave is propagating in the positive z-direction of the optical waveguide structure and assuming the field distribution is know in a plane at $z = z_k$, the task is to determine the field at $z = z_{k+1} = z_k + \Delta z$. Repeating this calculation stepwise from the input to the output, the whole component can be (approximately) analysed. Most commercially available BPM-packages now employ a finite difference-based approach for that purpose, but they can also rely on the method of lines, finite elements or Lanczos reduction.
The main restriction in all BPM-solvers is the choice of the propagation axis. The propagation must remain within a certain angle from this axis and the maximum allowable angle depends on the algorithm used.
BPM simulations are usually very time consuming. For this reason, the effective index method is used in many cases to reduce the full 3D problem to a 2D approximation. Also the full-vectorial problem may be reduced to a semi-vectorial problem by disregarding the coupling terms in the Helmholz equation but still taking into account the electrical field discontinuities at dielectric interfaces, which depend on the polarisation of the propagating field. For weakly guiding structures, the semi-vectorial problem may be further simplified to a scalar one by also ignoring the field discontinuities at interfaces.
An important issue is how to limit the calculation domain by a suitable boundary condition. A 'hard' boundary condition, like a metallic wall, will act as a mirror and reflect all radiated power back into the calculation window. To avoid this, absorbing or transparent conditions have to be employed.
In the context of this work, we used *BPM_CAD*, a commercial package from *Optiwave Corporation* [31]. It is a finite-difference based package allowing the modelling of both 2D and 3D structures. The specific nature of the spot-size converters studied in this work, with modes expanding in both lateral and transverse dimension and in some cases shifting to a lower semiconductor layer, necessitate the use of the 3D-BPM. The propagation in this 3D-BPM is paraxial (no wide-angle propagation implemented) with a choice between a scalar or semi-vectorial model. The calculation domain is bounded with transparent boundary conditions.

2.3.4 Design of spot-size converters: *BPM_CAD* versus *Fimmprop-3D*

Fimmprop-3D is without any doubt the most versatile tool for the design of spot-size converters. One of the main reasons why *BPM_CAD3D* was soon ruled out was the required simulation time, which easily exceeded 20 hours. Knowing that the design of a spot-size converter is mainly an iterative process with sometimes several parameters to be adjusted, it is soon concluded that this is not a very productive way of working. In addition, the choice of the (uniform) grid spacing and the propagation step Δz is a fuzzy process and it is not always clear whether unexpected simulation results are to be attributed to the device design or to the choice of the discretisation parameters.
The settings of the simulation parameters in *Fimmprop-3D* are more transparent, also thanks to the 'calculation diagnostic' tool. The possibility of visualising the local modes at any position along the device has the advantage of providing physical insight in the device behaviour. The optical device is fully characterised by the local modes and the coupling matrices between them. Once these coupling matrices are constructed, a wide variety of information can be obtained without performing further heavy calculations. For instance –and this is very important during the design procedure– a stretching of the device does not mean that the device has to be re-simulated from scratch as is the case in BPM simulations and the new results are calculated in a matter of seconds.

References

[1] K. DeMesel, I. Moerman, R. Baets, B. Dhoedt, P. VanDaele, and J. Stulemijer, "Spot size converters for low cost PICs," in *Proceedings European Conference on Integrated Optics '99*, no. 5, (Torino, Italy), pp. 253–258, April 1999. paper ThC2.

[2] I. Moerman, P. VanDaele, and P. Demeester, "A review on fabrication technologies for the monolithic integration of tapers with III-V semiconductor devices," *IEEE J. Sel. Top. Quant. Electron.*, vol. 3, pp. 1308–1320, December 1997.

[3] J. Buus, W. Stewart, J. Haes, J. Willems, and R. Baets, "Spot size expansion for laser-to-fiber coupling using an integrated multimode coupler," *IEEE J. Lightwave Techn.*, vol. 11, no. 4, pp. 582–588, 1993.

[4] T. Nakamura, C. Xu, X. Li, and W. Huang, "A spot-size transformer based on mode interference," *IEEE Photon. Technol. Lett.*, vol. 10, no. 3, pp. 325–327, 1998.

[5] C. Wei, F. Groen, M. Smit, I. Moerman, P. VanDaele, and R. Baets, "Integrated optical elliptic couplers: modeling, design and applications," *IEEE J. Lightwave Techn.*, vol. 15, no. 5, pp. 906–912, 1997.

[6] M. Spühler, B. Offrein, G. Bona, R. Germann, I. Massarek, and D. Erni, "A very short planar silica spot-size converter using a nonperiodic segmented waveguide," *IEEE J. Lightwave Techn.*, vol. 16, no. 9, pp. 1680–1685, 1998.

[7] J. Walpole, Z. Liau, J. Missaggia, and D. Yap, "Diode lasers with cylindrical mirror facets and reduced beam divergence," *Appl. Phys. Lett.*, vol. 50, no. 18, pp. 1219–1221, 1987.

[8] D. Tishinin, K. Uppal, I. Kim, and P. Dapkus, "1.3μm polarization insensitive amplifiers with integrated-mode transformers," *IEEE Photon. Technol. Lett.*, vol. 9, no. 10, pp. 1337–1339, 1997.

[9] R. Zengerle, O. Leminger, W. Weiershausen, K. Faltin, and B. Hübner, "Laterally tapered InP-InGaAsP waveguides for low-loss chip-to-fiber butt coupling: A comparison on different configurations," *IEEE Photon. Technol. Lett.*, vol. 7, pp. 532–534, 1995.

[10] R. Zengerle, B. Hübner, C. Gréus, H. Burkhard, H. Janning, and E. Kuphal, "Monolithic integration of spot-size transformers for highly efficient laser fiber coupling," *Electron. Lett.*, vol. 31, no. 14, pp. 1142–1143, 1995.

[11] M. Kamp, J. Hofmann, A. Forchel, F. Schäfer, and J. Reithmaier, "Low-threshold high-quantum-eficiency laterally gain-coupled InGaAs/AlGaAs distributed feedback lasers," *Appl. Phys. Lett.*, vol. 74, no. 4, pp. 483–485, 1999.

[12] G. Vawter, R. Smith, H. Hou, and J. Wendt, "Semiconductor laser with tapered-rib adiabatic-following fiber coupler for expanded output-mode diameter," *IEEE Photon. Technol. Lett.*, vol. 9, no. 4, pp. 425–427, 1997.

[13] V. Vusirikala, S. Saini, R. Bartolo, S. Agarwala, R. Whaley, F. Johnson, D. Stone, and M. Dagenais, "1.55-μm InGaAsP-InP laser arrays with integrated-mode expanders fabricated using a single epitaxial growth," *IEEE J. Sel. Top. Quant. Electron.*, vol. 3, no. 6, pp. 1332–1343, 1997.

[14] T. Takiguchi, T. Itagaki, M. Takemi, A. Takemoto, Y. Miyazaki, K. Shibata, Y. Hisa, K. Goto, Y. Mihasi, S. Takamiya, and M. Aiga, "Selective-area MOCVD growth for 1.3μm laser diodes with a monolithically integrated waveguide lens," *J. Cryst. Growth*, vol. 170, pp. 705–709, 1997.

[15] S. Kitamura, H. Hatakeyama, K. Hamamoto, T. Susaski, K. Komatsu, and M. Yamaguchi, "Spot-size converter integrated semiconductor optical amplifiers for optical gate applications," *IEEE J. Quantum Electron.*, vol. 35, no. 7, pp. 1067–1074, 1999.

[16] I. Moerman, M. D'Hondt, W. Vanderbauwhede, P. VanDaele, P. Demeester, and W. Hunziker, "InGaAsP/InP-laser strained MQW laser with integrated mode size converter using the shadow masked growth technique," *Electron. Lett.*, vol. 31, no. 6, pp. 452–454, 1995.

[17] L. Mörl, C. Weinert, F. Reier, L. Stoll, and H.-P. Nolting, "Uncladded InGaAsP/InP rib waveguides with integrated thickness tapers for efficient fibre-chip coupling," *Electron. Lett.*, vol. 32, no. 1, pp. 36–38, 1996.

[18] T. VanCaenegem, *Development of monolithic integration techniques by means of selective area growth and MOVPE for the realisation of photonic integrated ciruits for telecom applications.* PhD thesis, Ghent University, 2001.

[19] G. Coudenys, I. Moerman, Z. Yu, F. Vermaercke, P. VanDaele, and P. Demeester, "Atmospheric pressure and low pressure shadow masked growth of InGaAs(P)/InP and (In)GaAs/(Al)GaAs heterostructures and quantum wells," *J. Electron. Mater.*, vol. 23, no. 2, pp. 227–234, 1994.

[20] T. Brenner, W. Hunziker, M. Smit, M. Bachmann, G. Guekos, and H. Melchior, "Vertical InP/InGaAsP tapers for low-loss optical fibre-waveguide coupling," *Electron. Lett.*, vol. 28, no. 22, pp. 2040–2041, 1992.

[21] T. Brenner and H. Melchior, "Integrated optical modeshape adapters in InGaAsP/InP for efficient fiber-to-waveguide coupling," *IEEE Phot. Technol. Lett.*, vol. 5, no. 9, pp. 1053–1056, 1993.

[22] J. Stulemeijer, H. Bakker, I. Moerman, F. Groen, and M. Smit, "Efficient InP-based integrable spot-size converter," in *Technical Digest Integrated Photonics Research 1998*, no. 5, (Victoria, Canada), pp. 192–194, 1998.

[23] T. Koch, U. Koren, G. Eisenstein, M. Young, M. Oron, C. Giles, and B. Miller, "Tapered waveguide InGaAs/InGaAsP multiple-quantum-well lasers," *IEEE Photon. Technol. Lett.*, vol. 2, no. 2, pp. 88–90, 1990.

[24] M. Yoshimoto, K. Kawano, H. Takeuchi, S. Kondo, and Y. Noguchi, "Highly efficient coupling semiconductor spot-size converter with an InP/InAlAs multiple-quantum-well core," *Appl. Opt.*, vol. 34, no. 6, pp. 1007–1014, 1995.

[25] G. Wenger, L. Stoll, B. Weiss, M. Schienle, R. Müller-Nawrath, S. Eichinger, J. Müller, B. Acklin, and G. Müller, "Design and fabrication of monolithic optical spot size transformers (MOST's) for highly efficient fiber-chip coupling," *J. Lightwave Technol.*, vol. 12, no. 10, pp. 1782–1790, 1994.

[26] J. Kim, J. Lee, S. Park, J. Yu, S. Lee, A. Choo, T. Kim, and Y. Lee, "Spot-size converter integrated polarization insensitive semiconductor optical amplifiers," *IEEE Photon. Technol. Lett.*, vol. 11, no. 8, pp. 967–969, 1999.

[27] H. Jeon, A. Mathur, and M. Ziari, "High power narrow divergence DFB laser diode at 1.55μm," *Electron. Lett.*, vol. 34, no. 13, pp. 1313–1315, 1998.

[28] J. Stulemeijer, H. Blok, I. Moerman, and M. Smit, "Use of coupled mode approach to the design of spot-size converters," in *Proceedings European Conference on Integrated Optics '99*, no. 5, (Torino, Italy), pp. 361–364, 1999. paper ThI3.

[29] G. Guekos, *Photonic devices for telecommunications: how to model and measure*. Berlin, Germany: Springer-Verlag, 1999.

[30] Photon Design, "http://www.photond.com."

[31] Optiwave Corporation, "http://www.optiwave.com."

Chapter 3

Adiabaticity criterion

3.1 Introduction

A simple adiabaticity criterion for the design of tapered waveguides has been derived by Love et al. [1]. This reference has been cited many times with the comment that the criterion is intuitively clear, but we believe that a rather strong intuition is needed for a full understanding. We therefore derive and discuss this criterion here in a rigorous way by using the theory of coupled local modes.
This chapter is based on theory found in literature [1][2], but we try to bring it here in a more structured and didactic form and add some examples for illustration. We will first introduce the concept of local modes and then set up the coupled local mode equations from which the adiabaticity criterion will be constructed.

3.2 Local modes

The behaviour of electro-magnetic radiation is described by Maxwell's equations. The complexity of this set of coupled differential equations makes the treatment of a general waveguide a tough task requiring several numerical techniques to come to aid.
One of the simplest problems is the waveguide which is invariant in the direction of propagation z (figure 3.1). When monochromatic radiation with a $e^{j\omega t}$ time dependence is considered we can write the fields in such a waveguide in the separable form

$$\begin{aligned} \mathbf{E}(x,y,z) &= \mathbf{E}(x,y)\ e^{-j\beta z} \\ \mathbf{H}(x,y,z) &= \mathbf{H}(x,y)\ e^{-j\beta z} \end{aligned} \tag{3.1}$$

These are the so-called 'eigenmodes' of the structure, which are the field profiles that can propagate in the absence of any sources and that maintain their general shape during propagation. They are characterised by their propagation constant,

$$\beta = \frac{2\pi}{\lambda}\, n_{eff} \tag{3.2}$$

with λ the wavelength of the light and n_{eff} the effective index of the mode.

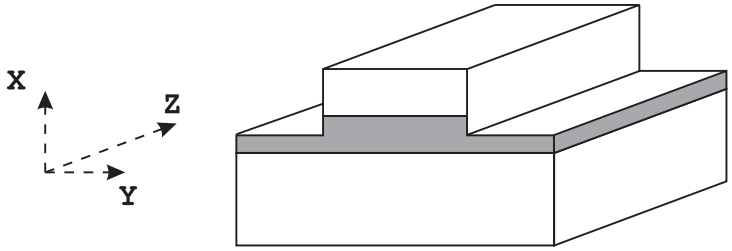


Figure 3.1: A z-invariant structure.

A spot-size converter, on the other hand, is a structure with a z-dependent cross-section. A general solution of Maxwell's equation does not exist for such a structure and modes as described in (3.1) are no longer supported.
We can, however, approximate the structure by a series of z-independent sections as shown in figure 3.2. Eigenmodes of the form (3.1) are again the solutions of Maxwell's equations within each section. We call them 'local modes' because they are the solutions of an infinitely long z-invariant waveguide with the same cross-section $n_i(x,y)$ as the local cross-section $n(x,y,z_i)$. The exact fields at a position $z = z_i$ are expressible as a superposition of the complete set of forward and backward propagating local modes:

$$\mathbf{E}(x,y,z_i) = \mathbf{E}_+(x,y,z_i) + \mathbf{E}_-(x,y,z_i) \tag{3.3}$$

$$\begin{aligned}\mathbf{E}_+(x,y,z_i) &= \sum_{j=0}^{n-1} b_{+j}(z_i)\,\widehat{\mathbf{e}}_{j+}(x,y,\beta_j(z_i)) \\ &\quad + \int_{\beta_{rad}} b_+(\beta(z_i))\,\widehat{\mathbf{e}}_+(x,y,\beta(z_i))d\beta(z_i)\end{aligned} \tag{3.4}$$

$$\begin{aligned}\mathbf{E}_-(x,y,z_i) &= \sum_{j=0}^{n-1} b_{-j}(z_i)\,\widehat{\mathbf{e}}_{j-}(x,y,\beta_j(z_i)) \\ &\quad + \int_{\beta_{rad}} b_-(\beta(z_i))\,\widehat{\mathbf{e}}_-(x,y,\beta(z_i))d\beta(z_i)\end{aligned} \tag{3.5}$$

with $\mathbf{E}_+(x,y,z_i)$ the total forward and $\mathbf{E}_-(x,y,z_i)$ the total backward propagating field. Both contributions consist of a discrete sum of the n guided local modes and an integral over the continuum of local radiation modes. The factor $b(z_i)$ contains the amplitude and phase of the particular mode.
The field profiles $\widehat{\mathbf{e}}_j$ and $\widehat{\mathbf{h}}_j$ are also a function of the z-coordinate and hence of the local propagation constant $\beta_j(z)$. The $\widehat{\ }$ indicates that the orthogonal modes are normalised such that they carry unit power. The orthonormality condition for non-absorbing waveguides is expressed as

$$\frac{1}{2}\int_{S_\infty} (\widehat{\mathbf{e}}_j \times \widehat{\mathbf{h}}_k)\cdot\hat{\mathbf{z}}\,dS = \frac{1}{2}\int_{S_\infty} (\widehat{\mathbf{e}}_k \times \widehat{\mathbf{h}}_j)\cdot\hat{\mathbf{z}}\,dS = \delta_{jk} \tag{3.6}$$

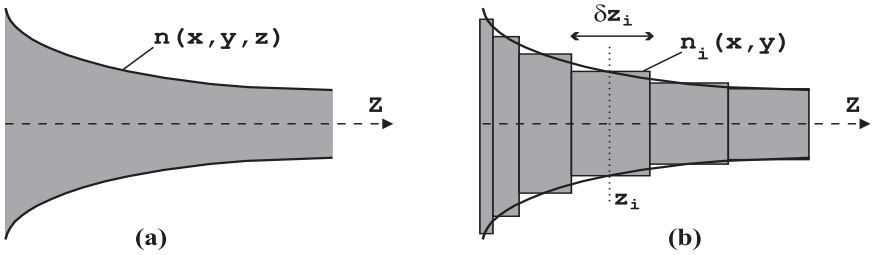


Figure 3.2: (a) A z-dependent structure and (b) the approximation as a series of z-independent sections.

with S_∞ the infinite cross-sectional plane, $\widehat{\mathbf{z}}$ the unit vector along the propagation axis and δ_{ij} the Kronecker-δ.
The phase of mode j increases over a section as the product of the section length δz_i with the local propagation factor $\beta_j(z_i)$. For slowly varying devices like the adiabatic mode converters, the power in each local mode remains virtually constant as the mode propagates (there is practically no interaction with the other modes). The phase of a particular mode at an arbitrary position along the structure is in this case a sum of such products. If the sections are chosen infinitesimally small, then this phase can be replaced by the integral $\int_0^z \beta_j(\xi)d\xi$.

All these considerations lead, for a perfectly adiabatic modeconverter, to the following notation of the local modes at position z:

$$\begin{aligned}
\mathbf{E}_j(x,y,z) &= b_j(z)\,\widehat{\mathbf{e}}_j(x,y,\beta_j(z)) \\
&= a_j\, e^{i\int_0^z \beta_j(\xi)\,d\xi}\ \widehat{\mathbf{e}}_j(x,y,\beta_j(z)) \\
\mathbf{H}_j(x,y,z) &= b_j(z)\,\widehat{\mathbf{h}}_j(x,y,\beta_j(z)) \\
&= a_j\, e^{i\int_0^z \beta_j(\xi)\,d\xi}\ \widehat{\mathbf{h}}_j(x,y,\beta_j(z)) \qquad (3.7)
\end{aligned}$$

with a_j real. Note that the coefficients a_j become a function of z in the case of a non-adiabatic device (see section 3.3).
The local mode concept as presented above already gives a basic description of the field evolution through an adiabatic spot-size converter. The local mode profile is described here as a field that gradually changes its shape as it propagates, without loss of power and without interaction with other modes.

3.3 Local mode coupling

The local mode theory describing modes that propagate without change in power is a good approximation of a slowly varying waveguide, but it is not an exact solution. Only for a truly z-invariant structure will the modal power remain constant. A z-dependent device will always exhibit an interaction between local modes, re-

sulting in power being coupled from one mode to another. A correction to the local mode approximation that takes this complicated behaviour of power exchange into account is given by the coupled local mode equations and will be discussed now.
Each local mode will propagate with constant power through a section until the interface with the next section is reached. The latter has a slightly different geometry with corresponding different mode profiles, resulting in power being coupled from the incident field to a whole set of forward propagating local modes in the next section. Also reflected fields will be generated at the transition (figure 3.3). This mode coupling has been neglected in the local mode description of section 3.2, but will now be taken into account and will lead to the coupled local mode equations.
Consider the j^{th} forward propagating local mode with amplitude $a_j(z)$ which is incident on a section at z. The change in $b_j(z)$ across the section follows from (3.7) and is given by

$$db_j = i\beta_j\, b_j\, dz + da_j\, e^{i\int_0^z \beta_j(\xi)d\xi} \tag{3.8}$$

The first term on the right accounts for the change in phase and the second for the change in amplitude across the section. The change in amplitude da_j originates from coupling of power from mode j to other modes (both backward and forward propagating) and coupling from the other modes to mode j.
To evaluate da_j we first examine the transmitted and reflected fields when the k^{th} forward-propagating mode is incident on the section (figure 3.3). This will yield an expression for the fraction of the incident power that is coupled from the k^{th} to the j^{th} mode. The sum of all such contributions of the different incident modes will then lead to the total da_j.

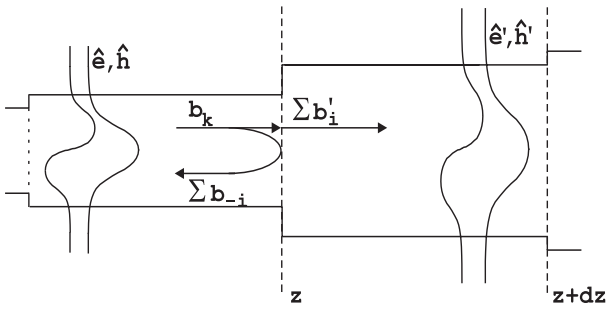


Figure 3.3: Fields at an interface

We know from Maxwell's equation $\nabla \cdot \mathbf{D} = 0$ that the total transverse electric field is continuous across the interface: $\widehat{\mathbf{E}}_t = \widehat{\mathbf{E}}_t'$. $\widehat{\mathbf{E}}_t$ is composed of the fields of the k^{th} forward propagating local mode and all backward propagating modes at z, while $\widehat{\mathbf{E}}_t'$ is composed of all forward propagating local modes at z+dz. For simplicity we only take the guided modes into account, but coupling to the radiation field is

readily included. The continuity relation can be written as

$$b_k\,\widehat{\mathbf{e}}_{tk} + \sum_i b_{-i}\,\widehat{\mathbf{e}}_{t-i} = \sum_i b_i^{'}\,\widehat{\mathbf{e}}_{ti}^{'} \tag{3.9}$$

Positive subscripts indicate forward propagating local modes, negative subscripts indicate backward-propagating local modes. The 't'-subscript denotes the transverse component of the field, corresponding to the field notation with the transverse and longitudinal component separated, i.e. $\widehat{\mathbf{e}} = \widehat{\mathbf{e}}_t + e_z \cdot \widehat{\mathbf{z}}$ and $\widehat{\mathbf{h}} = \widehat{\mathbf{h}}_t + h_z \cdot \widehat{\mathbf{z}}$. It can be shown from Maxwell's equations that the fields of the backward propagating modes can be written in terms of the forward propagating modes in the following way:

$$\begin{aligned} \widehat{\mathbf{e}}^{-} &= \widehat{\mathbf{e}}_t^{+} - e_z^{+}\widehat{\mathbf{z}} \\ \widehat{\mathbf{h}}^{-} &= -\widehat{\mathbf{h}}_t^{+} + h_z^{+}\widehat{\mathbf{z}} \end{aligned} \tag{3.10}$$

This means that $\widehat{\mathbf{e}}_{t-i}$ can be replaced by $\widehat{\mathbf{e}}_{ti}$ and that (3.9) can be rewritten as a function of the forward propagating field components only:

$$b_k\,\widehat{\mathbf{e}}_{tk} + \sum_i b_{-i}\,\widehat{\mathbf{e}}_{ti} = \sum_i b_i^{'}\,\widehat{\mathbf{e}}_{ti}^{'} \tag{3.11}$$

We cross-multiply (3.11) with $\widehat{\mathbf{h}}_j^{'}/2$ and integrate over the infinite cross-section. Making use of the orthonormality condition of (3.6), we find

$$b_j^{'} = \frac{b_k}{2}\int_{S_\infty}\left(\widehat{\mathbf{e}}_k \times \widehat{\mathbf{h}}_j^{'}\right)\cdot\widehat{\mathbf{z}}\,dS + \sum_i \frac{b_{-i}}{2}\int_{S_\infty}\left(\widehat{\mathbf{e}}_{-i} \times \widehat{\mathbf{h}}_j^{'}\right)\cdot\widehat{\mathbf{z}}\,dS \tag{3.12}$$

The transverse field components of (3.11) have been replaced by the complete field vectors in (3.12). This can be done because the longitudinal components do not contribute to the integral.
The field $\widehat{\mathbf{h}}_j^{'}$ can in a first order approximation also be written as

$$\widehat{\mathbf{h}}_j^{'} = \widehat{\mathbf{h}}_j + \frac{\partial\widehat{\mathbf{h}}_j}{\partial z}\,dz \tag{3.13}$$

Substitution of (3.13) into (3.12) gives

$$\begin{aligned} b_j^{'} &= \frac{b_k}{2}\left(\int_{S_\infty}\left(\widehat{\mathbf{e}}_k \times \widehat{\mathbf{h}}_j\right)\cdot\widehat{\mathbf{z}}\,dS + \int_{S_\infty}\left(\widehat{\mathbf{e}}_k \times \frac{\partial\widehat{\mathbf{h}}_j}{\partial z}dz\right)\cdot\widehat{\mathbf{z}}\,dS\right) + \\ &\quad \sum_i \frac{b_{-i}}{2}\left(\int_{S_\infty}\left(\widehat{\mathbf{e}}_{-i} \times \widehat{\mathbf{h}}_j\right)\cdot\widehat{\mathbf{z}}\,dS + \int_{S_\infty}\left(\widehat{\mathbf{e}}_{-i} \times \frac{\partial\widehat{\mathbf{h}}_j}{\partial z}dz\right)\cdot\widehat{\mathbf{z}}\,dS\right) \end{aligned} \tag{3.14}$$

and further application of (3.6) yields

$$\begin{aligned} b_j^{'} &= b_k\,D_{kj}\,dz + b_{-j} + \sum_i b_{-i}\,D_{-ij}\,dz \qquad j \neq k, \\ b_j^{'} &= (1 + D_{jj}\,dz)\,b_j + b_{-j} + \sum_i b_{-i}\,D_{-ij}\,dz \qquad j = k \end{aligned} \tag{3.15}$$

with $D_{ij} = \frac{1}{2}\int_{S_\infty} (\widehat{\mathbf{e}}_i \times \frac{\partial \widehat{\mathbf{h}}_j}{\partial z}) \cdot \widehat{\mathbf{z}}\, dS$. Note that $D_{-ij} = D_{ij}$ since $\widehat{\mathbf{e}}_{t-i} = \widehat{\mathbf{e}}_{ti}$ (see (3.10)). The b-factors contain the modal amplitude and phase. The phases of the incident, reflected and transmitted modes, however, are equal as a result of the flat phase fronts in the non-absorbing waveguides. This means that (3.15) can be rewritten in terms of the amplitude coefficients:

$$\begin{aligned} a_j^{'} &= a_k\, D_{kj}\, dz + a_{-j} + \sum_i a_{-i}\, D_{-ij}\, dz \qquad j \neq k, \\ a_j^{'} &= (1 + D_{jj}\, dz)\, a_j + a_{-j} + \sum_i a_{-i}\, D_{-ij}\, dz \qquad j = k \end{aligned} \tag{3.16}$$

A second set of equations is obtained in a similar manner by writing down the continuity of the tangential magnetic field across the interface (follows from $\nabla \cdot \mathbf{B} = 0$). The corresponding equation is

$$b_k\, \widehat{\mathbf{h}}_{tk} - \sum_i b_{-i}\, \widehat{\mathbf{h}}_{ti} = \sum_i b_i^{'}\, \widehat{\mathbf{h}}_{ti}^{'} \tag{3.17}$$

Note that $\widehat{\mathbf{h}}_{t-i}$ has been replaced by $-\widehat{\mathbf{h}}_{ti}$ (see (3.10)). A similar reasoning as for the tangential electric field yields a second pair of equations:

$$\begin{aligned} a_j^{'} &= -a_k\, F_{kj}\, dz - a_{-j} - \sum_i a_{-i}\, F_{-ij}\, dz \qquad j \neq k, \\ a_j^{'} &= (1 - F_{jj}\, dz)\, a_j - a_{-j} - \sum_i a_{-i}\, F_{-ij}\, dz \qquad j = k \end{aligned} \tag{3.18}$$

with $F_{ij} = \frac{1}{2}\int_{S_\infty} (\widehat{\mathbf{h}}_i \times \frac{\partial \widehat{\mathbf{e}}_j}{\partial z}) \cdot \widehat{\mathbf{z}}\, dS$ and $F_{-ij} = -F_{ij}$ since $\widehat{\mathbf{h}}_{t-i} = -\widehat{\mathbf{h}}_{ti}$ (see (3.10)). If we add the corresponding equations from eqs. (3.16) and (3.18), we find

$$\begin{aligned} 2a_j^{'} &= a_k\, (D_{kj} - F_{kj})\, dz + \sum_i a_{-i}\, (D_{-ij} - F_{-ij})\, dz \quad j \neq k, \\ 2(a_j^{'} - a_j) &= \sum_i a_{-i}\, (D_{-ij} - F_{-ij})\, dz \qquad\qquad j = k \end{aligned} \tag{3.19}$$

The left hand sides of the equations are equal to $2da_j$ in both cases ($a_j = 0$ for $j \neq k$). If we introduce again the phase factors (which are the same for all modes, as explained in the comments below (3.15)), we can write

$$2\, da_j\, e^{i\int_0^z \beta_j(\xi)\, d\xi} = (D_{kj} - F_{kj})\, b_k\, dz + \sum_i (D_{-ij} - F_{-ij})\, b_{-i}\, dz \tag{3.20}$$

(3.20) expresses the change in a_j originating from the incident k^{th} local mode. The total change in a_j is then found by summing over k the contributions of all incident modes. Substitution of the resulting total da_j in (3.8) finally leads to the coupled local mode equations

$$\frac{db_j}{dz} - i\beta_j b_j = \sum_k \left(C_{jk} b_k + C_{j-k} b_{-k}\right) \tag{3.21}$$

for forward propagating modes. A similar equation can be derived for backward propagating modes:

$$\frac{db_{-j}}{dz} + i\beta_j b_{-j} = -\sum_k \left(C_{-jk} b_k + C_{-j-k} b_{-k}\right) \tag{3.22}$$

with

$$C_{jk} = \frac{1}{4}\int_{S_\infty} \left(\widehat{\mathbf{h}}_j \times \frac{\partial \widehat{\mathbf{e}}_k}{\partial z} - \widehat{\mathbf{e}}_j \times \frac{\partial \widehat{\mathbf{h}}_k}{\partial z}\right) \cdot \widehat{\mathbf{z}}\, dA \tag{3.23}$$

Note that $C_{jj} = 0$ because it is proportional to the z-derivative of the unit normalisation of (3.6). It is also worth mentioning that $C_{jk} = -C_{kj}$[1]. Indeed, when the amplitude of mode j *in*creases because of coupling of power from mode k, then the amplitude of mode k will *de*crease with the same amount. It can also be shown from Maxwell's equations that the components of $\widehat{\mathbf{e}}$ and $\widehat{\mathbf{h}}$ in a non-absorbing waveguide can always be chosen such that $\widehat{\mathbf{e}}_t$ and $\widehat{\mathbf{h}}_t$ are real and e_z and h_z are purely imaginary. We will adopt this convention from now on, implying that the coupling coefficients of (3.23) are real.
Our intuitive feeling that a rapidly changing waveguide exhibits strong coupling between local modes is clearly confirmed through the presence of the modal derivatives in (3.23). The expression for the coupling coefficients can be further transformed into an alternative equation incorporating the derivative of the refractive index profile [2]:

$$C_{jk} = \frac{k_0}{4}\left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2} \frac{1}{\beta_j - \beta_k}\int_{S_\infty} \widehat{\mathbf{e}}_j \cdot \widehat{\mathbf{e}}_k \frac{\partial n^2}{\partial z}\, dS, \quad j \neq k \qquad C_{jj} = 0 \tag{3.24}$$

with $k_0 = \frac{2\pi}{\lambda}$ the free-space wavenumber, ε_0 the free-space dielectric constant, μ_0 the free-space permeability and $n^2 = \varepsilon_r$ the relative dielectric constant of the material.

3.4 Adiabaticity criterion

The coupled local mode equations form the basis for the construction of a criterion for slow variation, i.e. a delineation of 'how fast' the cross section of a spot-size converter is allowed to change before power is lost from the mode of interest.
An approximate solution of the coupled local mode equations can be found for the regime of weak power transfer between local modes. This is the regime that applies to an adiabatic spot-size converter.
Consider the case where only the l^{th} order mode is excited at the beginning of the device (z=0). The power in all other modes is assumed to be zero. The lowest order description of the field evolution through the device is given by the local mode theory as presented in (3.7), implying that coupling between modes is being

[1] This can easily be deduced by differentiating (3.6) with respect to z.

neglected and that the power in each local mode remains constant. The coupling coefficients in (3.21) are zero in this case and the solution to the equation is

$$b_l(z) = b_l(0)\, e^{i\int_0^z \beta_l(\xi)d\xi}; \qquad b_j(z) = 0, j \neq l \tag{3.25}$$

The more accurate solution with first order coupling taken into account is found by substituting (3.25) into (3.21). Higher order coupling effects are neglected, meaning that only coupling from the l^{th} order mode to the other modes is considered. The equations to be solved are then

$$\frac{db_{\pm j}(z)}{dz} \mp i\, \beta_j(z)\, b_{\pm j}(z) = \pm b_l(0)\, C_{\pm jl}(z)\, e^{i\int_0^z \beta_l(\xi)d\xi} \tag{3.26}$$

(3.26) can also be written as

$$e^{\pm i\int_0^z \beta_j(\xi)d\xi}\left[\frac{d\Big(b_{\pm j}(z)\, e^{\mp i\int_0^z \beta_j(\xi)d\xi}\Big)}{dz}\right] = \pm b_l(0)\, C_{\pm jl}(z)\, e^{i\int_0^z \beta_l(\xi)d\xi} \tag{3.27}$$

The solution for $b_l(0) = 1$, satisfying $b_{\pm j}(0) = 0$, is given by

$$b_{\pm j}(z) = \pm e^{\pm i\int_0^z \beta(\xi)d\xi} \int_0^z C_{\pm jl}(\xi)\, e^{i\overline{\delta\beta}(\xi)\,\xi} d\xi \tag{3.28}$$

with $\overline{\delta\beta}$ the average value of the accumulated difference in propagation constants:

$$\overline{\delta\beta}(z) = \frac{1}{z}\int_0^z \Big\{\beta_l(\xi) \mp \beta_j(\xi)\Big\} d\xi \tag{3.29}$$

With (3.28) and (3.29) we are now in a position to calculate the power in the j^{th} order mode at position z. In a true adiabatic device this power should remain zero, but in a real spot-size converter there will always be some power lost from the l^{th} order mode to the other modes. We will now come to the criterion for 'adiabatic' mode conversion by requiring that the power in the j^{th} order mode remains small. The power flow in a non-absorbing waveguide is parallel to the waveguide axis and the power density I_j of local mode j is given by the magnitude of the time-averaged Poynting-vector:

$$I_j = \frac{1}{2}\, Re\Big\{\mathbf{E}_j \times \mathbf{H}_j^* \cdot \widehat{\mathbf{z}}\Big\} \tag{3.30}$$

with $\mathbf{E}_j$ and $\mathbf{H}_j$ given by (3.7) and * indicating the complex conjugate. The total power P_j in the forward propagating mode j is found by integrating I_j over the infinite cross-section. Keeping the orthonormality condition of (3.6) in mind, we find

$$\begin{aligned} P_j &= \frac{1}{2}\,|b_j|^2 \int_{S_\infty} \widehat{\mathbf{e}}_j \times \widehat{\mathbf{h}}_j^* \cdot \widehat{\mathbf{z}}\, dS \\ &= |b_j|^2 \end{aligned} \tag{3.31}$$

In the derivation of (3.28) we assumed $b_l(0) = 1$, so that the fraction of power coupled to the j^{th} forward-propagating mode follows from (3.31) and (3.28) as

$$P_j(z) = \left(\int_0^z C_{jl}(\xi) e^{i\overline{\delta\beta}(\xi)\xi} d\xi \right) \left(\int_0^z C_{jl}(\xi) e^{-i\overline{\delta\beta}(\xi)\xi} d\xi \right); \quad j \neq l \tag{3.32}$$

The waveguide is slowly varying so that $\overline{\delta\beta}(z)$ and $C_{jl}(z)$ are also weak functions of z. This allows us to neglect $\partial\overline{\delta\beta}(z)/\partial z$ so that (3.32) can be rewritten as

$$P_j(z) = \left(\frac{1}{i\overline{\delta\beta}(\xi)} \int_0^z C_{jl}(\xi) de^{i\overline{\delta\beta}(\xi)\xi} \right) . \left(c.c. \right) \tag{3.33}$$

with 'c.c.' indicating the complex conjugate of the first factor. Partial integration yields

$$P_j(z) = \left(\frac{1}{i\overline{\delta\beta}(z)} \left[C_{jl}(z) e^{i\overline{\delta\beta}(z)z} - C_{jl}(0) \right] - \frac{1}{i\overline{\delta\beta}(z)} \int_0^z e^{i\overline{\delta\beta}(\xi)\xi} \frac{\partial C_{jl}(\xi)}{\partial \xi} d\xi \right) . \left(c.c. \right) \tag{3.34}$$

The integral in (3.34) can be dropped since the amplitude of the rapidly oscillating integrand, $\left(\partial C_{jl}(z)/\partial z\right).\left(1/\overline{\delta\beta}(z)\right)$, is negligibly small as explained above. (3.34) then finally reduces to

$$P_j(z) \simeq F^2(z) + F^2(0) - 2F(z)F(0)cos(\overline{\delta\beta}(z)z); \quad F(z) = \frac{C_{jl}(z)}{\overline{\delta\beta}(z)} \tag{3.35}$$

It can be concluded from (3.35) that the power fraction in mode j remains negligibly small if $|F(z)| \ll 1$, or $C_{jl}(z) \gg \overline{\delta\beta}(z)$, everywhere along the waveguide. $C_{jl}(z)$ is most sensitive to the change in waveguide profile with z (3.24) and $\overline{\delta\beta}(z)$ determines the average beat length $2\pi/\overline{\delta\beta}(z)$ between the j^{th} and l^{th} mode. This brings us to the following conclusion, which is also known as the adiabaticity criterion:

> ***A local mode will propagate with negligible loss to another local mode provided that the waveguide changes over a distance that is large compared to the local beat length between both modes.***

3.5 Practical design criterion

The general adiabaticity criterion can be elaborated for the more specific case of an adiabatic spot-size converter that transforms the fundamental guided mode. The criterion states that the waveguide cross-section should not change significantly over a distance equal to the local beat length. The beat length to be taken into account is the the longest beating phenomenon occurring in the device, i.e. beating between the guided mode and the mode with the closest propagation constant β_{max}. The corresponding beat length z_b is given by $z_b = 2\pi/(\beta_g - \beta_{max})$, where

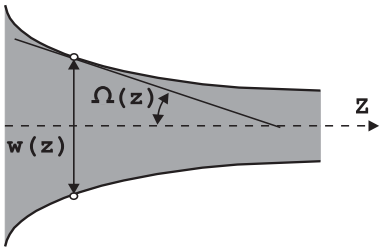


Figure 3.4: A typical spot-size converter with indication of the local taper angle $\Omega(z)$.

β_g is the propagation constant of the guided mode.
Consider a spot-size converter with a gradually narrowing core as sketched in figure 3.4. The adiabaticity criterion translates in this case to

$$\left|\frac{\delta w}{w}\right| \ll 1 \tag{3.36}$$

with δw the width variation corresponding to $\delta z = z_b$. Accordingly, (3.36) can be written as

$$\left|\frac{1}{w}\,\frac{dw}{dz}\,z_b\right| \ll 1\,, \quad \text{or} \quad \left|\frac{1}{w}\,\frac{dw}{dz}\,\frac{2\pi}{(\beta_g - \beta_{r,max})}\right| \ll 1 \tag{3.37}$$

Knowing that $dw(z)/dz = 2\,tan\Omega(z) \approx 2\,\Omega(z)$, with $\Omega(z)$ the local taper angle relative to the z-direction, we conclude that the following condition must be fulfilled to ensure an adiabatic mode transformation:

$$\Omega(z) \ll \frac{w(z)\,(\beta_g - \beta_{max})}{4\pi} \tag{3.38}$$

The adiabaticity criterion is only a qualitative criterion that provides insight into the required slowness in variation of the cross-section: it determines the shape, but an exact length-scale is not given because it depends on the interpretation of '$\ll$' in (3.38), or, on the transformation loss one allows. This will be made clear in the following examples.

Example 1: Design of an adiabatic multimodal spot-size converter

In this first example we calculate the shape of a broadening multi-mode slab waveguide. The modes in such a broad waveguide can be approximated by pure sine functions, as sketched in figure 3.5. As the waveguide is broadened, power is coupled from the fundamental local mode (the 0^{th} order mode) to the higher order local modes, and predominantly to the mode with the closest propagation constant.
There is no coupling between a laterally symmetric and asymmetric local mode in a laterally symmetric waveguide. This is explained by the presence of $\widehat{\mathbf{e}}_j \cdot \widehat{\mathbf{e}}_k$ in the integrand of (3.24). Knowing that a laterally symmetric local mode is

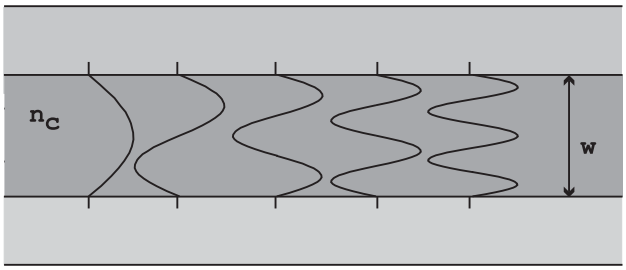


Figure 3.5: A multimode slab waveguide and the sine approximations of the exact fields.

launched at the start of the device, we can conclude that only laterally symmetric local modes will play a role in the conversion process. The mode with the 'closest' propagation constant, to be taken into account in the calculation of the taper shape, is therefore the second order guided mode with propagation constant $\beta_2(w)$.

The propagation constant of the m^{th} order guided mode in the sine approximation is given by

$$\beta_m = \sqrt{\left(\frac{2\pi}{\lambda_0} n_c\right)^2 - \left(\frac{(m+1)\pi}{w}\right)^2} \tag{3.39}$$

With λ_0 the vacuum wavelength, n_c the refractive index of the core and w the core width. The propagation constant difference $\Delta\beta_{0m}$ between the fundamental mode and the m^{th} order mode is then easily calculated as

$$\begin{aligned} \Delta\beta_{0m} &= \beta_0 - \beta_m \\ &= \frac{k_0^2 n_c^2 \pi^2}{2w^2} \left[\, m(m+2)\right] \end{aligned} \tag{3.40}$$

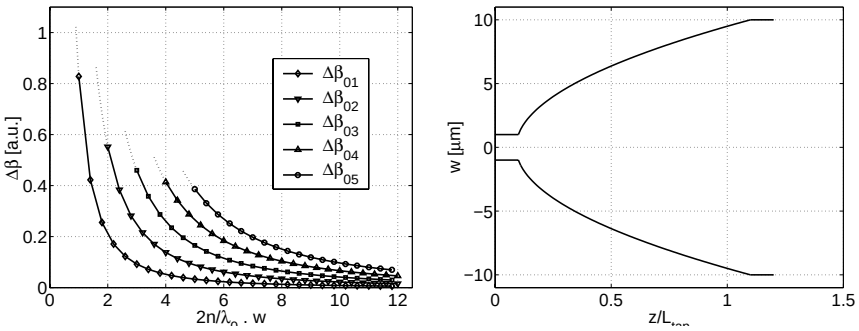
The evolution of $\Delta\beta_{0m}$ is plotted in figure 3.6(a) for the lowest order modes. The propagation constant differences are inversely proportional to w^2, implying that the taper angle will have to be reduced as the waveguide becomes wider. Insertion of (3.40) in (3.36) yields the following evolution of the width w(z) along the taper:

$$w(z) = \sqrt{w_i^2 + \alpha z} \tag{3.41}$$

with w_i the initial taper width and α a constant. In other words: a broad adiabatic multi-mode taper should follow a *parabolic* evolution. The shape of a parabolic taper making a transition between a width w_i and w_f can also be written more specifically as:

$$w(z) = w_i \sqrt{1 + \left(\frac{w_f^2 - w_i^2}{w_i^2}\right) \frac{z}{L}} \tag{3.42}$$

with L the length of the device.



(a) (b)

Figure 3.6: (a) Evolution of the propagation constant difference between the fundamental and higher order guided modes; (b) Calculated taper shape ('parabolic' taper).

It has been mentioned before that the criterion does not provide a strict length scale for the device because it depends on the interpretation of the '$\ll$'-sign in (3.38). In other words, it depends on the maximum loss one can allow. The final device geometry is determined by stretching or compressing the component (by adjusting the parameter α) and calculating the power in the fundamental local mode at the output by means of a BPM tool or a mode expansion technique.

Remark 1 One might argue that the adiabaticity criterion as stated in the conclusion of section 3.4 is only a very rough formulation of the requirement that the term $F(z) = C_{jl}(z)/\overline{\delta\beta}(z)$ in (3.35) should remain negligibly small, because it overlooks the *exact* form of the term F(z). The adiabaticity criterion is on its turn translated rather freely into equation (3.36), which states that the *relative* change of the cross-section should remain small over a distance equal to the beat length. For short, one might question whether the calculated 'ideal' shape is that ideal after all. We will now have a closer look at this matter and reconsider the design of the broadening multimode waveguide of example 1 by making an exact derivation of the taper shape, starting from (3.35).

When an approximately adiabatic mode evolution is pursued, F(z) must remain negligibly small for every z, or:

$$\frac{C_{02}(z)}{\Delta\beta_{02}(z)} \ll 1 \tag{3.43}$$

The expression (3.24) for the coupling coefficient C_{02} states that

$$C_{02} \sim \frac{1}{\Delta\beta_{02}} \int_{S_\infty} \widehat{\mathbf{e}}_0 \cdot \widehat{\mathbf{e}}_2 \frac{\partial n^2}{\partial z}\, dS \tag{3.44}$$

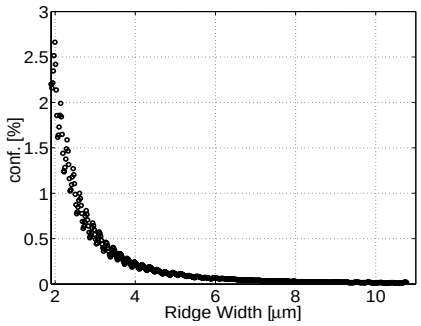


Figure 3.7: Evolution of the optical intensity at the edges of the core region.

The refractive index profile only changes at the edges of the core, meaning that the derivative in the integrand is zero everywhere, except at the edges. This means that the integral is proportional to $\{\widehat{\mathbf{e}}_0 \cdot \widehat{\mathbf{e}}_2\}\frac{\partial w}{\partial z}$, with the scalar product to be evaluated at the edges of the core. The term $\widehat{\mathbf{e}}_0 \cdot \widehat{\mathbf{e}}_2$ is not known in a closed form expression and requires a numerical calculation. We only know that this factor decreases rapidly as the waveguide is widened as a consequence of the stronger confinement in the core region.
To determine the evolution of $\widehat{\mathbf{e}}_0 \cdot \widehat{\mathbf{e}}_2$ we refer to section 4.4 where the realisation of broadening waveguides is discussed. For reasons mentioned there, we have simulated the evolution of the optical power of the fundamental slab mode in a narrow (50nm) region at the edge of the core. The evolution is also shown in figure 3.7 for convenience. Even though this curve is proportional to $\widehat{\mathbf{e}}_0 \cdot \widehat{\mathbf{e}}_0$ and not to $\widehat{\mathbf{e}}_0 \cdot \widehat{\mathbf{e}}_2$, we know that the latter product will show a similar evolution because the confinement of all modes increases rapidly for widening cores. A fit to the simulated data learns that the curve follows a w^{-3}-relationship ($w^{-2.9937}$ to be more precise). All these considerations combined lead to the following formulation of (3.43):

$$\frac{1}{\Delta\beta_{02}^2}\frac{1}{w^3}\frac{\partial w}{\partial z} \ll 1 \tag{3.45}$$

or, with $\Delta\beta_{02} \sim 1/w^2$ (see (3.40))

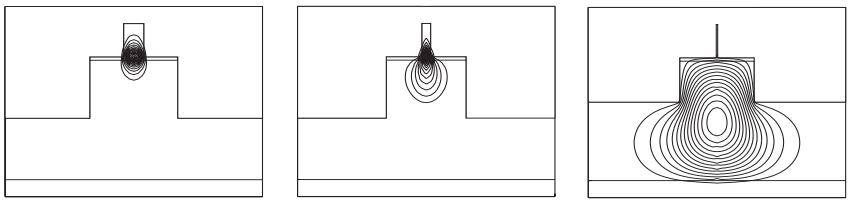
$$w\frac{\partial w}{\partial z} \ll 1 \tag{3.46}$$

In other words, the exact taper shape corresponds with the parabolic evolution calculated from the adiabaticity criterion and given in (3.41) or (3.42) !

Remark 2 The shape of the spot-size converter, according to the adiabaticity criterion, is found by setting F(z)=Constant for every position z. (3.35) then reduces to

$$P_j(z) = F(0)^2\left(1 - \cos(\overline{\delta\beta})\right) \tag{3.47}$$

meaning that power is being coupled resonantly between the fundamental mode and mode j, as is the case in an asymmetric directional coupler. The analogy with



(a) (b) (c)

Figure 3.8: Cross-sections of the spot-size converter: (a) start; (b) middle; (c) end.

the asymmetric coupler, where a constant disturbance of the input waveguide is obtained through the presence of a second straight waveguide, leads to the conclusion that the adiabaticity criterion ensures a constant degree of disturbance along the spot-size converter. The amplitude of the resonantly coupled power is controlled by stretching the device and hence reducing the factor F(0).
A calculation of the device loss would require the investigation of the coupling between all modes (including the radiation modes), which is an impossible task. All that can be said is that, when the strongest coupling (i.e. to the 'closest' mode) is kept low (F(z) $\ll 1$), a quasi-adiabatic behaviour is obtained.

Example 2: Design of an adiabatic spot-size converter
In this example we demonstrate the application of the adiabaticity criterion for the design of a *monomodal* adiabatic spot-size converter. Power will now be lost from the guided mode to the continuum of radiation modes. This also implies that β_{max} is now $\beta_{r,max}$, the propagation constant of the 'closest' radiation mode.
The device in question consists of a broad mesa with a smaller waveguide placed on top[2]. Figure 3.8 shows the cross-sections and the corresponding fundamental local modes at the start, in the middle and at the end of the device. The upper ridge is broad at the start and defines a small and strongly confined mode. This upper waveguide is gradually narrowed towards the end of the spot-size converter and the layer structure is designed such that this causes the fundamental mode to expand and to shift into the large underlying mesa.

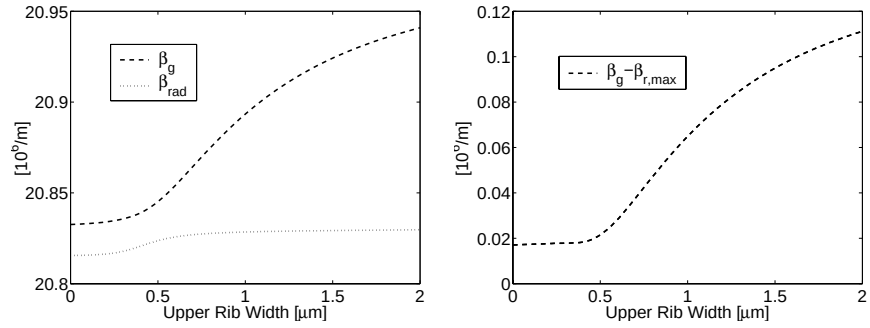
The shape of the spot-size converter is calculated by solving (3.38) in an iterative way (after replacing '$\ll$' by '='). A fixed step Δz is chosen and the corresponding permitted width variation Δw_i at position z_i is then calculated as

$$\Delta w_i = \frac{w_i}{4\pi}\left(\beta_g(w_i) - \beta_{r,max}(w_i)\right)\Delta z \tag{3.48}$$

and

$$\begin{aligned} w_{i+1} &= w_i - \Delta w_i \\ z_{i+1} &= z_i + \Delta z \end{aligned} \tag{3.49}$$

[2]This type of spot-size converter was first reported by Vawter et al. in [3].



(a) (b)

Figure 3.9: (a) Evolution of the propagation constant β_g of the guided local mode and $\beta_{r,max}$ of the closest local radiation mode; (b) Evolution of the propagation constant difference ($\beta_g - \beta_{r,max}$).

The evolution of the propagation constants β_g of the guided local mode and $\beta_{r,max}$ of the 'closest' radiation mode as a function of the upper rib width is calculated with the *Fimmwave* mode solver and is shown in figure 3.9. From this plot we can immediately draw the qualitative conclusion that for a monomodal device a rapid tapering is allowed for wide ridges while a very slow narrowing is required at the end. The calculated longitudinal profile of the spot-size converter is shown in figure 3.10 (a)[3]. Simulations with the *Fimmprop-3D* mode expansion program show that the device should be at least $430\mu m$ long if less than 5% (0.22dB) loss is required. The power in the fundamental local mode along the device and the evolution of the optical field intensity are shown in figure 3.10 (b)-(d).
The same quantities are displayed in figure 3.11 for a $900\mu m$ long linear spot-size converter. Although the device is twice as long as the designed non-linear component, only 85% (0.7dB loss) of the input power is found in the desired mode at the output.
The superior behaviour of the non-linear device appears clearly when the corresponding plots of figure 3.10 and 3.11 are compared. The optical mode in the non-linear converter changes its shape from the very beginning of the component and the transformation occurs continuously over its entire length. The first section of the linear device, on the other hand, shows no mode transformation at all, while the ridge is narrowing too fast at the end.

[3]The shape as shown in figure 3.10 (a) is not entirely correct for narrow ridge widths because the confinement factors in the upper ridge of the guided and the radiation mode of interest are very low. This implies that the upper ridge has practically no further influence on these modes (this can also be seen in the flat evolution of the propagation constants in figure 3.9) and that the ridge can be tapered very sharply or even stopped at $w = 0.3\mu m$.

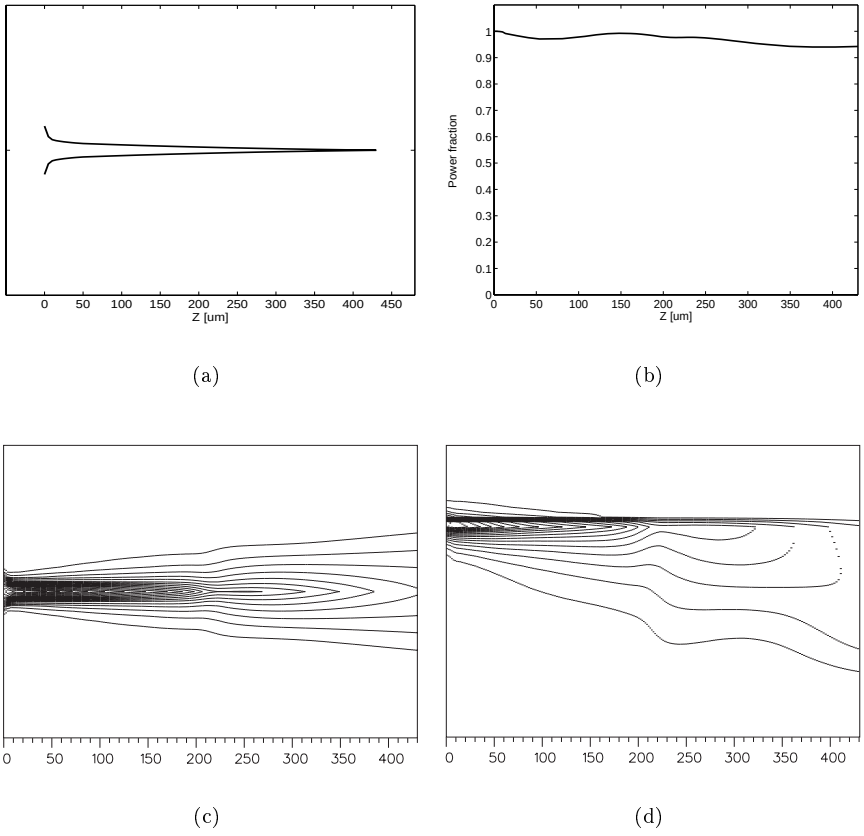


Figure 3.10: (a) Calculated shape of the adiabatic spot-size converter; (b) Evolution of the fundamental local mode power; (c) Top view showing the lateral evolution of the optical field intensity; (d) Side view showing the transverse evolution of the optical field intensity.

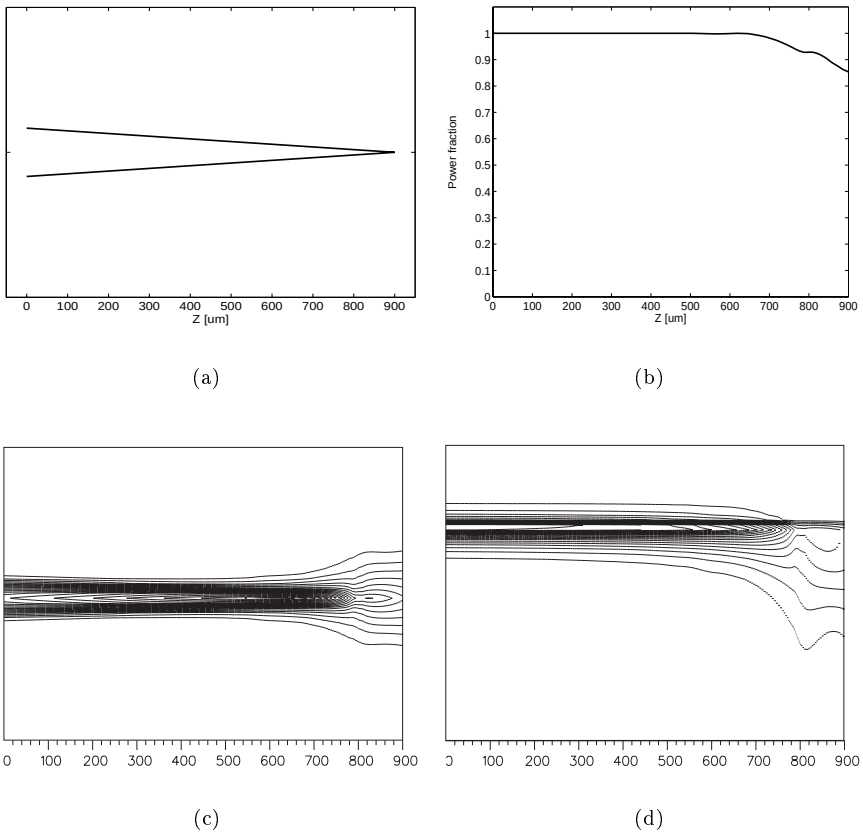


Figure 3.11: (a) Shape of a $900\mu m$ long linear spot-size converter; (b) Evolution of the power in the fundamental local mode; (c) Top view showing the lateral evolution of the optical field intensity; (d) Side view showing the transverse evolution of the optical field intensity.

References

[1] J. Love, W. Henry, W. Stewart, R. Black, S. Lacroix, and F. Gonthier, "Tapered single-mode fibres and devices. part 1: Adiabaticity criteria," *IEE Proc.-J*, vol. 138, no. 5, pp. 343–354, 1991.

[2] A. Snyder and J. Love, *Optical waveguide theory.* No. 6, London, England: Chapman and Hall Ltd., 1983.

[3] G. Vawter, R. Smith, H. Hou, and J. Wendt, "Semiconductor laser with tapered-rib adiabatic-following fiber coupler for expanded output-mode diameter," *IEEE Photon. Technol. Lett.*, vol. 9, no. 4, pp. 425–427, 1997.

Chapter 4

InP-based spot-size converters

4.1 Introduction

We have studied spot-size converters of the 'buried' and 'ridge' type, two waveguide implementation schemes that are being used in the InP-based world of integrated optics, both in research and for industrial device fabrication.

Ridge-type waveguide devices are fabricated in a single epitaxial growth, followed by one or more etching steps. Typical cross-sections of a passive and an active device are shown in figures 4.1(a) and 4.1(b) respectively. Active devices demand, next to the definition of a waveguide, also an adequate electrical (current) confinement. This confinement is created by the deposition of non-crystalline materials (like SiO_2 or Al_2O_3) or an insulating polyimide, through which an opening is etched and the metallisation can make physical contact with the device.
Buried structures involve additional epitaxial growth after the waveguide etching, resulting in waveguide geometries as shown in figure 4.1(c) and 4.1(d). A passive buried structure requires a single regrowth step to deposit an InP cladding layer after the ridge etching, while active devices involve two additional growth runs. The second epitaxy deposits an n-p junction next to the etched laser waveguide. Under normal bias conditions of the p-i-n layer stack, this n-p junction will be reversely biased and current will flow only through the active p-i-n heterostructure. A third epitaxial step finally adds the upper (p-doped) InP cladding layer.

The buried technology is in most cases preferred for the fabrication of active devices. They generally show a superior performance thanks to a better electrical confinement and lower internal losses. The lower electrical confinement in ridge lasers is attributed to the fact that the etching is usually stopped above the quantum wells. This avoids excessive surface recombination of the injected carriers, but also results in a spreading of the current as it flows towards the active region. The lower internal loss is related to a lower scattering loss at the etched ridges

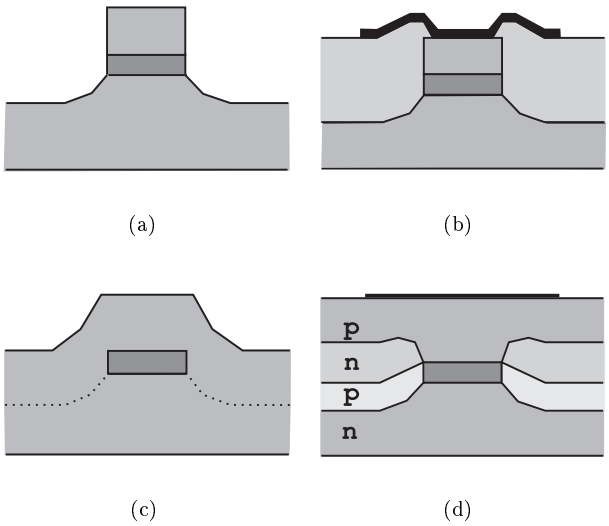


Figure 4.1: a) Passive and b) active ridge-type waveguide; c) passive and d) active buried structure.

and is ascribed to the fact that these ridges are completely surrounded by the regrown InP cladding that creates a lower refractive index contrast and hence a less efficient scattering of the propagating light.
Ridge-type waveguides are the obvious choice for the realisation of passive components since they present a far higher lateral refractive index contrast, which is essential to reduce bend radii and reduce the chip size.

It is obvious that the differences between the buried and ridge waveguide schemes also necessitate totally different approaches to the implementation of spot-size converters.
Buried spot-size converters mostly evolve towards a weakly guiding output waveguide by narrowing the core region in the lateral and/or transverse dimension (see e.g. figure 2.9(a)). The additional growth of a cladding layer makes that the etched core is entirely surrounded with InP and results in a more symmetrical and better expanded output mode. Excellent coupling characteristics have been reported for this type of devices, but the extra growth step complicates the fabrication process [1].
Spot-size converters for ridge-type structures demand a different approach to create a fibre-adapted output mode. A large and symmetric mode can only be produced by forcing the optical field to shift and expand into a large ('oversized') underlying waveguide (see e.g. figure 2.3). A traditional approach to such a device would involve the growth of a thick secondary core layer between the upper core and the substrate. This thick layer would have to be composed of InGaAsP with a refractive index between the index of the upper core and the index of the

InP substrate[1], which may require a very accurate control on the composition of the lattice-matched quaternary. Even when the growth process can be calibrated with the required accuracy, it cannot be guaranteed that the process parameters will remain steady during a run of several hours. In section 4.3 we have elaborated a solution to these problems that has resulted in a patent application. We discuss the concept of vertical Anti-Resonant Reflecting Optical Waveguides (ARROW) and how this concept can be been implemented in spot-size converters.
The design, fabrication and characterisation of the buried taper that was studied in this work is presented in section 4.2. The device is designed to operate at the more uncommon wavelength of 1.06μm.

So far we only mentioned the use of spot-size converters to obtain a highly-efficient fiber-chip coupling. Section 4.4 concentrates on 'on-chip' tapers to realise low-loss integrated waveguides. The principle relies on the fact that the fundamental mode of a broad waveguide exhibits lower scattering losses than the guided mode of traditionally used monomodal guides. Adiabatic tapers are then required to make the transition to the broad structure and ensure a monomodal excitation.

4.2 Buried spot-size converters

4.2.1 Background

The spot-size converter discussed in this section was designed to be integrated on a PIC that serves as a signal-combiner in a phased array antenna system. This phased array antenna is to make part of a free-space laser communication channel between two satellites[2].

'Aperture synthesis' through the use of a phased array antenna is a well-known concept that has been established at microwave frequencies. It uses a combination of several small sub-apertures, arranged in close proximity in a planar array. This array acts as a single large aperture by coherent combination of the different subsignals. Introduction of an appropriate phase shift between the different subantenna signals allows to point the composite antenna in a certain direction in an inertia-free way. However, microwave beam-forming networks cannot be installed on satellites because of their high weight and volume. Optical phased arrays, on the other hand, are more compact and lighter and do qualify for use in inter-satellite communication systems.
The operation wavelength is 1.06μm. This wavelength is chosen because of the availability of high-quality solid state lasers (e.g. the Nd-YAG laser) with narrow line width and high output power.

[1] The thick layer serves as a cladding for the upper waveguide defining the strongly confined 'on-chip' mode. At the output facet of the taper it serves as a waveguide core with the InP-substrate as cladding.

[2] This work was conducted in the context of a project for the European Space Agency (ESA), targeting the development of a fully packaged and pigtailed PIC that realises the phase shifting and combining of four optical signals of a phased array antenna [2].

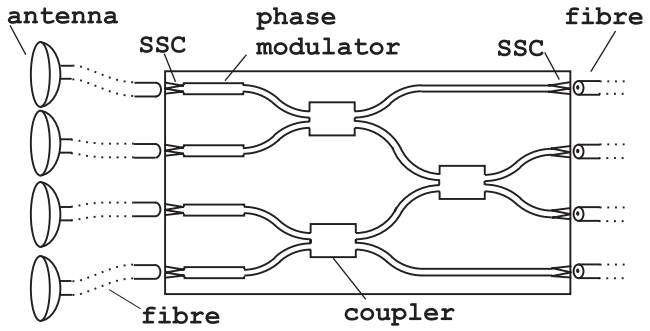


Figure 4.2: Schematic layout of a 4-channel phased array signal combiner.

A schematic layout of the module is given in figure 4.2[3]. Polarisation maintaining single-mode fibres guide the light from the sub-antennas to the PIC. The signals entering the PIC are combined in 3dB Multi Mode Interference couplers (MMIs). Phase modulators preceding the couplers regulate the phases of the tributary signals in order to obtain constructive interference in the 'signal' output port of the MMI. The power emerging from the other port, the 'control' port, is minimised and fed to a control unit that generates the steering signals for the phase modulators. An appropriate choice of the phase difference between the different MMI input signals allows to electronically fine-point the compound antenna into another direction. By using suitably fast effects to realise the phase control, also compensation of platform vibrations can be achieved.

One of the most demanding PIC specifications is the requirement that the module insertion loss should be less than 10dB. This loss can be divided into on-chip loss (waveguide propagation loss, absorption loss in the modulators,..) and fibre-chip coupling loss. The latter amounts up to 7.5dB per facet in a directly butt-coupled situation, making the integration of spot-size converters at the input- and output ports mandatory.

4.2.2 General considerations

A comparative study pointed out that the best suited material system to realise the PIC is the InP/InGaAsP semiconductor system [3][4]. In recent years, the InP material system has been studied extensively, mainly because of its excellent characteristics for applications at the fibre transmission wavelengths of 1.3 and 1.55μm, resulting in a very good knowledge of the material parameters at these wavelengths [4]. Since guided wave communication at 1.06μm has not been studied much in the past, few material parameters are readily known at this wavelength. InP belongs to the group of III-V semiconductors, being composed out

[3] The study of the complete PIC formed the subject of the PhD thesis of Giovanni Flamand [3].

[4] In fact, both the InP/InGaAsP and GaAs/AlGaAs semiconductor systems are an appropriate choice for the realisation of the PIC. However, also detectors will be integrated on the final PIC and InP-based detectors happen to show a better resistance to cosmic radiation.

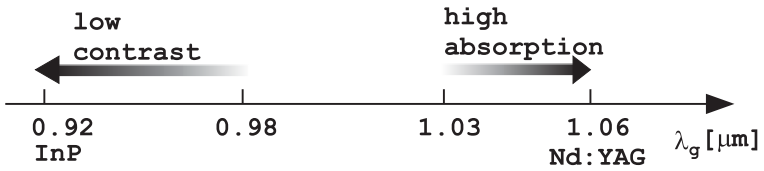


Figure 4.3: InGaAsP: usable range of materials.

of a group III (In) and a group V (P) element and has a bandgap wavelength of 0.917μm. By adding other group III (Ga, Al) and group V (As) elements, it is possible to obtain ternary or quaternary compounds with a different bandgap wavelength but the same lattice constant as InP. The most frequently used quaternary, $In_{1-x}Ga_xAs_yP_{1-y}$, can cover the bandgap wavelength range from 0.917μm (InP) to 1.65μm ($In_{0.532}Ga_{0.468}As$), while remaining lattice matched to InP.
The layer structure of the PIC is mainly determined by the phase modulators, which require a strong confinement of the InGaAsP core and the InP cladding layers. On the other hand, the guiding layer must have a bandgap wavelength sufficiently below the wavelength of the guided light (Nd:YAG, $\lambda = 1064nm$) to avoid high absorption. Thus, the available compositional range for the $In_{1-x}Ga_xAs_yP_{1-y}$ compound which will be used for the guiding layer will be very limited. These considerations narrow the range down to the 0.98 - 1.03μm region (figure 4.3). This implies also that a high degree of accuracy is required on the exact composition of this guiding layer. Table 4.1 illustrates this by showing the theoretical compositions of lattice-matched InGaAsP with a bandgap wavelength between 0.98μm and 1.03μm. Obviously, $In_{1-x}Ga_xAs_yP_{1-y}$ materials in the required range of composition have small Ga- and As-contents, and are rather sensitive to even small fluctuations in these contents. In addition, because of the low cracking efficiency of the P-precursor gas (PH_3), the As- and P-incorporation shows a non-linear dependence on the respective AsH_3 and PH_3 fraction in the gas phase. The required AsH_3/PH_3 ratio in the gas phase is extremely low for the required InGaAsP materials, complicating its control. The reproducible MOVPE-growth of the needed high-quality $In_{1-x}Ga_xAs_yP_{1-y}$ materials therefore proved to be a rather demanding task. The final layer structure, as required for the phase modulators, waveguides and MMI couplers, is summarised in table 4.2.

$\lambda_g[\mu m]$	**x (Ga)[%]**	**y (As) [%]**
0.98	5.2	11.4
0.99	6.0	13.2
1.00	6.8	15.0
1.01	7.6	16.7
1.02	8.4	18.4
1.03	9.2	20.2

Table 4.1: Composition of $In_{1-x}Ga_xAs_yP_{1-y}$, lattice matched to InP.

Identification	Thickness	Material
Contact	170nm	InGaAs
Upper cladding	1400nm	InP
Core	700nm	InGaAsP ($\lambda_g = 1.03\mu$m)
Lower cladding / substrate	-	InP

Table 4.2: PIC layer structure.

4.2.3 Design

Figure 4.4 shows a taper design that is consistent with the above presented layer structure [1]. As discussed in chapter 2, a larger mode size can be obtained by increasing the waveguide dimensions or by sufficiently narrowing or thinning the core. The device presented here uses both effects to create a two-dimensional mode expansion. The upper part of the InGaAsP core is gradually narrowed down to zero width, resulting in a thin and vertically weakly guiding structure at the end of the device. The lower core section is at the same time broadened to increase the lateral dimension of the guided mode. As already noticed in section 2.2.3 the thinning of the core is obtained by a pure lateral tapering of the device, avoiding the need for complicated growth or etching techniques in the fabrication process. The output cross-section was fixed by optimising the coupling efficiency to a 1064nm polarisation maintaining single mode fibre (PMF). A 100nm thick, $6\mu m$ wide waveguide with an upper cladding of $2.75\mu m$ was finally found to give a coupling loss of 1.75dB (68% coupling efficiency), Fresnel reflection losses not included. This has to be compared with 7.5dB coupling loss (17.8% coupling efficiency) when no taper is integrated.
Keeping the adiabaticity criterion of chapter 3 in mind, we have reduced the total taper length by defining a piecewise linear longitudinal taper evolution (see figure

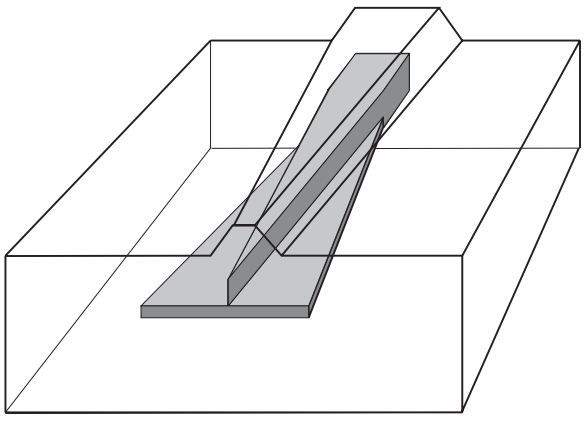


Figure 4.4: Taper integrated on the PIC. The upper part of the core is tapered down, while the lower section is broadened towards the end of the device.

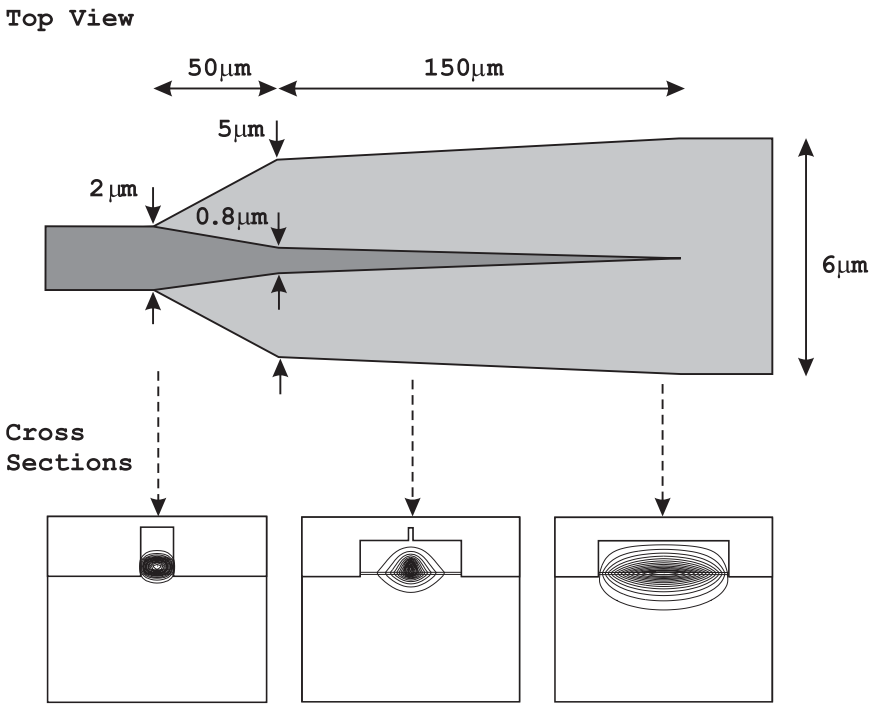


Figure 4.5: Final design of the piecewise linear taper.

4.5) and allowing a larger taper angle in the first section where the shape of the fundamental local mode changes only slightly. The behavior of the device was verified with very time-consuming 3D-BPM (Beam Propagation Method) simulations. The final design with a total length of 200μm is shown in figure 4.5, together with some field profiles illustrating the evolution of the optical field along the taper.

4.2.4 Fabrication

Wafer growth The fabrication of the taper starts with the planar MOCVD-growth of the lower InP cladding layer, the InGaAsP core layer and a 100nm thin InP cap layer to facilitate the later regrowth of the upper cladding (see further). Figure 4.6 shows the evolution of the cross-section at the beginning, in the middle and at the end of the taper during the fabrication process.

Etching of the taper tip The first processing step consists of etching the narrowing taper tip through the InP cap layer and 600nm into the guiding layer, using the following processing steps. After PECVD-deposition of a SiO_x-layer on top of the layer structure, the mask pattern is defined using a photolithography step. After dry etching (Reactive Ion Etching, RIE) of the dielectric layer (using

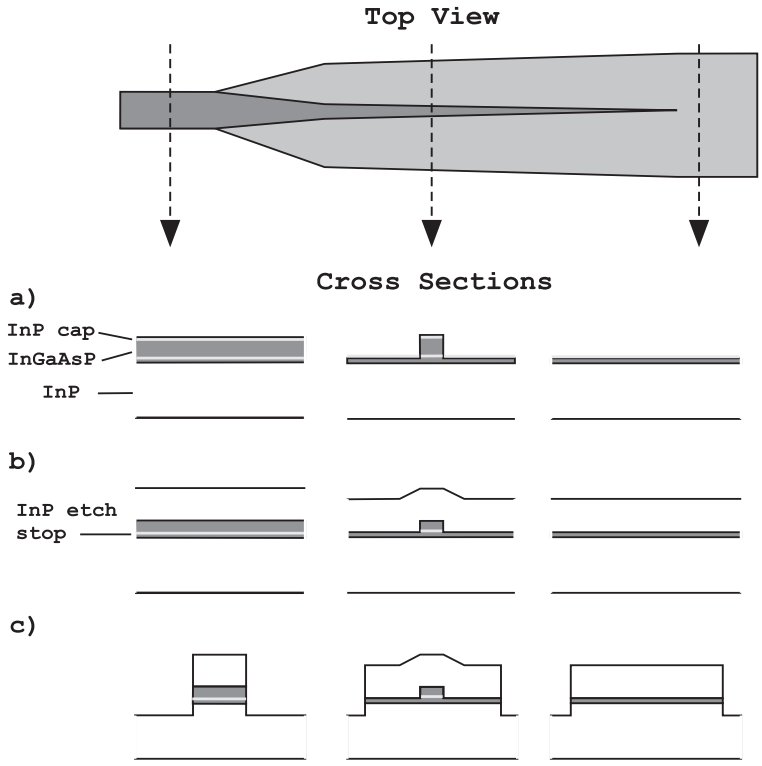


Figure 4.6: Input, intermediate and output cross-section of the taper during the fabrication process, a) 'after growth of bottom cladding, core and InP cap and first etch step; b) after regrowth of top cladding layer and c) after etching of waveguide ridges and uptapered section.

$CF_4 : O_2$) and removal of the resist layer in an O_2 plasma etch, the actual etching of the taper tip is performed using a combination of RIE and wet etching. Dry etching is preferred because it creates vertical sidewalls and limited underetching, which is desirable for the definition of the taper tip. An eventual underetching, typical for wet etching, would strongly narrow or even completely remove the narrowest section of the downtapered tip and disturb the adiabatic mode conversion. The disadvantage of RIE, on the other hand, is the uncertainty on the etch depth. Because the vertical dimension of the output mode is critically dependent on the remaining thickness of the core layer at the output, it was decided to insert a 30nm thin InP etch-stop layer in the core and to etch the final tens of nanometers by means of a selective wet etch. The etch-stop layer is so thin that it doesn't influence the waveguiding properties of the other components on the PIC. This procedure combines the advantages of dry and wet etching and waveguides with vertical edges are fabricated with an etch depth accuracy of about 10nm. Finally, the remaining oxide layer is removed using $HF : H_2O$.

Regrowth of the upper cladding Following the etching of the taper tip, the top cladding layer is grown. This regrowth step is preceded by a short dip in H_2SO_4, which has a smoothing and cleansing effect on the sample, enhancing the regrowth behaviour. The interface with the regrown InP cladding layer can only be of a good quality if the underlying layer is InP and not InGaAsP. For this reason, a thin InP cap layer was grown on top of the quaternary core. The thickness of the top cladding layer deposited in the regrowth is 1.4μm instead of the specified 2.75μm of the design, to avoid problems with the next etch step. The reduction of the upper cladding thickness obviously has an impact on the coupling characteristics to fibre and is found to increase the theoretical coupling loss from 1.75dB to 2.3dB (58% coupling efficiency instead of 68%).
The combination of a dry and wet etch for the definition of the taper tip is not only important for a nice ridge shape and a correct etch depth, but it is also essential to obtain a good quality of the regrown cladding layer. In a first experiment the taper tip was etched only with RIE. Considerable problems were encountered during the following regrowth of the InP cladding since growth occured primarily along preferential crystallographic directions, giving rise to the formation of so-called 'rabbit ears' (figure 4.7). These 'rabbit ears' hamper the subsequent etching of the broad waveguide (see next paragraph). The fabrication of these structures was retaken with modified growth conditions during the regrowth, in order to reduce the growth along undesired crystallographic planes, but this offered no solution.
In order to overcome the growth problems, it was decided to use a combination of RIE with a variety of wet etching techniques:

a) RIE etching followed by a dip in a $Br_2 : CH_3OH$ mixture;

b) RIE etching followed by a dip in a $HBr : Br_2 : CH_3OH$ mixture;

c) complete wet etching of the taper tips using a $H_2SO_4 : H_2O_2 : H_2O$ (1:1:3) mixture;

d) 600nm RIE etching followed by 100nm wet etching in a $H_2SO_4 : H_2O_2 : H_2O$ (1:1:3) mixture.

Only method d) resulted in a nice planar growth without introducing defects or creating any other complications. The severe underetch in option c) caused, as expected, a very poor definition of the taper tips and options a) and b) generated the formation of a lot of defects, which strongly limit the performance and yield.

Etching of the broad waveguide Finally, the broadening part of the taper is etched using a second mask, which has to be carefully aligned with the already formed taper tips. The procedure is identical to the one described above, except for the fact that only RIE is used for the 2.2μm deep etch.
An important issue in the fabrication of tapers is the sensitivity of taper performance to the possible misalignment between the two masks. As illustrated in figure 4.9, a misalignment of only a few hundred nanometers can seriously alter the shape of the taper, resulting in an abruptly changing cross-section between straight waveguide and taper. The effects of these possible misalignments were

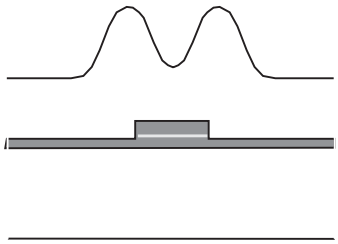


Figure 4.7: Formation of 'rabbit ears' during the regrowth on dry etched taper tips.

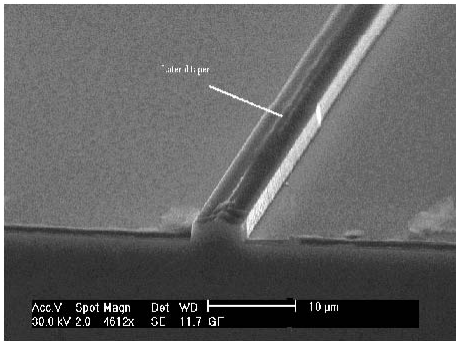


Figure 4.8: SEM photograph of a taper with one regrowth step.

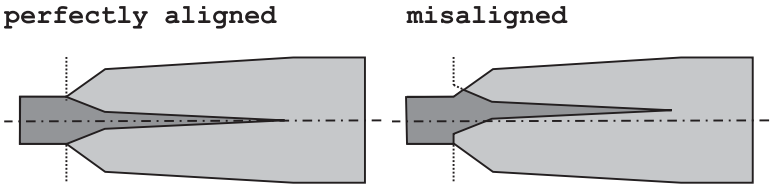


Figure 4.9: Illustration of taper deformation by mask misalignment.

reduced by inserting an additional converging section on the first mask layer, as illustrated in figure 4.10.

4.2.5 Characterisation of the tapered waveguides

Two important parameters that characterise the quality of the spot-size converter are the excess loss (the extra loss introduced by a taper as compared to a straight waveguide of the same length) and the taper-fibre coupling loss.
The excess loss is determined by means of Hakki-Paoli measurements (see appendix) on a straight waveguide terminated by a taper at both ends. This measurement yields the total propagation loss, i.e. all loss experienced by a mode

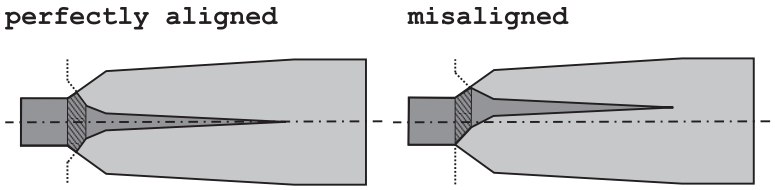


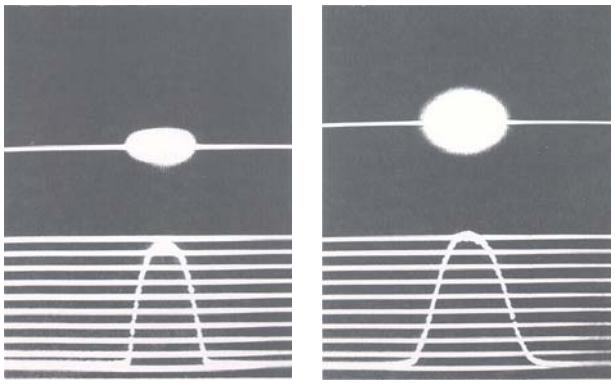
Figure 4.10: Reduction of misalignment sensitivity by addition of a converging section.

traveling from one facet to the other. By comparing this loss with the values obtained for an untapered waveguide of the same length, the excess taper loss is found to be about 1.15dB. It should also be remarked that the measurements showed a very good uniformity, indicating a good yield for the fabrication scheme discussed above. The excess taper loss can be accounted to the limitations of the photolithography used to define the taper tip. The minimal dimension which can be defined using this technology is approximately 800nm, such that the taper tip is not gradually tapered down to zero, but stopped abruptly somewhere along. This means that a non-adiabatic transition occurs at the end of the tip, introducing an excess taper loss.
The fibre-chip coupling loss has been measured at ETH Zürich, since a dedicated set-up was not yet available at INTEC at that time. It was found that the tapered waveguides exhibit an average coupling loss of 2.8dB, which is slightly higher than the calculated value of 2.3dB. The origin of this higher than expected fibre-chip coupling loss can be identified by looking at the near-field distribution at the tapered waveguide output facet (figure 4.11). According to the simulations, the taper output mode should be about 6μm wide and 4.5μm high. However, the actual observed near-field distributions indicate that the real height is only approximately 3.5μm, resulting in a worse fibre-chip coupling efficiency than predicted by theory[5]. This deviation is caused by a little too thick guiding layer in the output waveguide section or a small compositional error of the quaternary core layer, resulting in a better transversal confinement and hence smaller mode height than intended.
In conclusion, the integration of the spot-size converters still reduce the fibre-to-fibre losses of the PIC with about 7dB, despite the fact that there is a taper excess loss of 1.15dB and the coupling loss is somewhat higher than expected.

4.2.6 Improved design

The excess propagation loss of the taper described above turns out to be higher than expected due to technological limitations. It was therefore decided to alter the taper fabrication procedure such that narrower tips can be fabricated. This is accomplished by using the so-called 'knife-etch' technique for the definition of the

[5]This extra loss agrees with the value calculated from (1.7).



(a) (b)

Figure 4.11: Measured near-field patterns of (a) the taper mode and (b) the fibre mode.

taper tips.

The principle of the knife-etch technique is shown in figure 4.12. The narrow tip is defined in two photolithographic steps, using two different masks. Only the part of the sample, covered by both masks will not be etched when positive photoresist is used. The fabrication steps for the realisation of the downtapered tip are the following:

1. deposition of a SiO_x-layer;
2. spinning of resist;
3. photolithography with first knife-etch mask;
4. dry etching of the dielectric layer (using $CF_4 : O_2$);
5. removal of resist (using O_2 plasma etch);
6. spinning of resist;
7. photolithography with second knife-etch mask;
8. dry etching of the dielectric layer (using $CF_4 : O_2$);
9. removal of resist (using O_2 plasma etch);
10. actual etching of the taper tip, using a combination of 600nm ($CF_4 : O_2$) RIE and 100nm wet etching in a $H_2SO_4 : H_2O_2 : H_2O$ (1:1:3) mixture.

The subsequent fabrication steps are identical to those used for the convential taper fabrication procedure as explained before: the etching of the taper tip is

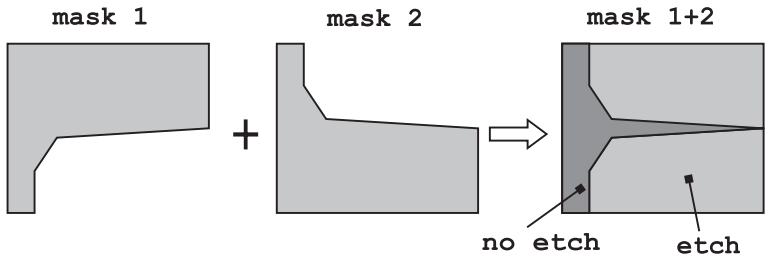


Figure 4.12: Definition of the taper tips by 'knife etching': mask 1, mask 2 and both masks together.

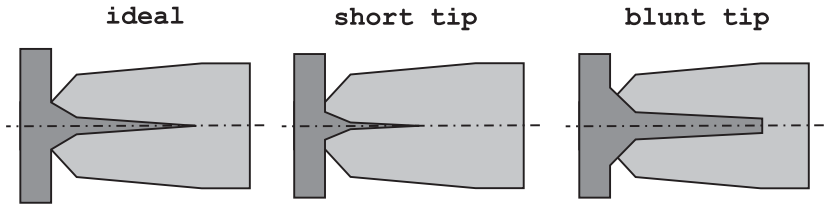
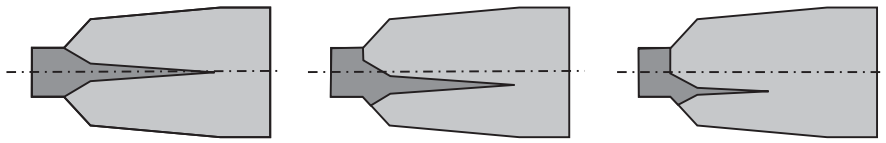


Figure 4.13: Effect of a misalignment between knife-etch masks.

followed by a regrowth of the top cladding layer and a second etch step defines the waveguide ridges and the lateral uptapered taper section.
It has already been illustrated in figure 4.9 how a misalignment between the different taper masks can result in abrupt transitions for the conventional taper fabrication procedure, and how this can be reduced by inserting an additional converging section. The situation becomes more complex for the knife-etch fabrication technique since an extra mask and mask alignment step is required.
Figure 4.13 illustrates the effect of a misalignment between the two knife-etch masks. This kind of misalignment clearly leads to a too short or a blunt taper tip, depending on the orientation of the misalignment. An abrupt transition is created at the start of the taper in the case of a short taper tip and the slowly tapered section, which should start at a tip width of 0.8μm to satisfy the adiabaticity criterion, starts at a smaller tip width, leading to a non-adiabatic taper and hence additional losses. Also a blunt tip is unwanted, since this results in an abrupt transition at the end of the taper and introduces excess losses, comparable to what was observed for the conventional taper design.

Of course, also the third mask, defining the broad outer ridge, can be misaligned with respect to the knife-etch masks and the resulting taper is even more catastrophic. This is illustrated in figure 4.14. The mask misalignments result in a taper tip that is placed asymmetrically with respect to the broadening ridge, possibly in combination with an increase of the abrupt transitions at the beginning of the taper. Special care was therefore paid to these problems when designing the different mask sets.

The adverse effects caused by the alignment errors are largely reduced by ap-



(a) (b) (c)

Figure 4.14: Effect of a misalignment between the three taper masks: a) perfect alignment; b) misalignment of the broad outer ridge mask and c) misalignment of all three masks.

plying the taper design as shown in figure 4.15. This design ensures that, even for worst case misalignment combinations, the correct small taper angle is kept along the critical taper section (tip width $\leq 0.8\mu$m), that no abrupt transitions occur and that blunt taper tips are avoided. This is achieved by considering the following arrangements:

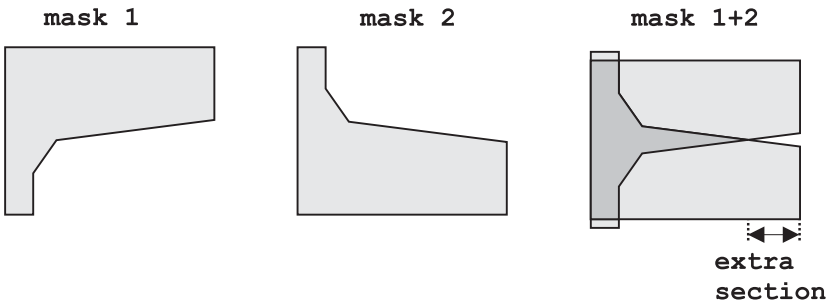
- the slowly tapered section is started at a width of 1.3μm instead of 0.8μm, so that the minimum and maximum starting widths (caused by mask displacements) are 0.8μm (=1.3μm - 0.5μm) and 1.8μm (=1.3μm + 0.5μm) respectively. This ensures that the critical taper section is always tapered under the correct small angle. The price to be paid for this security is an increased taper length. The design length is now 294μm (instead of the original 200μm); the minimum and maximum lengths are 200μm and 388μm respectively. The only consequence of this increased length, however, is a slight increase in propagation loss of 0.1dB;

- the knife-etch mask is extended with an extra section to eliminate blunt tips (see figure 4.15);

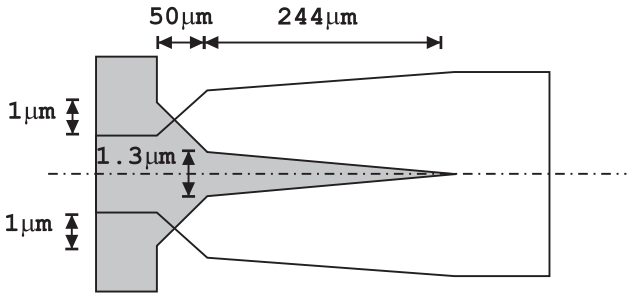
- the initial width of the tip is increased by 1μm on both sides to avoid any abrupt transistion at the beginning of the taper. A similar arrangement was also present in the conventional taper mask set. Simulations revealed that the initial ridge broadening, caused by this intervention, does not influence the device behavior (the effects look more dramatic than they are in reality because of the scales used in the figures).

The robustness of this new taper design is illustrated in figure 4.16 by showing the worst case misalignments (compare this with the situation of e.g. figure 4.14). The real taper shape will be situated somewhere in between these extreme situations. Figure 4.17 shows a SEM photograph of sharp taper tips fabricated using the knife etch technique.

This new spot-size converter design promises to yield better device performances than the (already satisfactory) conventional design described before, but unfortunately, we encountered large problems with the regrowth of the top cladding



(a)



(b)

Figure 4.15: Improved taper design reducing the influence of misalignments on the taper behavior: a) knife etch masks and b) complete optimised taper layout.

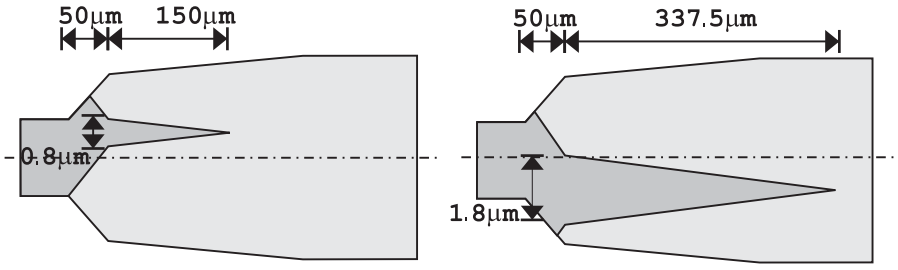


Figure 4.16: Safe taper shapes resulting from worst case misalignments of the optimised knife-etch masks that eliminate abrupt transitions at the taper start and ensure a slow tapering in the critical taper section.

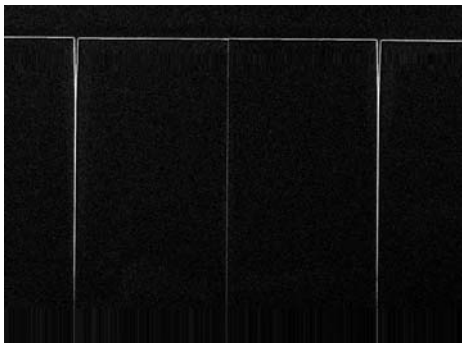


Figure 4.17: SEM photograph of taper tips, fabricated with the knife-etch technique.

layer. A highly non-uniform regrowth was observed on top of the downtapered taper tips, in spite of using the same taper etching process and optimised regrowth parameters as used for the functioning tapers presented in section 4.2.5. The most probable origin of these regrowth problems is the RIE etching process of the taper tips. It is not straightforward to reproduce the processing parameters used previously, since the RIE system is switched over every few weeks for the etching of various material systems, which require different etching gasses and conditions.

4.3 Ridge-type spot-size converters with transverse ARROW confinement

4.3.1 Introduction

This section discusses the design of a laser with a monolithically integrated ridge-type spot-size converter operating at a wavelength of 1.55μm. The device is based on a structure that was first proposed by Vawter et al. in the GaAs/AlGaAs semiconductor system [5] and later rendered into the InP/InGaAsP system by Vusirikala et al. [6]. Figure 4.18 shows the device as proposed in those articles.

The operation principle was already explained in chapter 2, but we shortly recapitulate here for convenience. The component comprises a narrow active ridge waveguide, which is placed on top of a large mesa that carries a fibre-adapted mode at the output of the taper. The upper ridge is relatively broad in the laser section and carries a strongly confined optical field in its active core. As the ridge is narrowed towards the end of the device, the mode is forced to expand and shift into the large underlying mesa. The design is such that the mode transformation is complete before very narrow tips are reached to avoid problems with the lithography of the ridge. This device is particularly interesting because of its simple fabrication in comparison with most other reported designs (see chapter 2). It

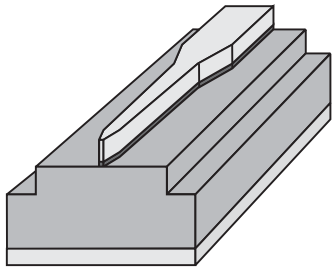


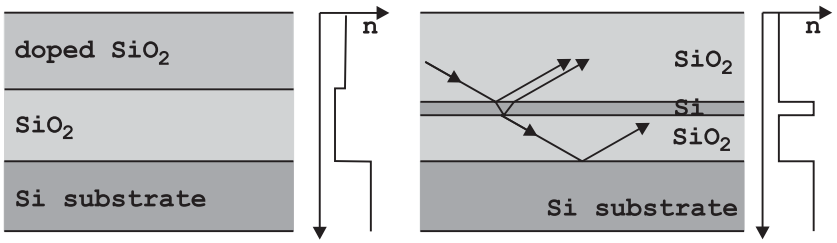
Figure 4.18: Ridge-type spot-size converter.

involves only a single planar epitaxial growth step and conventional etching.

The coupling efficiency to a standard single mode fibre is determined by the size of the large mesa. A good coupling therefore involves the growth of a thick layer of lattice matched InGaAsP with a refractive index between that of the upper core and the InP substrate (see section 4.1). In other words, a quaternary material with a low Ga- and As-fraction is needed. However, growing such a thick layer with a constant composition to remain lattice matched to InP is not obvious. The use of a diluted layer, i.e. a composite layer that creates an average index by stacking a number of thin (in the order of tens of nm) InP/InGaAsP layer pairs, to avoid this problem of material composition has been proposed and demonstrated in [6]. The best solution, however, would be to substitute the InGaAsP in the thick layer by the easy-to-grow InP. This is at first sight not applicable because of the required refractive index contrast with the InP substrate. However, if a Bragg-like structure is provided under the InP layer, then an optical field can be confined in the thick InP layer. The mode is then no longer a guided mode, but a 'leaky' mode that experiences total internal reflection at the upper semiconductor-air interface and a strong reflection from the lower mirror structure. The leaky nature of the mode implies that power is radiated into the substrate, but a proper design of the mirror can limit this loss to the order of 0.1dB/cm.

It appeared that this principle had already been demonstrated in the 1980s in the silica-on-silicon material system (be it for other reasons), where such waveguides are known as '**A**nti-**R**esonant **O**ptical **W**aveguides' or 'ARROWs' for short [7]. The next sections describe our proposal to introduce the ARROW-concept in InP-based ridge-type spot-size converters to relax the epitaxial growth specifications. This resulted in the patent application 'Semiconductor optical component utilizing leaky structures to match the mode of rib waveguides to that of the fiber' [8]. For a more profound discussion and more details on the device design, fabrication and characterisation, we refer to [9][6].

[6] It was soon clear that the ARROW-concept offered a broad range of applications, going from transverse confinement in tapers and a combined transverse and lateral confinement to application in bent waveguides for the reduction of the lateral radiation loss. Also buried structures were found to benefit from introducing the ARROW-concept. The study of the ARROW principle



(a) (b)

Figure 4.19: Silica-on-silicon waveguides: a) conventional design; b) with ARROW-confinement.

4.3.2 Principle and design of ARROWs

The ARROW concept was introduced in the silica-on-silicon material system to reduce the waveguide absorption losses [7]. Up to then, silica-on-silicon waveguides consisted of a SiO_2 cladding that was deposited on the high-index Si-substrate, as shown in figure 4.19(a). On top of the cladding came a doped SiO_2 core layer with a slightly higher refractive index than the undoped SiO_2 cladding. The doping of the core layer, however, also makes the material more lossy than pure SiO_2. In addition, very thick cladding layers are needed to avoid radiation loss into the Si substrate. The ARROW structure proposed in [7] is shown in figure 4.19(b) and circumvents the use of lossy SiO_2 and thick cladding layers. A thin poly-crystalline Si cladding layer is sandwiched between the SiO_2 core and a second SiO_2 cladding layer. The thickness of these cladding layers is chosen such that the reflections occuring at successive Si-SiO_2 interfaces interfere constructively, resulting in a strong overall reflectivity (>99.9%).

Implementation of the ARROW-concept in the spot-size converter leads to a layout as sketched in figure 4.20 [7]. The ARROW-confinement is only of importance at the end of the device, where the mode has already largely shifted to the thick InP-layer and where the upper narrow ridge has practically no further influence on the modal behaviour. The slab waveguide to be considered for the design of the ARROW is thus the broad central slab, without the narrow active ridge on top, as shown in figure 4.21.

When the core thickness is large enough ($d_c \gg \lambda_0/2n_2$), the fundamental mode will propagate with a glancing incidence into the thin quaternary ARROW-layer

and its variety of applications form the subject of the PhD thesis of Marko Galarza [9]. The work done in the context of the thesis at hand relates to design and measurement of the early ARROW-based ridge-type tapers.

[7] Koch et al. were the first to demonstrated the ARROW-principle in the InP/InGaAsP material system at 1.5μm [10].

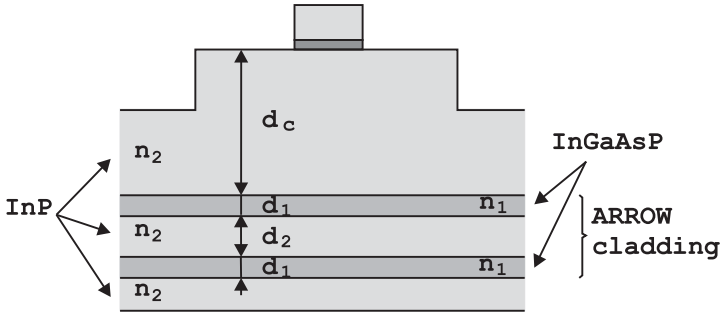


Figure 4.20: Spot-size converter with transverse ARROW-confinement of the output mode.

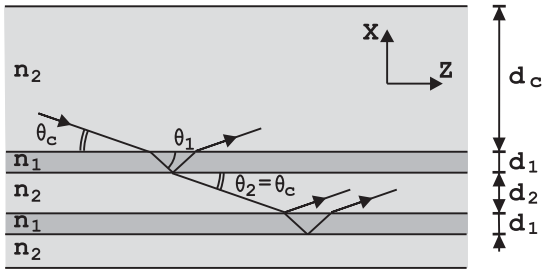


Figure 4.21: ARROW slab waveguide with indication of the design parameters.

and the air on top. The glancing incidence and the large refractive index contrast imply that the penetration depth into the air is negligible. Also the penetration depth in the interference cladding is very small when the high reflectance condition is met. This allows to write the transverse resonance condition of the fundamental leaky mode in the InP core as

$$2d_c\, k_0 n_2 \sin\theta_c \approx 2\pi \quad \text{or} \quad \sin\theta_c \approx \frac{\lambda_0}{2n_2 d_c} \tag{4.1}$$

The high reflectance condition of the ARROW cladding layers corresponds to the *anti*-resonance condition of a Fabry-Pérot resonator. This means that the InP ARROW cladding layer should be half as thick as the core and that the two high-index quaternary layers should satisfy

$$2d_1\, k_0 n_1 \sin\theta_1 = (2N+1)\pi \qquad \text{with N = 0,1,2,...} \tag{4.2}$$

Using Snell's law $n_1 \cos\theta_1 = n_2 \cos\theta_2$, we find

$$d_1 = \frac{\lambda_0}{4n_1}(2N+1)\sqrt{1 - \frac{n_2^2}{n_1^2} + \frac{\lambda_0^2}{4n_1^2 d_c^2}} \qquad \text{with N = 0,1,2,...} \tag{4.3}$$

ARROW waveguides are characterised by a strong discrimination against higher order modes. All odd order modes experience very high leakage losses (in the or-

der of 100dB/cm). In fact, they are resonant in the ARROW cladding layers and enhance transmission to the substrate. Higher order even modes again satisfy the anti-resonance condition, but impinge with larger angles on the lower structure and experience more bounces per unit length as well. Typical loss values for the second order mode are in the order of a few tens of dB/cm .

With respect to the design of polarisation-independent devices, like e.g. tapered semiconductor optical amplifiers (SOAs), it is important to mention that the physical behaviour of the TE and TM polarised waves is similar [11]. In general the attenuation is lower for TE than for TM [8], but the difference decreases rapidly as the waveguide becomes thicker (several μm) and a more glancing incidence is obtained.

4.3.3 Design of the spot-size converter

The final design of the laser and the integrated spot-size converter with transverse ARROW-confinement of the output mode, optimised for coupling to a standard 1.55μm single mode fibre, is given in figure 4.22. Overlap calculations were used to determine the dimensions of the large InP-mesa, bearing in mind the growth-imposed limitation that the overall layer stack thickness may not exceed 10-11μm. A coupling efficiency of 52% (2.8dB loss) was calculated for the device of figure 4.22 (reflection loss not included), improving the coupling by 6.6dB.

The ideal adiabatic taper shape was approximated by a piecewise linear evolution as shown in figure 4.23. The optical field of the fundamental local mode is also displayed for the laser cross section, a cross section further down the taper and at the output facet. The evolution of the optical field and the adiabaticity of the design were examined by *BPM_CAD3D* simulations.

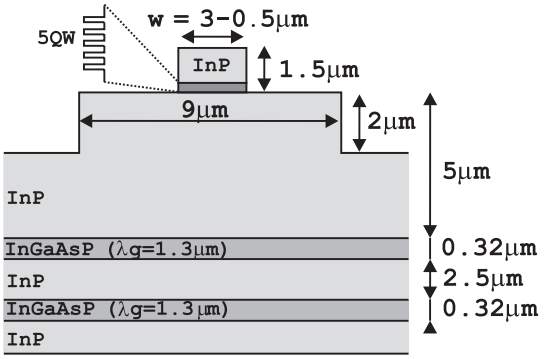


Figure 4.22: Final design of the tapered laser with transverse ARROW-confinement of the fibre-matched output mode (deeply etched design is shown).

[8] This can be attributed to the lower fresnel reflection coefficient for the TM polarised waves.

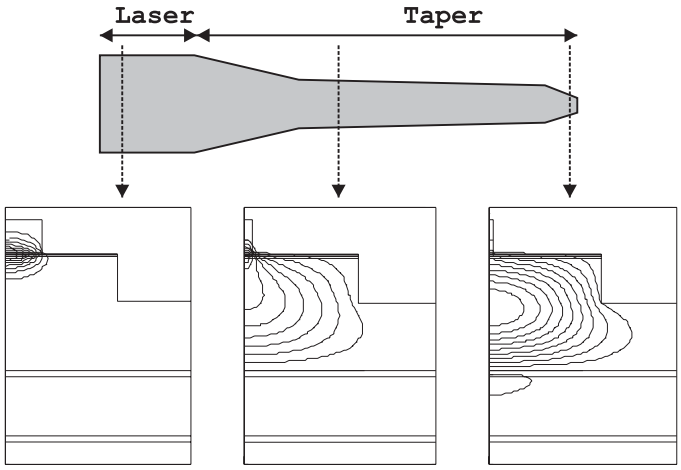


Figure 4.23: Piecewise linear approximation of the ideal adiabatic taper shape and evolution of the fundamental local mode.

4.3.4 Results

Two options for the etch depth of the upper laser ridge were implemented: deeply etched devices, which are etched completely through the active layer, half-deeply etched devices, where the etching is stopped above the quantum wells. The first option avoids absorption losses of the output mode in the unpumped active region next to the central ridge, but might suffer from increased surface recombination at the etched QW-walls. A lower surface recombination and a higher absorption in the unpumped active region (next to the central ridge) in the final taper section is expected for the half-deeply etched devices.
We will not discuss the observed differences between both options here in detail, since we only wish to present the results that illustrate the practical feasibility of introducing the ARROW-concept in spot-size converters. For a detailed comparison, we refer again to [9].
Figure 4.24 shows the room-temperature optical power versus current (PI) curves of deeply etched mode expanded lasers and an untapered reference laser. Both devices were 750μm long, including the taper section. The lasers were operated in continuous wave (CW) regime and fairly uniform threshold currents of 50mA were found for the tapered lasers, while the reference samples showed threshold currents around 30mA (corresponding with a threshold current density of 1.33kA/cm^2). The higher threshold current of the tapered laser is mainly explained by its larger active volume. Higher external efficiencies are observed for the tapered devices (0.15W/A versus 0.13W/A for the reference lasers), which can be attributed to the lower reflection coefficient of the tapered facet.
The farfield emission pattern was efficiently reduced from 30^{o}x40^o (lateral x transverse) for the reference laser to 10^{o}x27^o for the mode expanded device. These values are in close agreement with the simulations, as can be seen in figure 4.25.

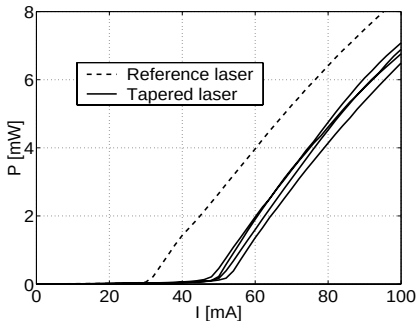
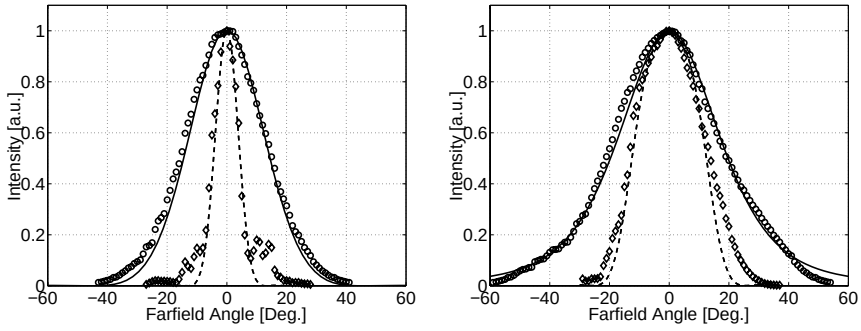


Figure 4.24: Typical PI-curves of a tapered and a reference laser with transverse optical ARROW confinement.



(a) (b)

Figure 4.25: Farfield emission patterns of a tapered and a reference laser with transverse optical ARROW-confinement: a) lateral farfield b) transverse farfield (∘ reference laser measured; ⋄ tapered laser measured, — laser simulated; - - taper simulated).

4.3.5 Relaxed ARROW design

The introduction of the ARROW concept solves the problem of growing thick layers of quaternary material, but increases the overall layer stack thickness significantly (especially the InP cladding layer with thickness $d_c/2$ is responsible for this)[9]. However, one might question whether it is really necessary to design the ARROW-layers such that they are in a perfect anti-resonance state, especially because the response of a Fabry-Pérot resonator is considerably flat and tolerant to

[9]The growth of thick InP layers is not necessarily a problem in an industrial production environment, when the process is optimised for this purpose. We work, however, in a research environment, with a MOCVD process that is concentrated on the growth of layers of a few hundreds of nanometers, the more common layer thickness encountered in integrated optics.

variations of the structure. In addition, the expanding taper mode is only confined by the ARROW structure over the last hundred micrometers of the device. It is clear that in view of this short propagation distance, a leakage loss of a few dB/cm can be tolerated.
Assuming that the expanded mode is effectively guided by the ARROW-structure over a distance of 250μm (this is an overestimate), we can allow leakage losses of 3dB/cm and still keep the resulting additional taper loss below 0.1dB. Calculations with the transfer matrix method reveal that the total ARROW cladding thickness for the 5μm thick InP core can be considerably reduced from 3.15μm for the perfectly anti-resonant design of section 4.3.2, to 1.3μm for the relaxed design that introduces an additional taper loss of less than 0.1dB.
3D BPM simulations showed a slightly higher taper excess loss of 0.08dB for the taper incorporating the relaxed ARROW design, which agrees well with the expectations. Simulations also confirm that the coupling efficiency to fibre and the alignment tolerances are not influenced by substituting the ARROW for the relaxed design.

4.3.6 Lateral ARROW confinement

The ARROW principle can also be implemented to create a lateral optical confinement by etching a set of ridges next to the central laser ridge (figure 4.26). The laser mode is still guided in the central ridge, but the output mode is confined by the lateral ARROW effect. This lateral ARROW confinement can roughly be understood by recognising that an effective index difference is created by the presence of the ribs. However, the design of the lateral ARROW is not as straightforward as was the case for the transverse ARROW and requires a modal investigation of the complete structure [9]. An important restriction on the use of the lateral ARROW is the fact that the lateral ridges loose their impact on the expanded mode if the thick coupling layer becomes too thick. This limits the thickness of the coupling layer to about 3.5μm, which is still sufficient to create a coupling efficiency of 43% to single mode fibre (3.7dB coupling loss, or an improvement of

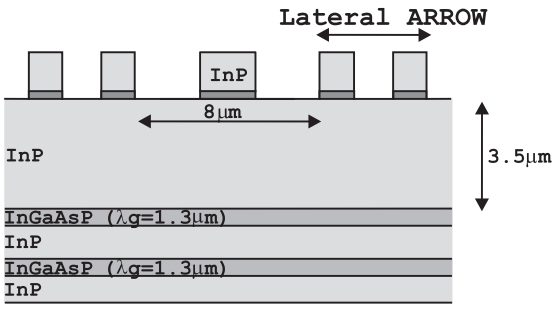


Figure 4.26: Schematic drawing of a tapered laser with both transverse and lateral optical ARROW confinement.

4.7 dB when compared to a direct laser-fibre coupling).
Even though the achievable improvement in coupling efficiency is not as high as for the design where the output mode is confined by an *etched* mesa, it must be stressed that the implementation of this device does not complicate the fabrication or add any processing steps as compared to the processing of an untapered laser. This is the simplest tapered laser scheme ever demonstrated on InP: only a single planar epitaxial step and one conventional etch are needed.
Some continuous wave (CW) PI-curves of tapered and reference lasers are shown in figure 4.28. Figure 4.29 shows the simulated and measured farfield patterns of tapered and reference lasers. A good agreement is observed between theory and measurements (except for the fact that the fabricated devices appear to be somewhat narrower than expected).

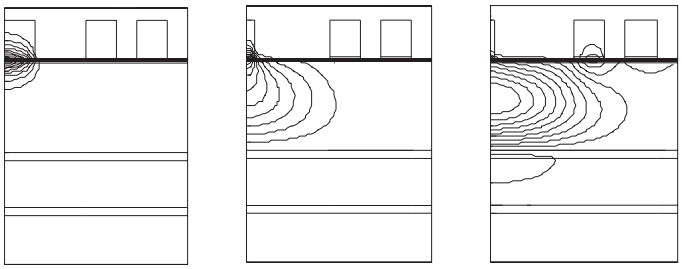


Figure 4.27: Evolution of the optical field along the taper: a) in the laser section; b) somewhere along the taper and c) at the output facet.

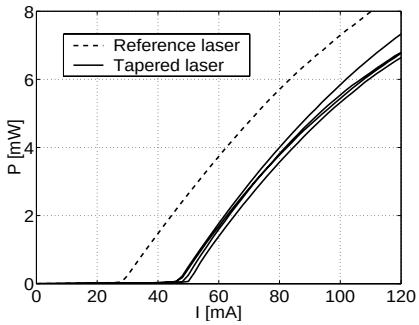
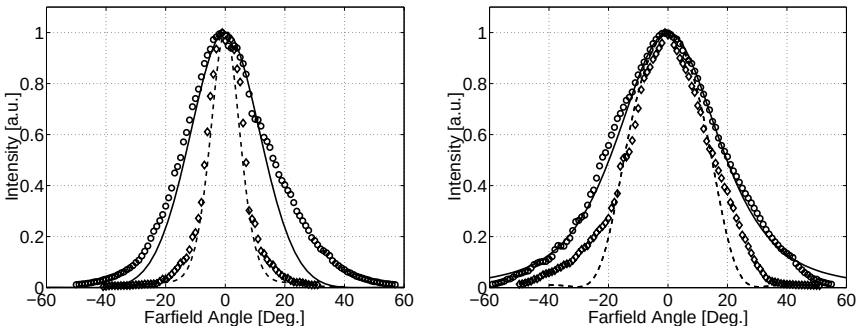


Figure 4.28: Typical PI-curves of a tapered and a reference laser with transverse and lateral optical ARROW confinement.



(a) (b)

Figure 4.29: Farfield emission patterns of a tapered and a reference laser with both transverse and lateral optical ARROW-confinement: a) lateral farfield b) transverse farfield (∘ reference laser measured; ⋄ tapered laser measured, — laser simulated; - - taper simulated).

4.4 Low-loss monomodal operation of multimode waveguides

4.4.1 Introduction

In theory, perfectly straight waveguides with an invariant refractive index distribution in the propagation direction guide optical modes that propagate without experiencing any loss. In real straight waveguides, however, there are several loss mechanisms that create either absorption or scattering of the propagating field[10]. The propagation constant of the mode becomes complex and can be written as $\beta_i = k_0 n_{\text{eff,i}} + j\alpha_i/2$, with k_0 the free-space wavenumber and α_i the power attenuation constant of mode i.

Intrinsic material absorption arises when a layer in the waveguide is composed of a semiconductor with a bandgap that is smaller than or equal to the photon energy of the propagating light. This absorption can be avoided by choosing an appropriate material composition, with a sufficiently wide bandgap. Absorption can also be caused by transitions between defect energy levels, like the levels introduced by dislocations, and the conduction or valence band. This loss is also negligible with the current epitaxial growth quality. A third absorption process is free carrier absorption, whereby the photon energy is transferred to a free carrier that is excited to occupy a higher energy level. This loss mechanism is negligible in passive waveguides, but important in active devices with intentionally doped cladding layers.

[10]Bent waveguides and abrupt transitions between waveguides with different cross-sections generate additional radiation losses. We only consider loss mechanisms in straight waveguides.

Figure 4.30: Top view on a typical RIE-etched waveguide, showing the surface roughness.

The other loss mechanism, scattering, is caused by inhomogeneities in the waveguide. Volume scattering is generated by fluctuations of the refractive index in the bulk of the layers. Surface scattering is due to irregularities at interfaces between two regions with different refractive indices. Volume scattering and scattering at interfaces between two material layers are again negligible with the current status of growth techniques.
The dominant loss mechanism in integrated optical waveguides is scattering at the etched walls of the ridges. Most waveguides on PICs are etched with a Reactive Ion Etching (RIE) process[11], which has the unpleasant feature of creating a rough vertical surface (see figure 4.30). The high refractive index contrast between the semiconductor and the surrounding air makes that the irregularities become efficient scatterers. Surface scattering is responsible for practically the entire 2dB/cm that is typically observed in InP-based waveguides.

Narrow monomodal waveguides are typically used in InP-based PICs to avoid interference effects between different guided modes. The scattering at the local disturbances of the waveguide cross-section cause a coupling between the guided mode and radiation modes. Marcuse has studied the scattering process in detail by using a coupled mode theory [12][13]. We will not repeat the theory here, but will restrict ourselves to a formulation of the conclusions:

- The loss increases with increasing refractive index difference:
  $\alpha \sim \left(n^2_{\text{eff,core}} - n^2_{\text{eff,clad}}\right)^2$, with $n_{\text{eff,core}}$ and $n_{\text{eff,clad}}$ the effective indices of the core and the cladding slab waveguides respectively. This means that deeply etched waveguides ($n_{\text{eff,clad}} = 1$) will suffer a larger scattering loss then shallowly etched structures;

- Obviously, a higher loss is introduced when the amplitude of the roughness is higher;

- The loss is proportional to the intensity of the guided mode at the rough ridge wall. This is a very important conclusion that forms the basis for the low-loss waveguiding principle as discussed below.

[11] Reactive Ion Etching is needed because it allows to fabricate ridges with straight walls, irrespective of the crystal orientation. This is important for bends, where the orientation of the ridge changes continuously with respect to the crystal and where wet etching cannot be used by consequence.

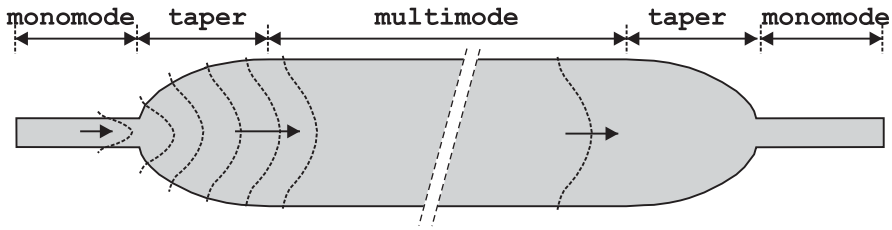


Figure 4.31: Configuration for low-loss propagation: the guided mode of a monomodal waveguide is adiabatically transformed into the fundamental guided mode of the broad waveguide.

The final conclusion suggests to reduce the intensity of the optical field at the ridge edges to lower the scattering loss. This can easily be realised by using broad waveguides. Indeed, for wider waveguides a larger fraction of the optical power is confined in the bulk of the waveguide at the expense of the power at the edges. Problems with the multi-modality of the broad guide can be eliminated by integrating an adiabatic taper between a (narrow) mono-modal waveguide and the broad section (figure 4.31). The adiabaticity of the design ensures that only the fundamental mode of the broad waveguide is excited. The design of the tapers and the characterisation of the devices were part of the M.Sc. thesis [14].

4.4.2 Design

It was decided to test the proposed principle on deeply etched waveguides because they exhibit the highest scattering losses and a reduction of the propagation loss will be more pronounced by consequence. A cross-section of the deeply etched waveguide is shown in figure 4.32. The layer stack is a typical passive waveguide structure that has already been used frequently at our the department.
The question remains about how broad the low-loss waveguides really have to be chosen before the scattering loss is sufficiently reduced. To decide on this we investigated the evolution of the effective index of the fundamental mode as a function of the guide width (figure 4.33). The evolution of the effective index is well-known from waveguide theory and typically strives asymptotically to the effective index of the central waveguide slab. The flat evolution of the effective index for wide ridges is an indication of the strong confinement in the central waveguide and hence a low intensity at the edges.
Also the confinement of the fundamental mode in a 50nm broad section at the edge of the waveguides was calculated as a function of the waveguide width. The width of the narrow slab was chosen as indicative for the expected amplitude of the wall roughness. Figure 4.34 shows the evolution of the confinement as a function of the width of the broad waveguide. It is difficult to conclude when the waveguide is broad enough, but the rapidly decaying curve ($\sim 1/w^3$) clearly shows that the intensity at the etched walls decreases drastically. Deeply etched waveguides with a width of 4, 6 and 10μm and 1.9μm wide monomodal reference waveguides have been investigated to observe the evolution of the propagation loss.

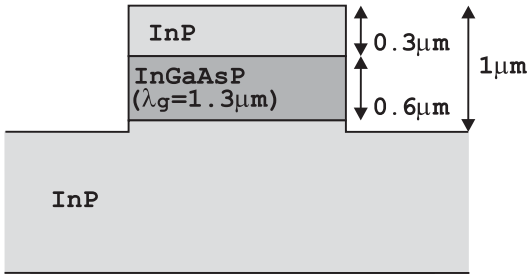


Figure 4.32: Cross-section of the deeply etched waveguides.

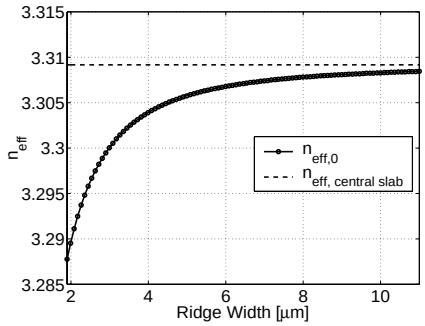


Figure 4.33: Effective index evolution of the fundamental mode of the structure shown in figure 4.32.

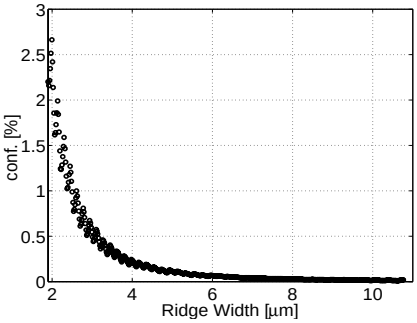


Figure 4.34: Evolution of the confinement in a 50nm wide region at the edges of the ridge.

The tapers that are needed to make the transition to the broad, multimoded waveguide have already been studied in the examples in section 3.5. It was found that the adiabaticity criterion leads to a parabolic evolution of the guide width along the taper. However, here we have chosen to implement exponential tapers instead. This option was chosen for the practical reason that the BPM tool that

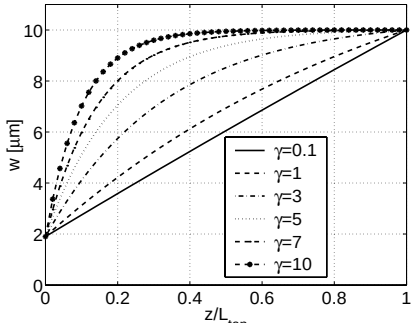


Figure 4.35: Some exponential taper shapes.

was used to verify the adiabaticity of the devices has a built-in model to define such exponential shapes (*Fimmprop* was not yet available at that time). The BPM program is a 2D tool, implying that the 2D waveguides first have to be reduced to a 1D slab by using the effective index method[12]. The exponential width evolution in the BPM program is defined by:

$$w(z) = w_{\text{in}} + \left(w_{\text{out}} - w_{\text{in}}\right)\frac{e^{-\gamma z/L_{tap}} - 1}{e^{-\gamma} - 1} \tag{4.4}$$

The shape of the exponential is determined by the choice of the damping factor γ[13]. A high γ-value results in a rapidly changing shape at the start of the taper and a slow evolution at the end. A linear evolution is obtained for γ approaching zero[14]. Some exponential shapes are shown in figure 4.35.

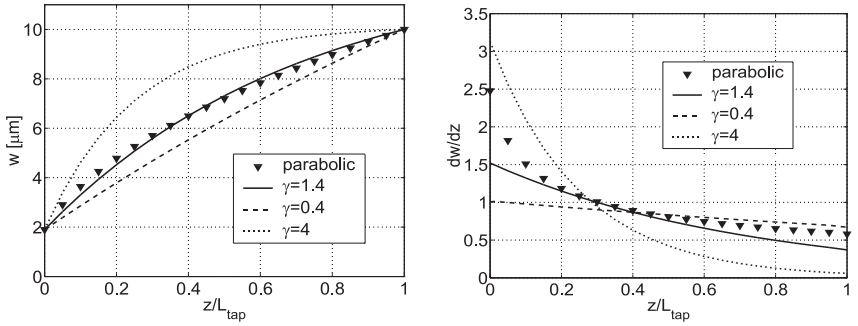
The fact that the exponential shape deviates from the parabolic evolution does not necessarily mean that a less efficient device is obtained. Indeed, the parabolic shape was found as the component with a constant disturbance along its length[15]. The only parameter left to be chosen was the length, which followed from the maximum allowed mode transformation loss. The exponential design, on the other hand, has two degrees of freedom: the damping factor γ and the length L_{tap}. The extra degree of freedom allows to choose where the loss should be situated predominantly. A high γ locates the loss mainly in the first section of the taper, while a low γ shifts the loss to the end. This is illustrated in figure 4.36. This figure shows the calculated parabolic shape that makes a transition between our monomodal waveguide (1.9μm) and a 10μm broad structure. Also the exponential curve that best approximates the parabola is shown ($\gamma = 1.4$). In order to obtain at least the same degree of adiabaticity, we stretched the device

[12] The effective index of the deeply etched slabs next to the central ridge can be replaced by 1.

[13] We only consider positive γ-values here since they have a decreasing slope for wider ridges (convex shape) as desired. A negative γ yields a concave shape.

[14] $\dfrac{e^{-\gamma z/L_{tap}} - 1}{e^{-\gamma} - 1} \approx \dfrac{-\gamma z/L_{tap}}{-\gamma}$

[15] This is what the adiabaticity criterion imposes for the taper shape by setting F(z)=Constant (see remark 2 in section 3.5)



(a) (b)

Figure 4.36: A parabolic taper and the best matching exponential ($\gamma = 1.4$) for a transition between a monomodal (1.9μm wide) waveguide and a 10μm wide guide. Also an exponential that concentrates the loss at the beginning of the taper ($\gamma = 4$) and one concentrating the loss at the end ($\gamma = 0.4$) are shown: a) shapes and b) slope of the shapes.

until the slope of the shape is lower then or equal to the slope of the parabolic evolution everywhere. The price to be paid is the fact that a 18% longer device is needed[16]. To demonstrate the influence of the choice of the damping factor γ, we also plotted the shapes and the slopes corresponding with $\gamma = 0.4$ and $\gamma = 4$ (all exponentials having the same length). For $\gamma = 4$ the shape tapers faster than the parabolic design at the beginning of the device, while a slower tapering is seen at the end. The loss will thus be situated mainly at the beginning of the taper. For $\gamma = 0.4$ the opposite behaviour is observed.
The example shown in figure 4.36 bridges a rather large width difference. The ideal parabolic shape tends to a linear evolution as $\Delta w \to 0$ (see (3.42)). Also the best exponential evolution will evolve towards a linear design ($\gamma \to 0$).

We will not list all the specific variations that were implemented on the test mask, but limit ourselves to mentioning that a variation of exponential shapes, with different damping factors and lengths, have been defined. Also linear tapers were implemented to confirm the superior behaviour of the exponential shapes. The taper angle of the linear device should always be larger than or equal to the lowest local taper angle of the exponential design. This implies that the angle corresponding with the broadest ridges has to be chosen (see section 3.5).

[16]This does not mean that the exponential shape is inferior to the parabolic evolution. The extra length is a result of the fact that the slope is at the highest equal to the parabolic slope. Simulations are needed to determine the length required for an identical degree of adiabaticity.

4.4.3 Characterisation and results

As mentioned in the previous section, we designed a mask that contains a variety of structures. In summary, we defined straight monomodal reference waveguides (1.9μm wide), and structures to determine the propagation loss of the fundamental mode of 4, 6 and 10μm wide multimode waveguides. Several versions, with different linear and exponential tapers, were implemented.

The measurement technique that was used here, and which is also the standard choice for loss measurements on passive waveguides, is the so-called Hakki-Paoli technique (see appendix). This method is based on the fact that the passive waveguide structure forms a cavity with the cleaved facets of the chip serving as mirrors. The contrast of the interference fringes that are observed in the output power when e.g. the wavelength of the incident light is changed, is dependent on the total loss encountered by the guided mode as it propagates from one facet to the other[17,18]. The Hakki-Paoli technique is a so-called relative characterisation method, meaning that only the ratio of the input power to the output power plays a role in the calculation of the propagation loss. In other words, the measurement does not depend on the exact amount of power coupled into the waveguide, or also, it is independent of the coupling efficiency obtained in a particular measurement. This is a very important issue, since absolute measurements require consecutive measurements with *exactly the same* coupling efficiency. A well-known absolute characterisation technique is for instance the 'cut-back' method. In this method, light is coupled into a waveguide and the power emerging from the other side of the waveguide is registered. Then, a section is cleaved from the waveguide and the measurement is repeated on the short waveguide. From these two measurements (sometimes more steps are used), the waveguide loss factor (dB/cm) can be calculated, provided that the coupling conditions are identical in both cases. If we bear in mind the discussion that was given in chapter 1 on the coupling problem and the sensitivity of the coupling efficiency to small misalignments, it is soon concluded that absolute methods are to be avoided.

The loss measured with the Hakki-Paoli technique is the total loss α encountered by the guided mode as it propagates from one facet to the other: $\alpha = \int_0^L \alpha(z')dz'$. The broad low-loss waveguides are always implemented together with monomodal sections and tapers (see figure 4.31). The measured loss will hence be a linear combination of the loss in the monomodal input- and output section, the taper excess loss and the propagation loss of the broad waveguide: $\alpha = \alpha_{\text{mono}} \cdot L_{\text{mono}} + 2.\alpha_{\text{tap}} + \alpha_{\text{broad}} \cdot L_{\text{broad}}$. The following structures were defined on the testchip, so that these different loss factors can be isolated by solving a set of two linear equations with α_{tap} and α_{broad} as unknowns[19] (see also figure 4.37):

[17] We used a 1.55μm DFB laser with a built-in temperature controller. A tuning of the temperature also results in a variation of the wavelength.

[18] More details on the Hakki-Paoli characterisation technique can be found in the appendix.

[19] The loss of the monomodal waveguides is known from measurements on the monomodal reference waveguide.

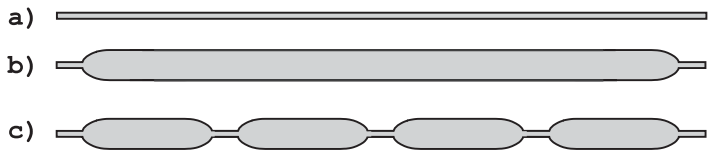


Figure 4.37: Structures implemented on the test mask to extract the taper loss and the propagation loss of the monomodal and the broad waveguide: a) straight monomodal waveguides; b) monomode-taper-multimode-taper-monomode sequence and c) a succession of the sequence given in b).

- monomodal waveguides;
- a first structure with broad waveguides, consisting of a monomodal section at the input, a taper to make the transition to the broad waveguide, a broad section and a taper to go back to the monomodal output waveguide;
- a concatenation of the mono-taper-multi-taper-mono sequence (five times repeated).

Table 4.4.3 and figure 4.38 summarise the average losses as extracted from the Hakki-Paoli measurements. The broad waveguides are found to reduce the propagation loss quite impressively from 5dB/cm for the monomodal reference waveguides to a few tenths of dB/cm for the broad devices. The 5dB/cm loss of the monomodal reference waveguides is rather high, but underlines the successful application of broad waveguides for low-loss interconnections even stronger. The broad waveguides, excited by a linear taper exhibit a slightly higher loss than the structures with exponential tapers. This might be attributed to the fact that the linear tapers are a bit less efficient, resulting in a small fraction of the incident power being transferred to higher order modes in the broad section.

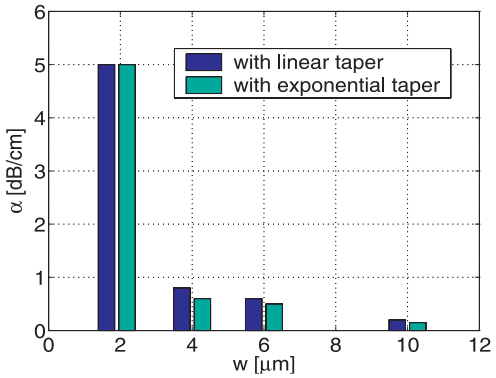


Figure 4.38: Measured average loss of monomodal reference waveguides and broad low-loss waveguides.

Width	**Configuration with exponential taper**	**Configuration with linear taper**
1.9 μm	5 dB/cm	
4 μm	0.6 dB/cm	0.8 dB/cm
6 μm	0.5 dB/cm	0.6 dB/cm
10 μm	0.15 dB/cm	0.2 dB/cm

Table 4.3: Measured average loss of monomodal reference waveguides and broad low-loss waveguides.

These higher order modes are characterised by larger losses because of their higher intensity at the waveguide edges. The different modes propagating in the broad section all have a different effective index, making that they arrive out of phase at the output taper that makes the transition back to the monomodal output waveguide. This effect results in a higher measured waveguide loss. Also the fact that the linear tapers remain narrow over a longer distance contributes to the higher taper loss for linear devices.

The loss introduced by the tapers is very low (0.05dB or less) for both linear and exponential devices. Figure 4.39 shows the average loss of the linear devices. The higher loss for the 4μm devices is odd, but the decreasing loss for longer tapers is clearly visible. It should be mentioned, however, that differences between the extracted loss values of different device designs are very small and might be well below the characterisation accuracy. This is also the case for the exponential devices. Figure 4.40 summarises the measured average taper loss of several exponential tapers making a transition to a 10μm wide waveguide. As a general conclusion, we can state that, in comparison with linear tapers, similar performances are obtained for shorter taper lengths.

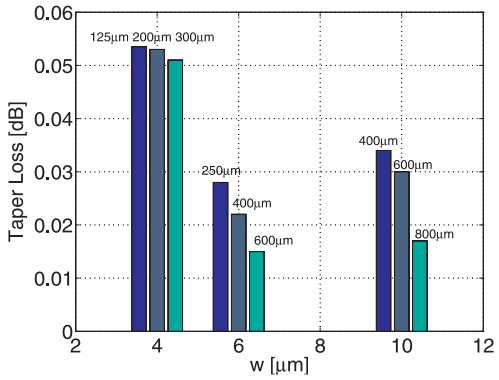


Figure 4.39: Measured average loss of linear tapers as a function of their length for transitions to a 4, 6 and 10μm wide waveguide.

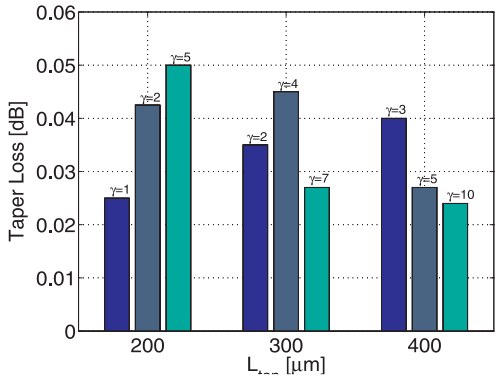


Figure 4.40: Measured average loss of exponential tapers as a function of their length (tapers making a transition to a 10μm wide waveguide are shown) for different exponential damping factors γ.

References

[1] R. Zengerle, O. Leminger, W. Weiershausen, K. Faltin, and B. Hübner, "Laterally tapered InP-InGaAsP waveguides for low-loss chip-to-fiber butt coupling: A comparison on different configurations," *IEEE Photon. Technol. Lett.*, vol. 7, pp. 532–534, 1995.

[2] "Development of a photonic integrated circuit for optical phase control at 1.06μm." ESTEC/Contract No. 11450/95/NL/PB, Work order No. 2.

[3] G. Flamand, *Development of a new generation of* I*n*P *photonic integrated circuits for optical phase control at 1.06μm.* PhD thesis, Ghent University, 2000.

[4] INSPEC, "Properties of Indium Phosphide," 1991.

[5] G. Vawter, R. Smith, H. Hou, and J. Wendt, "Semiconductor laser with tapered-rib adiabatic-following fiber coupler for expanded output-mode diameter," *IEEE Photon. Technol. Lett.*, vol. 9, no. 4, pp. 425–427, 1997.

[6] V. Vusirikala, S. Saini, R. Bartolo, S. Agarwala, R. Whaley, F. Johnson, D. Stone, and M. Dagenais, "1.55-μm InGaAsP-InP laser arrays with integrated-mode expanders fabricated using a single epitaxial growth," *IEEE J. Sel. Top. Quant. Electron.*, vol. 3, no. 6, pp. 1332–1343, 1997.

[7] M. Duguay, Y. Kokubun, and T. Koch, "Antiresonant reflecting optical waveguides in SiO_2-/Si multilayer structures," *Appl. Phys. Lett.*, vol. 49, no. 1, pp. 13–15, 1986.

[8] M. Galarza, K. DeMesel, C. Aramburu, and R. Baets, "Semiconductor optical component utilizing leaky structures to match the mode of rib waveguides to that of the fiber."

[9] M. Galarza, *PhD thesis in preparation.* PhD thesis, Ghent University.

[10] T. Koch, U. Koren, G. Boyd, P. Corvini, and M. Duguay, "Antiresonant reflecting optical waveguides for III-V integrated optics," *Electron. Lett.*, vol. 23, no. 5, 1987.

[11] J. Kubica, "Numerical analysis of InP-InGaAsP ARROW waveguides using transfer matrix approach," *IEEE J. Lightwave Technol.*, vol. 10, no. 6, pp. 767–771, 1992.

[12] D. Marcuse, *Theory of dielectric optical waveguides.* New York: Academic Press, 1974.

[13] D. Marcuse, "Mode conversion caused by surface imperfections of a dielectric slab waveguide," *The* B*ell system technical journal*, vol. 48, no. 10, pp. 3187–3215, 1969.

[14] M. Vanwolleghem, "Optimalisatie van verliezen in InP-gebaseerde golfgeleider-IC's," Master's thesis, Ghent University, 1999.

Chapter 5

Selective wet oxidation

5.1 Background

From the beginning of the extensive study of III-V semiconductor compounds, the Al-containing binaries have been listed as unstable. This unstability is due to the extremely reactive nature of the Al, especially in a moist environment. Through the years, atmospheric water vapour reacts with the buried AlAs layers via the crystal edges and through pinholes in the upper cap. The oxides that are formed expand and cause the upper layers to peel off. Also $Al_xGa_{1-x}As$ with a high Al-content x is known to show this behaviour. The stability of the $Al_xGa_{1-x}As$ layers is an important issue for the reliability of semiconductor laser structures, where it is generally being used as a cladding layer material. For this reason, Dallesasse et al. started a thorough investigation of this slow environmental degradation by inspection of wafers that were grown several years before and that were kept under normal room conditions [1][2]. This examination revealed that thicker ($\geq 0.1\mu$m) $Al_xGa_{1-x}As$ layers are unstable for compositions x$\geq$0.85. The degree of instability depends strongly on the alloy composition x. Defects are observed for x$\approx$0.85 on some crystals, while a clear deterioration is seen for x$\approx$0.95. Thicker AlAs layers even undergo a complete damaging.

An attempt was made to accelerate the degradation process by passing water vapour of the samples and raising the temperature as is customary for the oxidation of Si. However, instead of creating a destructive hydrolisation, a dense and transparent native oxide of high quality was obtained [3]. The properties of this new type of oxide and the environment in which it is formed, are completely different from the atmospheric oxidation. A comparison between both is given in table 5.1.

Atmospheric oxide	'Steam' oxide
Forms at $<100^oC$	Forms at $\geq 370^oC$ in steam environment
Forms over a period of years	Forms within tens of minutes to hours
Chemically and mechanically unstable	Chemically and mechanically stable
Expands upon formation	Shrinks upon formation

Table 5.1: A comparison between atmospheric and 'steam' oxidation.

5.2 The oxidation setup

In the following we will consider *lateral* oxidation of buried layers (figure 5.1), meaning that structures are etched in the sample to expose the high Al-content layers to the steam environment. The oxide is then formed from the exposed edge and the oxidation front moves deeper into the structure as the process proceeds.
A typical oxidation set-up is shown in figure 5.2. A N_2 carrier gas is bubbled through H_2O that is kept at a temperature of 80-95^{o}C. The saturated water vapour is then conducted to the oven containing the sample and typically heated to 370-450^{o}C.

Figure 5.3 shows the set-up that was built at our department[1]. The oven is constructed with a Liquid Phase Epitaxy (LPE) quartz tube that will contain the sample and through which the water vapour will be sent. The tube can withstand temperatures up to 800^{o}C, which is largely higher than the temperatures used in the oxidation process. The temperature is measured by two thermocouples. A first one is placed nearby the sample and is connected to a control circuit that actuates the current through the heating spiral. The accuracy of the oven temperature is typically $\pm$ 0.1^{o}C. The second thermocouple is used for security and is connected to a switch that automatically turns the heating current off in case the oven temperature exceeds a certain maximum value.
N_2 was chosen as a carrier gas. It is passed through a bubbler that is filled with de-ionised water and placed in a thermostatic bath to regulate the water temperature. The concentration C(T) of water vapour in the wet nitrogen leaving the bottle is

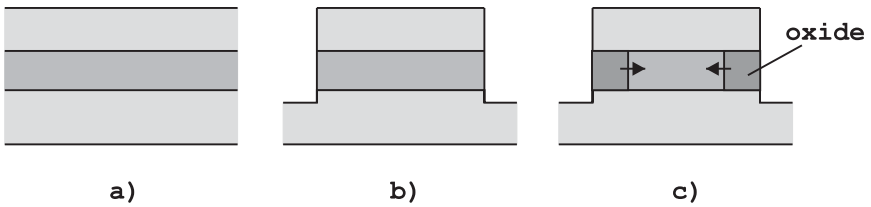


Figure 5.1: Lateral oxidation: the layer structure (a) is etched (b) to expose the high Al-content layers to the oxidising steam environment (c).

[1]The set-up was built by Carl Sys. The characterisation of the oxidation process was done cooperatively.

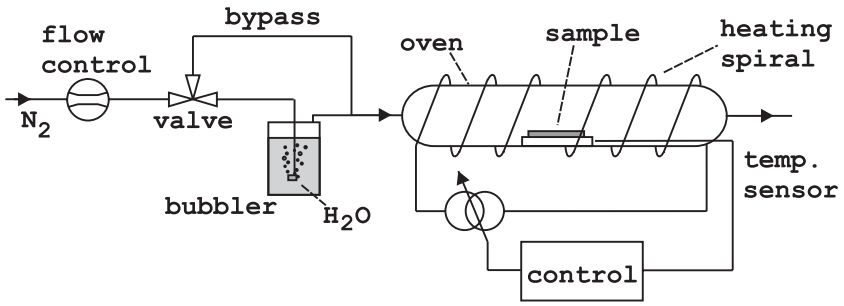


Figure 5.2: Schematic diagram of the oxidation system.

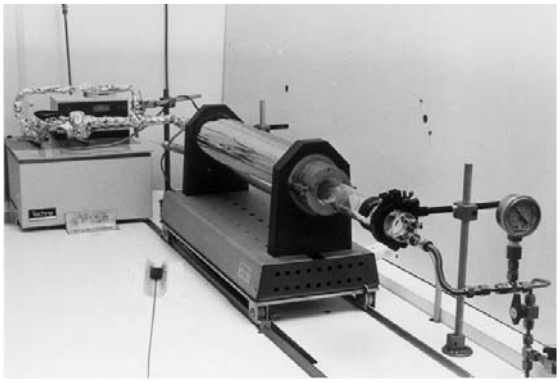


Figure 5.3: Photograph of the oxidation set-up used at INTEC.

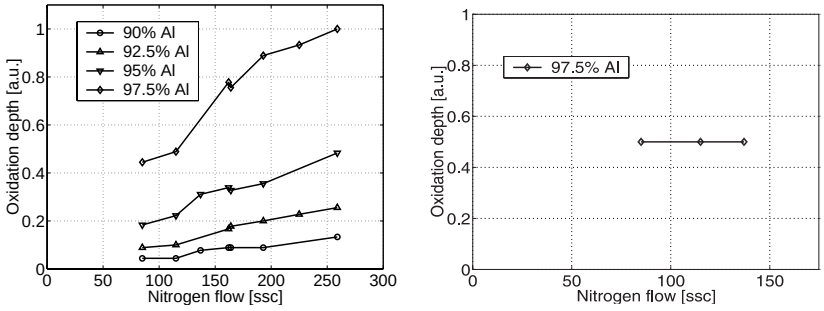
proportional to the vapour pressure of the water:

$$C(T) = \frac{p_{H_2O}(T)}{p_{amb}} \tag{5.1}$$

with $p_{H_2O}(T)$ the vapour pressure of the water at a temperature T and p_{amb} the atmospheric pressure. The strong temperature dependence of the vapour pressure of water also makes the water concentration C(T) in the wet nitrogen a strong function of temperature. When a flow F_{N_2} of dry N_2 enters the bubbler, then the flow of water vapour carried along to the oven is given by:

$$F_{H_2O} = \frac{C(T)}{1 - C(T)} F_{N_2} \tag{5.2}$$

The supply of water vapour to the oven is an important parameter in the oxidation process. A low F_{H_2O} will have a significant impact on the oxidation rate. It is therefore important to choose the bubbler temperature and the nitrogen flow high enough to ensure that the oven is fully saturated with water vapour. By doing so, we eliminate two parameters that affect the oxidation rate ($T_{bubbler}$ and F_{N_2}).



(a) (b)

Figure 5.4: Evolution of the oxide depth versus nitrogen flow for a constant bubbler temperature (=90^{o}C), a) without heating of the lines between bubbler and oven. b) with line heating.

Test samples were oxidised to find out the minimum temperature and nitrogen flow that fully saturate the oven. Figure 5.4(a) shows the evolution of the oxide depth after oxidation for a fixed time as a function of N_2-flow (bubbler temperature constant at 90^{o}C). It is seen that even for high flow values the oxide depth still doesn't saturate. This is attributed to condensation in the unheated lines that guide the wet nitrogen to the oven. As a consequence only a small fraction of the water vapour actually reaches the oven. The lines were therefore wrapped up with flexible heaters that keep the line temperature around 100^{o}C. The experiment was repeated with line heating and the oxide depth was found to be constant over the entire range studied (figure 5.4(b)). Figure 5.5 shows the influence of the bubbler temperature on the oxidation rate for a constant nitrogen flow (120 sccm[2]). The regime of interest is reached for bubbler temperatures of 70^{o}C and higher.
In conclusion, the oven is fully saturated with water vapour if a bubbler temperature of 80^{o}C is chosen in combination with a nitrogen flow of 120sccm. These will be the settings for all further oxidation runs.

A typical oxidation run proceeds as follows. First the thermostatic bath and the oven are switched on. During this heating period, dry nitrogen is forced through the oven to expell all oxygen. When the oven temperature has stabilised after about half an hour, the nitrogen flow is sent through the bubbler and the oven starts to fill with wet nitrogen. After five minutes the oven is fully saturated with water vapour (condensation is observed at the end of the quartz tube) and the cold sample is slid in the oven. The sample goes through a short transient heating period (typically a few minutes). Care is taken to follow the same procedure for every oxidation run to enhance the reproducibility of the oxidation.

[2] 'sccm' stands for standard cubic centimeter per minute.

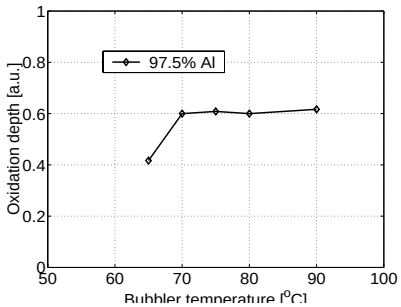


Figure 5.5: Influence of the bubbler temperature on the oxidation depth (for a constant nitrogen flow of 120 sccm).

5.3 The oxidation reaction

Identification by means of Raman spectroscopy of intermediate species formed during the oxidation process has been combined with thermochemical calculations to determine the reactions involved in the conversion of $Al_xGa_{1-x}As$ into oxide [4]. The orders of magnitude faster oxidation of high-Al content semiconductors with water in an oxygen-free environment than with pure oxygen is explained by a completely different reaction mechanism based on the oxidation of $Al_xGa_{1-x}As$ by H^+ instead of O_2. In a first step, the reaction between $Al_xGa_{1-x}As$ and water produces Al_2O_3 and the intermediate reaction product As_2O_3. The reaction for pure AlAs is written as

$$2AlAs + 6H_2O \rightarrow Al_2O_3 + As_2O_3 + 6H_2 \tag{5.3}$$

The As_2O_3 is further converted by the following reactions into volatile As that can escape from the oxidised film:

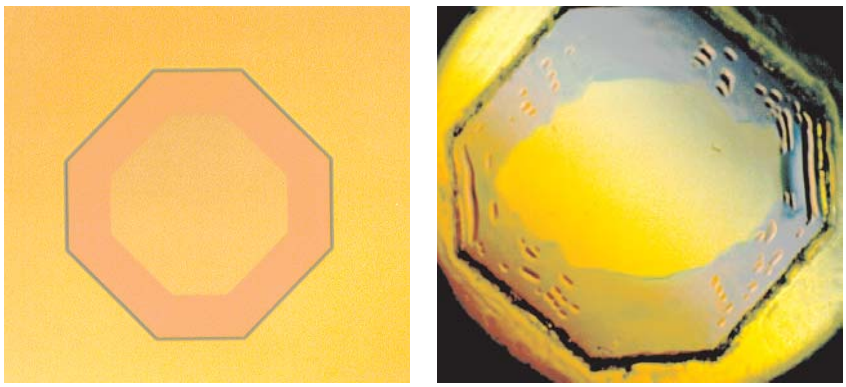
$$As_2O_3 + 3H_2 \rightarrow 2As(\uparrow) + 3H_2O \tag{5.4}$$

$$As_2O_3 + 6H \rightarrow 2As(\uparrow) + 3H_2O \tag{5.5}$$

Transmission electron microscope (TEM) studies revealed that the oxide layer consists of a polycrystalline Al_2O_3 with grain sizes down to 4nm and a dense amorphous reaction front [5] [6]. This dense region consists, according to [7], of Al_2O_3 and the intermediate reaction product As_2O_3.
Ga is not involved in the oxidation of $Al_xGa_{1-x}As$, but is incorporated in the oxide as $(Al_xGa_{1-x})_2O_3$. The incorporation of a small Ga-fraction also has a favourable influence on the mechanical stability of the oxide, as will be discussed in the next section.
When O_2 is present, the wet oxidation process is suppressed by the loss of hydrogen through reaction with O_2 to form H_2O in preference to the reduction of As_2O_3 to As. This also explains the observed dependence of oxidation rate on the type of carrier gas used. Inert gases like N_2 or Ar show the same rate, but addition of O_2 reduces the rate rapidly and no oxidation is observed for wet or dry O_2.

5.4 Oxide properties

Oxidation of $Al_xGa_{1-x}As$ results in a mechanically stable oxide, in contrast to AlAs where excessive strain causes delamination of the upper layers [8]. This mechanical instability is caused by the large shrinkage (12-13% [9][10]) of the oxide as compared to the AlAs semiconductor. Addition of a small fraction of Ga, however, is found to reduce the shrinkage significantly (6.7% reported in [5]) and to result in stable oxides. This is illustrated in figure 5.6 where a sample containing an AlAs-layer and a sample with an $Al_{0.92}Ga_{0.08}As$-layer are shown. The samples were oxidised together for 150 minutes at a temperature of 475^oC. The oxide in the $Al_{0.92}Ga_{0.08}As$ has formed uniformly from the edges, while this is clearly not the case in the AlAs-sample. Cracks can even be observed in the oxidised part.
The refractive index of the oxidised AlAs and AlGaAs is generally situated between 1.5 and 1.66. The scatter in data is possibly due to the influence of the fabrication method, differences in layer structures that result in different stress levels in the oxide or the use of different starting materials (AlAs versus AlGaAs) [11]. The refractive index contrast between oxides originating from $Al_{0.4}Ga_{0.6}As$(!)[3] and $Al_{0.8}Ga_{0.2}As$ was even sufficient to create fully oxidised waveguides [12]. The index of the oxidised $Al_{0.8}Ga_{0.2}As$ core was 1.59, while the oxide in the cladding had an index of 1.53. This contrast gives us an idea of the influence of the Ga-content on the refractive index of the oxide.



(a) (b)

Figure 5.6: Top view on two laterally oxidised structures containing (a) an $Al_{0.92}Ga_{0.08}As$-layer and (b) an AlAs-layer.

[3] This unusually low Al-fraction required a long oxidation (213 minutes) at a high temperature (500^oC).

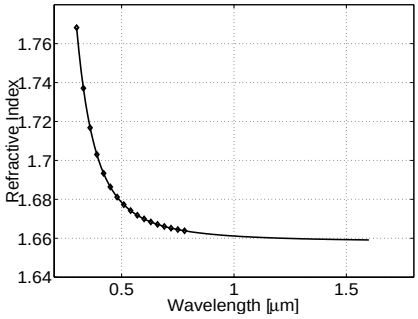


Figure 5.7: Refractive index of oxidised $Al_{0.98}Ga_{0.02}As$ as a function of wavelength (after [13]).

The refractive index of $Al_{0.98}Ga_{0.02}As$ was investigated by spectroscopic ellipsometry in the 300-800nm range and extrapolated to the 1.55μm region using a Cauchy model of the form $n(\lambda) = A + B/\lambda^2 + C/\lambda^4$ [13]. The results are reproduced in figure 5.7 and show that the index is practically constant in the 0.9-1.6μm range (n $\approx$ 1.66).

5.5 Kinetics and oxidation parameters

5.5.1 Dependence on time

The oxidation reaction can be modelled as a typical reaction/diffusion process, for which the temporal dependence of the oxide thickness is given by $t = \frac{x_{ox}^2}{k_p} + \frac{x_{ox}}{k_l}$ [14][15]. The parabolic constant k_p is proportional to the diffusion constant of the oxidant through the aluminium oxide that is formed. The linear constant k_l is proportional to the rate constant of the oxidation reaction at the oxidation front, assuming that the reaction is not limited by the transport of water vapour from the furnace atmosphere through the oxide.
A reaction limited process, with oxide being formed at a constant rate, is obtained when this diffusion process is able to supply the reactants faster to the oxide front than they can be consumed in the oxidation reaction. This is typically observed for small oxide depths, or hence for short oxidation times. Alternatively, when the oxidant diffusion across the oxide is the limiting mechanism, the oxidation follows a parabolic growth law and a diffusion-limited process is obtained. This is the case when the oxidant has to diffuse through long stretches of oxide, or else, at longer oxidation times.
The oxide growth is also found to be reaction limited at low temperatures. Conversely, diffusion across the oxide is the rate limiting mechanism at higher temperatures. This behaviour is explained by the fact that the oxidation reaction and the diffusion process are both thermally activated processes, each characterised

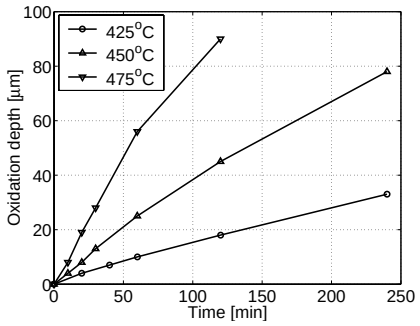


Figure 5.8: Evolution of the oxide depth in $Al_0.95Ga_0.05As$ as a function of time for different oxidation temperatures.

by an activation energy[4]. At lower temperatures, the activation energy of the oxidation reaction is higher than for the diffusion process. However, as the temperature increases, the reaction rate increases faster than the diffusion rate and the diffusion-limited regime is entered.
The transition from linear to parabolic growth is illustrated in figure 5.8. The oxidation is seen to proceed linearly for 425^{o}C, while the curve for 475^{o}C is clearly curved.

5.5.2 Dependence on temperature

According to figure 5.8 the lateral oxidation depth increases with increasing oxidation temperature T. An exponential dependence between oxidation rate and temperature has been reported [16].
This was confirmed in our tests. Samples containing $Al_xGa_{1-x}As$-layers with x=90% and 95% were oxidised for 30min at 425^{o}C, 450^{o}C and 475^{o}C. Figure 5.9 plots the the logarithm of the oxidation rate as a function of temperature[5]. The linear evolution of the curve indicates the exponential dependence of the oxidation rate on temperature.
The strong influence of the temperature on the oxidation rate necessitates a close control on the oven temperature.

5.5.3 Dependence on Al-content

The oxidation rate is strongly dependent on the Al-content of the material. This is an important feature with respect to the use of oxidation in optical components since it allows to *selectively* oxidise buried layers with a high Al-content (see section 5.6). Samples with an identical layer structure, but with a different Al-content for

[4]The rate r of a thermally activated process is given by $r = r_0.e^{-E_a/k_B.T}$, with E_a the activation energy and k_B Boltzmann's constant.

[5]A linear growth is supposed as a function of time (which is justified from figure 5.8 for the calculation of the oxidation rate).

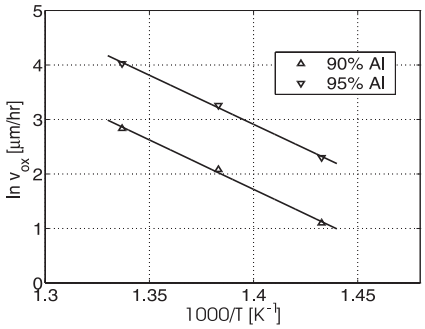


Figure 5.9: Logarithmic plot of the oxidation rate illustrating the exponential evolution of the oxidation rate as a function of temperature.

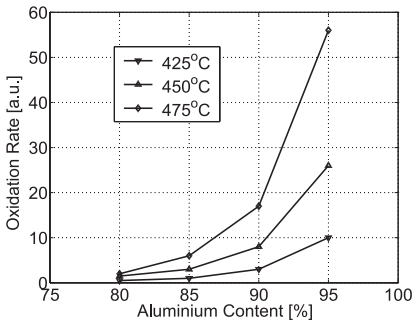


Figure 5.10: Normalised oxidation rate as a function of aluminium content for different temperatures.

the high-Al layer, were oxidised together. This was repeated for several temperatures. The resulting oxidation rates are shown in figure 5.10. The values of each set are normalised with respect to the rate of the sample with 95% aluminium for an easy comparison. No results are shown for pure AlAs because the oxidation proceeds in a very non-uniform way and no clear oxide length can be determined. Layers with less than 80% of aluminium are found to oxidise unpractically slow. These results agree very well with experiments reported in literature [16][17][18].

5.5.4 Dependence on layer thickness

The oxidation rate decreases rapidly as the layer thickness is reduced below 80nm and nearly stops at a thickness of 10-20nm [19][20]. A theoretical model of this phenomenon has been constructed through a thermo-dynamical study of the surface free energy of the oxide front [15]. It was concluded that the oxidation rate has a $e^{-d_0/d}$-dependence on the layer thickness d.
A layer structure containing nine $Al_{0.95}Ga_{0.05}As$ layers with different thicknesses (40, 50, 65, 85, 110, 140, 175, 215 and 260nm) and separated by 70nm GaAs

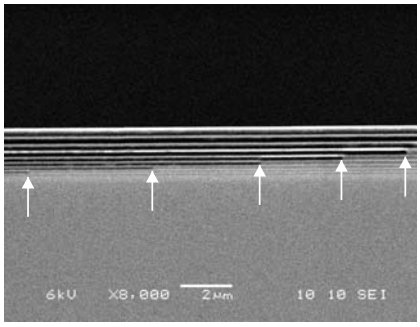


Figure 5.11: SEM-photograph of the structure to determine the influence of the layer thickness on the oxidation process. The arrows indicate the oxide front.

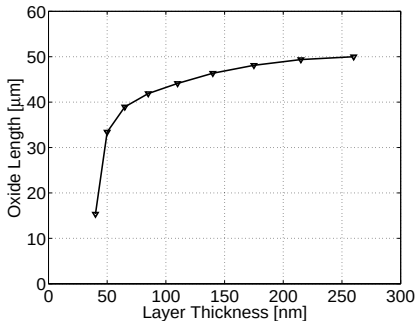


Figure 5.12: Influence of the layer thickness on the oxide length ($Al_{0.95}Ga_{0.05}As$, 70 minutes at 450°C).

barriers was grown to investigate this effect. The structure was oxidised for 70 minutes at 450^oC. Figure 5.11 shows a SEM-photograph of the structure. The oxide is clearly visible as the dark region extending from the left, with the oxide front indicated by the white arrows. The measured oxide length was found to decrease rapidly for layers thinner than 60-70nm (figure 5.12), again in agreement with literature.

5.5.5 Dependence on doping type

The conductivity type was also found to affect the oxidation rate, with p-type samples oxidising more rapidly than n-type samples [21]. The effect is not related to the type of dopant used and no difference was observed in the chemical structure of the oxide. This indicates that the doping type only affects the oxidation rate and not the composition of the oxide. The influence of the doping type on the oxidation rate is explained by the influence of charged defects on the diffusion of oxidant ions in the oxide. Also the oxidation reaction itself might be influenced by

the doping, in a similar way as the enhanced electrolyte etching of p-doped GaAs or AlGaAs.

5.6 Applications

Since the discovery of stable oxides of AlGaAs by Holonyak and coworkers the number of publications on the use of the selective wet oxidation process for the fabrication of optical components has grown exponentially. The applications are numerous and significant improvements in device characteristics and simplifications in processing schemes have been obtained. The success of the wet oxidation is mainly attributed to:

- the high selectivity of the process, allowing to selectively oxidise buried high-Al layers without affecting the surrounding material;
- the low refractive index and hence the high index contrast with the surrounding semiconductor material;
- the fact that the oxide is an electrical isolator.

It was immediately recognised that the high-quality oxide could serve as a passivation to protect the AlAs-layers in semiconductor lasers from further destructive oxidation [3].
Shortly after that, a first proposal to create current apertures for stripe-geometry gain-guided lasers was presented (figure 5.13(a)) [22]. Stripes were etched in the thin GaAs cap layer to expose the underlying $Al_{0.8}Ga_{0.2}As$ upper cladding next to the unetched GaAs. Subsequent oxidation formed a 150nm thick oxide next to (and also partly under) the ridge, starting from the top of the layer. The sample

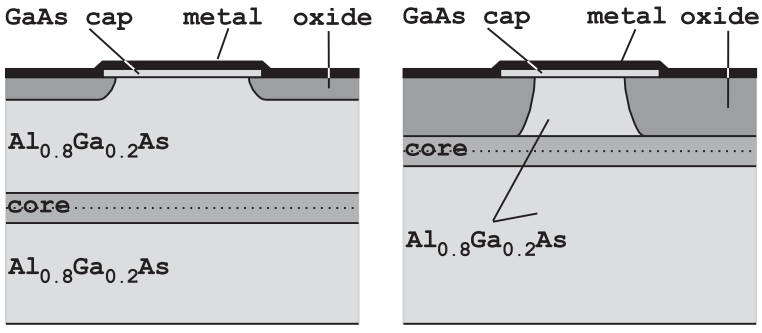


Figure 5.13: a) Gain-guided laser with oxide-defined current window and b) index-guided laser with both optical and current oxide confinement. Both devices are defined by vertical oxidation.

was then metallised across the native oxide and onto the GaAs cap layer. The metal is reported to adhere much better on the native oxide than on usually deposited dielectrics. The electrical isolation provided by the dielectric oxide thus creates a current aperture that constrains the current injection to the GaAs contact.

Oxide-confined gain-guided lasers are simple to fabricate but current spreading and lateral carrier diffusion in the active layer adversely influence its characteristics and limit the practical applications of these devices. It was an obvious evolution to also use the low refractive index of the oxide to form an optical waveguide in the lateral direction, by bringing the oxide closer to the core region. The first index-guided oxide-confined lasers were similar to the gain-guided structure discussed above (except for the thinner upper cladding) but was completely oxidised down to the active region (see figure 5.13(b)) [23][24]. The use of the oxide for both electrical and optical confinement automatically aligns the current aperture with the laser mode. Continuous-wave (CW) threshold currents of 8mA and a quantum efficiency of 60% were obtained for 4μm wide and 550μm long laser cavities operating at room-temperature.

Lateral oxidation, which makes uses of the high selectivity of the oxidation process to create buried current apertures, was introduced in a laser design comprising a 300nm thick oxide layer above and below the active core region, as shown in figure 5.14 [25]. The double oxide configuration was chosen for its higher lateral effective index contrast (Δn = 0.03) and the strong optical confinement of the laser mode in the quantum well. The wafer was patterned with 10μm wide Si_3N_4 stripes and etched in a wet etching solution (H_2O:H_2O_2:HSO_4) until the top of the lower high-Al layer was exposed. CW threshold currents of 8mA and quantum efficiencies of 61% were found for 2.5μm wide and 305μm long cavities.
Since this first publication on laterally oxidised lasers, several other groups have reported on devices incorporating single and double oxide windows with oxide thicknesses ranging from 30nm to 100nm [26][27][28]. The oxide layers were typically thinner than 300nm to avoid a possible excess stress. Threshold currents down to 3.5mA and external quantum efficiencies of 82% have been reported for 1.8μm wide and 400μm long cavities [26].

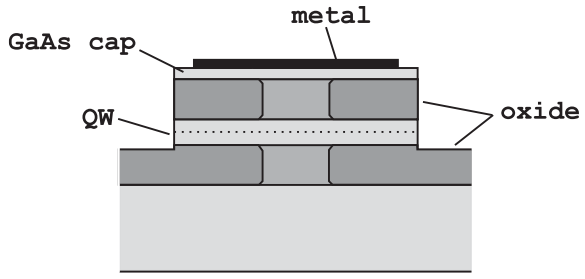
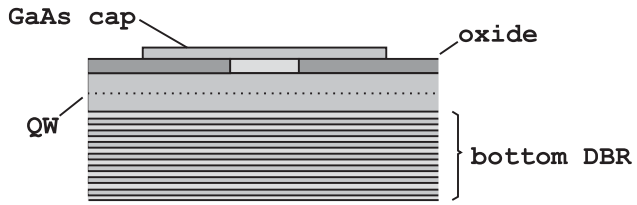


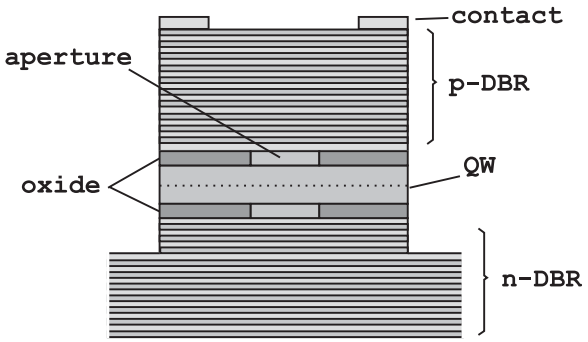
Figure 5.14: A laterally oxidised laser with a double oxide window.

The most significant improvements realised by introducing the selective wet oxidation technique have without any doubt been obtained in the domain of vertical cavity surface emitting laser diodes (VCSELs), where the oxide is used to create buried current apertures and high contrast distributed Bragg reflectors (DBRs). The use of native oxide to form a current aperture was introduced in [29] for a hybrid VCSEL incorporating a top dielectric DBR. The selective oxidation allows to define laser cavities with very small diameters, helping to reduce VCSEL threshold currents [30]. Also monolithic VCSELs with semiconductor DBRs were soon reported [31]. Both designs are shown in figure 5.15.

The high refractive index step between AlGaAs and the native oxide also permits to fabricate monolithic high-contrast GaAs/Al_2O_3 Bragg mirrors [10][32]. In traditional semiconductor DBR mirrors, the low refractive index contrast necessitates the use of many material pairs to achieve a reflectivity of more than 99%. Furthermore, the reflectivity and spectral bandwidth are very sensitive to the thickness of the layers. The use of two materials with a much larger refractive index difference



(a)



(b)

Figure 5.15: a) Oxide-confined current aperture before the deposition of the contacts and the top dielectric DBR mirror and b) buried oxide-confined current aperture in combination with monolithic DBR mirrors.

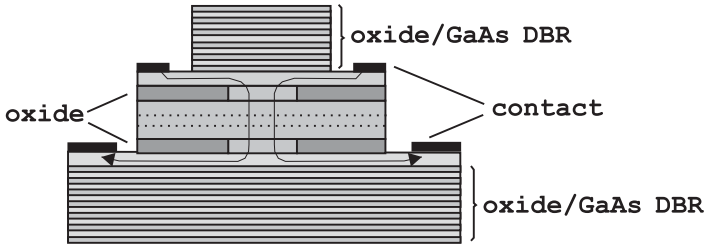


Figure 5.16: VCSEL with an oxide-defined current aperture in combination with high-contrast GaAs/Al_2O_3 DBR mirrors.

relaxes this constraint, reduces the required number of mirror pairs and increases the spectral bandwidth at the same time.

In [33] the first electrically pumped VCSELs incorporating both an oxide defined current aperture and two oxidised DBR mirrors were reported, exhibiting threshold currents of 160μA for a 8μm diameter aperture (figure 5.16). Of course, intracavity contacts are needed for current injection to bypass the the isolating Al_2O_3 in the DBR mirrors [34].

The AlGaAs oxidation has also been introduced successfully in a variety of other applications. We already mentioned the fully oxidised waveguides of [12] in section 5.4. Planar waveguide bends have been demonstrated in [35]. In [36], the oxidation technique was introduced to create a vertical fibre-coupling structure.
Low loss single mode optical waveguides with metal electrodes were fabricated with Al_2O_3 layers as upper cladding [37]. Metal electrodes in close proximity to optical waveguides are known to lead to excess optical propagation loss. A possible solution is to increase the thickness of the semiconductor upper cladding layer, but this would require a layer thickness of about 2μm to reduce the loss sufficiently. A semiconductor cladding would thus involve longer growth times, an excess applied voltage and planarisation problems during fabrication. Many low-index materials could be deposited under the contact as an alternative, but the use of oxidised AlGaAs is attractive because of the monolithic approach. A 400nm thick oxide layer was found to be sufficient to obtain a low loss coefficient.
Multilayered (Al)GaAs/Al_2O_3 waveguides presenting a large birefringence as a consequence of the high index contrast between GaAs and Al_2O_3 were also demonstrated [38]. This birefringence is sufficient to phase-match nonlinear frequency conversion in the near and mid-infrared.
Vertical oxidation of graded AlGaAs top cladding layers has been reported for the fabrication of 2D photonic lattices [39]. It forms an interesting alternative to deeply-etched 2D air/semiconductor gratings produced by electron-beam writing and dry etching methods, since it results in a planar-surface grating as needed for the deposition of metallisation or other layers on top. An anisotropic oxidation to create vertical oxide pillars was promoted providing a grading in the Al-content of the high-Al layer (decreasing Al-content towards the surface). A similar compositional grading approach has been reported for profile control of vertically oxidised upper confinement layers in oxide-confined stripe diode lasers [40].

References

[1] J. Dallesasse, P. Gravilovic, N. Holonyak, R. Kaliski, D. Nam, E. Vesely, and R. Burnham, "Stability of AlAs in $Al_xGa_{1-x}As$-AlAs-GaAs quantum well heterostructures," *Appl. Phys. Lett.*, vol. 56, no. 24, pp. 2436–2438, 1990.

[2] J. Dallesasse, N. El-Zein, N. Holonyak, K. Hsieh, R. Burnham, and R. Dupuis, "Environmental degradation of $Al_xGa_{1-x}As$-GaAs quantum-well heterostructures," *J. Appl. Phys.*, vol. 68, no. 5, pp. 2235–2238, 1990.

[3] J. Dallesasse, N. Holonyak, A. Sugg, T. Richard, and N. El-Zein, "Hydrolyzation oxidation of $Al_xGa_{1-x}As$-AlAs-GaAs quantum well heterostructures and superlattices," *Appl. Phys. Lett.*, vol. 57, no. 26, pp. 2844–2846, 1990.

[4] C. Ashby, J. Sullivan, K. Choquette, K. Geib, and H. Hou, "Wet oxidation of AlGaAs: the role of hydrogen," *J. Appl. Phys.*, vol. 82, no. 6, pp. 3134–3136, 1997.

[5] R. Twesten, D. Follstaedt, K. Choquette, and R. Schneider, "Microstructure of laterally oxidized $Al_xGa_{1-x}As$ layers in vertical-cavity lasers," *Appl. Phys. Lett.*, vol. 69, no. 1, pp. 19–21, 1996.

[6] A. Sugg, E. Chen, N. Holonyak, J. Hsieh, J. Baker, and N. Finnegan, "Effects of low-temperature annealing on the native oxide of $Al_xGa_{1-x}As$," *J. Appl. Phys.*, vol. 74, no. 6, pp. 3880–3885, 1993.

[7] C. Ashby, M. Bridges, A. Allerman, B. Hammons, and H. Hou, "Origin of the time dependence of wet oxidation of AlGaAs," *Appl. Phys. Lett.*, vol. 75, no. 1, pp. 73–75, 1999.

[8] K. Choquette, K. Geib, H. Chui, B. Hammons, H. Hou, and R. Hull, "Selective oxidation of buried AlGaAs versus AlAs layers," *Appl. Phys. Lett.*, vol. 69, no. 10, pp. 1385–1387, 1996.

[9] T. Takamori, K. Takemasa, and T. Kamijoh, "Interface structure of selectively oxidized AlAs/GaAs," *Appl. Phys. Lett.*, vol. 69, no. 5, pp. 659–661, 1996.

[10] M. MacDoughal, H. Zhao, P. Dapkus, M. Ziara, and W. Steier, "Wide-bandwidth distributed Bragg reflectors using oxide/GaAs multilayers," *Electron. Lett.*, vol. 30, no. 14, pp. 1147–1148, 1994.

[11] P. Heremans, M. Kuijk, R. Windish, J. Vanderhaegen, H. DeNeve, R. Vounckx, and G. Borghs, "Angular spectroscopic analysis: An optical characterization technique for laterally oxidized AlGaAs layers," *J. Appl. Phys.*, vol. 82, no. 10, pp. 5265–5267, 1997.

[12] Y. Luo, D. Hall, O. Blum, and H. Luo, "$Al_xGa_{1-x}As$ heterostructure native oxide planar waveguides," in *Conference on lasers and electro-optics/Quantum electronics and laser science conference*, (Seoul, Korea), August 1999. paper CTuK44.

[13] A. Bek, A. Aydinli, J. Champlain, R. Nane, and N. Dagli, "A study of wet oxidized $Al_xGa_{1-x}As$ for integrated optics," *IEEE Photon. Technol. Lett.*, vol. 11, no. 4, pp. 436–438, 1999.

[14] M. Ochiai, G. Guidice, H. Temkin, J. Scott, and T. Cockerill, "Kinetics of thermal oxidation of AlAs in water vapor," *Appl. Phys. Lett.*, vol. 68, no. 14, pp. 1898–1900, 1996.

[15] R. Naone and L. Coldren, "Surface energy model for the thickness dependence of the lateral oxidation of AlAs," *J. Appl. Phys*, vol. 82, no. 5, pp. 2277–2280, 1997.

[16] T. Langenfelder, S. Schröder, and H. Gröthe, "Lateral oxidation of buried $Al_xGa_{1-x}As$ layers in a wet ambient," *J. Appl. Phys.*, vol. 82, no. 7, pp. 3548–3551, 1997.

[17] H. Nickel, "A detailed experimental study of the wet oxidation kinetics of Al_xGa_{1-x} layers," *J. Appl. Phys.*, vol. 78, no. 8, pp. 5201–5203, 1995.

[18] K. Choquette, R. Schneider, K. Lear, and K. Geib, "Low threshold voltage vertical-cavity lasers fabricated by selective wet oxidation," *Electron. Lett.*, vol. 30, no. 24, pp. 2043–2044, 1994.

[19] J. Kim, D. Lim, K. Kim, G. Yang, K. Lim, and H. Lee, "Lateral wet oxidation of $Al_xGa_{1-x}As$-GaAs depending on its structures," *Appl. Phys. Lett.*, vol. 69, no. 22, pp. 3357–3359, 1996.

[20] T. Yoshikawa, H. Saito, H. Kosata, Y. Sugimoto, and Kenichi, "Self-stopping selective-oxidation process of AlAs," *Appl. Phys. Lett.*, vol. 72, no. 18, pp. 2310–2312, 1998.

[21] F. Kish, S. Maranowski, G. Höfler, N. Holonyak, and S. Caracci, "Dependence on doping type (p/n) of the water vapor oxidation of high-gap $Al_xGa_{1-x}As$," *Appl. Phys. Lett.*, vol. 60, no. 25, pp. 3165–3167, 1992.

[22] J. Dallesasse and N. Holonyak, "Native-oxide stripe-geometry $Al_xGa_{1-x}As$-GaAs quantum well heterostructure lasers," *Appl. Phys. Lett.*, vol. 58, no. 4, pp. 394–396, 1991.

[23] F. Kish, S. Caracci, N. Holonyak, J. Dallesasse, K. Hsieh, S. Smith, and R. Burnham, "Planar native-oxide index-guided $Al_xGa_{1-x}As$-GaAs quantum well heterostructure lasers," *Appl. Phys. Lett.*, vol. 59, no. 14, pp. 1755–1757, 1991.

[24] S. Caracci, F. Kish, N. Holonyak, S. Maranowski, S. Smith, and R. Burnham, "High-performance planar native-oxide buried-mesa index-guided AlGaAs-GaAs quantum well heterostructure lasers," *Appl. Phys. Lett.*, vol. 61, no. 3, pp. 321–323, 1992.

[25] S. Maranowski, A. Sugg, E. Chen, and N. Holonyak, "Native oxide top- and bottom-confined narrow stripe p-n $Al_yGa_{1-y}As$-GaAs-InGaAs quantum well heterostructure laser," *Appl. Phys. Lett.*, vol. 63, no. 12, pp. 1660–1662, 1993.

[26] Y. Cheng, P. Dapkus, M. McDoughal, and G. Yang, "Lasing characteristics of high-performance narrow-stripe InGaAs-GaAs quantum-well lasers confined by AlAs native oxide," *IEEE Phot. Technol. Lett.*, vol. 8, no. 2, pp. 176–178, 1996.

[27] J. Heerlein, S. Gruber, M. Grabherr, R. Jäger, and P. Unger, "Highly-efficient laterally oxidized $\lambda = 950$nm InGaAs-AlGaAs single-mode lasers," *J. Sel. Top. Quant. Electron.*, vol. 5, no. 3, pp. 701–706, 1999.

[28] G. Vawter, O. Blum, A. Allerman, and Y. Gao, "Highly-efficient laser with self-aligned waveguide and current confinement by selective oxidation," in *Proceedings 12th annual meeting* IEEE L*asers and* E*lectro-*O*ptics* S*ociety*, (San Francisco, USA), pp. 531–532, 1999.

[29] D. Huffaker, D. Deppe, K. Kumar, and T. Rogers, "Native-oxide defined ring contact for low threshold vertical-cavity lasers," *Appl. Phys. Lett*, vol. 65, no. 1, pp. 97–99, 1994.

[30] B. Demeulenaere, P. Bienstman, B. Dhoedt, and R. Baets, "Detailed study of AlAs-oxidized apertures in VCSEL cavities for optimized modal performance," *IEEE J. Quant. Electron.*, vol. 35, no. 3, pp. 358–367, 1999.

[31] K. Choquette, K. Lear, R. Schneider, K. Geib, J. Figiel, and R. Hull, "Fabrication and performance of selectively oxidized vertical-cavity lasers," *IEEE Phot. Technol. Lett.*, vol. 7, no. 11, pp. 1237–1239, 1995.

[32] M. Ries, T. Richard, S. Maranowski, N. Holonyak, and E. Chen, "Photopumped room-temperature edge- and vertical-cavity operation of AlGaAs-GaAs-InGaAs quantum-well heterostructure lasers utilizing native oxide mirrors," *Appl. Phys. Lett.*, vol. 65, no. 6, pp. 740–742, 1994.

[33] M. MacDoughal, G. Yang, A. Bond, C. Lin, D. Tishinin, and P. Dapkus, "Electrically-pumped vertical-cavity lasers with AlO_x-GaAs reflectors," *IEEE Phot. Technol. Lett.*, vol. 8, no. 3, pp. 310–312, 1996.

[34] D. Huffaker and D. Deppe, "Intracavity contacts for low-threshold oxide-confined vertical-cavity surface-emitting lasers," *IEEE Phot. Technol. Lett.*, vol. 11, no. 8, pp. 934–936, 1999.

[35] S. Caracci, M. Krames, N. Holonyak, C. Herzinger, A. Crook, T. DeTemple, and P. Besse, "Native-oxide-defined low-loss AlGaAs-GaAs planar waveguide bends," *Appl. Phys. Lett.*, vol. 63, no. 16, pp. 2265–2267, 1993.

[36] D. Taillaert, W. Bogaerts, P. Bienstman, T. Krauss, P. VanDaele, S. Verstuyft, K. DeMesel, and R. Baets, "An out-of-plane grating coupler for efficient butt-coupling between compact planar waveguides and single-mode

fibers." accepted for publication in IEEE J. of Quantum Electron., Vol. 38, No. 7, July 2002.

[37] A. Bek, A. Aydinli, J. Champlain, R. Naone, and N. Dagli, "Use of oxidized $Al_xGa_{1-x}As$ layers for low loss metal clad waveguides and polarization splitters," in *Proceedings European Conference on Integrated Optics '99*, (Torino, Italy), pp. 491–493, 1999.

[38] A. Fiore, V. Berger, E. Rosencher, S. Crouzy, N. Laurent, and J. Nagle, "δn=0.22 birefringence measurement by surface emitting second harmonic generation in selectively oxidized GaAs/AlAs optical waveguides," *Appl. Phys. Lett.*, vol. 71, no. 18, pp. 2587–2589, 1997.

[39] C. Smith, L. Hobbs, T. Krauss, R. DeLaRue, and M. Dawson, "1μm-scale oxide/semiconductor photonic lattice fabrication by wet oxidation methods," in *Proceedings 11th annual meeting* IEEE L*asers and* E*lectro-*O*ptics* S*ociety*, (Orlando, Florida), pp. 32–33, 1998.

[40] P. Evans and N. Holonyak, "Planar anisotropic oxidation of graded AlGaAs for high resolution vertical-wall current and light guiding in laser diodes," *Appl. Phys. Lett.*, vol. 71, no. 2, pp. 261–263, 1997.

Chapter 6

GaAs-based spot-size converters

6.1 Introduction

The field of application of GaAs-based active components is situated in the short wavelength range from 700nm to 1.1μm. Laser sources emitting at 850-870nm, the emission wavelength of GaAs, are typically used in short distance optical communication systems that run over multimode optical fibres. The coupling from a laser to a multimode fibre is not as problematic as is the case for long-wavelength long-distance single-mode systems incorporating InP-based transmitters. Multimode fibre coupling results in a higher coupling efficiency with the incident power being distributed among the many guided fibre modes. The large core diameter (50 or 62.5μm) also implies that a laser-fibre misalignment in the order of a few μm can be tolerated. The use of spot-size converters is, in other words, not as imperative as for InP-based transmitters. This explains why the number of publications on GaAs-based spot-size converters is rather limited.

Another important class of GaAs-based devices are lasers and amplifiers incorporating compressively strained InGaAs active layers and emitting at a wavelength of 980nm. These devices are used as high-power lasers to optically pump Erbium Doped Fibre Amplifiers (EDFAs). It is important for pump lasers to operate in a single-mode regime to avoid spectral hole burning, since this causes kinks in the light-current characteristic. Small single-mode waveguides with a strongly confined optical field are required by consequence. However, the Catastrophic Optical Damage (COD) level (typically 1-10MW/cm^2) limits the output power of the strongly confined mode to rather modest levels. One solution to this problem is to work with a uniform single mode waveguide that supports a large optical mode. This offers a simple and straightforward means of realising a narrow divergence laser beam with a reduced power density at the facets, but it is accomplished at the expense of an increase in threshold current. A better solution is to work with a

conventional waveguide structure with strong optical confinement, monolithically integrated with an adiabatic spot-size converter that gradually expands the mode near the output facet. In this way, a high-performance laser with high output power is obtained.

Fabrication schemes that involve the regrowth of semiconductor materials are not applicable to the GaAs/AlGaAs material system because of the problematic nature of regrowth on AlGaAs [1]. This limits the applicable spot-size converter concepts to ridge-type structures.
Transversely tapered lasers emitting beams with a 15^ox29^o (LxT) divergence have been fabricated using the shadow masked growth technique [2]. Most devices, however, are based on a lateral tapering scheme. Pump lasers at 980nm incorporating a laterally narrowing ridge have been reported with a beam divergence of 10^ox15^o [3]. In order to overcome the lack of regrowth, a polyimide was spun on the sample to surround the narrow ridge with a higher refractive index and allow a better mode expansion. Also laterally broadening ridge lasers have been demonstrated, not only for an improved coupling to fibre [4], but also optimised for free-space communications [5]. The latter end in a very broad output aperture (several hundreds of μm) that emits an extremely narrow beam ($<0.5^o$) carrying powers up to 4W.

An interesting design that was proposed by Vawter et al. comprises a laser structure that is grown on top of a fibre-adapted mesa[1] [6]. A narrowing of the laser ridge forces the mode to expand and shift into the large underlying waveguide. A coupling loss to long-wavelength single mode fibre as low as 0.9dB has been measured at 980nm, with farfield FWHM emission patterns of 5.6^ox7.4^o (LxT). These excellent values are attributed to the large dimensions of the fibre-adapted mesa (10μm wide and 8μm high). The component creates a two-dimensional mode expansion by a pure lateral tapering of the laser ridge, thereby avoiding the need for regrowth. The fabrication scheme is reduced to a single planar growth and two conventional etching steps.

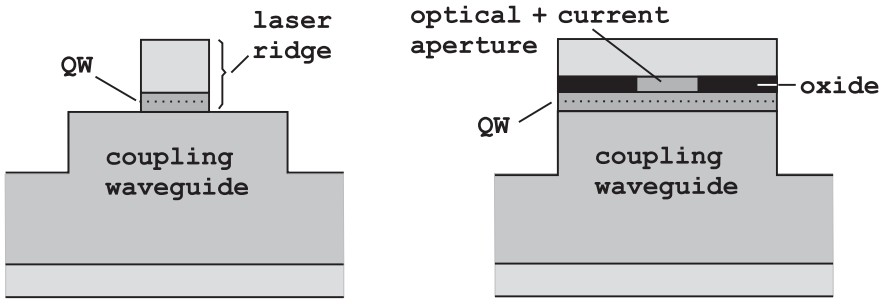
In this chapter we will discuss how we have further simplified the processing scheme and improved the device performance by taking advantage of the opportunities offered by the selective wet oxidation process. It will be shown how the deliberate insertion of a thin oxide layer can create an optical and current confinement and at the same time direct the modal behaviour to obtain an expanded optical field towards the end of the device. In this way, the processing is further reduced to a single planar growth step and one conventional etch, eliminating the need for crucial mask alignments. In addition, the photolithography of the mask patterns is further relaxed since only coarse features (>8μm) are to be defined.

[1] This is the structure that was also used to introduce the principle of ARROW-confinement (see section 4.3).

6.2 Principle

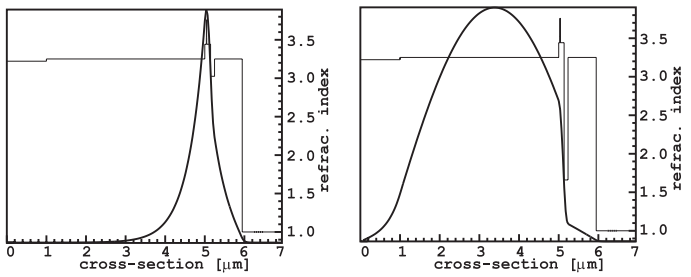
The selective wet oxidation process offers some interesting opportunities to realise an improved device performance. This has already been discussed in chapter 5, but here we exploit the high selectivity of the process and properties of the native oxide a step further to direct the modal evolution in a tapered laser.
Figure 6.1 shows both the 'conventional' design as proposed by Vawter et al. and the structure proposed in this work. In the conventional design both the laser ridge and the broad mesa are defined by etching. This requires the careful alignment of two lithographic masks. The oxide confined device, on the other hand, requires only one photolithography-and-etch step to define a broad mesa. The refractive index contrast that is needed to define the laser waveguide is created by the selective wet oxidation of a 100nm thin AlGaAs-layer that is positioned in close proximity of the active core. In addition, the oxide aperture also creates a self-aligned current aperture and reduces the spreading of the current as it flows towards the active layer. The feasibility of this approach has already been demonstrated for untapered lasers, but here we design the layer structure such that the mode in the central, unoxidised slab is well-confined in the active core, while the oxidised slab carries a mode that is situated in the thick underlying waveguide. Figure 6.2 illustrates the drastic change of the slab modes.

Figure 6.3 gives a top view on the tapered oxide-confined laser and also sketches the evolution of the optical field in the taper. The ridge is broadest in the laser section, where the oxide defines a central window of about 2μm. This results in a laser mode that is well confined in the active region.
The ridge is gradually narrowed in the tapered section, bringing the stretches of oxide closer together and finally closing the oxide window towards the output facet. This forces the optical mode to expand and shift into the underlying fibre-adapted guide. It is important to notice that in this way, a tapering to zero width is created



(a) (b)

Figure 6.1: Cross-section of a) the conventional taper design and b) the oxide confined tapered laser.



(a) (b)

Figure 6.2: Fundamental slab mode of a) the central unoxidised slab and b) the oxidised slab, showing the shift from the active layer to the thick underlying core.

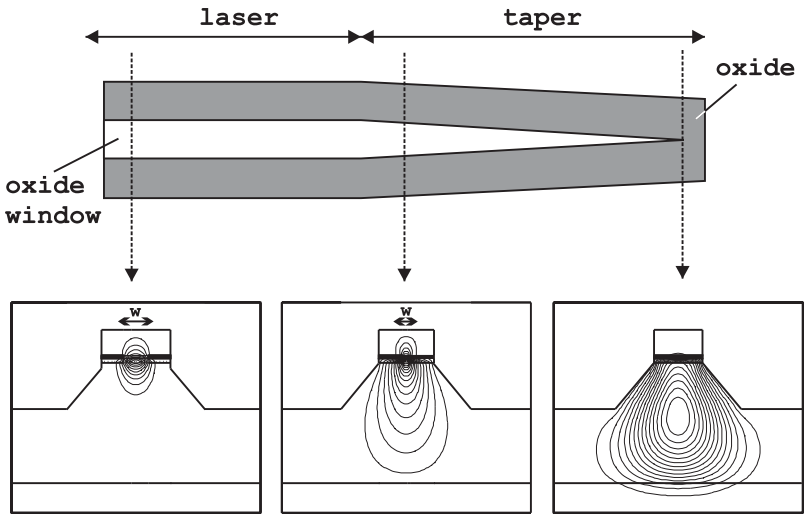


Figure 6.3: Top view on the tapered oxide confined laser and evolution of the optical field.

without involving the definition of extremely narrow features. The mask pattern is narrowed from 10μm in the laser to 8μm at the end of the taper, thereby reducing the width of the oxide window (= the upper waveguide) from 2μm to 0μm.
The slanted edges of the cross-sections sketched in figure 6.3 are typical for the underetch appearing in wet-etched devices like we fabricated (see section 6.4). The slanted edges help to push the expanded mode deeper into the underlying core, but are not essential for a good taper operation. Figure 6.4 illustrates this by showing the evolution of the fundamental local mode to a fully expanded output mode for a typical dry-etched cross-section.

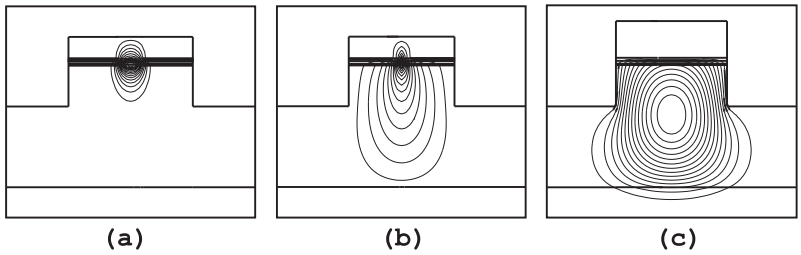


Figure 6.4: Evolution of the fundamental local mode in a dry-etched device: a) in the laser section; b) along the taper and c) at the end of the taper.

In conclusion, the device behaves just like the conventional design, but has the advantage that a more relaxed fabrication scheme is obtained. Of course, an oxidation step is needed to replace the second etch step of the conventional processing, but section 6.3 will discuss how an accurate control on the oxide length is obtained by reducing the Al-content and oxidising at a reduced temperature.

6.3 Design

6.3.1 Design of the cross-section

Before starting the actual discussion on the device design, it should be mentioned that this first implementation aims primarily at a demonstration of the practical feasibility of the above discussed spot-size converter concept. It was not our intention to immediately strive for the shortest component realising the highest achievable coupling efficiency. Instead, a robust design that is tolerant to potential process variations is presented.

The essence of the design of the layer stack is to generate a slab mode that changes its shape as drastically as possible when the high-aluminium layer is selectively converted into native oxide. On the one hand, the mode in the unoxidised slab should be tightly confined in the active laser core, while on the other hand the oxidised slab mode should be situated as deeply as possible in the underlying fibre coupling waveguide, away from the laser core. It can already be understood intuitively that this will imply the use of relatively high aluminium concentrations for all layers in the device, since a change in refractive index from approximately 3 (AlGaAs) to 1.6 (oxide) will have the strongest impact when the refractive indices of the other materials are relatively close to the index of the unoxidised high-aluminium layer. The design is found as a compromise between the opposed requirements of a high confinement in the active laser section and a well expanded mode at the taper output. Factors enhancing the confinement (e.g. thicker laser core, higher index contrast between laser core and underlying coupling waveguide,..) tend to counteract the mode expansion and vice versa.

The design procedure of the tapered laser is an iterative process, with the first step embracing a modal investigation of the structure by means of the 2D mode solver *Fimmwave*. These simulations yield the confinement of the laser mode in the quantum well, the coupling efficiency between the output section and the fibre and the critical section in the taper where the mode changes its shape most rapidly. Once a satisfactory *local mode* behaviour is obtained, the propagation of the field through the taper is investigated with the mode expansion tool *Fimmprop-3D*. The simulated device behaviour may then suggest some further modifications to the layer structure, bringing us back to the start of the design cycle.

We assumed a thickness of 100nm for the high-aluminium layer. Thicker layers were avoided to eliminate potential problems with stress induced by the shrinking oxide. On the other hand, we also decided not to work with thinner layers to eliminate the strong dependence of oxidation rate on the layer thickness for thin layers (see section 5.5.4). The composition of the high-aluminium layer was reduced to 90% Al to lower the oxidation rate and hence obtain a better control on the oxide length (longer oxidation times).
The active layer structure is based on a standard 980nm separate confinement heterostructure (SCH) laser design that has already been used frequently at our department. This standard design incorporates a symmetrically placed 6nm thick $In_{0.2}Ga_{0.8}As$ quantum well in a GaAs core. Because of the strong transverse asymmetry of our oxide-confined active waveguide, the QW position was shifted downward in the active core to enhance the confinement. The GaAs core material has also been changed to $Al_{0.15}Ga_{0.85}As$ to obtain a sufficient change in slab mode profile between the oxidised and unoxidised slab. It was also found that a satisfactory modal behaviour is obtained when the thick coupling waveguide is composed of $Al_{0.5}Ga_{0.5}As$. This is also the composition of the upper laser cladding.
The taper output facet is optimised for coupling to a single-mode 980nm fibre with a mode field diameter of 6.2μm (fibre from *Wave Optics*). For ease of simulation, we chose to investigate a dry etched device as sketched in figure 6.4 instead of the wet etched component. The dimensions of the output section were determined by investigating the evolution of the coupling efficiency as a function of the thickness and width of the broad mesa (figure 6.5). Also the influence of the etch depth was simulated (figure 6.6). From these curves it was decided to implement a 4μm thick and 8μm wide mesa, defined by 3μm etching (this corresponds with etching 1.5μm into the thick coupling waveguide). This results in a coupling efficiency of 75.5%, or a coupling loss of 1.22dB. The coupling efficiency is seen to improve further for thicker core layers, but, as mentioned before, we only wish to demonstrate the feasibility of the oxide confined tapered laser principle. Thicker layers would require special attention to the MOCVD growth process. It is important to note that the coupling efficiency is not much influenced by the exact etch depth, meaning that the single conventional etch that is needed to fabricate the device is reduced to a non-critical processing step. The final layer structure is presented in table 6.1.

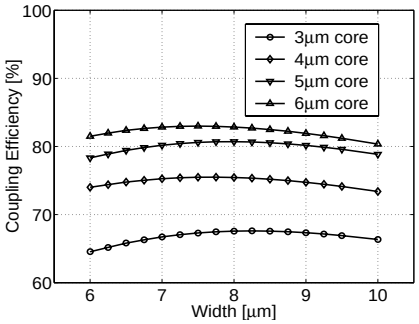


Figure 6.5: Evolution of the taper-fibre coupling efficiency as a function of the mesa width for a core thickness of 3, 4, 5 and 6 μm.

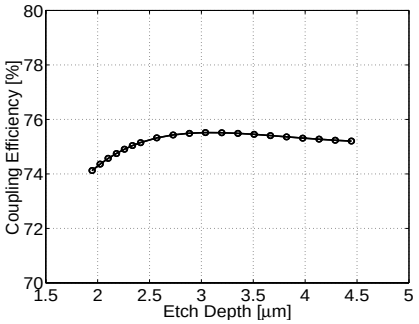
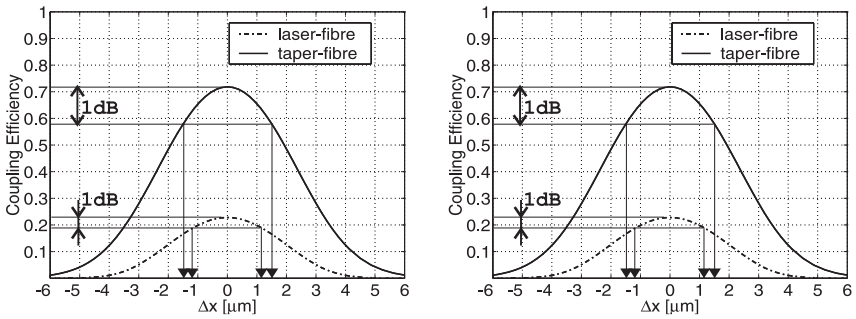


Figure 6.6: Influence of the mesa etch depth on the taper-fibre coupling efficiency for a 4μm thick and 8μm wide taper output waveguide.

Thickness [nm]	**Material**	**Doping [cm^{-3}]**	**Name**
50	GaAs	p=x.e19	contact
50	$Al_{0.125}Ga_{0.875}As$	p=5.e17	grading
50	$Al_{0.25}Ga_{0.75}As$	p=5.e17	grading
50	$Al_{0.375}Ga_{0.625}As$	p=5.e17	grading
1000	$Al_{0.5}Ga_{0.5}As$	p=5.e17	laser cladding
100	$Al_{0.9}Ga_{0.1}As$	p=5.e17	oxide
97	$Al_{0.15}Ga_{0.85}As$	uid	laser core
1.5	GaAs	uid	transition
6	$In_{0.2}Ga_{0.8}As$ (980nm)	uid	QW
1.5	GaAs	uid	transition
40	$Al_{0.15}Ga_{0.85}As$	uid	laser core
4000	$Al_{0.5}Ga_{0.5}As$	n=5.e17	coupling guide
1000	$Al_{0.55}Ga_{0.45}As$	n=5.e17	buffer

Table 6.1: Layer structure of the 980nm oxide-confined tapered laser.



(a) (b)

Figure 6.7: Simulated evolution of the laser-fibre and taper-fibre coupling efficiency for a) lateral misalignment and b) transverse misalignment.

Once the layer structure was fixed, we also verified the coupling efficiency and alignment tolerances from *wet etched* devices to single mode fibre. These results will be useful for comparison with the measured values. In figure 6.7 the simulated coupling behaviour of a reference laser and a tapered laser are compared. The expected maximum coupling efficiencies for laser-fibre and taper-fibre arrangements are 23.7% and 72.5% respectively (or a coupling loss of -6.25dB resp. -1.4dB). The alignment tolerances increase only marginally from 2.6x2.55μm (LxT) for the laser-fibre arrangement to 3.1x2.7μm for the taper-fibre case.

6.3.2 Longitudinal evolution

Three-dimensional mode expansion simulations were performed to calculate the power in the fundamental local mode at the taper output and to visualise the evolution of the optical field. The modal investigation discussed above revealed that the mode is actually being transformed for an oxide window width w_{ox} between 1μm and 0.3μm. A faster tapering is thus allowed for broader ridges. The design satisfying the adiabaticity criterion has a similar shape as calculated in section 3.5 for the conventional device. A quasi-adiabatic behaviour is obtained for a 450μm long device (only 2% mode transformation loss).

The philosophy behind the design of the oxide-confined tapers is, as mentioned before, to provide a robust device design that is tolerant to process variations rather than to immediately aim at an ultimate taper performance. In addition, the mask set for the definition of the devices should also be suited for other layer structures and longer wavelengths (up to 1.3μm, see section 6.7). Knowing that the calculated adiabatic taper shape is only 'ideal' for a specific layer structure and wavelength, we have decided to approximate the ideal shape by a piecewise linear device consisting of two linear sections.

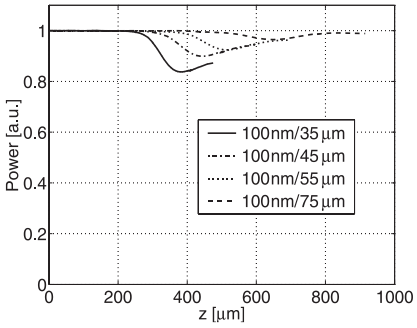
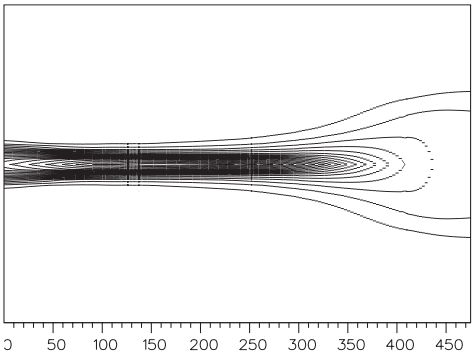


Figure 6.8: Evolution of the fundamental local mode power along the piecewise linear taper for different slopes of the slowly tapered section.

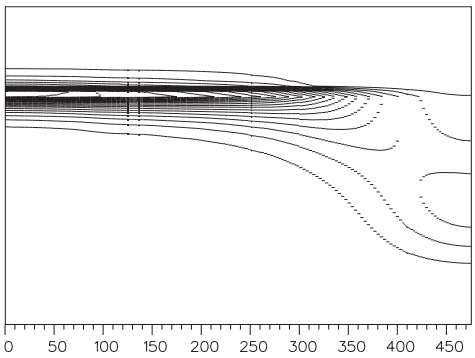
Simulations with *Fimmprop-3D* revealed that the critical mode transformation region is situated between w_{ox}=0.8μm and w_{ox}=0.3μm[2]. A safety margin is taken into account by starting the slowly tapered section in the piecewise linear design already at w_{ox}=1.1μm and extending it until the end of the device (i.e. until w_{ox}=0μm). The initial sharp taper section is narrowed at $\Delta w/\Delta z$=100nm/10μm, introducing less than 0.1% mode transformation loss. Different slopes have been considered for the slowly tapered section, ranging from $\Delta w/\Delta z$=100nm/35μm to 100nm/75μm.
Figure 6.8 shows the evolution of the power in the fundamental local mode for the different slopes implemented on the test mask (see section 6.3.3). The critical section is recognised as the region where the fundamental mode power fluctuates. A quasi-adiabatic behaviour is obtained for $\Delta w/\Delta z$=100nm/55μm (i.e. a 695μm long device), where about 97% of the input power is transferred to the fibre-matched output mode. The design with $\Delta w/\Delta z$=100nm/35μm has an overall length of 475μm, which is comparable with the ideal adiabatic design, but introduces about 15% transformation loss. The final design is sketched in figure 6.10. Figure 6.9 shows a top view and a side view on the evolution of the optical field in the shortest taper.
The sensitivity to variations in layer thickness and material composition has also been investigated to ensure a robust design. The results of this study served to refine the layer structure, finally leading to the design presented in table 6.1.
The thickness of the layers in the active core region were found to be a critical design parameter. However, the control on the layer thickness is known to be very accurate for MOCVD processes and should not be a point of concern here.

It is possible, however, that the composition of the AlGaAs material differs slightly from what is specified. An absolute deviation of maximally 2% on the aluminium content should be taken into account. This corresponds with a re-

[2]This corresponds roughly with the 1μm $\rightarrow$ 0.3μm region that was found from an inspection of the local mode evolution.



(a)



(b)

Figure 6.9: Evolution of the expanding optical field along the taper: a) top view and b) side view.

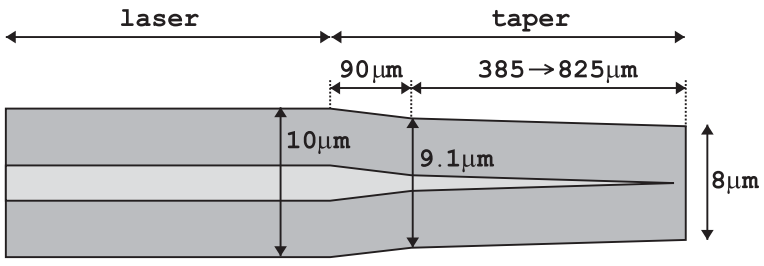


Figure 6.10: Top view on the final taper design.

fractive index variation of approximately 0.01. Fortunately, it appears that the deviations for different layers tend to be in the same direction and that the refractive index contrast between layers is not influenced significantly by consequence. We investigated the influence of a deviation from the expected refractive index *contrast* between the laser core layer and the upper cladding and lower waveguide core. When the contrast is increased by 2%, problems might be expected with the expansion of the mode. However, no significant influence was observed. On the other hand, when the contrast is decreased by 2%, the confinement in the laser quantum well is expected to be affected adversely, but also this effect seems to have a marginal impact. Finally, a shift of 2% on all aluminium contents in the same direction (both increase and decrease were simulated) was also found not to influence the device behaviour. This last result is not surprising, since the index contrast is not changed.

As a final remark we would like to add that devices with an oxide layer above and below the active core have been simulated as well. Such devices can provide a higher QW-confinement in the laser section and at the same time guarantee a taper output mode that is well concentrated in the fibre-coupling waveguide. Simulations showed, however, that the mode transformation is more critical and that it occurs over a smaller width range of the oxide window, requiring a slower tapering. Because of the more critical nature of this doubly oxide-confined device, we have chosen to continue with the single-oxide concept (even though it is not precluded that the double-oxide tapers might operate satisfactory).

6.3.3 Mask design

A mask set has been designed for the photolithographic definition of the components. The mask set comprises three mask levels. The first level defines the ridges, the second level is used to locate the openings in the polyimide isolation layer and the last mask defines the metallisation patterns. Figure 6.11 shows the metallisation mask.
The mask consists of six groups of devices. The two left-most columns (groups I-IV) contain the tapered structures. The difference between the four tapered groups is the slope of the slowly tapered section, as indicated on the mask (Δz= 35μm, 45μm, 55μm and 75μm respectively for Δw = 100nm). The initial section, narrowing from 2 to 1.1μm, is the same for all groups and runs over a length of 90μm. Within each group, two identical tapers are placed in a back-to-back position with a straight laser section in between. They will later be separated by cleaving at the desired laser length. To anticipate a potential uncertainty on the oxide length, it was decided to put not only the designed structures on the mask but also patterns that are 200nm and 400nm narrower and broader. This way at least one subgroup, containing five identical devices, should work. The additional width is also indicated on the mask.
Groups V and VI contain untapered lasers that serve as a reference for comparison with the tapered devices.

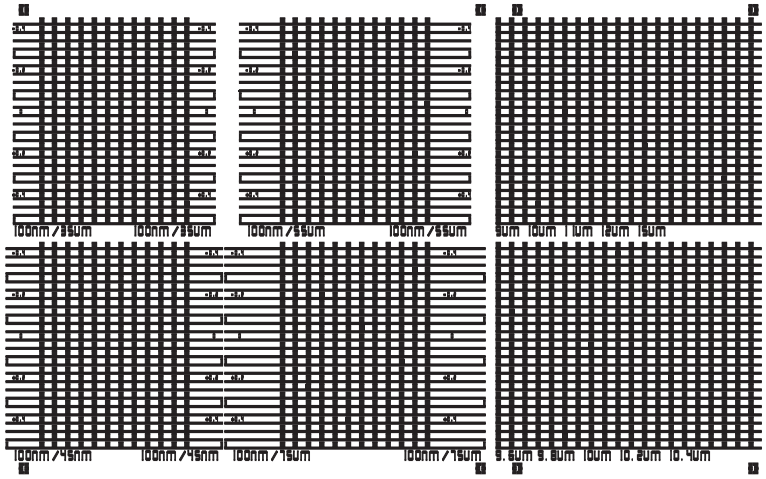


Figure 6.11: Layout of the designed mask set for the fabrication of the tapered lasers (metallisation mask shown).

It is important to mention that the metallisation is chosen to continue over the entire taper length. The choice of a completely active taper has its influence on the threshold current and the efficiency, but on the other hand it avoids the risk of an excessive absorption loss in the passive section of the taper.

6.4 Fabrication

One of the main advantages of the oxide-confined tapered laser is the simple processing scheme that is employed for its fabrication. The layer structure is grown by means of conventional planar MOCVD growth. The mask patterns are transferred by means of photolithographic exposure into a resist layer that is spun on top of the sample. Subsequent standard wet etching ($\approx 3\mu$m deep) in a H_2SO_4:H_2O_2:H_2O (1:1:18) solution creates a ridge as shown in figure 6.12. Because of the underetching under the resist mask, the waveguide top width was reduced from 10μm to $\approx 5\mu$m. Only 1.5μm of oxide is thus needed to create a 2μm wide laser waveguide. A good control on the exact oxide length is obtained by combining a reduced oxidation temperature (410°C) with the lowered aluminium content (90%) in the high-aluminium layer. This results in a practical oxidation time of 45 minutes. The oxide and the central window are clearly visible in figure 6.12. After the oxidation, a polyimide isolation/planarisation layer is spun over the sample, openings are etched above the ridges, a Ti/Au (40nm/150nm) contact is deposited and the contacts are thickened by electrochemical plating. Next the sample is thinned to $\approx 150\mu$m and a AuGe/Ni (150nm/60nm) back contact is applied.

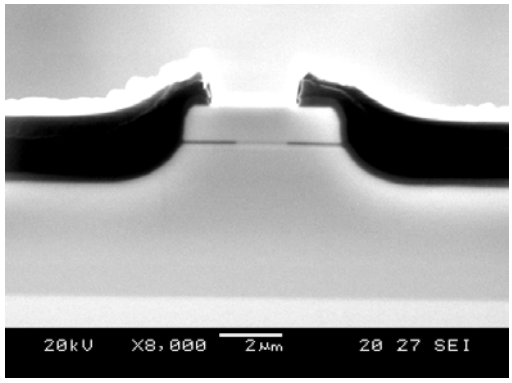


Figure 6.12: SEM photograph of the oxidised tapered laser (cross-section taken from the laser section).

6.5 Characterisation

A variety of measurements has been performed to characterise the various aspects of the tapered lasers. Untapered reference lasers (L=800μm, w=1.8μm) and tapered lasers with different taper slopes, each incorporating a 800μm long laser section (see figure 6.10), were cleaved for characterisation.

Figure 6.13 shows some representative PI-curves. The increasing threshold current for more slowly tapered devices is explained by the fact that the metallisation also covers the spot-size converters, as explained in section 6.3.3. The reference lasers exhibit threshold currents of 15mA and external differential quantum efficiencies η_{ext} of 21%/facet, while the threshold current of the tapered lasers ranges from 25mA (shortest tapers, L_{tap}=475μm) to 34mA (longest tapers, L_{tap}=915μm). The external differential quantum efficiencies amount up to 23% (shortest tapers) and 16% (longest tapers). This surprisingly high value for the shorter devices indicates that the radiation loss from the taper region is low. The higher efficiency can be explained by the lower reflection coefficient of the tapered facet with respect to a normal laser facet[3]. The lower efficiency of the longest tapered lasers is presumably due to the longer section where the optical field has not yet fully shifted to the fibre coupling waveguide and is hence still moderately confined in the quantum well, which is, for the narrow (submicron) oxide windows in that region, not efficiently pumped.

Figure 6.14 shows typical horizontal and transverse farfield emission patterns of a reference laser and a tapered laser. A significant narrowing of the emitted beam

[3]This can be understood qualitatively by recognising that the spatial frequency spectrum of a taper output field is much narrower than for a reference laser facet, as is also seen in the farfield emission patterns. The 'plane waves' that constitute the taper output field make, on average, a smaller angle with the normal on the taper facet than the plane waves of the reference laser field, and experience a lower reflection by consequence.

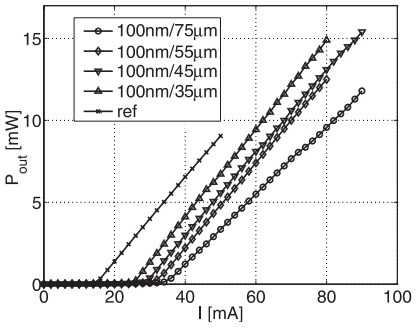
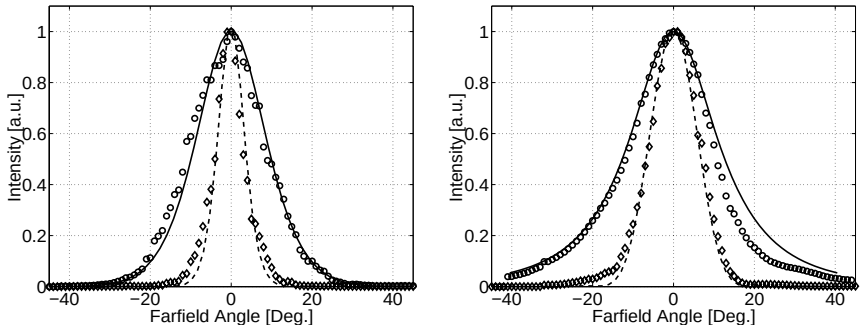


Figure 6.13: Typical PI-curves of a reference laser and different tapered lasers.



(a) (b)

Figure 6.14: Farfield emission patterns of a tapered and a reference laser: a) lateral farfield b) transverse farfield (∘ reference laser measured; ⋄ tapered laser measured; — laser simulated; - - taper simulated) .

is observed, indicating a successful expansion of the laser mode. A good agreement is observed between the measured and the simulated curves. The FWHM-values are strongly reduced from 21^ox23^o (LxT) for the untapered lasers to 7.5^ox13.5^o for the tapered lasers.

Another important issue is the stability of the emitted beam for increasing currents. Figure 6.15 shows the evolution of the measured farfield patterns for currents increasing from below the threshold value to more than three times the threshold ($I_{th} \approx$ 18mA and 34mA for the reference laser and the tapered device respectively). An excellent beam stability is observed for the tapered device. The plots for the reference laser also show the data measured below threshold. The breakthrough of the laser mode is clearly seen from the constriction of the emitted field.

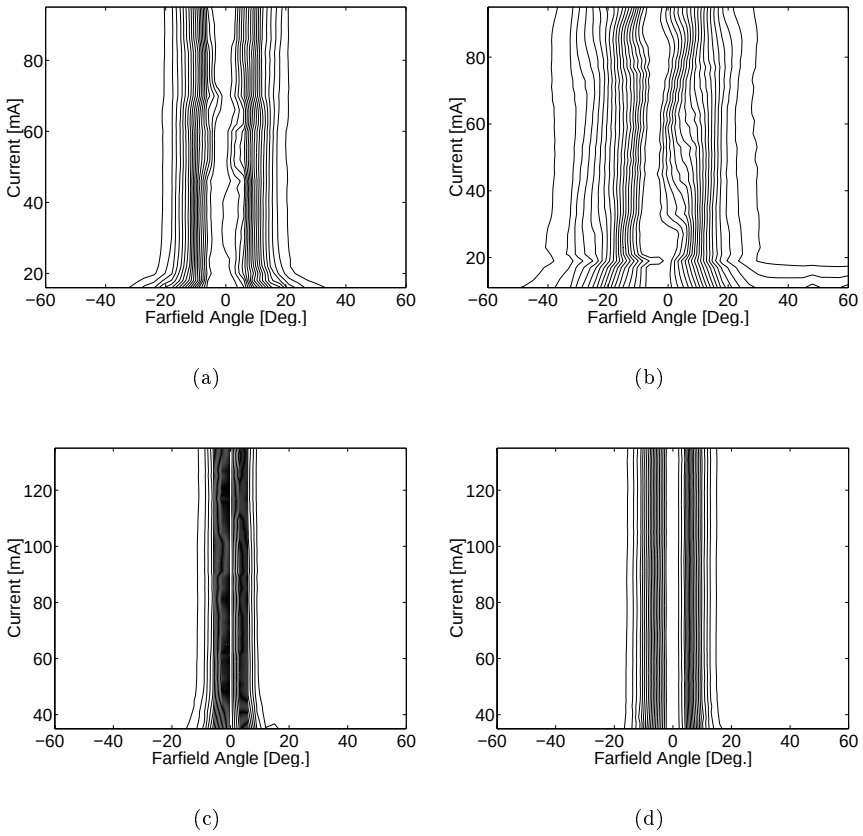
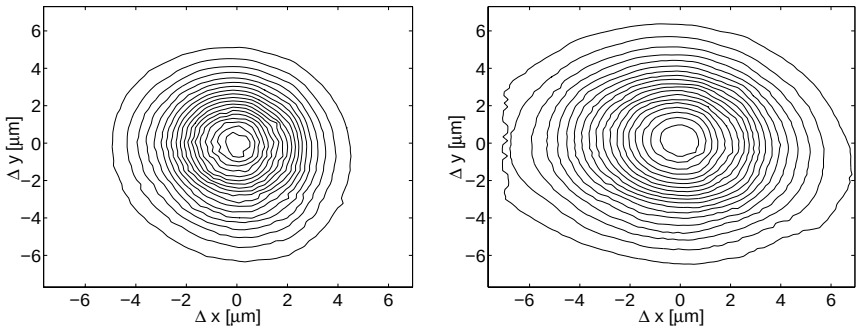


Figure 6.15: Evolution of the farfield emission pattern as a function of current: a) horizontal and b) transverse reference laser; c) horizontal and d) vertical tapered laser.

We also investigated the coupling characteristics from reference and tapered lasers to fibre. For this purpose an automated setup was arranged to scan the position of the fibre in a plane parallel to the chip facet. A *LabView* application, '3DAct', was written for this purpose. It actuates the piezo-electric micro-controllers and reads the power, coupled in the single-mode fibre, from the power meter. Figure 6.16 shows the result of such a measurement for a reference laser and a tapered laser. This distance between two contours corresponds with 5% difference in coupling efficiency.
The maximum coupling efficiency and the horizontal and transverse -1dB alignment tolerances have been extracted from such alignment scans. Table 6.2 summarises the results and compares with the simulated values from section 6.3 and

	Laser-Fibre	
	Simulated	**Measured**
Coupling Eff.	23.7% (-6.25dB)	23.2% (-6.34dB)
-1dB Align. Tol. (LxT)	2.6 x 2.55μm	2.35 x 2.8μm
	Taper-Fibre	
	Simulated	**Measured**
Coupling Eff.	72.5% (-1.4dB)	71% (-1.49dB)
-1dB Align. Tol. (LxT)	3.1 x 2.7μm	3 x 3.2μm

Table 6.2: Comparison between the simulated and measured coupling efficiencies and -1dB alignment tolerances of typical laser-fibre and taper-fibre coupling arrangements. A reflection loss of 4% (0.18dB) has been taken into account.



(a) (b)

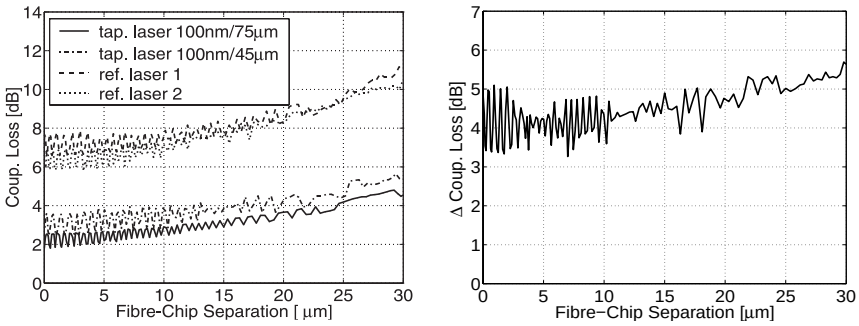
Figure 6.16: Measured fibre coupling efficiency in a plane parallel to the facet: a) a laser-fibre and b) a taper-fibre coupling arrangement.

shows an amazingly good agreement between theory and measurement[4,5] The tolerance of the fibre coupling efficiency to a longitudinal misalignment was investigated by manually increasing the chip-fibre distance and optimising the coupling in the plane parallel to the chip facet. The results of this measurement are shown in figure 6.17(a). The oscillations are attributed to the Fabry-Pérot resonances that occur in the air gap between the chip and the fibre facet[6]. This figure clearly illustrates the initial ($\Delta z \approx 0$) higher coupling efficiency for the tapered laser in comparison with the reference laser and a slower increase of the coupling loss as the fibre-chip distance is increased. This latter effect is even better visible in figure 6.17(b), which shows the difference between the fibre-laser and the fibre-taper coupling loss for an increasing fibre-chip separation.

[4]The values reported in the table have been corrected for the reflection losses occurring at the fibre facet. A reflection loss of 4% or 0.18dB has been subtracted from the measured values.

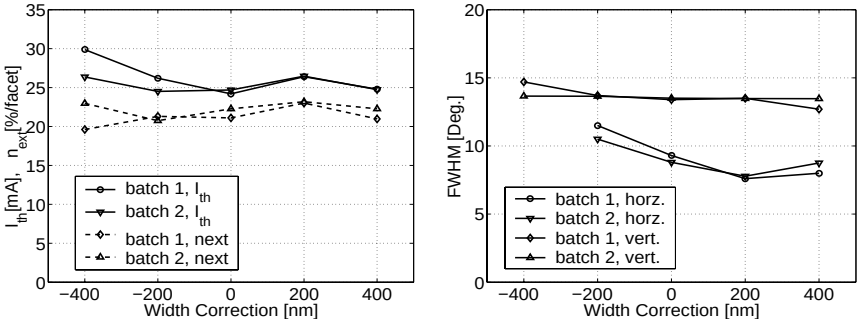
[5]The values listed in the table are for the best performing devices.

[6]This was verified by checking the resonance period.



(a) (b)

Figure 6.17: a) Evolution of the coupling loss as a function of the fibre-chip distance Δz and b) Evolution of the difference between the taper-fibre and reference laser-fibre coupling loss, showing a slower coupling degradation for the taper-fibre configuration.



(a) (b)

Figure 6.18: Comparison of the characteristics of components resulting from different processing runs, illustrating the reproducibility of the device fabrication scheme: a) threshold currents and external efficiencies and b) comparison of the FWHM-values. The 'width correction' is an indication of the device subgroup defined on the mask, and corresponds with the additional device width (see section 6.3.3.)

6.6 Reproducibility of the fabrication

When a new device concept or fabrication scheme is proposed, it is important to verify the robustness of the design and the reproducibility of the fabrication process. For this reason a second set of devices has been processed from the same wafer by following exactly the same procedure (etching, oxidizing and contacting) as used for the first batch of components. A comparison of the characteristics of this new set of devices shows a nearly identical behaviour for both processing runs (figure 6.18).

6.7 Long-wavelength devices on GaAs

6.7.1 Background

Demands on data communication networks are increasing dramatically. The standard wavelength for data communication lasers in optical fibre systems are 850nm and 1.3μm. GaAs-based emitters are being used at 850nm, while the 1.3μm devices are fabricated in the InP semiconductor system. Several technical aspects greatly favour 1.3μm as the operating wavelength of data communication systems. The dispersion minimum of the standard single mode fibre at this wavelength allows higher transmission rates and longer transmission distances due to the much lower absorption losses as compared to 850nm. 1.3μm lasers also offer an improved eye safety, since at this wavelength the human eye tolerates about fifteen times the optical power that would be dangerous at 850nm.

Until now, 1.3μm and 1.55μm emitters were the domain of the InP material system, but recently Kondow et al. proposed to use the new GaInNAs quaternary as an active material to extend the emission wavelength of GaAs-based devices to 1.3μm and beyond [7]. The new compound may be the ideal material to replace InP-based devices and the problems associated with them. This would lead to a significant cost reduction and improvement of the device characteristics:

- InP-based devices are characterised by poor temperature characteristics, meaning that the threshold current and the external efficiency strongly depend on the operating temperature. This is attributed to the shallow potential wells, especially in the conduction band, so that electrons easily disperse to higher quantum levels and overflow to the barrier layers. One of the main advantages of the GaAs-based lasers is a much deeper potential well at the conduction band that prevents this carrier overflow. This superior temperature stability of GaAs-based devices is a very important issue and may greatly affect the module cost by avoiding the need for costly device cooling;

- The low refractive index contrast of materials that are lattice matched to InP leads to difficulties for the production of high-quality DBRs for VCSELs. Either large numbers of DBR layers need to be grown, or novel approaches such as wafer fusion are required. All this adds to cost, complexity and reliability.

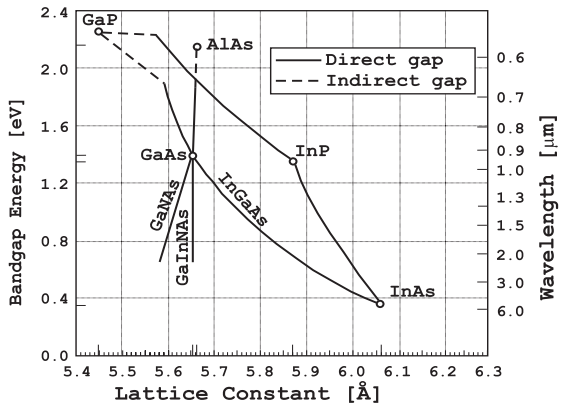


Figure 6.19: Bandgap energies and lattice constants of several semiconductor materials at room temperature.

In addition, these materials exhibit relatively low thermal conductivities, making the heat removal a problematic issue;

- Of course, the availability of the selective wet oxidation process helps to enhance the device characteristics and simplify fabrication schemes, as discussed before;
- GaAs substrates are cheaper and of a higher quality than InP substrates. In addition, larger wafer diameters are available (important for mass production).

Figure 6.19 shows the bandgap energies and the lattice constants of various semiconductor materials at room temperature. When Indium is added to GaAs, the emission wavelength increases, together with the lattice constant. The strain introduced by the increasing lattice constant limits the upper emission wavelength of InGaAs active layers to 1.1μm, or 1.2μm if special strain-balancing structures are introduced [8]. Addition of a small fraction of nitrogen ($\sim$1%) to the InGaAs was found to reduce the lattice constant and further increase the emission wavelength at the same time. The GaInNAs material that is obtained in this way can hence be grown lattice matched to the GaAs substrate and extend the wavelength range. During the last years, the GaInNAs material has attracted the attention of many research groups and is still being studied intensively at the moment. At the beginning of this thesis, the material was not evaluated very much. By participating in two projects, OPTIVAN and GIFT[7], we got access to the material to extend the oxide-confined laser principle to 1.3μm. The other partners involved in the projects were both research institutes and industrial partners, from suppliers of source materials for the growth processes to manufacturers of complete telecommunication systems.

[7]OPTIVAN (**OPT**ical **I**nterconnects using **V**CSELs based on **A**s/**N** compounds), a Brite-Euram project, and GIFT (**G**aAs-based emitters for **F**iber-optical data and **T**elecommunication), an IST-project, were financially supported by the European Union.

The work that we performed in the context of these projects consisted of the design and processing of basic electro-optic devices and also to characterise these components as a feedback for the optimisation of the material growth process. In a later stage we also designed advanced components like oxide-confined tapered lasers operating at or near 1.3μm.

6.7.2 Characterisation of GaInNAs-based components

Material that was supplied by the partners who were concentrating on the GaInNAs growth process was processed into basic broad area and ridge-type laser structures. A variety of measurements was then performed on these components in order to assess basic device and material parameters. The results were then fed back for evaluation and a further optimisation of the growth process. In the following paragraphs we shortly summarise the different types of measurements.

PI/IV measurements
Power-versus-current (PI) and current-versus-voltage (IV) measurements are basic measurements that yield the threshold current I_{th}, the external efficiency η_{ext}, the junction voltage V_j and the series resistance R_s of the device.
The superior temperature behaviour of GaAs-based lasers was mentioned as one of its main advantages. Therefore we have also investigated this aspect. The threshold current of a semiconductor laser increases with temperature by a factor e^{T/T_0}, where T_0 is the characteristic temperature and is used for evaluating the temperature characteristics. The T_0 of conventional InGaAsP/InP lasers is typically between 50K and 70K. To obtain coolerless operation, T_0-values of 150K are required. 980nm InGaAs/GaAs lasers on GaAs substrates have shown a T_0 of over 200K. Figure 6.20 shows the measured evolution of the threshold current for a GaInNAs-laser, operated in pulsed regime to avoid device heating. An exponential fit to these measurements then yields the characteristic temperature. Values between 100K and 130K were typically measured. A further improvement can be expected by using other barrier materials.

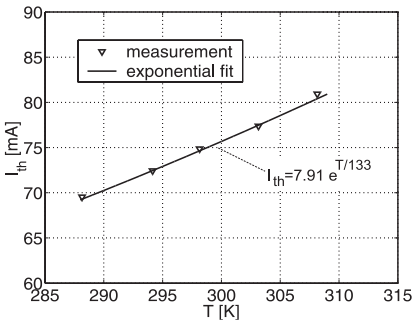


Figure 6.20: Measured evolution of the threshold current versus temperature and exponential fit for the extraction of T_0.

Gain curves
A lot of information can be deduced from the spontaneous emission spectrum of the laser, operated below threshold. As can be seen in figure 6.21, the spontaneous emission spectrum contains Fabry-Pérot interference fringes caused by the interference of amplified spontaneous emission that bounces back and forth in the laser cavity. The depth or 'contrast' of the fringes is determined by the net loss or gain in the cavity at that particular wavelength, as explained in the appendix. This means that the modal gain spectrum[8] g(λ,I) can be calculated from the spontaneous emission spectrum. Some calculated modal gain spectra are shown in figure 6.22 for different currents. The increasing absorption for shorter wavelengths is clearly visible. For long wavelengths the curve is seen to saturate. The saturation value corresponds with the internal loss factor α_i of the cavity, since the gain has become negligible at these wavelengths and the semiconductor is nearly transparent. In between these extremes, the gain reaches its peak value. The evolution of the peak gain, shown in figure 6.23(a), evolves from negative values over the transparency current I_0 (modal gain=0) to the threshold value where the lasing action sets in. The value of the threshold gain corresponds with the mirror losses it has to compensate[9]. The gain remains clamped at this value when the current is further increased, as suggested by the dotted line in the graph. The differential gain dg/dI, which is an important parameter that determines the modulation speed, is then easily calculated from this curve. The values of the differential gain, expressed as dg/dN with N the carrier density, were found to be similar to those for InGaAsP/InP devices, promising a similar performance under high frequency modulation. Finally, the evolution of the gain peak wavelength is shown in figure 6.23(b).

6.7.3 Tapered laser designs

Several oxide-confined tapered lasers incorporating GaInNAs active layers have been designed at wavelengths ranging from 1.2 to 1.3μm. The layer structures are defined such that the existing mask set that was used for the fabrication of the short wavelength devices can be reused. This is accomplished by ensuring that the critical mode transformation region is situated in the slowly tapered section. Also the dedicated tolerant mask design is essential for its reuse at longer wavelengths and/or other layer structures (see section 6.3.3).
A few structures are summarised in tables 6.3-6.5. Each of these structures has been designed to be in compliance with the boundary conditions imposed by the partners. It is obvious that each partner prefers to work with his own specific stan-

[8] The gain that is calculated from this spectrum is $\Gamma\alpha_{mat}$, where Γ is the confinement factor of the laser mode and α_{mat} is the material gain. After determining the confinement factor Γ by means of mode solver simulations, the material gain can be calculated from the modal gain.

[9] A threshold gain of 22 cm^{-1} is measured in this case, in agreement with the 22.8cm^{-1} that is calculated from a mirror reflection coefficient of 32% and a cavity length of 500μm.

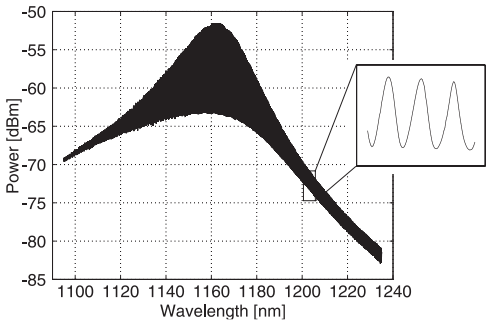


Figure 6.21: Spontaneous emission spectrum of a laser operated below threshold.

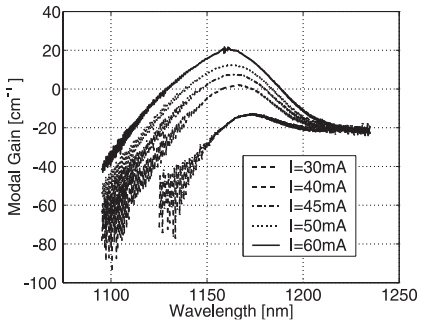
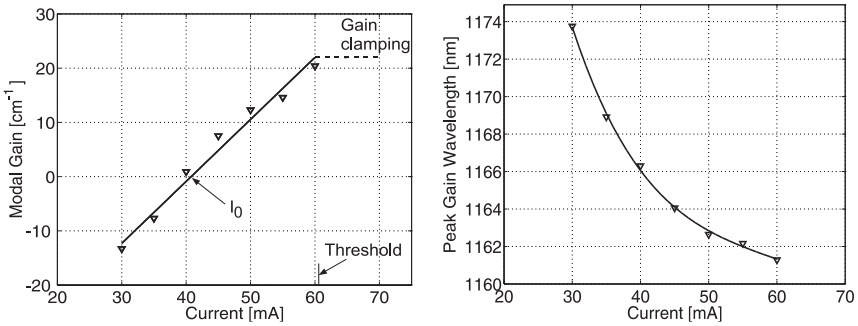


Figure 6.22: Modal gain curves calculated from the spontaneous emission spectrum.



(a) (b)

Figure 6.23: Evolution of a) the peak modal gain and b) the peak gain wavelength versus current.

dard active layer structure and wishes to grow material compositions for which the growth condition are well characterised. The designs and material compositions clearly illustrate that the oxide-confined tapered laser principle is applicable to a wide diversity of layer structures.
Several wafers have been grown by the project partners, but unfortunately, none of them could be processed into working devices. The problems are, however, not related to the design or device concept but are a consequence of unsatisfactory material quality (it should not be forgotten that the GaInNAs growth process is still in a research phase).

Thickness [nm]	**Material**	**Doping [cm^{-3}]**	**Name**
150	GaAs	p=x.e19	contact
1600	$Al_{0.30}Ga_{0.70}As$	p=5.e17	upper cladding
100	$Al_{0.9}Ga_{0.1}As$	p=5.e17	oxide layer
105	GaAs	uid	laser core
6	GaInNAs (1200nm)	uid	QW
20	GaAs	uid	barrier
6	GaInNAs (1200nm)	uid	QW
20	GaAs	uid	barrier
6	GaInNAs (1200nm)	uid	QW
35	GaAs	uid	laser core
4000	$Al_{0.30}Ga_{0.70}As$	n=5.e17	coupling guide
1000	$Al_{0.35}Ga_{0.65}As$	n=5.e17	buffer

Table 6.3: Layer structure of an oxide-confined tapered laser emitting at 1200nm, compatible with the standard *Infineon* laser structure.

Thickness [nm]	**Material**	**Doping [cm^{-3}]**	**Name**
150	GaAs	p=x.e19	contact
1600	$Al_{0.45}Ga_{0.55}As$	p=5.e17	upper cladding
100	$Al_{0.9}Ga_{0.1}As$	p=5.e17	oxide layer
110	GaAs	uid	laser core
6	GaInNAs (1300nm)	uid	QW
60	GaAs	uid	barrier
100	$Al_{0.55}Ga_{0.45}As$	uid	laser core
4000	$Al_{0.45}Ga_{0.55}As$	n=5.e17	coupling guide
1000	$Al_{0.55}Ga_{0.45}As$	n=5.e17	buffer

Table 6.4: Layer structure of an oxide-confined tapered laser emitting at 1300nm, based on the standard laser structure of *KTH*.

Thickness [nm]	**Material**	**Doping [cm^{-3}]**	**Name**
100	GaAs	p=x.e19	contact
1600	$Al_{0.34}Ga_{0.66}As$	p=5.e17	upper cladding
100	$Al_{0.9}Ga_{0.1}As$	p=5.e17	oxide layer
100	GaAs	uid	laser core
7	GaInNAs (1300nm)	uid	QW
16	GaAs	uid	barrier
7	GaInNAs (1300nm)	uid	QW
50	GaAs	uid	laser core
4000-5000	$Al_{0.34}Ga_{0.66}As$	n=5.e17	coupling guide
1000	$Al_{0.45}Ga_{0.55}As$	n=5.e17	buffer

Table 6.5: Layer structure of an oxide-confined tapered laser emitting at 1300nm, based on the standard laser structure of *Nanoplus*.

References

[1] M. Kamp, J. Hofmann, A. Forchel, F. Schäfer, and J. Reithmaier, "Low-threshold high-quantum-eficiency laterally gain-coupled InGaAs/AlGaAs distributed feedback lasers," *Appl. Phys. Lett.*, vol. 74, no. 4, pp. 483–485, 1999.

[2] G. Vermeire, F. Vermaercke, I. Moerman, J. Haes, R. Baets, P. VanDaele, and P. Demeester, "Monolithic integration of a SQW laser diode and a mode-size converter using shadow masked MOVPE growth," *J. Cryst. Growth*, vol. 145, pp. 875–880, 1994.

[3] K. Shinoda, K. Hiramoto, M. Sagawa, T. Toyonaka, and K. Uomi, "Circular beam, high power operation of 0.98μm InGaAs/InGaAsP lasers with a tapered waveguide spot-size expander," *Electron. Lett.*, vol. 32, no. 12, pp. 1101–1102, 1996.

[4] M. Sagawa, K. Hiramoto, T. Toyonaka, T. Kikawa, and K. Uomi, "High-power and highly-reliable operation of 0.98-μm lasers with an exponential-shaped flared stripe," in *Proceedings 9th annual meeting* IEEE L*asers and* E*lectro-*O*ptics* S*ociety*, (Boston, USA), pp. 42–43, 1996.

[5] E. Kintzer, J. Walpole, S. Chinn, C. Wang, and L. Missaggia, "High-power, strained-layer amplifiers and lasers with tapered gain regions," *IEEE Phot. Technol. Lett.*, vol. 5, no. 6, pp. 605–608, 1993.

[6] G. Vawter, R. Smith, H. Hou, and J. Wendt, "Semiconductor laser with tapered-rib adiabatic-following fiber coupler for expanded output-mode diameter," *IEEE Photon. Technol. Lett.*, vol. 9, no. 4, pp. 425–427, 1997.

[7] M. Kondow, K. Uomi, A. Niwa, T. Kitatani, S. Watahiki, and Y. Yazawa, "GaInNAs: A novel material for long-wavelength-range laser diodes with excellent high-temperature performance," *Jpn. J. Appl. Phys.*, vol. 35, pp. 1273–1275, 1996.

[8] M. Hetterich, M. Dawson, A. Egorov, D. Bernklau, H. Riechert, S. Bland, J. Davies, and M. Green, "Comparison of GaInNAs/GaAs and strain-compensated InGaAs/GaAsP quantum wells for 1200-1300nm diode lasers," in *Proceedings 12th annual meeting* IEEE L*asers and* E*lectro-*O*ptics* S*ociety*, (San Francisco, USA), pp. 368–369, 1999.

Chapter 7

Conclusions and perspectives

In this chapter we will briefly recapitulate the main results obtained in this work. The central theme of this thesis is spot-size converters to improve the fibre-chip coupling. We studied different types of devices on both InP and GaAs substrates and successfully demonstrated two new device concepts: the introduction of ARROW-confinement in InP-based components and the use of the selective wet oxidation process for GaAs-based oxide-confined spot-size converters, with the former concept resulting in a patent application. The main goal was to obtain relaxed fabrication schemes, since this is a general point of concern with most designs reported so far.

7.1 Summary and conclusions

Optical telecommunications have known a long history of crucial inventions to bring the technology to what it is today. Chapter 1 sketches how it has evolved from short-distance point-to-point communication channels to complete optical networks carrying huge amounts of traffic over hundreds of kilometers, requiring complex photonic integrated circuits (PICs) to handle the optical signals. A common obstacle is the problematic coupling to single mode optical fibre, which constitutes about 80% of the total cost and hinders a further breakthrough of optical modules. It is explained that integrated spot-size converters are the best solution to relax the coupling problem, but unfortunately, most approaches complicate the device fabrication process. We tried to produce some order out of the large variety of designs by presenting a classification scheme in chapter 2. Several examples were discussed with their reported performance to sketch the state-of-the-art.

An important class of devices are the so-called 'adiabatic' designs, which gradually transform the tightly confined input mode into a more fibre-adapted field profile with a negligible loss of power. A criterion for the design of such devices has been published by Love et al. [1] and has cited many times. In chapter 3, we derived the criterion in a rigorous way by combining bits and pieces of theory on

local modes and mode coupling. We illustrated the criterion with some examples and also made a critical remark on its validity.

The work on InP-based devices has been grouped in chapter 4. After some general remarks on ridge-type and buried technologies, we discussed the design, fabrication and characterisation of a spot-size converter for monolithic integration with a signal-combining PIC that is part of a phased-array antenna system operating at the (for the InP material system) unusual wavelength of 1.06μm. Accurate fibre-coupling measurements, performed at ETH Zürich, revealed a coupling loss of 2.8dB, reducing the fibre-to-fibre loss with 7dB. In our search for taper designs that can be fabricated by using standard growth and processing steps, we proposed the use of ARROW-structures for both the transverse and lateral confinement of the expanded taper mode. The ARROW-concept reduces the epitaxial growth requirements and eventually leads to the most simple taper fabrication scheme ever reported. This simple scheme only involves a single planar growth step and a single conventional etch, or, it only requires the same processing as for a simple straight ridge waveguide. We also designed on-chip spot-size converters that guarantee a mono-modal excitation of broad multimode waveguides for low-loss propagation. The losses of deeply etched waveguides were successfully reduced from 5dB/cm for monomodal structures to 0.15dB/cm for 10μm wide waveguides, with total taper losses of less than 0.05dB.

As an introduction to the GaAs-based spot-size converters, chapter 5 discusses the selective wet oxidation of AlGaAs-layers with a high aluminium composition. The different parameters influencing the oxidation reaction, like oxidation temperature, material composition, layer thickness,... were discussed. We also gave an overview of the wide variety of applications that have already been demonstrated. This oxidation technique forms the basis of the new GaAs-based spot-size converter concept that forms the subject of chapter 6. In that chapter we discussed how the high selectivity of the oxidation process and the properties of the oxide (low refractive index and electrical isolation) have been exploited to direct the modal behaviour in a spot-size converter and to create a buried current aperture that strongly reduces the current spreading towards the QWs. This taper concept once more results in a very simple fabrication scheme that only involves a single planar epitaxial growth and a non-critical conventional etch step. The measured characteristics of the device showed a very close agreement with the simulated values, while the high external efficiency suggests a low mode transformation loss. Coupling efficiencies of 71.2% (1.49dB coupling loss) to a single-mode 980nm fibre were measured, realising an improvement of 4.9dB as compared to a direct laser-fibre coupling scheme. Also the alignment tolerances (3.1x2.7μm (LxT)) show a fairly good agreement with theory. Characterisation of devices from an identical processing of a second batch of devices proved the reproducibility of the device fabrication scheme. Chapter 6 further summarises the work that was done in the context of two European projects, OPTIVAN and GIFT, that concentrated on the development of the novel active material GaInNAs. This new material allows to extend the application range of GaAs-based devices to 1.3μm and beyond. A

variety of oxide-confined spot-size converters incorporating GaInNAs and complying with the specific needs of the different partners involved in the projects, were designed. This illustrates the versatility of the spot-size converter design and its applicability to a wide variety of layer structures and wavelengths.

7.2 Perspectives

The introduction of the ARROW concept in the design of ridge-type spot-size converters has triggered the use of this optical confinement technique in a variety of other applications, such as for buried spot-size converters or for the reduction of the radiation loss in bent waveguides. These aspects are currently being investigated in the Ph.D. work of Marko Galarza.

The oxide-confined tapered lasers have been demonstrated at 980nm and the device designs have been extended to the longer wavelength range for devices incorporating the new GaInNAs active material. Even though different designs have been provided to the different partners involved in the development of the new material, and attempts have been made to grow a wafer, we have not yet realised a device at these wavelengths. The problems are, however not related to the device concept, but rather to the quality of the grown material. New material is being grown at the moment and will be processed and characterised in the near future.

The potential of the selective wet oxidation is still being discovered. Several applications have already been proposed and demonstrated and more is to be expected. An interesting application might be to use the oxidation technique to create low-loss waveguides at the telecommunication wavelengths. It was explained in section 4.4 that scattering loss at the etched waveguide edges is responsible for almost the entire propagation loss in optical waveguides. By defining waveguides by using the low refractive index of the oxide (as was done for the definition of the narrowing laser ridge in the oxide-confined taper) we are convinced that the scattering loss can be largely reduced. The reason for this belief is the fact that the side wall roughness may be expected to be smoothed in the oxidation front. A similar technique has recently been proposed for Si/SiO_2 waveguides with submicron dimensions [2]. In fact, the oxidation technique might open the way to a complete GaAs-based integration technology for low-loss PICs for telecommunication applications, thereby replacing the commonly used InP technology and many of its disadvantages (see section 6.7.1). The oxide-confined taper concept could be introduced as an active-passive integration enabler, similar to a proposal for InP-based circuits presented in [3], where the active layer stack is grown on top of the passive structure and a transition between both is made by a device that is similar to the oxide-confined spot-size converter that was demonstrated in this work.

References

[1] J. Love, W. Henry, W. Stewart, R. Black, S. Lacroix, and F. Gonthier, "Tapered single-mode fibres and devices. part 1: Adiabaticity criteria," *IEE Proc.-J*, vol. 138, no. 5, pp. 343–354, 1991.

[2] K. Lee, D. Lim, L. Kimerling, J. Shin, and F. Cerrina, "Fabrication of ultralow-loss Si/SiO_2 waveguides by roughness reduction," *Opt. Lett.*, vol. 26, no. 23, pp. 1888–1890, 2001.

[3] B. Mersali, A. Ramdane, and A. Carenco, "Optical-mode transformer: A III-V circuit integration enabler," *IEEE J. Sel. Top. Quant. Electron.*, vol. 3, pp. 1321–1331, December 1997.

Appendix:
Hakki-Paoli measurements

The Hakki-Paoli measurement technique is a relative characterisation method to determine the total gain or loss of a waveguide that is terminated by a mirror at both ends [1]. These mirrors are formed by the facets that are obtained by cleaving the waveguides from the processed wafer. The high refractive index contrast between the semiconductor at one side of the facet and the air at the other side results in a power reflection coefficient R of $\approx 32\%$ for a TE-polarised field in a typical InP or GaAs waveguide.
Light propagating from one facet to the other will experience gain in case of a sufficiently pumped active waveguide or loss in case of a passive waveguide or an insufficiently pumped active structure. The light that impinges on the output facet will be partly transmitted and partly reflected back towards the other facet. The total power emerging from the output facet will be the interference of this first contribution and higher order contributions from light that has bounced several times back and forth in the cavity. A constructive or destructive interference is observed at the output, depending on the phase difference between the interfering fields. A schematical representation is given in figure 7.1.

We will now discuss the method for a passive waveguide, where a laser field E_i is incident on the input facet. The total output field E_o, emerging from the output

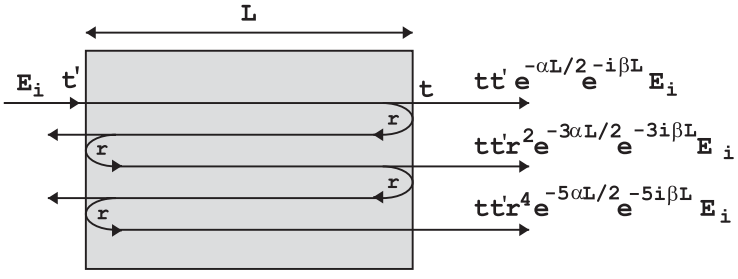


Figure 7.1: Fabry-Pérot interference in a waveguide cavity.

facet, is found by adding all contributions mentioned above:

$$\begin{aligned} E_o &= E_i t' t e^{\alpha/2L} e^{-i\beta L} \sum_{j=0}^{\infty} r^{2j} e^{-j\alpha L} e^{-j(2i\beta L)} \\ &= E_i \frac{(1-r^2)e^{-\alpha L/2} e^{-i\beta L}}{1 - r^2 e^{-\alpha L} e^{-2i\beta L}} \end{aligned} \tag{7.1}$$

with t' and t the field transmission coefficients from air to semiconductor and from semiconductor to air respectively, r=$\sqrt{R}$ the field reflection coefficient from semiconductor to air, β the propagation constant of the guided mode and L the length of the optical cavity.
The power transmission coefficient is then found as:

$$T = \left| \frac{E_o}{E_i} \right|^2 = \frac{(1-R^2)e^{-\alpha L}}{(1 - Re^{-\alpha L})^2 + 4e^{-\alpha L} R \sin^2 \phi} \tag{7.2}$$

By changing the optical pathlength of the cavity (e.g. by varying the device temperature) or changing the wavelength of the incident light (by using a tunable laser), the interference is seen to change from destructive to constructive. Figure 7.2 shows some measured interference fringes obtained by tuning the wavelength. The contrast of the fringes, $K = P_{max}/P_{min}$, depends on the total loss $\alpha = \int \alpha(z) dz$ experienced by the optical mode as it propagates from the one facet to the other:

$$K = \frac{P_{max}}{P_{min}} = \frac{1 + Re^{-\alpha L}}{1 - Re^{-\alpha L}} \tag{7.3}$$

The power loss or gain factor α can now be deduced from (7.3) as:

$$\alpha = \frac{1}{L} \left(\ln R + \ln \left(\frac{\sqrt{K} + 1}{\sqrt{K} - 1} \right) \right) \tag{7.4}$$

The expression for α is no longer dependent on the coupling efficiency at the input since it is obtained from taking the ratio of two transmitted powers. This makes the Hakki-Paoli technique the obvious method to determine waveguide losses. It should be mentioned that the method is limited to *monomodal* waveguides, since each mode in a multimodal structure will propagate with its own propagation constant and produce a set of interference fringes of its own. All these contributions add to each other and hinder an interpretation of the results.

In the case of an active waveguide, no light should be coupled into the cavity. Instead, a current is sent through the device and light is generated through the spontaneous emission process. As the light propagates through the cavity it will be amplified by stimulated emission and also lose a fraction of power because of the waveguide propagation loss. By inspecting the spontaneous emission spectrum, the Fabry-Pérot fringes are again observed and the modal gain $g = \Gamma G - \alpha_i$ around that wavelength can be calculated, leading to the gain curves as discussed in section 6.7.2 (G is the material gain, Γ the confinement factor of the mode in the active material and α_i the internal loss factor).

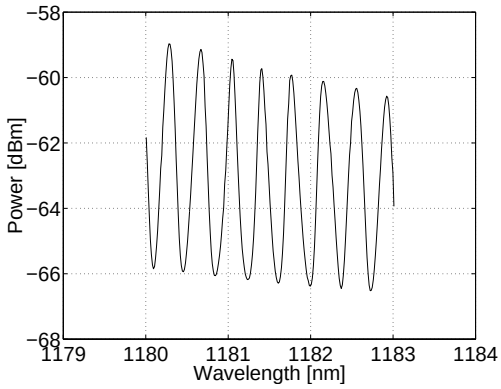


Figure 7.2: Observed Fabry-Pérot interference fringes by changing the wavelength of the incident light.

References

[1] B. Hakki and T. Paoli, "Gain spectra in GaAs double-heterostructure injection lasers," *J. Appl. Phys.*, vol. 46, pp. 1299–1305, 1975.