

Stellingen behorend bij het proefschrift

## Multidisciplinary Design Optimization of a Second-Generation Supersonic Transport Aircraft using a Hybrid Genetic / Gradient-Guided Algorithm

A.H.W. Bos

Daar de rekeninspanning voor een multivariabel/multidisciplinair optimalisatieprobleem buitengewoon groot is, onafhankelijk van de gebruikte optimalisatiemethode, moet de belangrijkste vooruitgang op dit gebied worden geboekt in het reduceren van de benodigde rekentijd.

Omdat de snelheid van nieuwe processoren zich in een hoger tempo ontwikkelt dan de wetenschappelijke kennis, zal het moment waarop de meest geavanceerde analysemethoden binnen acceptabele rekentijd kunnen worden toegepast in combinatie met directe zoekmethoden uiteindelijk aanbreken.

Het commerciële falen van de Concorde kan voor een belangrijk deel worden toegeschreven aan Amerikaans protectionisme. Frans en Brits nationalisme heeft echter in hoge mate bijgedragen aan het technische succes.

Indien het verbod op het boven land produceren van "geluidsknallen" door supersoon kruisende vliegtuigen van kracht blijft, moet terwille van de geloofwaardigheid ook worden opgetreden tegen "geluidsknallen" veroorzaakt door stereo-installaties in auto's.

Pas als de politiek wereldwijd overeenstemming kan bereiken over een stabiel pakket van milieueisen waaraan een supersoon verkeersvliegtuig van de tweede generatie moet voldoen, zal het mogelijk zijn een economische vergelijking te maken tussen dit concept en subsone alternatieven.

Het toepassen van genetische algoritmen en neurale netwerken is een logische stap als men bedenkt dat de antwoorden op veel technische problemen in de natuur kunnen worden gevonden.

De theorieën van Charles Darwin worden wel zeer sterk onderbouwd door het feit dat zelfs het meest eenvoudige model van de genetische mechanismen die men in de natuur aantreft, kan worden toegepast als optimalisatie algoritme.

De discussie rond de vraag of de evolutietheorie van Darwin deel uit moet maken van het Centraal Schriftelijk Eindexamen, toont dat wetenschappelijke theorieën door de kerk vaak worden verward met geloof; de wetenschap past echter haar theorieën aan zodra de feiten daartoe aanleiding geven, de kerk heeft daar eeuwen voor nodig.

Hoewel veel ontwikkelingen die in de luchtvaart op het ogenblik als *high-tech* worden beschouwd al miljoenen jaren in de natuur worden toegepast, heeft het kijken naar de natuur de luchtvaart in de tijd van de pioniers aanvankelijk op een verkeerd been gezet.

Het combineren van de disciplines Landbouw, Visserij en Natuurbeheer in één Ministerie is een contradictio in terminis.

Er vanuit gaande dat extreem rechtse politieke partijen hun groei voor een groot deel aan proteststemmen danken, zullen verontwaardigde reacties en overmatige aandacht vanuit de media en meer gematigde partijen zulk een groei slechts in de hand werken.

Het doel van nieuws- en actualiteitenprogramma's is niet het voorlichten van de bevolking, maar het realiseren van hoge kijkcijfers.

De invloed die de Westerse maatschappij (zelfscheppend, expanderend) door de eeuwen heen op meer primitieve culturen (jagend, verzameland) heeft uitgeoefend is niet per definitie ontstabiliserend [Farb], maar meer een verstoring van een overigens inherent stabiel systeem. Het is de Westerse maatschappij zelf die instabiel is.

Farb, P. *Man's Rise to Civilization-The Cultural Ascent of the Indians of North America*

Het jury-rechtssysteem doet geen recht aan het beginsel dat een verdachte onschuldig is zolang zijn schuld niet is bewezen.

Het door de christelijke organisaties zo gepropageerde respecteren van de zondagsrust dient in de eerste plaats van henzelf uit te gaan door het luiden van kerkklokken te beperken.

De intolerante, agressieve en ongenuanceerde houding die men bij veel "vredesactivisten" aantreft, toont duidelijk dat conflicten niet door wapens maar door mensen worden veroorzaakt.

De gedachte dat het nooit te laat is om van richting te veranderen kan iemand doen volharden in zijn oude koers.

God schiep niet de mens maar de mens schiep god naar zijn evenbeeld.

Het afbeelden van engelen met vleugels suggereert dat er een atmosfeer is in de hemel.

De Wet van Murphy dient een waarschuwing en geen excuus te zijn.

Het verlenen van ere-doctoraten aan beroemdheden moet eerder worden gezien als reclame voor de betreffende universiteit dan als een eerbetoon aan de begunstigde.

Los van beschouwingen over zijn karakter en politieke ambities kan alleen al uit de bewegingen van de 7<sup>e</sup> Cavalerie voorafgaand aan de Slag bij de Little Big Horn (25 juni 1876) worden afgeleid dat generaal Custer nooit van plan is geweest de colonne van generaal Gibbon af te wachten.

Graham, W.A. *The Story of the Little Big Horn*

Miller, D.H. *Custer's Fall-The Native American Side of the Story*

De historische woorden van veel indiaanse opperhoofden, met name de beroemd geworden toespraken van Seattle van de Duwamish en Young Joseph van de Nez Percé (Choppunish), kunnen nog steeds als een waarschuwing aan de mensheid worden beschouwd.

---

# Multidisciplinary Design Optimization of a Second-Generation Supersonic Transport Aircraft using a Hybrid Genetic / Gradient-Guided Algorithm

666569

3190089

TR diss 2781

---

---

# Multidisciplinary Design Optimization of a Second-Generation Supersonic Transport Aircraft using a Hybrid Genetic / Gradient-Guided Algorithm

## Proefschrift

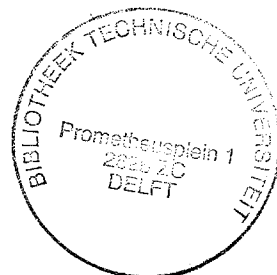
ter verkrijging van de graad van doctor  
aan de Technische Universiteit Delft,  
op gezag van de Rector Magnificus, Prof. ir. K.F. Wakker,  
in het openbaar te verdedigen ten overstaan van een commissie  
aangewezen door het College van Dekanen,  
op donderdag 27 Juni 1996 te 13:30 uur

door

Adolf Hendrik Willem BOS

vliegtuigbouwkundig ingenieur

geboren te Winschoten





Dit proefschrift is goedgekeurd door de promotor:

Prof. ir. E. Torenbeek

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter

Prof. ir. E. Torenbeek, promotor,

Prof. C. Fleury,

Prof. dr. ir. H.W.M. Hoeijmakers,

Prof. ir. E. Obert,

Prof. ir. G.J.J. Ruijgrok,

Prof. ir. W.B. de Wolf,

Dr. ir. R.A. Vingerhoeds,

Technische Universiteit Delft

Université de Liège

Technische Universiteit Delft

Technische Universiteit Delft

Vrije Universiteit Brussel

Technische Universiteit Delft

Technische Universiteit Delft

Published and distributed by:

Delft University of Technology

Faculty of Aerospace Engineering

P.O. Box 5058

2600 GB Delft

The Netherlands.

tel: (015)2782058

fax: (015)2781822

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Bos, Adolf Hendrik Willem

Multidisciplinary Design Optimization of a Second-Generation Supersonic Transport Aircraft using a Hybrid Genetic/Gradient-Guided Algorithm / A.H.W. Bos

Delft: Delft University of Technology. -Ill.

Delft: Thesis Delft University of Technology. -With ref. -With summary in Dutch.

ISBN 90-5623-041-7

NUGI 841

Keywords: Multidisciplinary Design Optimization / Supersonic Transport Aircraft / Genetic Algorithms

Copyright © 1996, by A.H.W. Bos.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the author.

*Front cover:*

*The "Man-in-the-Maze" symbol is frequently encountered throughout the deserts of the Southwestern United States. Referring to a legend of the Tohono O'odham -or Papago- people about the origin of their nation, it depicts the path taken by Elder Brother climbing the top of a sacred mountain in Arizona.*

*Where today are the Pequot ? Where are the Narragansett, the Mohican, the Pokanoket and many other once powerful tribes of our people ? They have vanished before the avarice and the oppression of the white man as snow before a summer sun.*

Chief Tecumseh of the Shawnee in a speech to the Choctaw of present-day Mississippi and Alabama prior to the War of 1812

*I had never seen a wasichu<sup>1</sup> (..) and did not know what one looked like, but every one was saying that the wasichus were coming and that they were going to take our country and rub us all out and that we should all have to die fighting (..). I remember once that I asked my grandfather about this. I said: "When the scouts come back from seeing the prairie full of bison somewhere, the people say the wasichus are coming; and when strange men are coming to kill us all, they say the wasichus are coming. What does it mean ?" And he said: "That they are many."*

Black Elk  
Oglala Sioux Medicine Man  
Neihardt, J.G. *Black Elk Speaks*

---

<sup>1</sup> word to designate a white man, however bearing no reference to skin color

# Summary

The purpose of this dissertation is to investigate methods for multivariate and multidisciplinary **optimization**<sup>1</sup> of conceptual aircraft designs, using a second-generation supersonic transport plane (Supersonic Transport, SST, also known as High Speed Civil Transport, HSCT) as a test case. In this sense, the work described in this dissertation is twofold: available methods suitable for the optimization of large, complex systems are presented and evaluated and an **analysis** procedure for supersonic transport planes is developed.

The major problem concerned with multivariate optimization is how to obtain the necessary information to guide the optimization procedure toward the optimum in a reasonable time, given the fact that a multivariate **design evaluation** with sufficient accuracy and predictive fidelity is extremely time-consuming. The presently known optimization methods get this information either by analyzing the gradients of the objective function and constraints with respect to the independent **design variables** or by performing a very large number of design evaluations.

Because of the large number of design variables and constraints, the number of required gradients can become quite large whereas these gradients must be recalculated for each improved design until the optimum is attained.

Since the calculations for an aircraft design analysis are rather complex and often include iterative solution procedures, no closed-form analytic relations are available for the gradients needed. Instead, they must be obtained by finite-differencing (perturbing one variable at a time and recalculating the complete design) which requires many design evaluations (number of variables + 1 for every optimization step in case of simple forward differencing). It was established that the accuracy of gradients thus obtained is not very high.

Some methods to decompose either the process of calculating the gradients or the optimization problem itself were presented in the 1980's and tested in the present work. Instead of calculating global derivatives by perturbing and recalculating the complete design, partial derivatives are generated by finite-differencing of subsystems. These methods of **decomposition** are especially well suited for parallel computation and it is here that their main advantage lies. The use of methods of decomposition in combination with second-order optimization algorithms was not completely successful, however. Also, the problem of inaccurate gradients remains.

The use of non-gradient-guided algorithms, notably the genetic algorithms that enjoy a fast rise in popularity today, was investigated in the present work as well. The large number of design evaluations required and the problem of premature convergence are the main disadvantages identified for this type of optimization. Furthermore, as opposed to gradient methods there is no guarantee that the optimum is attained or even will be found at all. On the other hand, the method is likely to locate the global

---

<sup>1</sup> boldly printed terms are explained in the Glossary at page 349

---

optimum instead of converging on a local one. Therefore, the robustness of the method, unless premature convergence occurs, is higher than that for gradient-guided methods.

Since the calculation effort required for a multivariate/multidisciplinary optimization problem is excessive regardless of the method of optimization employed, the main improvement is to be obtained by a reduction of the computational time, either by using high-speed computers (expensive) or by reducing the time needed for a single design evaluation. The latter can be achieved either by arranging the analysis in a smart way or by using a fast and simple **approximate analysis** instead of the slow detailed one. The first option of arranging the analysis in a smart way includes minimizing feed-backs that require iterative solution, using smart estimates for values needed at the beginning of the analysis that are only known exactly at the end, without re-iteration, in short, trading accuracy for speed, within reasonable limits of course.

The second option of using a simple approximating analysis instead of the elaborate analysis is extensively used in the present work. Two kinds of approximate analysis exist. The first kind is a local approximation, at the current design point, which has to be updated by the exact analysis after each optimization step. The most simple example of this kind of approximate analysis are the first-order derivatives of the current design with respect to its **independent variables**, as used by every gradient-guided optimization algorithm.

The second kind is global approximation by **regression surfaces**, in the present work represented by second-order polynomials. This concept has many advantages. Not only does it shorten the required computational time -which as explained before is advantageous to any method of optimization- also the approximation can be generated in advance and does not need to be updated after every optimization step. The approximate analysis may be used to include results from disciplines that are analyzed by separate programs that cannot be implemented within an optimization algorithm or that simply are too time-consuming to consider using them in an optimization cycle. Global approximation is also a tool to include analyses that are performed at a geographically remote location (a different department or company). Last but not least, the use of second-order polynomials yields the relation between the objective function, the constraints and the independent design variables as simple, closed-form functions. Therefore, gradient information is calculated fast and easy.

In the present work a hybrid optimization method is employed, consisting of a global search using a genetic algorithm to decrease the design space, followed by a gradient-guided search to locate the exact optimum. The global search uses the exact analysis in which all redundant iterative loops are omitted, using approximate analysis to include some disciplines. The gradient-guided search uses only approximate analyses generated for the reduced design space containing the optimum as obtained from the global search. In this way the good characteristics of gradient methods (efficient, exact) and global search methods (robust) are combined.

---

The optimization procedure is applied to the design of a second-generation supersonic transport plane (SST). The main features of the analysis method are:

- a design sensitive wing weight prediction method
  - the use of a panel method to obtain subsonic and supersonic aerodynamic characteristics of high-speed configurations
  - application of the Polhamus leading-edge suction analogy and Carlson attainable thrust concept to account for partially separated leading-edge flow and the resulting leading-edge vortices, important for highly swept wings at angle of attack
  - a procedure to predict the off-design performance of bypass engines designed for supersonic flight, including a detailed intake routine to provide optimal intake efficiency at the design point and to include spillage air and powerplant drag in the analysis
  - an estimation of the environmental impact of a second-generation SST
-



# Table of Contents

Summary .....	i
---------------	---

Table of Contents .....	v
-------------------------	---

## Part I

### **Presentation and Evaluation of Methods Suitable for Multidisciplinary Design Optimization**

1. Introduction .....	1
1.1 References .....	3
2. Introduction to Optimization Theory .....	5
2.1 Summary .....	5
2.2 Notation .....	5
2.3 A Survey of Methods for Constrained Optimization .....	6
2.4 Treating Nonlinear Constrained Problems .....	15
2.5 Direct Methods .....	19
2.6 References .....	22
3. Methods to Decompose Large Optimization Problems .....	23
3.1 Summary .....	23
3.2 Notation .....	24
3.3 Decomposition of Hierarchic Systems .....	26
3.4 Decomposition of Non-hierarchic Systems (Networks) .....	28
3.5 Application to Multidisciplinary Optimization of a Conceptual Aircraft Design .....	33
3.6 Results .....	39
3.7 Conclusions .....	43
3.8 References .....	44
Appendix A Overview of the Variables and Parameters used in the Decomposition Scheme .....	46
Appendix B Accuracy of Sensitivity Derivatives .....	49
Appendix C Initial and Final Design .....	55
4. Optimization of a Conceptual Aircraft Design using Approximate Analysis .....	57
4.1 Summary .....	57
4.2 Notation .....	57
4.3 Introduction .....	58
4.4 Comparing the Results of Approximate and Exact Analysis .....	63
4.5 Conclusions .....	71
4.6 References .....	75

---



---

5.	The Applicability of Regression Surfaces for Approximating Purposes . . . . .	77
5.1	Summary	77
5.2	Notation	77
5.3	Response Surfaces	78
5.4	References	90
6.	Direct Search Methods with Special Reference to Genetic Algorithms . . . . .	91
6.1	Summary	91
6.2	Notation	91
6.3	A Summary of Direct Search Optimization Methods	92
6.4	Results	107
6.5	Conclusions	111
6.6	References	113

## Part II

### The Design and Analysis of a Second-Generation Supersonic Transport Aircraft

7.	Introduction . . . . .	117
7.1	Summary	117
7.2	The Development of a Second-Generation Supersonic Transport	117
7.3	References	125
	Appendix      A Summary Of Concorde Accidents	126
8.	The Mass Prediction Model . . . . .	129
8.1	Summary	129
8.2	Notation	130
8.3	Wing Weight Estimation	131
8.4	Maximum Take-off Weight	132
8.5	References	136
9.	The Aerodynamics Model . . . . .	137
9.1	Summary	137
9.2	Notation	137
9.3	Introduction	139
9.4	Leading-edge Suction Analogy	140
9.5	Attainable Thrust Concept	147
9.6	Application to Double-delta Wings	157
9.7	Results	159
9.8	Implementation of the Method	167
9.9	Conclusions	172
9.10	References	173
	Appendix      Calculation of Spanwise Coordinates for Mean Attainable Thrust Coefficient	174

---

---

10.	Using the NLRAERO Panel Method in Combination with the Attainable Thrust Concept in Multivariate Optimization .....	177
10.1	Summary	177
10.2	Notation	177
10.3	The NLRAERO Program	179
10.4	Accuracy of Aerodynamic Approximations	183
10.5	The Influence of Optimal Camber on the Cruise Glide Ratio	187
10.6	Conclusions	202
10.7	References	203
	Appendix      Compressibility Correction for Friction Drag Coefficient	203
11.	The Propulsion Model .....	207
11.1	Summary	207
11.2	Notation	207
11.3	Introduction	209
11.4	Some Typical Off-design Problems of Supersonic Propulsion	211
11.5	Intake Model	220
11.6	Design Point Procedure	227
11.7	Off-design Procedure	229
11.8	References	256
12.	The Environmental Impact of a Supersonic Transport Aircraft .....	259
12.1	Summary	259
12.2	Notation	259
12.3	Pollutant Emissions	260
12.4	Take-off Noise	268
12.5	Sonic boom	277
12.6	Conclusions	280
12.7	References	281
13.	The Mission Analysis .....	283
13.1	Summary	283
13.2	Notation	283
13.3	Conditions for Climb with Minimal Fuel Consumption	284
13.4	Mission Analysis Procedure	286
13.5	References	294
14.	The Analysis Program .....	295
14.1	Summary	295
14.2	Notation	295
14.3	Design Variables and Constraints	297
14.4	Geometry	307
14.5	Determination of Take-off Field Length	308
14.6	Cruise Analysis	308
14.7	Climb Performance Analysis	309
14.8	Stability Analysis	309
14.9	References	313

---

15. The Optimization Procedure .....	315
15.1 Summary .....	315
15.2 Notation .....	315
15.3 Description of Optimization Procedure .....	316
15.4 Approximate Analysis .....	317
15.5 Optimization Results .....	319
15.6 References .....	334
Appendix A Final Design Characteristics .....	334
Appendix B Evaluation of the Mixing Process .....	338
Appendix C Evaluation of the Fan Design .....	341
16. Conclusions and Recommendations .....	343
16.1 Summary .....	343
16.2 Multidisciplinary Design Optimization .....	343
16.3 The Design of a Second-Generation Supersonic Transport Aircraft .....	345
16.4 References .....	348
Glossary .....	349
Samenvatting .....	351
Curriculum Vitae .....	355

---

---

## **Part I**

# **Presentation and Evaluation of Methods Suitable for Multidisciplinary Design Optimization**

---

# 1. Introduction

Modern aircraft design is a multidisciplinary process in which a large number of design variables is involved. Due to the ever increasing complexity of this process, which causes the number of couplings between the different disciplines to increase as well, it is often hard to predict the impact of changes to a certain subsystem on the system as a whole. To cope with the great amount of variables, in engineering practice the design process develops sequentially and iteratively. During this process many variables are chosen according to engineering judgement or experience and kept constant in order to reduce the dimensionality of the problem. At the end of each iteration cycle, the design is evaluated, to see whether it performs according to the specifications. If not, changes are made to the design, although certain subsystems often will be kept unchanged. When, at a later stage in the design process, improved values are obtained for assumptions made in the initial stages, this new information sometimes cannot be used because certain subsystems are already frozen and cannot be changed any more because of time or budget limits. Sometimes the effect that changing them could have on other subdesigns or performance **parameters** cannot be foreseen. This causes in effect a reduction of the degrees of freedom of the design problem and therefore leads to a suboptimal design. The apparent paradox between the increasing number of participating disciplines and the decreasing number of free variables that can be manipulated by the designer, was first addressed in [ref.1] by Sobieszczanski-Sobieski e.a.

Today, there is a tendency towards increasing the dimensionality of the design problem, in order to generate better designs. This calls for the development of a Multidisciplinary Design Optimization (MDO) procedure that enables different subdesigns to be modified without this causing other subdesigns to violate their constraints. Besides the obvious benefit of producing the best possible design from the point of view of industrial competition, other reasons to employ MDO techniques can be forwarded.

Especially in civil aviation, possibly with the exception of the development of new materials, large "technology jumps" today can no longer be relied upon to facilitate major product improvements (as was the case with the introduction of the jet engine, the bypass engine, the swept wing, the supercritical wing etc.). Aircraft in a specific category seem to converge to more or less equal solutions, and optimization might be the only means of improvement.

Since aircraft design is by definition a compromise between conflicting requirements, Multidisciplinary Design Optimization is a valuable technique to find the best possible compromise, thus minimizing the company's development risk. Especially in case of projects where the requirements imposed on the design are particularly conflicting, like for instance in the design of stealth fighters (aerodynamic and propulsion efficiency versus radar cross section and infrared signature) or supersonic cruising aircraft (subsonic versus supersonic performance), MDO techniques are a necessity to produce a feasible design.

In practical engineering, many expenses are caused by feed-backs in the design process after evaluation of the constraints. This might be a result of either inaccuracies

---

in the prediction methods, or of the fact that certain disciplines only enter the design process during the final phase. By taking into account as many design aspects as practicable at an early and relatively inexpensive stage of the project, most feed-backs will take place here and not at a later moment when production has started already. The latter would not only cause great expenses, it would most certainly lead to a suboptimal design. This is the essence of a modern trend in industry, known as *Concurrent Engineering* (CE). Concurrent Engineering implies incorporating as many relevant disciplines as practicable at an early stage of the design process (the author is aware of the fact that different definitions for the term concurrent engineering are in use). This implies, for instance, that not only disciplines like aerodynamics or performance, but also manufacturing and economics, are taken into account at an early stage of the design. To be able to reduce the number, the length and the seriousness of the feed-backs, it is important to have the possibility to implement more advanced methods of analysis while still taking into account, at this higher design level, the couplings between disciplines. In current engineering practice, these couplings are generally preserved as long as zero-order methods are being used, but later on, as the design process turns toward more sophisticated methods, these interactions between subsystems are usually lost. This, again, could lead to a suboptimal design.

The main reasons to employ Multidisciplinary Design Optimization techniques are therefore to increase the dimensionality in order to obtain better designs and to increase the number of disciplines involved to reduce expensive feed-backs in the design process. Of course both actions are coupled: to increase the dimensionality generally more disciplines will have to be taken into account, whereas taking more disciplines into account will increase dimensionality. As indicated before, it is very important that the level of sophistication of the analysis methods employed is as high as possible without the resulting increase in design variables becoming excessive. The reason is that an optimum often lies at the vertex of several constraints, which implies that an inaccurately predicted constraint value could render the final design useless.

The complexity of the analysis (and possibly the fact that it is carried out at different, sometimes even geographically separated locations) will make it very difficult to carry out the analysis in a single program within the optimization routine. This calls for an evaluation of the communication between the optimization algorithm and the analysis program. Several solutions like decomposition of the optimization problem or the use of approximate analysis are discussed in this work.

Increasing the number of independent design variables causes a number of problems, the solution of which is the main goal of Multidisciplinary Design Optimization research. First of all, the ability of the human designer to grasp a design/optimization problem in which more than three independent variables feature is very limited. This almost automatically calls for the use of Computer Aided Design techniques. Further, the more disciplines are taken into account, the more complex and time-consuming the design evaluation will be. This complexity has a direct consequence for the optimization.

---

The presently used methods of optimization can be divided into two major classes, being the gradient-guided algorithms and the **direct search methods**. The difference between the two types of methods is the way in which the information is gathered that guides the algorithm towards the optimum. The gradient-guided algorithms use the derivatives of the constraints and merit function with respect to the independent variables, in order to define a search direction that reduces the merit function while staying in the feasible region of the design space. Especially in a multivariate design space where the relations between the design parameters and variables are susceptible to noise, the gradient-guided algorithms tend to get stuck in the first local optimum they encounter. This is a major drawback of gradient-guided algorithms. The main advantage, however, is that they give exact results.

Since the calculations for an aircraft design analysis are generally rather complex and often include iterative solution procedures, no closed-form analytic relations are available for the gradients needed. Instead, they need to be obtained by finite-differencing (perturbing one variable at a time and recalculating the complete design), which requires many design evaluations (number of variables +1 for every major optimization iteration in case of simple forward differencing). Since the number of design evaluations necessary to attain the optimum will increase with the number of independent variables, the increased computational time required for a single design evaluation will render the calculation effort prohibitively large.

The direct search methods do not make use of gradients. This type of optimization algorithm merely generates a large number of designs to compensate for the lack of information about the nature of the design space which would normally be provided by the derivatives. For a complex design analysis as associated with conceptual aircraft design, the computational effort of this class of optimization methods will be excessive too. Direct search methods on the one hand do not display the tendency to get stuck in local optima, but on the other hand they are not exact. Further characteristics of both methods will be discussed in this work.

Summarizing, the major problem that can be identified with respect to multivariate optimization is how to obtain the necessary information to guide the optimization procedure toward the optimum in a reasonable time, given the fact that a multivariate design evaluation with sufficient accuracy and predictive fidelity is extremely time consuming. Since the calculation effort required for a multivariate/multidisciplinary optimization problem is excessive, regardless of the method of optimization employed, the main improvement is to be obtained by a reduction of the computational time, either by using high-speed computers, which is quite expensive, or by reducing the time needed to carry out a single design evaluation.

## 1.1 References

1. Sobieszcanski-Sobieski, J.; Barthelémy, J.-F.M. and Giles, G.L.; *Aerospace Engineering Design by Systematic Decomposition and Multilevel Optimization*; ICAS-84-4.7.3
-

---



## 2. Introduction to Optimization Theory

### 2.1 Summary

In this chapter, a general introduction to optimization theory is given and some gradient-based methods of optimization are summarized. The purpose of the following paragraphs is not to give a complete mathematical treatise of optimization but to provide a basic introduction to readers not familiar with the subject. Therefore, the -often tedious- proofs required for a sound mathematical analysis are omitted here. However, they can be found in any elementary textbook on classic optimization theory. For more information, the reader is referred to the literature at the end of this chapter, on which most of the following is based.

### 2.2 Notation

A	matrix of constraint gradients
a	element of matrix of constraint gradients
b	active value of constraint
C	constraint constant
F	objective function
g	nonlinear constraint
H	Hessian matrix of objective function (or approximate Hessian) (matrix of second derivatives)
I	unit matrix
i	index
L	Augmented Lagrangian Function lower bound matrix
m	number of constraints
P	penalty (function)
Q	quadratic approximation to objective function
q	Jacobian matrix of objective function (matrix of first derivatives)
r	multiplier
S	search direction vector
t	number of active constraints
U	upper bound matrix
$\bar{x}$	vector of design variables
$\bar{y}$	vector of design variables
$\alpha$	steplength
$\lambda$	Lagrange multiplier tensor
$\bar{\lambda}$	vector of Lagrange multipliers
$\Phi$	penalty function

---

### superscripts

\*            optimal condition  
k            iteration number

### abbreviations

SLP        Sequential Linear Programming  
SQP        Sequential Quadratic Programming  
SUMT      Sequential Unconstrained Minimization Techniques

## 2.3 A Survey of Methods for Constrained Optimization

### 2.3.1 First-Order Methods

In general, an optimization problem can be described as follows: calculate the vector of design variables  $\bar{x}$ , such that a certain function  $F(\bar{x})$ , called the objective function, is minimized subject to constraints  $g_i(\bar{x})$ ,  $i=1,m$ . The objective function and constraints are in general nonlinear functions of the vector of independent design variables. In the case of aircraft design, the problem could be to determine the geometric variables (like aspect ratio or wing area) that drive the maximum take-off weight or the direct operating costs to a minimum, while simultaneously satisfying the specifications concerning range, take-off field length etc. Extra constraints can appear as **bounds** imposed on the design variables themselves, either to keep their values realistic, or to prevent them from attaining values outside the range where the calculation methods are valid. The constraints in effect limit the design space as they reduce the degrees of freedom of the design problem. Therefore, at the optimum at least some of these constraints will be satisfied with equality. Such constraints are called "active". Mathematically the problem can be stated as follows:

$$\begin{array}{llll} \text{minimize} & F(\bar{x}) & & \\ \text{subject to} & g_i(\bar{x}) \geq 0, \quad i=1,m & \text{inequality constraints} & \\ & L \leq \bar{x} \leq U & \text{bounds on variables} & \end{array}$$

As the "optimality" of the design is measured by the objective function, special attention should be paid to its choice, because it determines the values of the design variables (an aircraft designed for minimum take-off weight will look different from an aircraft that was designed for minimum fuel consumption, see also [ref.2] and [ref.5]). Although the question of which parameter should become the objective function is an important one, in this work it was chosen arbitrarily.

Most optimization algorithms proceed in an iterative manner, computing a new vector of design variables at every iteration step according to

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \alpha S^{(k)} \quad (1)$$

such that

$$F(\bar{x}^{(k)} + \alpha S^{(k)}) \leq F(\bar{x}^{(k)}) \quad (2)$$

In this notation  $S$  is the vector that defines the search direction, while  $\alpha$  is the step taken in that direction which is positive by definition. The condition that every step must yield an improved, or at least an equal value for the objective function can be linearized as follows:

$$F(\bar{x}^{(k)}) + \alpha \nabla F(\bar{x}^{(k)}) S^{(k)} \leq F(\bar{x}^{(k)}) \quad (3)$$

$$\Rightarrow \nabla F(\bar{x}^{(k)}) S^{(k)} \leq 0 \quad (4)$$

This relation states that the maximum value of the inner product of the search direction vector and the gradient vector is zero, which implies that the search direction is perpendicular to the gradient of the objective function. The region for which the inner product is less than zero is referred to as the usable region. The above expression is a necessary condition for an optimum: if there is a vector  $S$  that satisfies this condition, then the corresponding vector of design variables does not define a minimum in the objective function. Apparently, the largest reduction of the objective function can be achieved if the inner product in equation (4) is minimized. This forms the basis for the steepest-descent method used in unconstrained optimization. In this method, the search direction is found according to:

$$S^{(k)} = -\nabla F(\bar{x}^{(k)}) \quad (5)$$

$$\Rightarrow \bar{x}^{(k+1)} = \bar{x}^{(k)} - \alpha \nabla F(\bar{x}^{(k)}) \quad (6)$$

Along the computed search direction a line search is performed in order to obtain a value for the steplength that provides the largest decrease in the objective function. For the new value of the vector of design variables  $\bar{x}$  a new gradient is calculated and the procedure is repeated. Several modifications to this method were developed in order to improve the rather poor convergence speed, which is associated with a phenomenon known as zigzagging. Once a search direction has been established, the

---

steplength towards the new iteration point follows from:

$$\frac{\partial}{\partial \alpha} F(\bar{x}^{(k)} + \alpha S^{(k)}) = 0 \quad (7)$$

which yields:

$$\nabla F(\bar{x}^{(k+1)}) S^{(k)} = 0 \quad (8)$$

This implies, that the gradient at the new iteration point is perpendicular to the previous search direction. Therefore, the convergence process proceeds along mutual orthogonal paths. If the optimum lies in a narrow valley, this is obviously not very efficient.

The steepest-descent method does not possess the property of quadratic termination, which implies that it cannot minimize a quadratic function with a positive-definite Hessian matrix in a finite number of iterations. Since the search proceeds along mutual orthogonal paths, the gradients of a point obtained in the  $k^{\text{th}}$  iteration and a point obtained in iteration  $k+2$  are parallel. Therefore, the following applies:

$$\nabla F(\bar{x}^{(k)}) = C \nabla F(\bar{x}^{(k+2)}) \quad (9)$$

If  $F$  is a quadratic function with a positive-definite Hessian, it can be written as follows (the Hessian  $H$  is the matrix of second-order derivatives):

$$F(\bar{x}) = \frac{1}{2} \bar{x}^T H \bar{x} - q^T \bar{x} \quad (10)$$

Then equation (9) can be written as:

$$H \bar{x}^{(k)} - q = C (H \bar{x}^{(k+2)} - q) \quad (11)$$

Since at the optimum  $(\bar{x}^*)$   $H \bar{x}^* = q$ :

$$H(\bar{x}^{(k)} - \bar{x}^*) = C H(\bar{x}^{(k+2)} - \bar{x}^*) \quad (12)$$

and because  $H$  is not singular:

$$\bar{x}^{(k)} = C(\bar{x}^{(k+2)} - \bar{x}^*) + \bar{x}^* \quad (13)$$

which shows that  $\bar{x}^{(k)}$ ,  $\bar{x}^{(k+2)}$  and  $\bar{x}^*$  are situated on one line. Therefore, the search will zigzag between two straight lines through the optimum and hence it will require an infinite number of iterations to attain the optimum, unless the Hessian is the unit matrix, in which case the contours of  $F(\bar{x})$  are circular.

A necessary condition for the optimum of  $F(\bar{x})$  is that  $\nabla F(\bar{x}^*)=0$ . However, this condition is by no means sufficient, since it can also define a maximum or a saddle-point. For  $\bar{x}^*$  to be a global minimum, it is necessary that the Hessian of the function  $F(\bar{x})$  is positive-definite, that is,  $S^T \nabla^2 F(\bar{x}) S > 0$ , for any  $\bar{x}$  and  $S$ .

A function is said to be convex if:

$$f(x + \lambda(y - x)) \leq \lambda f(y) + (1 - \lambda)f(x) \quad (14)$$

and it can be proven that a function with a positive semi-definite Hessian is a convex function ([ref.3]). Convex functions are the only functions for which the obtained minimum is truly a global minimum. This can be proven as follows. Imagine a point,  $w$ , on the line connecting the optimum  $\bar{x}^*$  and a point  $\bar{x}$ , which is not the optimum. If  $w$  is chosen sufficiently close to the optimum, then  $F(w) \geq F(\bar{x}^*)$  and:

$$w = \bar{x}^* + \lambda(\bar{x} - \bar{x}^*), \quad 0 < \lambda < 1 \quad (15)$$

Since  $F(\bar{x})$  is a convex function

$$F(w) \leq F(\bar{x}^*) + \lambda(F(\bar{x}) - F(\bar{x}^*)) \quad (16)$$

or:

$$F(\bar{x}) \geq \frac{F(w) - F(\bar{x}^*)}{\lambda} + F(\bar{x}^*) \geq F(\bar{x}^*) \quad (17)$$

which was to be proven.

If constraints are imposed on the problem, an extra condition with respect to the search direction should be fulfilled. This condition is expressed in the Farkas Lemma, which states:

given the vectors  $a_i$  ( $i=1,2,\dots,t$ ) and  $\nabla F(\bar{x})$ , there is no vector  $S$  satisfying both conditions

$$\nabla F(\bar{x}^{(k)})^T S^{(k)} < 0 \quad (18)$$

and

$$a_i^T S \geq 0, \quad i=1,2,\dots,t \quad (19)$$

if and only if  $\nabla F(\bar{x})$  is in the convex cone spanned by the vectors  $a_i$  ( $i=1,2,\dots,t$ ).

The proof, which will not be treated in detail here, follows from the fact that if  $\nabla F(\bar{x})$  is in the convex cone spanned by the vectors  $a_i$  ( $i=1,2,\dots,t$ ) it can be written in the shape:

$$\nabla F(\bar{x}) = \sum_{i=1}^t \lambda_i a_i, \quad \lambda_i \geq 0 \quad (20)$$

Thus,  $S$  cannot satisfy both conditions (18) and (19).

When the vectors  $a_i$  ( $i=1,2,\dots,t$ ) are interpreted as the first-order derivatives of the linear approximations of the  $t$  active constraints,  $\nabla g_i(\bar{x})$  ( $i=1,2,\dots,t$ ), the Lemma can be explained as follows: condition (18) -as explained before- defines a search direction that diminishes the objective function to be minimized. This direction is known as the usable direction. The second condition ((19)) states that the inner product of the search direction and the gradient vector of the active constraints must be positive, meaning that the search is made in a direction that increases the linearized constraints or keeps them active. This direction is referred to as the feasible direction. In the method of feasible directions these two conditions are used to define a search direction that is both usable and feasible. In case the constraints are nonlinear and form a convex set, a push-off factor prevents the search direction to become tangent to the constraint boundaries, which would result in a constraint violation.

The present notation is valid only for inequality constraints of the type  $g_i(\bar{x}) \geq 0$ . If the constraint values should be negative for feasibility, the sign of the second condition changes and the condition (20) becomes:

$$\nabla F(\bar{x}) = -\sum_{i=1}^t \lambda_i a_i, \quad \lambda_i \geq 0 \quad (21)$$

Now let  $\bar{x}$  be a point that satisfies the linearized constraints

$$a_i^T \bar{x} \geq b_i, \quad i=1,2,\dots,m \quad (22)$$

where just the first  $t$  constraints are satisfied as equalities. Then  $\bar{x}$  minimizes  $F(\bar{x})$  subject to the constraints only if  $\nabla F(\bar{x})$  can be expressed in the shape of equation (20) because then there does not exist a vector  $S$  satisfying both the usable and feasible conditions. If  $\nabla F(\bar{x})$  cannot be expressed in this shape, then such a search direction does exist. Since the linearized constraints in the next point can be written as:

$$a_i^T (\bar{x} + \alpha S) = a_i^T \bar{x} + \alpha a_i^T S \quad (23)$$

and because

$$a_i^T \bar{x} \geq b_i, \quad i=1,2,\dots,m \quad (24)$$

and condition (19) is valid, the following applies:

$$a_i^T (\bar{x} + \alpha S) \geq b_i \quad (25)$$

Thus, there is a feasible direction satisfying the usability condition. Then  $F(\bar{x})$  is not the optimum.

The relations that are fulfilled at the optimum can be expressed in a slightly different way:

$$\nabla F(\bar{x}^*) - \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}^*) = 0 \quad (26)$$

where  $\bar{x}^*$  is the optimum vector of design variables, which should be feasible, and:

$$\lambda_i g_i(\bar{x}^*) = 0, \quad \lambda_i \geq 0, \quad i=1, m \quad (27)$$

These relations are known as the Kuhn-Tucker relations necessary for an optimal point. In the same way as applied to the analysis of unconstrained problems, it can be shown that when the design space is convex, the Kuhn-Tucker relations are both necessary and sufficient to guarantee that the obtained point is a global optimum. The design space is said to be convex if the constraint boundaries form a convex set (meaning that all elements of a straight line drawn between two feasible points are feasible too) and if the objective function is convex as well (meaning that the Hessian is positive definite for all possible design points). The constraints form a convex set if the constraint functions are concave. Taking two arbitrary points,  $\bar{x}$  and  $\bar{y}$ , that comply with the concave constraint  $g_i(\bar{x})$ , it follows for all points on a line between  $\bar{x}$  and  $\bar{y}$ :

$$g_i(\bar{x} + \lambda(\bar{y} - \bar{x})) \geq g_i(\bar{x}) + \lambda(g_i(\bar{y}) - g_i(\bar{x})) = (1 - \lambda)g_i(\bar{x}) + \lambda g_i(\bar{y}) \geq 0 \quad (28)$$

since  $0 \leq \lambda \leq 1$ , which implies that any point in a convex set of constraints (which are concave by definition) is part of that set.

The Kuhn-Tucker relations can be interpreted as follows: when the optimum is reached, no search direction satisfying both the usability and the feasibility condition can be found, because the gradient of the objective function is canceled out by the vectorial sum of the gradients of the active constraints. The factors  $\lambda_i$  are referred to as the Lagrange multipliers which measure the change in the objective function caused by a change in the value of their corresponding constraint. If at the optimum certain constraints are not active, their Lagrange multipliers should be zero, as follows from (27).

If  $\bar{x}^*$  is a Kuhn-Tucker point of a convex problem, then it follows from (26) and  $\lambda_i \geq 0, i=1, m$  that the function:

$$F(\bar{x}) - \sum_{i=1}^m \lambda_i g_i(\bar{x}) \quad (29)$$



which is convex for a fixed  $\bar{\lambda}$ , has an unconstrained minimum in  $\bar{x}^*$ . Hence

$$F(\bar{x}) - \sum_{i=1}^m \lambda_i g_i(\bar{x}) \geq F(\bar{x}^*) - \sum_{i=1}^m \lambda_i g_i(\bar{x}^*) \quad (30)$$

Using equation (27) this yields:

$$F(\bar{x}) - \sum_{i=1}^m \lambda_i g_i(\bar{x}) \geq F(\bar{x}^*) \quad (31)$$

which implies  $F(\bar{x}) \geq F(\bar{x}^*)$  for  $\bar{x}$  complying with  $g_i(\bar{x}) \geq 0$ ,  $i=1, m$ . Therefore,  $\bar{x}^*$  is a minimum solution of  $F(\bar{x})$ .

This analysis shows, that in case the Lagrange multipliers  $\lambda_i$  at the optimum are known, the minimal solution  $\bar{x}^*$  of  $F(\bar{x})$ , subject to  $g_i(\bar{x}) \geq 0$ ,  $i=1, m$ , can be found by unconstrained minimization of the function

$$L(\bar{x}, \lambda) = F(\bar{x}) - \sum_{i=1}^m \lambda_i g_i(\bar{x}) \quad (32)$$

This function is known as the Lagrangian Function. If the problem is convex (convex objective function, concave constraint functions), the minus sign in the expression for the Lagrangian Function accomplishes that the Lagrangian itself is convex as well. Unfortunately, the optimal Lagrange multipliers are generally not known until the optimization is completed. In the Augmented Lagrangian Method, the optimal Lagrange multipliers are approximated by a recursive updating formula.

The optimum of a constrained function corresponds to a saddle point of the Augmented Lagrangian function (minimum in  $\bar{x}$ , maximum in  $\lambda$ ). This saddle point can be found either by first maximizing the Lagrangian over  $\lambda$  followed by minimizing over  $\bar{x}$  (this is known as the minimax, or primal problem, since it results in solving the original problem) or by first minimizing over  $\bar{x}$  followed by maximizing over  $\lambda$  (maximin, or dual problem). It can be proven that both methods yield the same optimal point.

### 2.3.2 Second-Order Methods

A second-order approximation to the objective function is:

$$F(\bar{x}) \approx F(\bar{x}^{(0)}) + \nabla F(\bar{x}^{(0)}) \delta \bar{x} + \frac{1}{2} \delta \bar{x}^T H(\bar{x}^{(0)}) \delta \bar{x} \quad (33)$$

where  $\delta \bar{x} = \bar{x} - \bar{x}^{(0)}$ . Differentiating with respect to  $\bar{x}$  yields:

$$\nabla F(\bar{x}) = \nabla F(\bar{x}^{(0)}) + H(\bar{x}^{(0)}) \delta \bar{x} \quad (34)$$

The gradient at the optimum should be zero, hence:

$$\delta \bar{x}^* = -H(\bar{x}^*)^{-1} \nabla F(\bar{x}^{(0)}) \quad (35)$$

or:

$$\bar{x}^* \approx \bar{x}^{(0)} - H(\bar{x}^{(0)})^{-1} \nabla F(\bar{x}^{(0)}) \quad (36)$$

Compared with the relation  $\bar{x}^{(k+1)} = \bar{x}^{(k)} + \alpha S^{(k)}$  this suggests a search direction ( $\alpha^*=1$ ):

$$S^{(k)} = -H(\bar{x}^{(k)})^{-1} \nabla F(\bar{x}^{(k)}) \quad (37)$$

This method for unconstrained optimization is called Newton's method. The use of second-order information in determining the search direction in nonlinear optimization is very efficient (a truly quadratic function would be optimized in exactly one step with steplength  $\alpha=1$ , therefore, Newton's method does possess the property of quadratic termination). However, it is very costly to obtain the Hessian matrix using finite-differencing methods. In most algorithms an approximate Hessian is used, which is updated after every iteration using a special recursive updating formula that uses gradient information gathered in previous steps. The procedure often starts by taking the unit matrix as an initial approximation (thus making the first step in the steepest-descent direction). Such methods are known as Quasi-Newton methods.

At every point, the Hessian must be in accordance with:

$$\left\{ \nabla F(\bar{x}^{(k+1)}) - \nabla F(\bar{x}^{(k)}) \right\} = H_k \left\{ \bar{x}^{(k+1)} - \bar{x}^{(k)} \right\} \quad (38)$$

Unfortunately, it is not possible to have the approximation at iteration  $k$  to comply

with this equation. However, it is endeavored to find an approximation to the Hessian after iteration  $k$ , to be used in iteration  $k+1$ , that complies with the equations:

$$\left\{ \nabla F(\bar{x}^{(i+1)}) - \nabla F(\bar{x}^{(i)}) \right\} = H_{k+1} \left\{ \bar{x}^{(i+1)} - \bar{x}^{(i)} \right\} \quad (39)$$

for all previously generated points ( $i=0,k$ ). Several updating equations are known, the most frequently used being the Davidon, Fletcher, Powell (DFP) and the Broyden, Fletcher, Goldfarb, Shanno (BFGS) equations ([ref.3]).

## 2.4 Treating Nonlinear Constrained Problems

Methods for solving nonlinearly constrained optimization problems can be divided into two classes:

1. methods that make use of so-called penalty functions that add a penalty to the objective function if constraints are violated. This results in an unconstrained problem, therefore these methods are known as Sequential Unconstrained Minimization Techniques, or SUMT.
2. methods that use the nonlinear constraints to calculate a feasible direction in order to minimize the objective function. These methods are known as **direct methods** (not to be confused with the direct search methods, which will be treated later on).

### 2.4.1 Sequential Unconstrained Minimization Techniques (SUMT)

The Sequential Unconstrained Minimization Technique creates a pseudo-objective function by imposing a penalty term for every constraint violation:

$$\Phi(\bar{x}, r_p) = F(\bar{x}) + r_p P(\bar{x}) \quad (40)$$

where  $F(\bar{x})$  is the original objective function and  $P(\bar{x})$  the imposed penalty function. The multiplier  $r_p$  determines the magnitude of the penalty. There are two kinds of SUMT, the exterior **penalty method** and the interior penalty method. The exterior penalty method imposes a penalty that can generally be described as:

$$P(\bar{x}) = \sum_{i=1}^m \left[ \min(g_i(\bar{x}), 0) \right]^2 \quad (41)$$

for inequality constraints of the type  $g_i(\bar{x}) \geq 0$ . Thus, this method imposes a penalty when the constraint is already violated. If a certain constraint violation is compensated by an improvement of the objective function larger than the penalty

---

imposed, the optimizer will accept this. Therefore, the exterior penalty method often yields an optimum solution that is in fact infeasible. By changing the multiplier  $r_p$ , such violations might be controlled, but on the other hand, too large a multiplier will result in a virtually discontinuous pseudo-objective function, which makes it impossible to use second-order methods. The principle of the exterior penalty function is illustrated in the following figures. The original objective function and the imposed constraints are shown in Figure 1, whereas the new pseudo-objective function is illustrated in Figure 2 for two different values of the multiplier  $r_p$ . It is clear that the optimum of the pseudo-objective function lies outside the feasible region, but by increasing the multiplier  $r_p$  the constraint violations can be reduced.

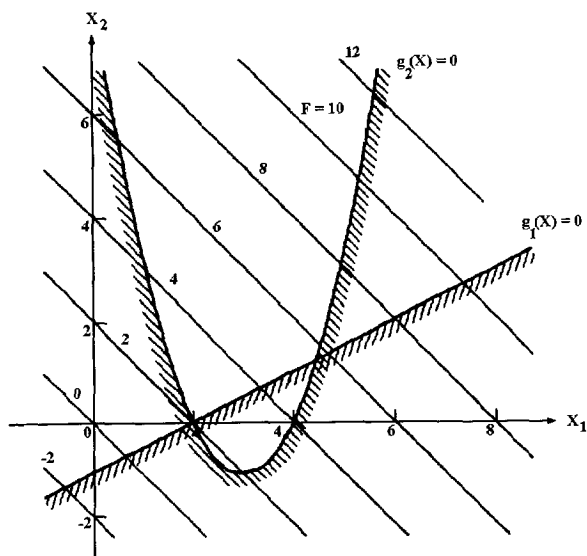


Figure 1 Original objective function and constraints ([ref.8])

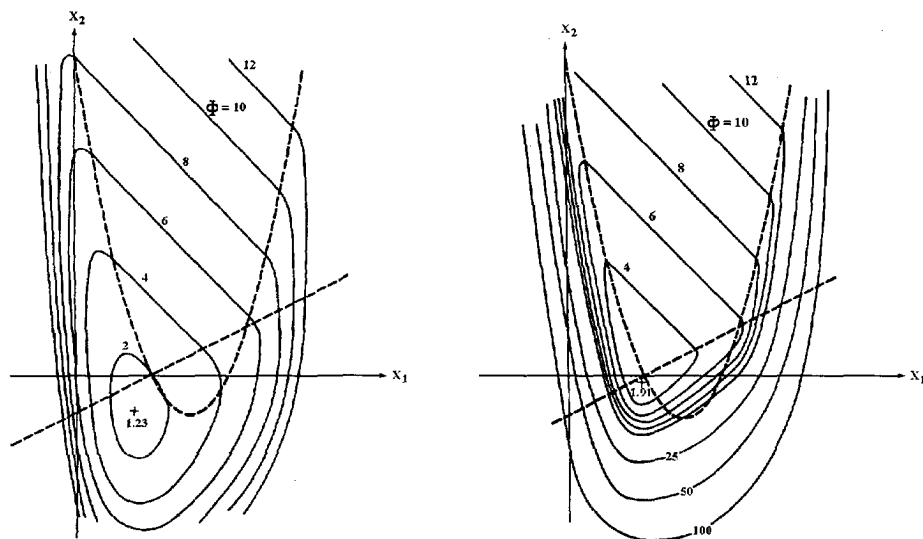


Figure 2 Pseudo-objective function  $r_p=0.05$  and  $r_p=1.0$  ([ref.8])

The problem of infeasible results does not occur with the interior penalty method. This method generally imposes a penalty of the kind:

$$P(\bar{x}) = \sum_{i=1}^m \frac{1}{g_i(\bar{x})} \quad (42)$$

or

$$P(\bar{x}) = - \sum_{i=1}^m \ln g_i(\bar{x}) \quad (43)$$

In this case, as  $g_i(\bar{x})$  is positive for feasibility, the penalty is a positive number that approaches infinity as a constraint becomes active.

There exists an important relation between Lagrangian functions and penalty

functions. Consider for instance the frequently used logarithmical barrier function:

$$B_r(\bar{x}) = F(\bar{x}) - r \sum_{i=1}^m \ln g_i(\bar{x}) \quad (44)$$

Assuming that the original (primal) problem is a convex one (convex objective function and concave constraint functions), the barrier function is a convex function as well. A necessary and sufficient condition for this function to be minimal is:  $\nabla B_r(\bar{x}) = 0$ . For the barrier function under consideration (although equivalent expressions can be derived for other penalty functions, either interior or exterior, provided that they are differentiable), this yields:

$$\nabla F(\bar{x}) - r \sum_{i=1}^m \frac{\nabla g_i(\bar{x})}{g_i(\bar{x})} = 0 \quad (45)$$

By putting:

$$\lambda_i = \frac{r}{g_i(\bar{x})} \quad (46)$$

the following expressions hold:

$$\nabla F(\bar{x}(r)) - \sum_{i=1}^m \lambda_i(r) \nabla g_i(\bar{x}(r)) = 0 \quad (47)$$

$$\lambda_i > 0 \quad (48)$$

$$\lambda_i(r) g_i(\bar{x}(r)) - r = 0 \quad (49)$$

Therefore, in any point within the constraint set, the first two equations are in accordance with the Kuhn-Tucker relations but for an equality sign in the second equation. For  $r$  approaching zero, the equality and the second Kuhn-Tucker relation will be satisfied as well. Thus, the use of penalty functions implies that dual-feasible solutions are generated.

As explained previously, in the Augmented Lagrangian Method, the Lagrange multipliers are approximated by a recursive formula. Depending on the values of the initial estimates of the Lagrange multipliers relative to their optimal values, the method converges from the feasible or the infeasible region. Theoretically, the initial

values could be chosen in such a way that the method would perform as an interior penalty function if infeasible steps are undesirable. In practice however, when using a standard routine from a program library like [ref.6], the choice of the initial values for the Lagrange multipliers is made in the routine independent of the problem to be solved. In that case, it will turn out during the optimization process whether the method acts as an interior or an exterior penalty function with respect to the active constraints.

For a more extensive survey of SUMT methods, the reader is referred to [ref.1], [ref.7] and [ref.8].

## 2.5 Direct Methods

### 2.5.1 Sequential Linear Programming (SLP)

The Sequential Linear Programming method, also known as the cutting plane method, approximates both the objective function and the constraints by a first-order Taylor expansion. The resulting linear constrained problem can be solved by standard linear optimization methods like for instance the Simplex method. At the obtained approximate optimum, the objective function and the constraints are linearized again and the procedure is repeated until convergence is attained. For convex constraints, the approximate optima that result from the linear programming problem could be infeasible. When the problem is relinearized at this point and reoptimized, the exact optimum finally would be obtained. However, if the procedure for any reason is stopped, there is a chance that the obtained design, although it will have an improved value for the objective function, is infeasible and therefore useless. In some applications this could be a disadvantage. An other problem can occur if the optimization problem is underconstrained, that is, when there are less active constraints at a point than there are design variables. Because the method proceeds toward the optimum by jumping to vertices in the design space, this could lead to an unbounded step when the problem is underconstrained. One way to solve this problem is to impose **move-limits** on the design variables, that constrain the step taken in the search direction. However, the magnitude of the move-limits has a direct influence on the search direction, as move-limits create artificial vertices. As shown in Figure 3, this can even make the difference between a certain variable being increased or decreased. The move-limits have another task in that they prevent the linearization from becoming inaccurate. Especially in the neighborhood of the optimum, linearization errors can prevent the optimizer from finding an improved value for the objective function. Therefore, the move-limits should be reduced during the optimization so that at the end the optimum is determined with an accuracy equal to the magnitude of the move-limits.

---

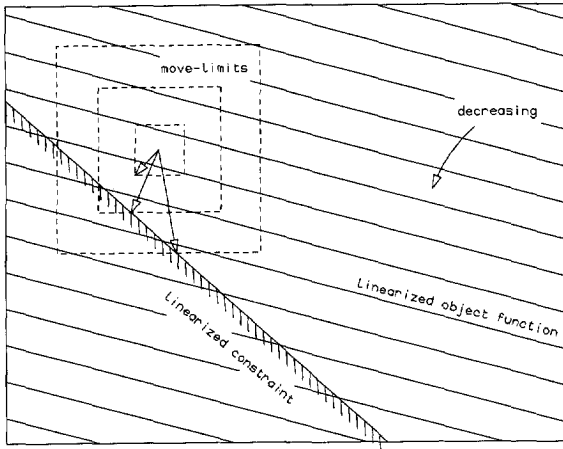


Figure 3 Effect of move-limits on search direction

### 2.5.2 Sequential Quadratic Programming (SQP)

The Sequential Quadratic Programming method creates a second-order approximation to the objective function and a linear approximation to the constraints:

$$\text{Minimize:} \quad Q(S) = F(\bar{x}) + \nabla F(\bar{x})S + \frac{1}{2}S^T H S \quad (\alpha=1)$$

$$\text{Subject to:} \quad \nabla g_i(\bar{x}) \cdot S + g_i(\bar{x}) \geq 0 \quad i=1,m$$

Where  $H$  is an approximation to the Hessian. The initial value for this approximation is generally the unit matrix. This is an optimization problem with the unbounded search vector as a variable. This problem could be solved using any available optimization routine. However, most quadratic programming routines use a so-called active set strategy. This implies that any constraint that becomes active (within certain tolerances) will be added to the active set. The projected gradient method is a method that solves the quadratic programming problem while keeping active linearized constraints active. This condition can be described by putting the feasibility condition to zero:  $\nabla g_i(\bar{x})S = 0$ . Writing the second-order approximation to the objective function as:

$$F(\bar{x}) = \frac{1}{2} \bar{x}^T H \bar{x} + q^T \bar{x} + \text{constant} \quad (50)$$

the condition for a global optimum (provided that  $H$  is positive definite for all



possible designs) becomes:

$$\nabla F(\bar{x}+S) = H(\bar{x}+S) + q = \sum_{i=1}^t \nabla g_i(\bar{x}) \lambda_i \quad (51)$$

Because the active constraints feature as equality constraints in the active set strategy, the Lagrange multipliers may be negative as well. Writing the gradients of the  $t$  active constraints in a matrix  $A$  and rewriting the Kuhn-Tucker conditions yields:

$$HS = A\bar{\lambda} - \nabla F(\bar{x}) \quad (52)$$

$$\Rightarrow A^T S = A^T H^{-1} A \bar{\lambda} - A^T H^{-1} \nabla F(\bar{x}) = 0 \quad (53)$$

$$\Rightarrow \bar{\lambda} = (A^T H^{-1} A)^{-1} A^T H^{-1} \nabla F(\bar{x}) \quad (54)$$

Thus, the search direction projected on the constraint boundaries becomes:

$$S = \{H^{-1} A (A^T H^{-1} A)^{-1} A^T H^{-1} - H^{-1}\} \nabla F(\bar{x}) \quad (55)$$

This is equivalent to the projected gradient method with unit steplength. If all Lagrange multipliers are non-negative, then the optimum is attained. Otherwise, if there are negative Lagrange multipliers the corresponding constraint with the largest negative multiplier is removed from the active set. When the search direction becomes zero, according to the Kuhn-Tucker relations the following expression holds for the  $t$  active constraints:

$$\nabla F(\bar{x}) = \sum_{i=1}^t \lambda_i a_i \quad (56)$$

where  $a_i$  are the elements of the  $i^{\text{th}}$  active constraint gradient. When the  $t^{\text{th}}$  active constraint is removed, the Kuhn-Tucker relations will not hold anymore, therefore there exists a vector  $S$  such that  $F(\bar{x}+S) < F(\bar{x})$ . Because  $a_t^T S = 0$ , substituting the above expression in the usability condition  $\nabla F(\bar{x}) S < 0$  gives:

$$\lambda_t a_t^T S < 0 \quad (57)$$

Because  $\lambda_t$  is negative and because  $a_t^T \bar{x} = b_t$ :

$$a_t^T (\bar{x} + S) > b_t \quad (58)$$

Then a search direction that is both usable and feasible can be calculated after deleting an arbitrary constraint with a negative Lagrange multiplier.

When the unbounded search direction  $S$  is found, a line search is started along this direction in order to find the steplength toward the new design point that will reduce the objective function as much as possible, while staying in the feasible region, according to (1). When a new constraint becomes active, this constraint is added to the active set.

The quadratic programming routine used in the present application, E04VCF [ref.6], uses the Augmented Lagrange Multiplier Method to perform the line search. After the line search has been completed, the approximate Hessian is updated using one of the previously discussed recursive updating formulas.

## 2.6 References

1. Gill, P.E. and Murray, W.; Numerical Methods for Constrained Optimization; Academic Press 1974
2. Jensen, S.C.; Rettie, I.H. and Barber, E.A.; *Role of Figures of Merit in Design Optimization and Technology Assessment*; Journal of Aircraft Vol.18, No. 2, also AIAA 79-0234R
3. Lootsma, F.A.; *Algorithms for Unconstrained Optimization*; Delft University of Technology, June 1986
4. Lootsma, F.A.; *Duality in Non-Linear Optimization*; Delft University of Technology, November/December 1991
5. Malone, B.; *High-Speed Civil Transport Study Using ACSYNT*; AIAA 93-4006
6. Numerical Algorithms Group; *Fortran Libraries Mk.12, Vol.4*
7. Piggot, B.A.M. and Taylor, B.E.; *Application of Numerical Optimization Techniques to the Preliminary Design of a Transport Aircraft*; RAE TR 71074
8. Vanderplaats, G.N.; Numerical Optimization Techniques for Engineering Design: With Applications; McGraw-Hill Inc. 1984

# 3. Methods to Decompose Large Optimization Problems

## 3.1 Summary

The concept of system decomposition was first introduced by Sobieszczanski-Sobieski ([ref.28]-[ref.37]) as a means of optimizing a design with many variables, subject to many constraints. The advantage of dividing the design process into various subsystems, while preserving the couplings between them, is that the calculation of the derivatives needed for a gradient-based optimization (and possibly the analysis and optimization of the subsystems themselves) can be performed concurrently and independently while taking into account the effects that changing a parameter in one subsystem has on parameters of other subsystems.

This lines up with procedures in industry (where decomposition is an organizational fact), gives good insight in the communications between subsystems and yields partial solutions as well. Furthermore, the concept of decomposition allows full flexibility in the choice of methods of analysis used in the different subsystems and in the choice of the dependent and independent variables. Also, it allows extra disciplines to be appended in a later stage of the project. Finally, the decomposition scheme provides sensitivity derivatives of subsystems, and therefore gives more information about the design space and the way in which the optimum is attained. In that way, the decomposition compensates -at least to some extent- for the disadvantage of optimization in comparison with a **parameter study**, that an optimization study only generates a single design (and possibly its sensitivity to design variable changes) whereas a parameter study gives some insight in the design space by systematically changing the design variables (or parameters). With increasing dimensionality, a parameter study becomes less practical and the information about the design space is lost. By applying the decomposition scheme, it is possible to obtain at every optimization step, the local sensitivities of the design parameters with respect to each other and of the (current) optimum with respect to design variables and parameters.

In the past, several software systems have been developed that are able to analyze different aircraft configurations as a function of a set of independent design variables. These systems invariably incorporate a series of self-contained programs or modules which provide the design parameters related to certain disciplines. Examples of such modules are weight prediction methods, aerodynamic prediction methods and performance calculation methods. Although in this way the design problem is decomposed into different disciplines, the decomposition scheme does not enable different subtasks to be carried out independently and concurrently, but merely consists of a number of subroutines called by a main coordination program.

Some of these computer-aided analysis systems are only capable of performing parameter studies by visualizing the effect of changing certain variables (IPAD, [ref.12] and [ref.13]), others incorporate optimizers as well (ACSYNT, [ref.39], ADAS, [ref.6] and the RAE program, [ref.18]). Where optimizers are used, the essential

---

gradient information is supplied by perturbing certain design variables and recalculating the whole design. This very inefficient method of calculating derivatives by finite-differencing of the complete design is known as the "brute-force" method. Since no true decomposition is carried out in these methods, it is hard to visualize changes within subsystems. Only end solutions are generated without knowledge on how these solutions are attained. Furthermore, because in some of these programs the methods of analysis that are implemented and the chosen independent variables cannot be changed, these design systems are often too rigid to be practical. Also, these programs use so-called zero-order methods of analysis, based on statistical or empirical data and the actual number of design variables that can be optimized is rather small. Therefore, these systems are applicable to a very limited class of design problems indeed.

The concept of system decomposition possesses two main advantages over the older generation of computer aided design systems that are important in multidisciplinary design and optimization. Firstly, it offers the possibility to perform certain calculations or analyses concurrently, thus saving much (computational) time. Secondly it enables calculations belonging to different disciplines to be carried out independently without the need to implement the complete design analysis into one program, which may not be possible because of the nature of the calculations. Furthermore, it leaves the responsibility for the calculations in the hands of those specialized and experienced in that particular field.

The way in which the communication between subsystems is carried out and how the optimization proceeds will be explained in this chapter. Furthermore, an optimization study is carried out in order to investigate the applicability of the decomposition method to large optimization problems.

For some recent publications on multivariate design and optimization systems employing the principle of decomposition, reference is made to [ref.2], [ref.8], [ref.19], [ref.20], [ref.21], [ref.24] and [ref.25].

### 3.2 Notation

A	subsystem name aspect ratio
B	subsystem name
C	subsystem name constraint constant
F	objective function
g	nonlinear constraint
i	index incidence
KS	Kresselmeir-Steinhauser function

---

k	index
L	lower bound matrix
M	Mach number
m	number of constraints
p	parameter
	index
r	responsibility factor
S	wing area
s	switch factor
t	permitted amount of constraint violation
U	upper bound matrix
$\bar{x}$	vector of design variables
Y	subsystem output
$\bar{y}$	vector of subsystem outputs
$\alpha$	angle of attack
$\Lambda$	sweep angle
$\lambda$	Lagrange multiplier
$q$	controlling factor
$\sigma_T$	engine thrust scale factor

**subscripts**

cr	cruise
le	leading-edge
w	wing

**abbreviations**

GSD	Global Sensitivity Derivatives
GSE	Global Sensitivity Equations
LSD	Local Sensitivity Derivatives
NSS	number of subsystems
SEP	Specific Excess Power
SLP	Sequential Linear Programming
SQP	Sequential Quadratic Programming
SUMT	Sequential Unconstrained Minimization Techniques

In accordance with tradition, weights are given in kilograms, although -according to the SI system- the unit of weight is Newton and the unit of mass is kilogram.

---

### 3.3 Decomposition of Hierarchic Systems

In the multilevel linear decomposition method introduced by Sobieszczanski-Sobieski the system is decomposed into a hierarchical tree structure. The couplings between the different subsystems are such that a certain subsystem gets its input only from subsystems at a higher level and there are no iterative relations between two subsystems. This kind of system breakdown is very well applicable to structural design analysis where a natural hierarchy can be identified in the way the structure is built up, with increasing detail toward the lower levels. Every subsystem is treated as a black-box, where only the inputs and outputs are important, but where it is not really necessary to know exactly which processes are carried out inside.

The optimization proceeds as follows. Each black-box has assigned to it a unique set of design variables (that is, these variables only feature in that particular black-box). Furthermore, it gets parameters from one or more higher level black-boxes as inputs (see Figure 4). These parameters are kept constant during a sublevel optimization. All constraints that are to be satisfied within one subsystem are put in one "cumulative constraint function" that measures the amount of constraint violations. There are several ways to formulate such a function (which acts as a penalty function, see chapter 2). Sobieszczanski-Sobieski favors the so-called Kresselmeir-Steinhauser function which is a continuous and differentiable function that follows the constraint boundaries at a distance controlled by the factor  $q$ :

$$KS = \frac{1}{q} \ln \left( \sum_{i=1}^m e^{q g_i} \right) \quad (59)$$

This cumulative constraint function features as the objective function that is to be minimized in each subsystem. When the optimal solution is found, the sensitivity derivatives of this optimum with respect to changes in the parameters (the inputs from higher level subsystems that were kept constant) are calculated. This calculation is performed using a special relation based on the Kuhn-Tucker equations, which should be satisfied at the optimum ([ref.4] and [ref.29]). A disadvantage of this method is, however, that inaccurate results might be obtained if the Kuhn-Tucker relations are not exactly satisfied at the optimum. It is assumed that the originally active constraints remain active after a slight perturbation of one of the parameters  $p$ :

$$\frac{d}{dp} \sum_{i=1}^m g_i(\bar{x}) = \sum_{i=1}^m \frac{\partial g_i(\bar{x})}{\partial p} + \sum_{i=1}^m \nabla g_i(\bar{x}) \frac{\partial \bar{x}}{\partial p} = 0 \quad (60)$$

Combination with the first Kuhn-Tucker equation ((26)), yields:

$$\nabla F(\bar{x}) \frac{\partial \bar{x}}{\partial p} = - \sum_{i=1}^m \lambda_i \frac{\partial g_i(\bar{x})}{\partial p} \quad (61)$$

Thus, the sensitivity of the optimal objective function value with respect to perturbances of the parameter  $p$  can be written as:

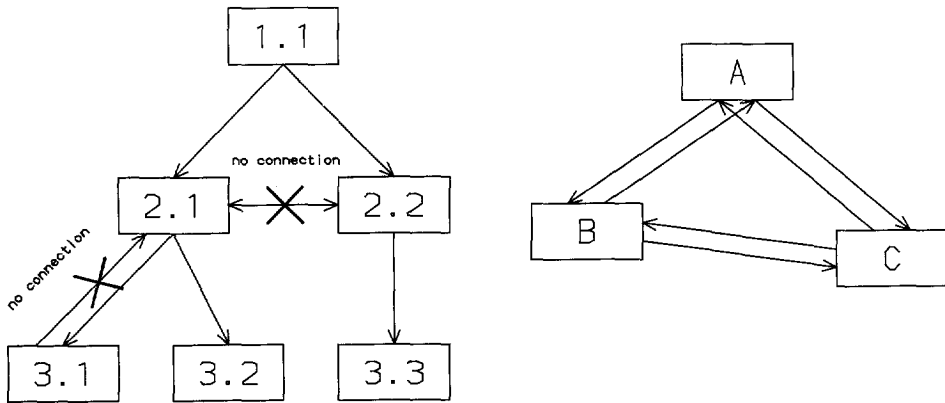
$$\frac{dF(\bar{x})}{dp} = \frac{\partial F(\bar{x})}{\partial p} + \nabla F(\bar{x}) \frac{\partial \bar{x}}{\partial p} = \frac{\partial F(\bar{x})}{\partial p} - \sum_{i=1}^m \lambda_i \frac{\partial g_i(\bar{x})}{\partial p} \quad (62)$$

Since at the optimum the Lagrange multipliers are usually known, the sensitivity derivatives of the objective function with respect to a parameter  $p$  only depend on the local sensitivities with respect to  $p$  of the objective function and the constraints.

The results of these **sensitivity analyses**, which are in fact linearizations of the lower level constraints with respect to the upper level variables, are taken upward to the next higher level where they feature in the cumulative constraints of the black-boxes at that level. To safeguard the accuracy of the constraint linearizations, move-limits are imposed on the free variables of the higher-level subsystems. This procedure enables the higher level subsystems to be optimized, while at the same time minimizing the constraint violations of the lower level subsystems. This sequence is repeated up to the highest level where the actual objective function is minimized subject to its cumulative constraint, which includes linearizations of all lower level constraints.

Next, the whole procedure is repeated with the newly obtained values for the parameters until convergence occurs. Although there are many applications of this method (mainly in the field of constructions and aero-elasticity, [ref.30] and [ref.33]) the method has some major drawbacks. As mentioned previously, it is necessary to assign a unique set of design variables to each subsystem. When applying the method to a complete conceptual aircraft design optimization, this may be very difficult -if not impossible- to accomplish. Also, the method is not applicable if interactions exist between two subsystems. Furthermore, if there is a relation between two subsystems at the same level or if one subsystem receives input from a lower level but provides input itself for a third subsystem at that same lower level, the method will not work. The former of these two problems might be solved by moving one of the two lower level subsystems to an "intermediate" level while preserving all its relations with other black-boxes, but the latter problem will be very hard if not impossible to cope with. In case one or more of the problems mentioned above occur, which prevents the system from being decomposable into a straightforward hierarchical tree-structure, the system is referred to as a network. In Figure 4 the difference in permitted interdisciplinary connections between a hierarchical decomposition scheme and a network is illustrated. It should be realized that the arrows in this figure only represent the data flow of parameters required for the analysis and not for the optimization, that is, the upward flow of optimum sensitivities with respect to higher-level parameters is not shown.

For the optimization of design problems that can only be decomposed in networks, a different procedure will be needed. The step from hierarchic to non-hierarchic systems was introduced in 1988 by Sobieszczanski-Sobieski ([ref.35], see also [ref.34], [ref.36])



**Figure 4** Data flow between black-boxes in hierarchical decomposition and network and [ref.7]).

### 3.4 Decomposition of Non-hierarchic Systems (Networks)

A network can be abstracted by the following equations (in this example the network consists of three subsystems denoted by the capitals A, B and C, see also Figure 4):

$$A(\bar{x}, \bar{y}_B, \bar{y}_C, \bar{y}_A) = 0 \quad (63)$$

$$B(\bar{x}, \bar{y}_A, \bar{y}_C, \bar{y}_B) = 0 \quad (64)$$

$$C(\bar{x}, \bar{y}_A, \bar{y}_B, \bar{y}_C) = 0 \quad (65)$$

The computations that take place within each subsystem (which will be treated as a black-box) are represented by a series of equations set to zero. In this notation  $\bar{x}$  represents the vector of independent design variables, while  $\bar{y}$  represents output from the black-box with the corresponding subscript. In practice not all elements of  $\bar{x}$  and  $\bar{y}$  will be used in every subsystem, but only certain subsets of these vectors. In this notation each subsystem is treated merely as a black-box that converts input (contained in the inner parentheses of the above equations) to output.

The influence of the design variables on the outputs of the various black-boxes can be



measured by means of the system sensitivity derivatives,  $\frac{dY_i}{dx_k}$ . By using the rules of chain differentiation, the following set of expressions for these Global Sensitivity Derivatives (GSD's) can be obtained:

$$\frac{dY_1}{dx_k} = \frac{\partial Y_1}{\partial x_k} + \frac{\partial Y_1}{\partial Y_2} \frac{dY_2}{dx_k} + \frac{\partial Y_1}{\partial Y_3} \frac{dY_3}{dx_k} \quad (66)$$

$$\frac{dY_2}{dx_k} = \frac{\partial Y_2}{\partial x_k} + \frac{\partial Y_2}{\partial Y_1} \frac{dY_1}{dx_k} + \frac{\partial Y_2}{\partial Y_3} \frac{dY_3}{dx_k} \quad (67)$$

$$\frac{dY_3}{dx_k} = \frac{\partial Y_3}{\partial x_k} + \frac{\partial Y_3}{\partial Y_1} \frac{dY_1}{dx_k} + \frac{\partial Y_3}{\partial Y_2} \frac{dY_2}{dx_k} \quad (68)$$

This set of equations, which clearly shows the coupling between the different subsystems (although not all couplings are necessarily present in practice), can be rewritten as follows:

$$\frac{dY_1}{dx_k} - \frac{\partial Y_1}{\partial Y_2} \frac{dY_2}{dx_k} - \frac{\partial Y_1}{\partial Y_3} \frac{dY_3}{dx_k} = \frac{\partial Y_1}{\partial x_k} \quad (69)$$

$$\frac{dY_2}{dx_k} - \frac{\partial Y_2}{\partial Y_1} \frac{dY_1}{dx_k} - \frac{\partial Y_2}{\partial Y_3} \frac{dY_3}{dx_k} = \frac{\partial Y_2}{\partial x_k} \quad (70)$$

$$\frac{dY_3}{dx_k} - \frac{\partial Y_3}{\partial Y_1} \frac{dY_1}{dx_k} - \frac{\partial Y_3}{\partial Y_2} \frac{dY_2}{dx_k} = \frac{\partial Y_3}{\partial x_k} \quad (71)$$

Or, in matrix notation:

$$\begin{pmatrix} I & -\frac{\partial Y_1}{\partial Y_2} & -\frac{\partial Y_1}{\partial Y_3} \\ -\frac{\partial Y_2}{\partial Y_1} & I & -\frac{\partial Y_2}{\partial Y_3} \\ -\frac{\partial Y_3}{\partial Y_1} & -\frac{\partial Y_3}{\partial Y_2} & I \end{pmatrix} \begin{pmatrix} \frac{dY_1}{dx_k} \\ \frac{dY_2}{dx_k} \\ \frac{dY_3}{dx_k} \end{pmatrix} = \begin{pmatrix} \frac{\partial Y_1}{\partial x_k} \\ \frac{\partial Y_2}{\partial x_k} \\ \frac{\partial Y_3}{\partial x_k} \end{pmatrix} \quad (72)$$

These equations, which are called the Global Sensitivity Equations, provide the Global Sensitivity Derivatives for a given vector of design variables  $\bar{x}$  (i.e. for a given design), when the System Sensitivity Matrix and the right-hand-side vector are known. The partial derivatives which feature in the System Sensitivity Matrix,

represent the sensitivities of the output values of each black-box to inputs from other black-boxes, while the right-hand-side vector represents the direct influence exerted on the output of each black-box by a certain design variable  $x_k$ . These partial derivatives can be calculated independently within each black-box, for a given design point. If these black-boxes are programs executed on a parallel processing computer or if they represent different engineering groups operating on their own, these local sensitivity analyses can be executed concurrently. The partial derivatives can be obtained either analytically or by finite-differencing within each black-box, that is, disturbing one of the input values and recalculating the outputs. If these calculations are executed concurrently for every black-box, this method means a significant reduction in CPU time, compared to finite-differencing of the complete analysis ("brute force method"). Another advantage of this method in comparison with the brute force method is, that good insight is obtained in what happens between black-boxes and what influence the change of a certain design variable exerts on the output of each black-box.

There are as many right-hand-side vectors as there are design variables, and the Global Sensitivity Equations must be solved for each right-hand-side vector. However, the Global Sensitivity Matrix is only generated once for every design point. The Global Sensitivity Derivatives measure the influence that each design variable exerts (either directly or indirectly via other black-boxes) on each black-box's output.

In [ref.35] two procedures are introduced for multivariate optimization, based on the explained System Sensitivity Analysis. In the first procedure (referred to as the method of Concurrent Subsystem Optimization) a unique subset of the vector of design variables is assigned to each subsystem. The allocation of a variable to a certain subsystem may be intuitive; the results from the System Sensitivity Analysis may be used as well. For each subsystem, the outputs that are subjected to constraints are represented by one cumulative constraint (for instance the Kresselmeir-Steinhauser function, equation (59)). Each subsystem now will be optimized independently to minimize the objective function of the system. If a certain subsystem contributes to the objective function indirectly, a first-order Taylor expansion will be used, as the total derivatives of the objective function with respect to each design variable are known from the System Sensitivity Analysis. If the cumulative constraint of the subsystem to be optimized is violated, this violation will be reduced at the least increase of the system objective function and if it is satisfied the system objective function will be minimized without violating the constraint.

Because of the coupling between subsystems, changes in the subset of variables of one subsystem will have influences on the outputs of other subsystems. Therefore, each subsystem will be optimized subject to its own cumulative constraint, plus linear extrapolations (based on the System Sensitivity Analysis) of the cumulative constraints of other subsystems. Using the results from the System Sensitivity Analysis it is possible to assign to each subsystem a factor measuring the "responsibility" for reducing violated constraints (including its own, since this constraint will partially be reduced by other subsystems as well). Also, if certain cumulative constraints are satisfied with equality, it may be beneficial to allow that constraint to become violated

---

in one subsystem optimization, as long as its violation is compensated in other subsystem optimizations. This could result in an overall improvement of the system objective function. The fact that a linear approximation is used in subsystems which indirectly contribute to the objective function, is consistent with linear programming, where the objective function is approximated by a first-order Taylor expansion instead of using exact analysis:

$$F(\bar{x}) = F(\bar{x}_0) + \nabla F(\bar{x}_0) \delta \bar{x} \quad (73)$$

To preserve the accuracy of the linear approximations, move-limits are imposed on the subsets of design variables to prevent the optimization from proceeding with too large steps. The formal notation of the problem as described above is [ref.35]:

minimize  $F(\bar{x}^k)$  subject to:

$$\begin{aligned} C^p &\leq C^{p0} s^p (1 - r_k^p) + (1 - s^p) t_k^p & p=1, \text{NSS} \\ L &\leq \bar{x}^k \leq U & (\text{move-limits}) \end{aligned}$$

(NSS is the number of subsystems)

In this notation the superscript  $k$  denotes the  $k^{\text{th}}$  subsystem optimization, while the superscript  $p$  corresponds to other subsystems. The factor  $r_k^p$  indicates the "responsibility" of the  $k^{\text{th}}$  subsystem optimization to reduce the violated constraint of subsystem  $p$ . The sum of all  $r_k^p$ 's must be one. The factor  $t_k^p$  measures the amount of the violation of the cumulative constraint of the  $p^{\text{th}}$  subsystem that is allowed in the  $k^{\text{th}}$  subsystem optimization. The sum of all  $t_k^p$ 's must be zero. The coefficient  $s^p$  is a switch that indicates whether the cumulative constraint of the  $p^{\text{th}}$  subsystem is violated at the beginning of the  $k^{\text{th}}$  subsystem optimization or not. If so,  $s^p$  is set to 1, thus enabling the factor  $r_k^p$ , in which case the  $p^{\text{th}}$  cumulative constraint is forced below a certain partition of its initial violation. If the  $p^{\text{th}}$  cumulative constraint is satisfied with equality, and only then, the switch is set to zero, enabling  $t_k^p$ . The  $C^p$ 's from other subsystems are approximated by:

$$C^p = C^{p0} + \sum_i \left( \frac{dC^p}{dx_i^k} \right) \Delta x_i^k \quad (74)$$

If the  $k^{\text{th}}$  subsystem optimization contributes indirectly to the objective function, then the influence exerted on the objective function is approximated by:

$$F = F^0 + \sum_i \left( \frac{dF}{dx_i^k} \right) \Delta x_i^k \quad (75)$$

The influence on inputs obtained from other black-boxes by the  $k^{\text{th}}$  subsystem optimization's subset of variables is approximated by a first-order Taylor expansion as well:

$$Y = Y^0 + \sum_i \left( \frac{dY}{dx_i^k} \right) \Delta x_i^k \quad (76)$$

The advantage of this method is that the overall optimization problem is reduced to a number of less complicated optimization problems that may be solved concurrently and independent of each other, while still preserving the couplings between them. When the subsystem optimizations are completed, the result is a new set of variables, new values for the outputs and a new, reduced value for the objective function. Because the optimizations were carried out for fixed values of the  $r_k^p$ 's and the  $t_k^p$ 's, the obtained optimum objective function depends on these values. Using the same procedure as in the hierarchic decomposition scheme (that is, based on the Kuhn-Tucker relations satisfied at the optimum, [ref.4] and [ref.29]), the sensitivity derivatives of the objective function to these factors can be calculated. These derivatives can be used for a linear extrapolation of the objective function with respect to the mentioned factors, to feature in a linear programming problem with the objective to calculate new values for the  $r_k^p$ 's and  $t_k^p$ 's that will further reduce the system objective function. This linear programming problem is subjected to the constraints that the  $r_k^p$ 's add up to one, and the  $t_k^p$ 's to zero. As in every linear programming problem, move-limits will have to be imposed in order to preserve the accuracy of the linear extrapolations. The results of this optimization are new values of the  $r_k^p$ 's and  $t_k^p$ 's that can be used in the next round of subsystem optimizations.

An alternative to the method explained above, is to use the Global Sensitivity Derivatives obtained from the System Sensitivity Analysis directly as an approximation for the exact analysis, thus creating a linear programming problem. Although this method has the disadvantage that the optimizer has to deal with the complete set of design variables simultaneously, it has all the advantages of the System Sensitivity Analysis in that the local sensitivities can still be calculated concurrently and independently. A further advantage is, that this alternative lacks the need to perform an optimum sensitivity analysis with respect to the factors  $r_k^p$  and  $t_k^p$  and an extra optimization with these factors as variables. Furthermore -as explained earlier- the calculation of this optimum sensitivity is based on the Kuhn-Tucker relations being fulfilled. If in one of the subsystem optimizations this is not the case, the procedure might render inaccurate results, or fail altogether.

A drawback of the simpler alternative, according to Sobieszczanski-Sobieski, is that it depends completely on approximate analysis, whereas the extensive procedure (Concurrent Subsystem Optimization) uses approximate analysis only to account for the couplings between subsystems. However, in a multidisciplinary design optimization (and especially in conceptual aircraft design), these interdisciplinary couplings by definition constitute a large part of all relations. Furthermore, even the influence of a subsystem's variables on its own output may be partly based on

approximate analysis. In the method of Concurrent Subsystem Optimization, only a partition of the vector of independent design variables is used in a certain subsystem optimization, while other variables that may also feature in that particular subsystem are kept constant. However, these variables will be changed in the subsystem optimization to whose subset they are allocated. The effect that this change has via the input from other subsystems on its own output (for instance the objective function) is modelled by a linear extrapolation. Therefore, the argument that this method relies on approximate analysis only in a very limited way, will not hold in many cases. This further implies that the method of Concurrent Subsystem Optimization -which in principle could employ any method of optimization (and even several different methods) to carry out the separate subsystem optimizations- will reduce in practical applications to a linear optimization method (just like the simpler alternative).

The Global Sensitivity Analysis takes its advantage mainly from the fact that the calculation of derivatives belonging to different subsystems can be performed concurrently, thus reducing computational time, vital in multidisciplinary optimization. If this is not possible, the method is far less efficient than the brute-force method that disturbs and reanalyzes the whole system. The question if the design evaluation itself can be carried out concurrently too depends entirely on the definition of the subsystems. Another advantage of the Global Sensitivity Analysis is of course that it provides information about the strength and the character of interactions between subsystems (but only if the existence of these interactions is recognized and incorporated in the System Sensitivity Analysis).

Applications of non-hierarchical optimization by decomposition can be found in [ref.1], [ref.3], [ref.9], [ref.15], [ref.22] and [ref.37].

### 3.5 Application to Multidisciplinary Optimization of a Conceptual Aircraft Design

The method of System Sensitivity Analysis was used to decompose the design synthesis of a conceptual aircraft design into several disciplines, in order to obtain local and global sensitivities of parameters with respect to the design variables. The Global Sensitivity Derivatives thus obtained were used directly in an optimization procedure.

The aircraft to be optimized is a forty-four passenger, short range airliner, powered by two General Electric CF-34 Turbofans.

The design parameters are divided into three categories: constants, **dependent variables** and independent variables. In the context of this work, the term parameters is used to indicate dependent variables, whereas the independent variables are often referred to as design variables (see Glossary). To keep the number of independent variables small for this application of multidisciplinary optimization, many design

---

parameters are kept constant. These constants include the number and type of engines, number of passengers and the complete fuselage design. The dependency of the parameters on each other is in some instances used to reduce the number of constraints. Many dependent variables feature as in- and outputs to and from the different subsystems, thus accounting for the interdisciplinary couplings. Also, the methods of analysis used in the different subsystems reduce the dimensionality of the problem, because these methods use certain assumptions or are only applicable to a specific kind of design problem (for instance, the use of Bréguet's range formula dictates a specific cruise technique).

Finally, seven variables remain as truly independent variables: wing area  $S$ , wing aspect ratio  $A$ , wing leading-edge sweep angle  $\Lambda_{le}$ , wing root incidence  $i_w$ , cruise Mach number  $M_{cr}$ , engine thrust scaling factor  $\sigma_T$  and cruise angle of attack  $\alpha_{cr}$ .

The optimization problem is to minimize the maximum take-off weight subjected to a number of constraints. Some of these constraints are bounds on the design variables, to prevent them from attaining values that are either unrealistic (for instance negative wing areas), or outside the range where certain methods of analysis apply (no forward swept wings or supersonic Mach numbers). Other bounds on the design variables are not explicitly set in the optimization problem, but are checked within the design analysis programs; violation of these constraints causes the program to terminate with a warning message (for instance if the combination of  $\alpha_{cr}$  and  $i_w$  is such as to cause the wing to stall). On one of the dependent variables, the wing span, a bound is imposed to prevent it from becoming unrealistically large.

The constraints that are used in the actual optimization problem are performance constraints, keeping the design within its specifications. In this application, the nonlinear inequality constraints are:

- 1) the difference between the available fuel fraction and the fuel fraction required to achieve a specified range, keeping 10% of the total amount of fuel as reserve.
- 2) the difference between the cruise thrust (95% Maximum Continuous Thrust) and the cruise drag, scaled with the cruise weight.
- 3) the difference between the specific excess power at zero altitude and a specified sea-level maximum rate of climb.
- 4) the difference between the specific excess power with one engine out at a specified one-engine-inoperative ceiling and the service ceiling rate of climb (0.5 m/s).
- 5) the difference between the calculated take-off field length with all engines operative and the specified take-off field length, scaled with the specified value.

With this definition, all constraints are of the kind  $g_i(\bar{x}) \geq 0$ ,  $i=1,5$ . The choice of the objective function (minimize the maximum take-off weight, or, since the number of passengers is kept constant, maximize the payload fraction) is arbitrary. However, it

must be realized that the choice of the objective function has a large influence on the optimum solution. In some cases it might be worthwhile to incorporate many features like weight, operating cost and fuel consumption in one objective function in which the relative importance of each feature is measured by means of scaling factors (see also [ref.5], [ref.10], [ref.17], [ref.23], [ref.24] and [ref.27]).

The decomposition scheme used is further illustrated in Appendix A. It is based on five separate disciplines or subsystems, being geometry, mass, aerodynamics, propulsion and performance. Each subsystem consists of a FORTRAN computer program, incorporating different kinds of system analysis. These programs get their input from a file containing the design variables, several files with constant parameters and files containing the output parameters from other subsystems. For a given vector of design variables these programs are executed in a specific order to fill the files of subsystem outputs. In some cases this requires re-executing programs between which an iterative relation exists. When the analysis of the current design point is completed, programs are executed that calculate the local sensitivities of each subsystem independently. These programs get as input the variables at the current design point and the output that was generated by the subsystem under consideration and the subsystems that supply input to it. The programs then perturb the input parameters and variables one at a time and calculate the local sensitivity derivatives by forward differencing. Incidentally, some local derivatives are more easily obtained analytically or by engineering judgement.

The obtained local sensitivity derivatives are sent to output files from where they are read by the program that performs the Global Sensitivity Analysis. This program stores the local derivatives in the Global Sensitivity Matrix and the right-hand-side vector of the Global Sensitivity Equations and solves them for each variable. Finally, the obtained Global Sensitivity Derivatives, which measure the influence, directly or indirectly, that each design variable exerts on each subsystem's output, are fed into an optimization program (see Figure 5).

Actually, only those derivatives that correspond to the objective function or the constraints will be used in this program. As can be seen in Figure 6, which shows the filling-up of the Global Sensitivity Matrix and the matrix of right-hand-side vectors, a total of  $339+144=483$  local sensitivity derivatives has been calculated in order to obtain  $7 \times 49=343$  global sensitivities. For the objective function and the six constraints (one geometric and five performance) only  $7 \times 7=49$  Global Sensitivity Derivatives are needed in the optimization program. Using the brute-force method, exactly this number would have been calculated. Although it is realized that the local sensitivities must be calculated concurrently to have an advantage in CPU time over the brute-force method at all, the difference between the number of derivatives calculated and the number of derivatives actually needed is quite large. On the other hand, the derivatives that are not directly needed in the optimizer still provide useful information about changes that occur within subsystems.

The optimization program calculates a new vector of design variables that should define a design with a reduced objective function (in this example a lower value of

---

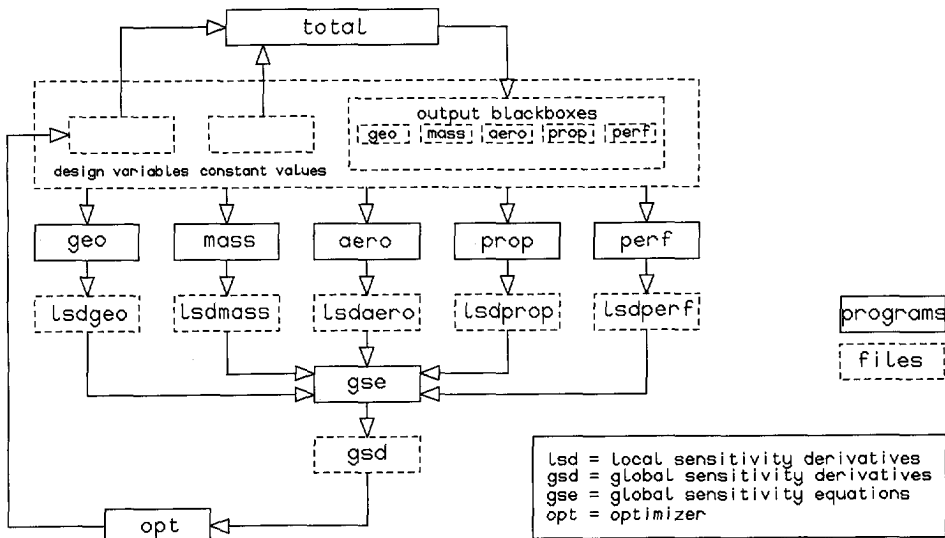


Figure 5 Decomposition scheme: programs and files

the maximum take-off weight) while still satisfying the constraints (although in some cases -as explained in chapter 2- constraints might be violated to some extent).

### 3.5.1 The System Sensitivity Analysis

When the analysis of the initial design is completed, the outputs from each black-box are written to their corresponding output files. After this, each subsystem analysis is again executed separately in order to obtain the local sensitivity derivatives by forward differencing. The results are written to local sensitivity files from which they are read by the program that solves the Global Sensitivity Equations. When the method is applied to a truly multivariate problem, the System Sensitivity Matrix can become quite large. To reduce the dimensions of the matrix, all constraints that are not input in other subsystems could be taken together in one cumulative constraint. This also reduces the feasible-direction finding problem in the optimization routine. To solve a large set of equations, Sobieszczanski-Sobieski proposes the use of the Gauss-Seidel iterative method. For a system as described in this chapter, with a sensitivity matrix that is for the greater part a lower triangular matrix, the solution is obtained immediately.

### 3.5.2 The Optimization Program

Two different optimization routines were used to solve the problem. Both routines are available from the Numerical Algorithms Group (NAG) Fortran Library ([ref.26]). The first algorithm used is a Sequential Linear Programming algorithm (E04NAF), the second one a Sequential Quadratic Programming algorithm (E04VCF).



	geometry	mass	aerodynamics	propulsion	performance	wing area	aspect ratio	sweep angle	angle of incidence	Machnumber	thrust scale factor	angle of attack
geometry	0	0	0	0	0	20	20	2	0	0	0	0
mass	0	12	0	0	0	2	0	0	0	0	2	0
aerodynamics	270	2	0	0	0	15	15	15	15	2	15	2
propulsion	0	3	7	1	2	3	0	0	0	2	5	0
performance	2	7	25	9	1	7	0	0	0	2	0	0

Figure 6 System Sensitivity Matrix and Right Hand Vectors

If every subsystem is optimized separately, the optimization program can be coupled directly with the subsystem analyses. This not only reduces the complexity of the problem, it also provides the possibility to use quasi-Newton optimization methods like the Sequential Quadratic Programming routine. In case only the system analysis is decomposed (to calculate the system sensitivity derivatives) but not the optimization itself, the optimization program operates separate from the analysis programs and gets only the system sensitivity derivatives as input. These sensitivities are used to linearize the problem, a procedure known as approximate analysis (see Figure 5). An implication of this method is that only one iteration at a time can be performed in the optimization program. After this, a new set of design variables is obtained and the whole sequence of system analysis and System Sensitivity Analysis is repeated before the optimization program is started again. This makes the application of quasi-Newton methods rather useless, because the Hessian cannot be updated after each iteration. When the optimization program operates independent of the analysis programs, no line search can be performed because the optimization program cannot carry out **function evaluations**. Thus, to facilitate the successful implementation of the Sequential Quadratic Programming method, the system analysis should be implemented in the optimization program. The gradient information will be supplied by the System Sensitivity Analysis, either implemented in the optimizer as well, or executed separately ([ref.9]). In case the system analysis is implemented in the optimizer, it would seem logical to do the same with the System

Sensitivity Analysis, thus enabling the optimization program to perform an arbitrary number of iterations, with an updated Hessian after each iteration. If it is not possible to execute the sensitivity analyses concurrently within the optimization program, the decomposition scheme should be dropped to save computation time.

Thus, when implementing a quasi-Newton optimization technique in the simplified alternative of Sobieszczanski-Sobieski's method of decomposition, two opposing interests appear: one to implement the total analysis in order to update the Hessian and perform a line search and another to preserve decomposition. In [ref.9] the analysis is directly coupled to the optimizer because of the simplicity of the analysis, but when the methods of analysis grow more complex, it may not be possible anymore to implement the complete analysis in one program. In this case some kind of approximate analysis more advanced than first-order Taylor approximations is needed, as it is meaningless to use a second-order optimization method to minimize a linearized objective function. The use of approximate analysis is discussed in [ref.14] and [ref.16] and in the following chapters.

In the present example problem, three different optimization methods were used:

- 1) Sequential Linear Programming  
(linear approximations to objective function and constraints)  
(one iteration at a time)

$$F(\bar{x}) = F(\bar{x}_0) + \nabla F(\bar{x}_0)(\bar{x} - \bar{x}_0)$$

$$g(\bar{x}) = g(\bar{x}_0) + \nabla g(\bar{x}_0)(\bar{x} - \bar{x}_0) \geq 0$$

problem statement:

$$\text{minimize} \quad F = \nabla F(\bar{x}_0)\bar{x} = C\bar{x}$$

$$\text{subject to:} \quad \nabla g(\bar{x}_0)\bar{x} \geq \nabla g(\bar{x}_0)\bar{x}_0 - g(\bar{x}_0)$$

$$\text{or:} \quad A\bar{x} \geq A\bar{x}_0 - g(\bar{x}_0) = \text{constant}$$

- 2) Sequential Quadratic Programming with separate sensitivity analysis by decomposition method  
(quadratic approximation to objective function, linear approximations of constraints)  
(one iteration at a time)

- 3) Sequential Quadratic Programming including sensitivity analysis by brute force method  
(quadratic approximation to objective function, linear approximations of constraints)  
(arbitrary number of iterations)

## 3.6 Results

### 3.6.1 Accuracy of Global Sensitivity Derivatives

Before the optimization was started, the Global Sensitivity Derivatives obtained by the decomposition scheme were compared to values obtained by the brute force method. Calculating sensitivity derivatives by means of finite-differencing implies that the accuracy of the results depends on the magnitude of the perturbations. If the system as a whole is being perturbed, for instance by perturbing one of the design variables with 1% of its initial value, the perturbation of the parameters will be proportional to their sensitivity with respect to that variable. However, if the local sensitivity derivatives are calculated for each subsystem independently, each parameter will be perturbed by the same 1% of its original value. The perturbations that occur within each subsystem when using the brute force method will be of quite a different order. Therefore, the accuracy of derivatives obtained by the decomposition method, will be different from that of the values obtained using the brute force method. This might result in large differences between the Global Sensitivity Derivatives calculated with the decomposition method and those calculated with the brute force method.

In table 1a of Appendix B, the Global Sensitivity Derivatives with respect to the wing area for the initial design are presented for both methods. The perturbation was 0.1% of the parameter's original value. Although there is very good agreement for many parameters, some derivatives seem to have quite considerable errors, notably the derivatives of the cruise drag and the related cruise thrust constraint, as well as those of the aerodynamic coefficients at take-off, the take-off thrust and the take-off field length constraint. Even worse, the signs of the latter two derivatives are wrong. Obviously, such errors in values that are supposed to supply the search direction for an optimization problem are intolerable. The same analysis was performed for perturbations of 10% (table 1b, Appendix B). Although now the agreement between both values of the take-off thrust and both values of the take-off field length has improved much and most other values are much the same for both methods, most aerodynamic derivatives have quite considerable errors. When finally the results of both methods are compared for perturbations of 1% it can be seen from table 1c of Appendix B that the agreement for all values is very good, but that there are again large discrepancies between both values for the take-off thrust and those for the take-off field length constraint.

From these results the following can be concluded: obviously the derivatives for take-off thrust and take-off field length are very susceptible to noise. Perturbations of 10%

---

show good agreement for these values, but yield inaccurate results for all other values. For perturbations of 0.1% other derivatives show a certain noise sensitivity too. Since perturbations of 1% give accurate derivatives for all except the two take-off parameters, it seems sensible to perturb all parameters by 1% and only the take-off thrust and the take-off field length constraint by 10%. The derivatives of the take-off thrust and the take-off field length constraint are not exactly equal to those obtained by the decomposition method with a 10% perturbation because these derivatives are related via the chain rule to derivatives which are based on a 1% perturbation.

### 3.6.2 Results of Sequential Linear Programming (SLP) Optimization

Twenty iteration steps were performed using the SLP method E04NAF. Apart from the bounds on the variables, move-limits were imposed that prohibited the variables to change more than 10% from their current values. As explained earlier in this dissertation, the SLP method can cause some constraints to become violated at the end of an iteration step. In some cases this could imply that the objective function is not being reduced in the next step.

If the problem is underconstrained, as in this case, the program will create artificial vertices by keeping some of the variables at temporary bounds. This has the disadvantage that the optimizer initially searches for a suboptimum, which makes convergence rather slow. The cruise Mach number and angle of attack are kept constant throughout the optimization, despite the fact that from the fourteenth step on, the incidence has reached its upper bound (which was imposed in order to prevent the sum of the angle of attack and the incidence to become so large as to cause the wing to stall). To prevent variables from being kept constant, the move-limits could be reduced to 1%. This indeed appeared to release both the cruise Mach number and the angle of attack from their temporary bounds. However, move-limits of 1% reduce the magnitude of the steplength to such an extent that this has a negative effect on the convergence speed as well.

Therefore it can be concluded that the SLP method, though it calculates new steps rather fast, is not well suited to perform the complete optimization, as it requires a large number of iteration steps, especially when the problem is underconstrained, as in this case. On the other hand, the SLP method provides a very fast, simple and cheap means to find a design point closer to the optimum than the initial design, which could very well be used as a starting point for a more effective method.

### 3.6.3 Results of Sequential Quadratic Programming (SQP) Optimization

To be able to use the SQP method, the system analysis program was incorporated in the optimization program (for more complex methods of analysis this should be done using some sort of approximate analysis). The decomposition was retained in order to calculate the Global Sensitivity Derivatives using the method of Global Sensitivity Equations. This meant that, like with the SLP method, only one iteration at a time could be performed by the optimization routine, before the System Sensitivity Equations had to be solved again. As explained earlier, the Hessian could therefore

---

not be updated after every iteration, causing the quadratic approximation of the objective function to be based on a unit quadratic coefficient.

After each major iteration (determination of a search direction) a minor iteration is performed in order to determine the steplength taken in the computed search direction (line search). The line search aims at finding the steplength that causes the largest reduction of a Lagrange Augmented Penalty Function (that performs as an interior penalty function or an exterior penalty function according to the values of the initial estimated Lagrange multipliers relative to the optimal ones). Thus, it is possible that the method generates infeasible designs. Especially when the design is still at a relatively large distance from the optimum, the Lagrange Augmented Penalty Function can be reduced a lot by reducing the part with the original objective function as much as possible, while accepting some constraint violations (these violated constraints will generally be the ones with the largest Lagrange multipliers). Towards the end of the optimization, when the objective function can not be improved upon much more, constraint violations will be avoided. The reason that a SUMT method is used to perform the line search is that -in theory- at the end of each iteration, infeasible designs might have been generated. This is caused by the fact that a projected gradient method is used to calculate the search direction. The only method that does not allow constraints to become violated is the method of feasible directions (see chapter 2). This method was probably not incorporated in the E04VCF program because it requires move-limits to be imposed on the variables, which would influence the general applicability of the program.

In this example, move-limits are imposed in order to obtain feasible designs while still improving on the objective function. Imposing move-limits, however, is not a guarantee for feasibility. As already explained, the projected gradient of a convex constraint set always points toward the infeasible region, no matter how small the move-limits are taken. Still, constraint violations can be limited by imposing move-limits. At the beginning of each iteration a feasible point is generated. From this point the projected gradient is calculated that defines the new search direction. When the derivatives needed for this calculation are provided from outside the optimization program, like in this case using the decomposition method, these derivatives are calculated from the result of the previous iteration. If this was an infeasible point, then the starting point for the next iteration will differ from this point. Therefore, imposing proper move-limits will reduce the error caused by the use of Global Sensitivity Derivatives that were calculated for an infeasible point.

The fact that the magnitude of the move-limits has to be adapted in every iteration makes convergence of this method very slow indeed. However, when the move-limits are chosen properly, the method will generate an (almost) feasible design, after each iteration step. This implies that whenever the program is terminated, at least a practical design has been created that is an improvement on the initial design (and all other previously generated designs). The problem that the move-limits make convergence rather slow, might be somewhat alleviated by assigning different move-limits to each variable, based on the sensitivity derivatives.

---

### 3.6.4 Sequential Quadratic Programming including System Sensitivity Analysis

To be able to perform an arbitrary number of iterations without having to stop the optimization program after each iteration to calculate the sensitivity derivatives elsewhere, the System Sensitivity Analysis was incorporated in the optimization program as well. This enables the program to update the approximate Hessian after each iteration. Because it was not possible to perform concurrent calculations within one program, the decomposition scheme was dropped in this case and the Global Sensitivity Derivatives were calculated using the brute force method. Because now both the system analysis and the System Sensitivity Analysis are incorporated in the optimization program, instead of using approximate analysis, the time to complete an iteration step increases dramatically. On the other hand, the program proceeds efficiently toward the optimum without requiring move-limits or keeping variables constant. Of course, dropping the decomposition scheme creates an optimization program like the older generation, which only delivers end-solutions while any information on how these solutions are attained is lost.

The program sometimes terminates because of ill-conditioned matrices (probably caused by some dependency between linearized constraints). Scaling the objective function and constraints greatly reduces this problem, but it does not completely solve it. However, when the maximum number of iterations is limited such that the ill-conditioning does not become critical, the program can be restarted from the final point. In this case the approximate Hessian will be taken as the unit matrix again. In optimization practice this phenomenon is well known. It seems that often the iteration process can be guided toward convergence by eliminating all information previously gathered by the Hessian recursive updating formula and starting off from the current design point by taking a steepest descent step.

When the final design point generated with the Sequential Linear Programming method was used as a starting point for the Quadratic Programming routine with full analysis, the program converged toward the optimum rather fast. Using the results of the thirteenth step generated by the SLP routine as a starting point, the optimum was found by executing the SQP program twice, with the maximum allowed number of iterations fixed at eight. The second time the program terminated after five iterations, having found the optimum. Thus, in a total of twenty-six iterations (thirteen using SLP and thirteen using SQP), the optimum solution was attained (see Appendix C).

At the optimum, three constraints are active; both cruise constraints and the constraint for sea level maximum rate of climb. Furthermore, the lower bound on the leading-edge sweep angle is active. Apparently, the program tends to keep the leading-edge sweep angle as small as possible in order to reduce wing weight. The cruise Mach number is reduced accordingly to avoid too large a drag rise. In the present example no constraint was put on the cruise speed or flight time and therefore the Mach number can become quite low. Also, no constructive constraints were imposed, other than the upper bound on the wing span. Therefore, the aspect ratio of the wing attains a very high value at the optimum. This again illustrates the importance of taking as many disciplines as possible into account.

---

It should be kept in mind that the obtained optimum is not only constrained by the active constraints and bounds, but also by the methods used. For instance, the effect that reducing the sweep angle has on stability has not been sufficiently accounted for. Also, the wing thickness ratio, which together with the wing sweep angle, lift coefficient and cruise Mach number influences the drag increase due to compressibility, has been kept constant.

Therefore, to obtain realistic results, as much disciplines, variables and constraints as possible should be taken into account. This is in accordance with Concurrent Engineering and only then optimal designs can be obtained that are suitable to be further developed toward a real design.

### 3.7 Conclusions

The example problem treated in this chapter is meant to illustrate the use of decomposition in optimization and to investigate its applicability. It can be concluded that the method of system decomposition provides a very useful tool for system analysis, System Sensitivity Analysis and optimization problems because it enables different tasks to be carried out independently and concurrently. The latter feature is important from the viewpoint of calculation time while the former feature makes the method especially suitable for use in industry, where different design tasks are divided over separate groups that sometimes are separated geographically as well.

The controversy between keeping the decomposition scheme on the one hand and being able to use more sophisticated optimization methods (quasi-Newton methods) successfully on the other hand, has been pointed out. It seems, that the complexity of available optimization routines must decrease, as the complexity of methods of analysis increases during the design process. This is the main reason why the decomposition concept was not further used in this work. Instead, a different way of obtaining gradient information in an efficient way is employed (see following chapters).

One should be very careful when using finite-differencing methods to calculate the sensitivity derivatives. If certain relations are highly nonlinear or sensitive to computational noise, finite-differencing may yield inaccurate results. Although finite-differencing seems to be ideally suitable for use in the decomposition method, as it requires no knowledge about the contents of each black-box, the inaccuracy and the fact that at least a second analysis has to be performed are serious drawbacks. Whenever possible, derivatives should be determined mathematically, or semi-mathematically (using approximate analysis). For simple relations this may be very practical, but for more complex relations like derivatives of aerodynamic parameters it may become very difficult indeed. Though some methods have been published to derive sensitivity derivatives of structural or aerodynamic parameters, these methods are all very problem dependent and therefore not universally applicable.

---

### 3.8 References

1. Abdi, F.; Ide, H.; Levine, M. and Austel, L.; *The Art of Spacecraft Design: a Multidisciplinary Challenge*; NASA CP-3031 part 1
  2. Balling, R.J. and Sobieszczanski-Sobieski, J.; *Optimization of Coupled Systems: a Critical Overview of Approaches*; AIAA 94-4330 CP
  3. Barnum, J.; Bathras, C.; Beene, K.; Bush, M.; Kaupin, G.; Lowe, S.; Sobieski, I.; Tingen, K. and Wells, D.; *Advanced Transport Design Using Multidisciplinary Design Optimization*; AIAA 91-3082
  4. Barthelemy, J.-F.M. and Sobieszczanski-Sobieski, J.; *Optimum Sensitivity Derivatives of Objective Functions in Nonlinear Programming*; AIAA journal Vol.21, No.6, June 1983
  5. Belegundu, A.D.; Murthy, D.V.; Salagame, R.R. and Constans, E.W.; *Multi-Objective Optimization of Laminated Ceramic Composites using Genetic Algorithms*; AIAA 94-4363 CP
  6. Bil, C.; *Development and Application of a Computer-based System for Conceptual Aircraft Design*; Ph.D. dissertation, ISBN 90-6275-484-8, Delft University Press, Delft 1988
  7. Bloebaum, C.L.; *Formal and Heuristic System Decomposition Methods in Multidisciplinary Synthesis*; NASA CR-4413, December 1991
  8. Braun, R.D.; Kroo, I.M. and Gage, P.J.; *Post-Optimality Analysis in Aerospace Vehicle Design*; AIAA 93-3932, August 1993
  9. Consoli, R.D. and Sobieszczanski-Sobieski, J.; *Application of Advanced Multidisciplinary Analysis and Optimization Methods to Vehicle Design Synthesis*; ICAS 90-2.3.4.
  10. Dovi, A.R. and Wrenn, G.A.; *Aircraft Design for Mission Performance using Nonlinear Multiobjective Optimization Methods*; Journal of Aircraft Vol.27, No.12, December 1990
  11. Edwards, B.; *The Use of Computer Based Optimization Methods in Aircraft Studies*; Royal Aircraft Establishment, Farnborough, England, 1979
  12. Fulton, R.E.; Sobieszczanski-Sobieski, J. and Landrum, E.J.; *An Integrated Computer System for Preliminary Design of Advanced Aircraft*; AIAA paper 72-796
  13. Fulton, R.E.; Sobieszczanski-Sobieski, J. e.a.; *Application of Computer-Aided Aircraft Design in a Multidisciplinary Environment*; Journal of Aircraft Vol.11, No.7, July 1974
  14. Haftka, R.T.; *Combining Global and Local Approximations*; AIAA journal Vol.29, No.9, September 1991
  15. Hajela, P.; Bloebaum, C.L. and Sobieszczanski-Sobieski, J.; *Application of Global Sensitivity Equations in Multidisciplinary Aircraft Synthesis*; Journal of Aircraft Vol.27, No.12, December 1990
  16. Hutchison, M.; Unger, E.; Mason, W.; Grossman, B. and Haftka, R.; *Variable-Complexity Aerodynamic Optimization of an HSCT Wing Using Structural Wing-Weight Equations*; AIAA 92-0212
  17. Jensen, S.C.; Rettie, I.H. and Barber, E.A.; *Role of Figures of Merit in Design Optimization and Technology Assessment*; Journal of Aircraft Vol.18, No.2, also AIAA 79-0234R
-



18. Kirkpatrick, D.L.I. and Larcombe, M.J.; *Initial Design Optimization of Civil and Military Aircraft*; AGARD CP-147, 1973
  19. Kroo, I.; *An Interactive System for Aircraft Design and Optimization*; AIAA 92-1190
  20. Kroo, I.; Altus, S.; Braun, R.; Gage, P. and Sobieski, I.; *Multidisciplinary Optimization Methods for Aircraft Preliminary Design*; AIAA 94-2543
  21. Kroo, I.; Altus, S.; Braun, R.; Gage, P. and Sobieski, I.; *Multidisciplinary Optimization Methods for Aircraft Preliminary Design*; AIAA 94-4325 CP
  22. Logan, T.R.; Sobieszczanski-Sobieski, J. and Abdi, F.F.; *Optimization of Aircraft Configurations in a Multidisciplinary Environment*; ICAS 90-2.3.2.
  23. Malone, B. and Mason, W.H.; *Aircraft Concept Optimization using the Global Sensitivity Approach and Parametric Multiobjective Figures of Merit*; AIAA 92-4221
  24. Malone, B.; *High-Speed Civil Transport Study using ACSYNT*; AIAA 93-4006
  25. Myklebust, A. and Gelhausen, P. *Putting the ACSYNT on Aircraft Design*; Aerospace America, September 1994, pp.26-30
  26. Numerical Algorithms Group; *Fortran Libraries Mk.12, Vol.3, Vol.4*
  27. Sliwa, S.M.; *Use of Constrained Optimization in the Conceptual Design of a Medium-Range Subsonic Transport*; NASA TP 1762, 1980
  28. Sobieszczanski-Sobieski, J.; *A Linear Decomposition Method for Large Optimization Problems - Blueprint for Development*; NASA TM 83248, 1982
  29. Sobieszczanski-Sobieski, J.; Barthelemy, J.-F.M. and Riley, K.M.; *Sensitivity of Optimum Solutions to Problem Parameters*; AIAA journal Vol.20, No.9, September 1982
  30. Sobieszczanski-Sobieski, J.; James, B. and Dovi, A.; *Structural Optimization by Multilevel Decomposition*; AIAA paper 83-0832
  31. Sobieszczanski-Sobieski, J.; Barthelemy, J.-F.M. and Giles, G.L.; *Aerospace Engineering Design by Systematic Decomposition and Multilevel Optimization*; ICAS 84-4.7.3.
  32. Sobieszczanski-Sobieski, J. and Barthelemy, J.-F.M.; *Improving Engineering System Design by Formal Decomposition, Sensitivity Analysis and Optimization*; International Conference on Engineering Design ICED 85, Hamburg, 26-28 August 1985
  33. Sobieszczanski-Sobieski, J.; James, B.B. and Riley, M.F.; *Structural Sizing by Generalized, Multilevel Optimization*; AIAA journal Vol.25, No.1, January 1987
  34. Sobieszczanski-Sobieski, J.; *Sensitivity Analysis and Multidisciplinary Optimization for Aircraft Design: Recent Advances and Results*; 16th Congress of the International Council of the Aeronautical Sciences (ICAS), Jerusalem, Israel, August 28-September 2, 1988, Journal of Aircraft Vol.27, No.12, December 1990
  35. Sobieszczanski-Sobieski, J.; *Optimization by Decomposition: a Step from Hierarchic to Non-hierarchic Systems*; NASA CP-3031 part 1, September 1988
  36. Sobieszczanski-Sobieski, J.; *Sensitivity of Complex, Internally Coupled Systems*; AIAA journal Vol.28, No.1, January 1990
  37. Sobieszczanski-Sobieski, J.; Bloebaum, C.L. and Hajela, P.; *Sensitivity of Control-Augmented Structure Obtained by a System Decomposition Method*; AIAA journal Vol.29, No.2, February 1991
  38. Torenbeek, E.; *Synthesis of Subsonic Airplane Design*; Delft University Press, Delft, 1982
  39. Vanderplaats, G.N.; *Automated Optimization Techniques for Aircraft Synthesis*; AIAA paper 76-909
-

## Appendix A Overview of the Variables and Parameters used in the Decomposition Scheme

### design variables (total: 7)

wing area  
wing aspect ratio  
wing leading-edge sweep angle  
wing root incidence  
cruise Mach number  
thrust scale coefficient  
cruise angle of attack

### constant parameters

tailplane sweep angles  
wing and tailplane taper ratios  
tailplane aspect ratios  
tailplane to wing quarter chord distance  
tailplane to wing vertical distance  
boat tail upsweep angle  
wash-out of wing and horizontal tailplane  
horizontal tailplane incidence  
fuselage length, width and height  
lengths of fuselage non-cylindrical nose- and tailsections  
distance from nose to leading-edge wing-fuselage intersection  
type of flaps  
horizontal tailplane position with respect to vertical tailplane  
wing vertical position  
type of engine  
number of engines  
areas of fuselage non-cylindrical nose- and tailsections in top and side view  
unscaled engine weight  
tailplane volume  
fuel density  
fan-section length and diameter (baseline engine)  
gasgenerator-section length and diameter (baseline engine)  
lifting surfaces thickness to chord ratios of tip and root airfoil sections  
nose radii and tailshapes of airfoil sections  
design lift coefficients of airfoil sections  
maximum lift coefficients of airfoil sections  
Mach number of fully expanded flow behind fan  
flap chord relative to wing chord  
aileron volume

---

area of rudder relative to vertical tailplane area

distance between engine cowling and fuselage surface in top view

**geometry parameters (dependent variables) (total: 20)**

- 1      wing span
- 2      wing root chord
- 3      horizontal tailplane root chord
- 4      vertical tailplane root chord
- 5      wing geometric chord
- 6      wing root thickness
- 7      wing mean aerodynamic chord
- 8      horizontal tailplane area
- 9      vertical tailplane area
- 10     horizontal tailplane span
- 11     vertical tailplane span
- 12     wing semi-chord sweep angle
- 13     wing quarter chord sweep angle
- 14     net wing area
- 15     horizontal tailplane mean aerodynamic chord
- 16     vertical tailplane mean aerodynamic chord
- 17-19   spanwise locations of mean aerodynamic chords
- 20     fuel weight

**mass parameters (dependent variables) (total: 2)**

- 21      maximum take-off weight
- 22      average cruise weight

**aerodynamic parameters (dependent variables) (total: 15)**

- 23      cruise lift coefficient
- 24      cruise drag coefficient  
(maximum lift coefficient take-off flap setting)  
(maximum lift coefficient climb flap setting)
- 25-27   drag polar coefficients ( $h = 0$  m climb, take-off flap setting)
- 28-30   drag polar coefficients ( $h = 4600$  m climb, climb flap setting)
- 31      drag coefficient due to asymmetric flight condition
- 32      engine windmilling drag coefficient
- 33      take-off lift gradient
- 34      take-off zero degrees angle of attack lift coefficient
- 35-37   drag polar coefficients (take-off, take-off flap setting)

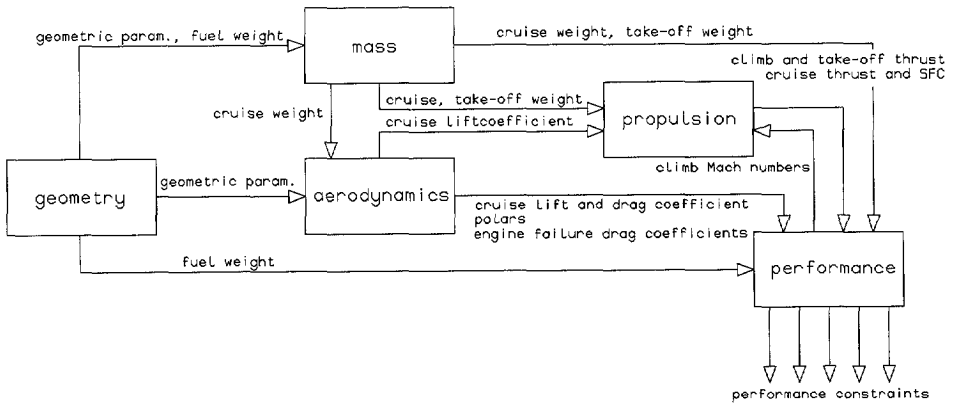
**propulsion parameters (dependent variables) (total: 5)**

- 38      cruise thrust
  - 39      cruise specific fuel consumption
-

- 40 thrust at sea level climb
- 41 thrust at one-engine-inoperative ceiling
- 42 average take-off thrust

**performance parameters (dependent variables) (total: 7)**

- 43 range constraint
- 44 cruise thrust constraint
- 45 sea level climb constraint
- 46 sea level climb Mach number
- 47 one-engine-out ceiling constraint
- 48 one-engine-out ceiling Mach number
- 49 take-off field length constraint



**Figure 7** Information flow between subsystems

Appendix B Accuracy of Sensitivity Derivatives

geometry parameters	
brute force method	decomposition method
1 0.256280	0.256202
2 0.036208	0.036223
3 0.040048	0.040070
4 0.054831	0.054848
5 0.023527	0.023542
6 0.007606	0.007609
7 0.025821	0.025834
8 0.374396	0.374559
9 0.254372	0.254382
10 0.130193	0.130220
11 0.047971	0.047972
12 0.000000	0.000000
13 0.000000	0.000000
14 0.894686	0.894887
15 0.028551	0.028576
16 0.048309	0.048317
17 0.000000	0.000001
18 0.000000	0.000001
19 0.000002	0.000001
20 196.4493	196.4801

mass parameters	
brute force method	decomposition method
21 270.5314	270.8242
22 182.1256	176.4947

aerodynamic parameters	
brute force method	decomposition method
23 -0.002024	-0.002221
24 -0.000582	-0.001006
25 -0.000354	-0.000259
26 -0.000102	-0.000083
27 0.000275	0.000282
28 -0.000470	-0.000469
29 -0.000086	-0.000085
30 0.000238	0.000237
31 -0.000149	-0.000149
32 0.000000	0.000000
33 -0.003783	-0.008681
34 -0.040121	-0.043883
35 -0.000405	-0.000249
36 -0.000047	-0.000033
37 0.000220	0.000239

Table 1a Global Sensitivity Derivatives with respect to S (perturbation 0.1%)

propulsion parameters	
brute force method	decomposition method
38 -37.43961	-35.83450
39 -0.008043	-0.007691
40 0.000000	0.000000
41 0.000000	0.000000
42 30.91787	-12.05427

performance parameters	
brute force method	decomposition method
43 0.008111	0.010719
44 -0.000927	-0.000381
45 -0.386014	-0.409952
46 0.000000	0.000000
47 -0.065775	-0.066116
48 0.000000	0.000000
49 -0.002990	0.117755

Table 1a (continued)

Global Sensitivity Derivatives with respect to S  
(perturbation 0.1%)

geometry parameters		
	brute force method	decomposition method
1	0.250329	0.250329
2	0.035365	0.035365
3	0.039586	0.039586
4	0.054195	0.054195
5	0.022987	0.022987
6	0.007427	0.007427
7	0.025209	0.025209
8	0.383606	0.383607
9	0.260584	0.260584
10	0.128653	0.128653
11	0.047420	0.047420
12	0.000000	0.000000
13	0.000000	0.000000
14	0.897442	0.897441
15	0.028218	0.028217
16	0.047743	0.047743
17	0.000000	0.000000
18	0.000000	0.000000
19	0.000000	0.000000
20	201.2352	201.2353

mass parameters		
	brute force method	decomposition method
21	276.6522	276.4500
22	186.0966	180.0291

aerodynamic parameters		
	brute force method	decomposition method
23	-0.002023	-0.001284
24	-0.000413	-0.000347
25	-0.000397	-0.000355
26	-0.000169	-0.000295
27	0.000362	0.000735
28	-0.000430	-0.000379
29	-0.000101	-0.000244
30	0.000264	0.000662
31	-0.000133	-0.000139
32	0.000004	0.000000
33	0.002127	0.004346
34	-0.036468	-0.021696
35	-0.000465	-0.000324
36	-0.000102	-0.000273
37	0.000269	0.000748

Table 1b    Global Sensitivity Derivatives with respect to S (perturbation 10%)

propulsion parameters	
brute force method	decomposition method
38 -29.72705	-51.07652
39 -0.006376	-0.009405
40 70.97826	41.67815
41 0.000000	-2.116536
42 30.58696	40.00017

performance parameters	
brute force method	decomposition method
43 0.006874	0.007573
44 -0.001283	-0.001329
45 -0.348647	-0.339726
46 -0.002415	0.001418
47 -0.067312	-0.103129
48 0.000000	0.000231
49 -0.004686	-0.005516

Table 1b (continued)      Global Sensitivity Derivatives with respect to S  
(perturbation 10%)



geometry parameters		
	brute force method	decomposition method
1	0.255797	0.255787
2	0.036138	0.036138
3	0.040014	0.040017
4	0.054785	0.054785
5	0.023488	0.023489
6	0.007589	0.007589
7	0.025761	0.025761
8	0.375338	0.375340
9	0.254961	0.254963
10	0.130056	0.130058
11	0.047935	0.047935
12	0.000000	0.000000
13	0.000000	0.000000
14	0.895193	0.895191
15	0.028524	0.028526
16	0.048263	0.048263
17	0.000000	0.000000
18	0.000000	0.000000
19	0.000000	0.000000
20	196.8937	196.8953

mass parameters		
	brute force method	decomposition method
21	271.0386	271.0558
22	182.4396	176.6330

aerodynamic parameters		
	brute force method	decomposition method
23	-0.001988	-0.001928
24	-0.000443	-0.000492
25	-0.000460	-0.000452
26	-0.000128	-0.000141
27	0.000282	0.000323
28	-0.000466	-0.000461
29	-0.000082	-0.000096
30	0.000231	0.000271
31	-0.000148	-0.000148
32	0.000000	0.000000
33	0.001926	0.002102
34	-0.036925	-0.037249
35	-0.000498	-0.000419
36	-0.000081	-0.000121
37	0.000233	0.000320

Table 1c    Global Sensitivity Derivatives with respect to S (perturbation 1%)

propulsion parameters	
brute force method	decomposition method
38 -37.58454	-41.69720
39 -0.008063	-0.008944
40 0.000000	0.000000
41 0.000000	0.000000
42 39.15458	7.567567

performance parameters	
brute force method	decomposition method
43 0.007231	0.007673
44 -0.001350	-0.001203
45 -0.357826	-0.361777
46 0.000000	0.000000
47 -0.065597	-0.070033
48 0.000000	0.000000
49 -0.001659	-0.019182

Table 1c (continued)

Global Sensitivity Derivatives with respect to S  
(perturbation 1%)

Appendix C   Initial and Final Design

In accordance with tradition, weights are given in kilograms, although -according to the SI system- the unit of weight is Newton and the unit of mass is kilogram.

design variables						
S [m <sup>2</sup> ]	A	$\Lambda_{ie}$ [rad]	$i_w$ [rad]	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$ [rad]
41.4	10.89	0.436	0.0349	0.60	1.00	0.000

objective function (maximum take-off weight [kg])
17527.15

constraints					
wing span	range	cruise thrust	zero altitude climb	one-engine-out ceiling	take-off field length
21.23 m	0.0377	0.0312	2.4155	0.2888	0.2546

cruise altitude = 6200 m

Table 2    Initial design

design variables						
S [m <sup>2</sup> ]	A	$\Lambda_{le}$ [rad]	$i_w$ [rad]	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$ [rad]
30.3	21.33	0.100	0.0712	0.55	0.71	0.0521

objective function (maximum take-off weight [kg])
13682.83

constraints					
wing span	range	cruise thrust	zero altitude climb	one-engine-out ceiling	take-off field length
25.44 m	-0.0006	-0.0003	-0.0005	1.0871	0.1915

cruise altitude = 11,500 m

decrease in maximum take-off weight: 21.9%  
 payload fraction increased from 23.8% to 30.5%  
 fuel-fraction decreased from 30.9% to 17.8%  
 range = 2498.6 km (specification: 2500 km)  
 cruise engine rating: 95.5% (specification 95%)

Table 3   Final design

# 4. Optimization of a Conceptual Aircraft Design using Approximate Analysis

## 4.1 Summary

This chapter describes the results of a conceptual aircraft design optimization using second-order polynomials as an approximation to the exact analysis methods. A comparison is made between the obtained optimal design and a design obtained by exact analysis in chapter 3. In the context of this chapter, the "exact" analysis is in fact a semi-empirical conceptual design method.

## 4.2 Notation

A	matrix of least squares coefficients aspect ratio
a	least squares coefficients (elements of A)
b	polynomial coefficients wing span
f	objective function
i	incidence
k	number of samples
M	Mach number
n	number of variables
S	wing area
$\bar{x}$	vector of design variables
y	response function
$\hat{y}$	approximate response function
$\alpha$	angle of attack
$\beta$	scale factor
$\epsilon$	error
$\epsilon_t$	constraint tolerance
$\hat{\epsilon}_t$	approximate analysis constraint tolerance
$\Lambda$	sweep angle
$\sigma_T$	thrust scale factor

### subscripts

0	initial (baseline)
a	analytical approximation
cr	cruise
d	detailed analysis
i	index

---

le	leading-edge
t	tolerance
w	wing

### abbreviations

MCC <sup>2</sup>	Multiple Correlation Coefficient Squared
SQP	Sequential Quadratic Programming

## 4.3 Introduction

As explained in chapter 3, to use sequential quadratic optimization routines successfully, it is necessary to:

1. incorporate the complete design analysis in the optimizer, thus enabling a line search to be performed
2. provide gradient information without having to stop the optimizer after each iteration step.

When dealing with complex methods of analysis, the first condition might be impossible to fulfil, especially when part of the design analysis is performed by finite element or panel methods. Such methods cannot be linked to other programs, and furthermore, they take a lot of computational time. Also, when the design analysis is performed by several different departments, possibly separated geographically, it is clearly impossible to implement the complete analysis into one single program. Still, it is very important to use methods of analysis that are as accurate as possible. In constrained optimization, the optimal design will always have some constraints active. If, however, the values of the active constraints are not predicted accurately, the eventual design might very well turn out to be infeasible, and therefore useless. Also, as optimal points often lie in the vertex of several constraint hyperplanes, a slight inaccuracy in the slopes of two nearly parallel constraint boundaries, might cause large variations in the location of the optimal point and have consequences for the predicted value of the objective function.

The second condition in fact prohibits decomposition to obtain derivatives when using quadratic optimization methods. As the decomposition scheme should be carried out within the optimization program, it is more efficient to calculate the derivatives by the brute force method (finite-differencing of the complete design analysis), because the decomposition method calculates many irrelevant derivatives as well (chapter 3). However, when dealing with multivariate problems, calculating derivatives by the brute force method will be a very cumbersome task, as a new design has to be re-analyzed for each and every variable, and again for every iteration step. If advanced and complex methods of analysis are used, this process will become prohibitively time-consuming. As the derivatives for large multidisciplinary problems practically always will have to be calculated by finite-differencing (because relations can no longer be expressed in simple differentiable functions), decomposing the system to

---

calculate gradients simultaneously will be vital. (In this context, the new developments in the field of automatic differentiation ([ref.1] and [ref.2]) are interesting. Automatic differentiation is carried out by a program that enhances the original code with instructions that recognize statements relevant for calculating a certain variable. The operations in those statements are then differentiated after which the global derivatives are computed by means of the chain rule of differentiation. The enhanced code then calculates the design parameters and their derivatives with respect to the design variables.)

These problems will force the designer, or so it seems, to use linear optimization routines, as they enable decomposition of the problem. Examples of such linear routines are Sequential Linear Programming, or, in case the optimization problem itself needs to be decomposed into smaller subproblems, the method of Concurrent Subsystem Optimization presented by Sobieszczanski-Sobieski, chapter 3, [ref.10] and [ref.11]. A major disadvantage of linear methods, however, is their slow convergence, because of the necessity to impose move-limits. (It is once more stressed here that the method of Concurrent Subsystem Optimization by Sobieszczanski-Sobieski is not inherently a linear programming method, since the subsystem optimizations in principle could be carried out by any optimization algorithm desired. However, since in practical conceptual airplane design many interdisciplinary couplings have to be approximated by linear extrapolation, the method will effectively reduce to a linear programming method.)

A possible solution to the problem described above, is the use of approximate analysis. In literature, methods using linear Taylor expansions, like all linear programming methods do, are referred to as approximate methods as well. In the context of this chapter, the term approximate analysis will imply higher order approximations of accurate but complex methods of analysis. Earlier results with advanced approximate analysis were reported in [ref.3], [ref.5], [ref.6] and [ref.9]. In these methods, the availability of two methods of analysis was assumed; a complex, accurate one, and a simpler empirical or statistical one. At the beginning of each iteration step, the results from both methods are examined for a series of designs, linearly varied between the current design and a second one. The approximation uses a scale factor  $\beta(\bar{x}_0) = f_d(\bar{x}_0) / f_a(\bar{x}_0)$  in which the subscript "d" stands for detailed analysis and the "a" for analytical method. At a certain design point  $\bar{x}$ , the scale factor is approximated linearly by:

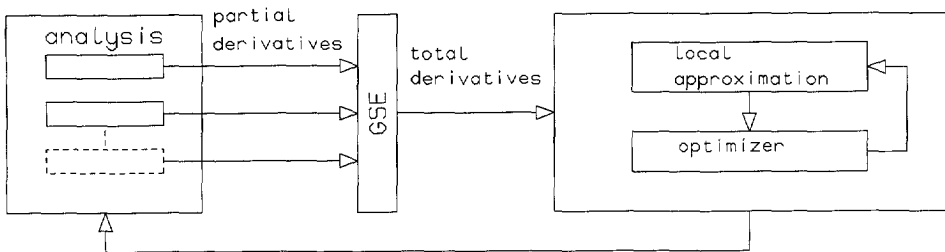
$$\beta(\bar{x}) \approx \beta(\bar{x}_0) + \nabla \beta_0 \Delta \bar{x} \quad (77)$$

This yields:

$$\beta(\bar{x}) = \beta(\bar{x}_0) \left[ 1 + (\bar{x} - \bar{x}_0) \left( \frac{f'_d(\bar{x}_0)}{f_d(\bar{x}_0)} - \frac{f'_a(\bar{x}_0)}{f_a(\bar{x}_0)} \right) \right] \quad (78)$$

The function value obtained with the simple method is multiplied with this scale factor. Clearly, this method has a major disadvantage in that the detailed analysis must be performed at every iteration step, and that derivatives of the detailed analysis are required. A simpler approach just keeps the scale factor constant at its value at the beginning of each iteration. Still, at each iteration the exact analysis has to be performed. Furthermore, move-limits have to be imposed in order to preserve the accuracy of the approximations, which makes convergence rather slow.

The method of [ref.5] uses gradient information of the previous design point to construct approximations to the objective function and constraints at the current design point. These approximations are either a linear or an inverse Taylor expansion or something in between, depending on the gradient information gathered at the previous design point. The optimizer then calculates a new optimal design point using the approximate analysis only, after which the detailed analysis is used again to calculate derivatives. Then a new approximation is generated and the process is repeated. In [ref.4] these derivatives are calculated independently and concurrently by means of Sobieszczanski-Sobieski's Global Sensitivity Equations. The process is depicted in Figure 8.



**Figure 8** Optimization of local approximate analysis

An alternative way of approximate analysis is to represent the relation between a certain parameter and its variables by a single function. Some results obtained by this method are described in [ref.8]. In this reference, second-order polynomials were used as approximating functions. To obtain such polynomials, it is necessary to perform a number of design evaluations and create a hypersurface fit through the obtained points, using the method of least squares.



A second-order polynomial hypersurface can be described by the following expression:

$$y = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i=1}^n \sum_{j=i}^n b_{ij} x_i x_j \quad (79)$$

in which  $y$  is the function to be approximated (the objective function, the constraints or a pseudo objective function to be used in an unconstrained problem). The factors  $b_i$  are the polynomial coefficients while the  $x$ 's are the  $n$  design variables. If a number of samples ( $k$ ) is taken across the design space the residual can be expressed as follows:

$$b_0 + \sum_{i=1}^n b_i x_i^{(k)} + \sum_{i=1}^n \sum_{j=i}^n b_{ij} x_i^{(k)} x_j^{(k)} - y^{(k)} = r^{(k)} \quad (80)$$

Written in a slightly different shape:

$$a_{k0} \cdot b_0 + a_{k1} b_1 + \dots + a_{kn} b_n + a_{k(n+1)} b_{n+1} + \dots + a_{k\frac{1}{2}n(n+3)} b_{\frac{1}{2}n(n+3)} - y_k = r_k \quad (81)$$

with:

$$\begin{aligned} a_{k0} &= 1 \\ a_{k1} \dots a_{kn} &= x_1^{(k)} \dots x_n^{(k)} \\ a_{k(n+1)} \dots a_{k\frac{1}{2}n(n+3)} &= x_1^{(k)2} \dots x_1^{(k)} x_j^{(k)} \dots x_n^{(k)2} \end{aligned}$$

Thus, the desired relation can be written as:

$$A \cdot \bar{b} = \bar{y} \quad (82)$$

with:

$$\begin{aligned} A = & \begin{matrix} 1 & x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} & x_1^{(1)2} & x_1^{(1)} x_2^{(1)} & \dots & x_n^{(1)2} \\ 1 & x_1^{(2)} & x_2^{(2)} & \dots & x_n^{(2)} & x_1^{(2)2} & x_1^{(2)} x_2^{(2)} & \dots & x_n^{(2)2} \\ \text{etc.} \end{matrix} \end{aligned}$$

The least squares solution with respect to the residuals can be obtained by solving the following equation:

$$A^T A \bar{b} = A^T \bar{y} \quad (83)$$

In case a second-order hypersurface approximation is desired for the objective function as well as for the constraints, the above equation has to be solved several times, each time with the same matrix  $A$ , but with a different right-hand-side vector  $y$ .

As indicated above, the total number of coefficients that describe a second-order polynomial hypersurface fit is equal to  $(n+1)(n+2)/2$ , in which  $n$  is the number of design variables. When using the brute force method to obtain derivatives, the number of designs generated will be  $i(n+1)$ , where  $i$  is the number of iteration steps needed to attain the optimal solution. Therefore, when the number of iteration steps exceeds  $(n+2)/2$ , the use of second-order polynomials will require less function evaluations than the use of the Sequential Quadratic Programming routine with the exact methods. In this simple analysis two effects are not included; on the one hand the number of function evaluations that the quadratic programming routine performs during the line search is not accounted for, and this number can become quite large. On the other hand, the number of function evaluations needed for an accurate surface fit actually must be greater than  $(n+1)(n+2)/2$ . It is clear, however, that this method could have very big advantages.

The approximations can be generated independently and concurrently prior to the optimization itself. Furthermore, they can be provided by the department specializing in the discipline under consideration, thus benefiting from experience. This also implies that analyses may be included that are carried out by separate programs or at remote locations or that are too time-consuming to consider using them in an optimization cycle. In this sense a kind of decomposition is inherent in this method of global approximation. Since the approximations are global they can be used throughout the optimization process, without the need to be updated after every iteration step. But most importantly, gradients can be calculated analytically. This means, that at each iteration step the optimization routine can perform function evaluations and gradient calculations in a very simple and fast way. This method would therefore be very suitable for use in quadratic programming routines. The process is depicted in Figure 9.

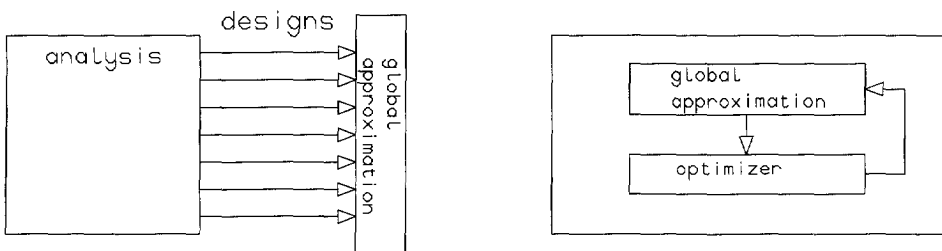


Figure 9 Optimization of global approximate analysis

There are some disadvantages, however. First of all, it might not be possible to obtain an accurate second-order approximation of a highly nonlinear function. Furthermore, a global second-order approximation (which is the essence of this method), might globally be quite accurate, but might produce large errors locally. For a gradient dependent optimizer, this is a drawback. In comparison, a quadratic optimization routine using the exact analysis, creates a local second-order approximation.

To validate the method, it was tried to reproduce the optimal aircraft design that was generated in chapter 3. This optimum was found using 13 iteration steps of Sequential Linear Programming followed by 13 iterations of Sequential Quadratic Programming. Thus, during the SQP optimization at least  $13 \cdot (7+1) = 104$  complete designs were calculated. Using second-order polynomials would have required  $(7+1)(7+2)/2 = 36$  design evaluations at least. When using the 13 SLP designs and the initial design, this minimal required number reduces to 22, that is, almost 5 times less than the 104 designs generated during the SQP optimization.

#### 4.4 Comparing the Results of Approximate and Exact Analysis

In order to obtain accurate approximations, at first 70 designs were generated. These designs were obtained by choosing the seven design variables at random in the design space defined by the variable boundaries:

$$\begin{aligned}
 10 &\leq S \leq 50 \\
 5 &\leq A \leq 25 \\
 0.1 &\leq \Lambda_{le} \leq 0.8 \\
 0 &\leq i_w \leq 0.1 \\
 0.5 &\leq M_{cr} \leq 0.9 \\
 0.5 &\leq \sigma_T \leq 2 \\
 0 &\leq \alpha_{cr} \leq 0.1
 \end{aligned} \tag{84}$$

The objective function depends on only four variables;  $S$ ,  $A$ ,  $\Lambda_{le}$  and  $\sigma_T$ , and therefore its polynomial has 15 coefficients. Both cruise constraints use all seven variables (36 coefficients) and the other constraints use five variables (without  $M_{cr}$  and  $\alpha_{cr}$ ), and therefore have 21 coefficients. The span constraint was not approximated as it depends directly on  $A$  and  $S$ :  $b = \sqrt{AS}$ . The accuracy of the polynomials was tested by calculating their squared Multiple Correlation Coefficients,  $MCC^2$ . This coefficient, which should be as close to 1 as possible, is defined as follows:

$$MCC^2 = \frac{\sum_{i=1}^k (\hat{y}_k - \bar{y})^2}{\sum_{i=1}^k (y_k - \bar{y})^2} \tag{85}$$

in which  $\bar{y}$  is the mean of the exact function values,  $y_k$  is the  $k^{th}$  exact function value

and  $\hat{y}_k$  is the approximation to  $y_k$ . As the sum of the residuals of the approximated values equals zero, the mean approximate function value equals the mean function value. This implies that the squared Multiple Correlation Coefficient is equal to the ratio of the variances of the exact function values and their approximations. For a more complete discussion on the accuracy of approximating polynomials, the reader is referred to chapter 5.

The calculated values for  $MCC^2$  are presented in the following table:

approximated function	$MCC^2$
objective function	0.998
range constraint	0.995
cruise thrust constraint	0.991
0 m climb constraint	0.999
engine failure constraint	0.969
take-off constraint	0.961

**Table 4**  $MCC^2$  values of approximating polynomials

For the initial design, the function values were compared with those obtained by using the exact method.

design variables (angles in rad)							
S [m <sup>2</sup> ]	A	$\Delta_{ie}$	$i_w$	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$	
41.4	10.89	0.436	0.0349	0.60	1.00	0.000	
objective function and constraints							
objective function [kg]	wing span [m]	range	cruise thrust	zero altitude climb	one-engine-out ceiling	take-off field length	
17527	21.23	0.0377	0.0312	2.4155	0.2888	0.2546	exact
17627	21.23	0.0398	-0.0166	2.4658	0.5264	0.2553	approximate

**Table 5** Comparison of exact and approximate analysis for initial design

Using the polynomials in a Sequential Quadratic Programming routine gives the following optimum after 28 iterations. The exact values for this design are given in

the last row:

design variables (angles in rad)							
S [m <sup>2</sup> ]	A	$\Lambda_{je}$	$i_w$	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$	
26.4	15.36	0.100	0.0645	0.55	0.70	0.070	
objective function and constraints							
objective function [kg]	wing span [m]	range	cruise thrust	zero altitude climb	one-engine-out ceiling	take-off field length	
12993	20.12	0.0000	0.0337	0.0000	0.7690	0.113	approximate
13007	20.12	-0.0213	-0.0047	0.0411	-0.0240	0.146	exact

Table 6 Comparison of exact and approximate analysis for optimum after 28 iterations

The optimum found by using quadratic programming and the exact analysis was (chapter 3):

design variables (angles in rad)						
S [m <sup>2</sup> ]	A	$\Lambda_{le}$	$i_w$	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$
30.3	21.33	0.100	0.0712	0.55	0.71	0.0521
objective function and constraints						
objective function [kg]	wing span [m]	range	cruise thrust	zero altitude climb	one-engine-out ceiling	take-off field length
13683	25.44	-0.0006	-0.0003	-0.0005	1.0871	0.192

Table 7 Best design as obtained by SQP optimization on exact analysis

As, apart from the wing area and wing aspect ratio, the variables are very close to those at the exact optimum, it was decided to generate another set of polynomials in a

reduced design space, according to:

$$\begin{aligned}
 20 &\leq S \leq 35 \\
 10 &\leq A \leq 25 \\
 0.1 &\leq \Lambda_{le} \leq 0.2 \\
 0.06 &\leq i_w \leq 0.07 \\
 0.5 &\leq M_{cr} \leq 0.6 \\
 0.6 &\leq \sigma_T \leq 0.8 \\
 0.04 &\leq \alpha_{cr} \leq 0.08
 \end{aligned} \tag{86}$$

Within this design space 100 new designs were generated.

The  $MCC^2$  values for these polynomials are:

approximated function	$MCC^2$
objective function	0.999
range constraint	0.999
cruise thrust constraint	0.998
0 m climb constraint	0.999
engine failure constraint	0.998
take-off constraint	0.994

Table 8  $MCC^2$  values of second set of approximating polynomials

Apparently, these polynomials are very accurate.

Using the last obtained optimum as a starting value, with these new polynomials a new optimum is obtained after 12 more iterations.

design variables (angles in rad)						
S [m <sup>2</sup> ]	A	$\Lambda_{le}$	$i_w$	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$
28.7	15.10	0.100	0.0706	0.56	0.72	0.0498
objective function and constraints						
objective function [kg]	wing span [m]	range	cruise thrust	zero altitude climb	one-engine-out ceiling	take-off field length
13543	20.81	0.0000	0.0000	0.0000	0.201	0.182
13534	20.81	-0.0013	0.0005	0.0183	0.198	0.182

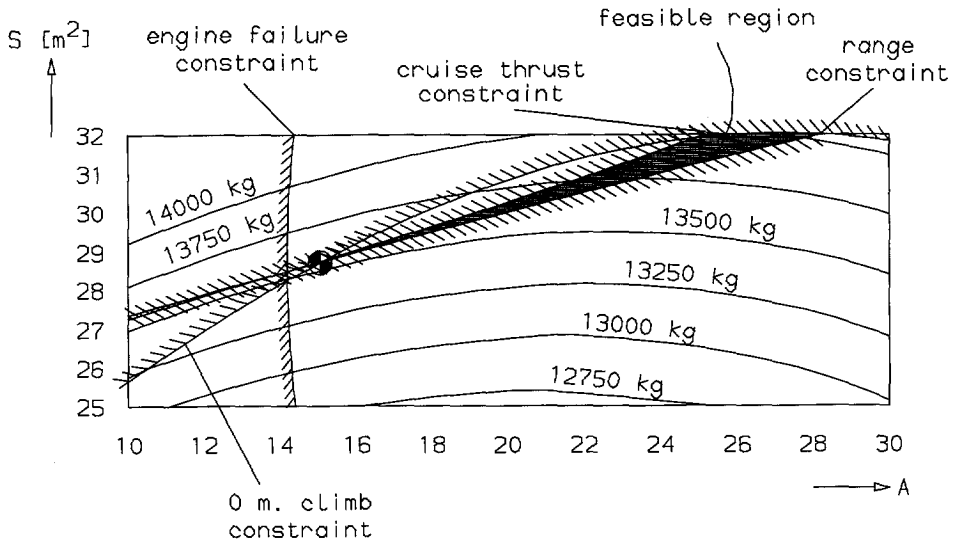
Table 9 Final optimal design

The approximate analysis of the initial design shows a violation of the cruise thrust constraint (value=-0.0166). To the approximate analysis, the initial design is therefore infeasible.

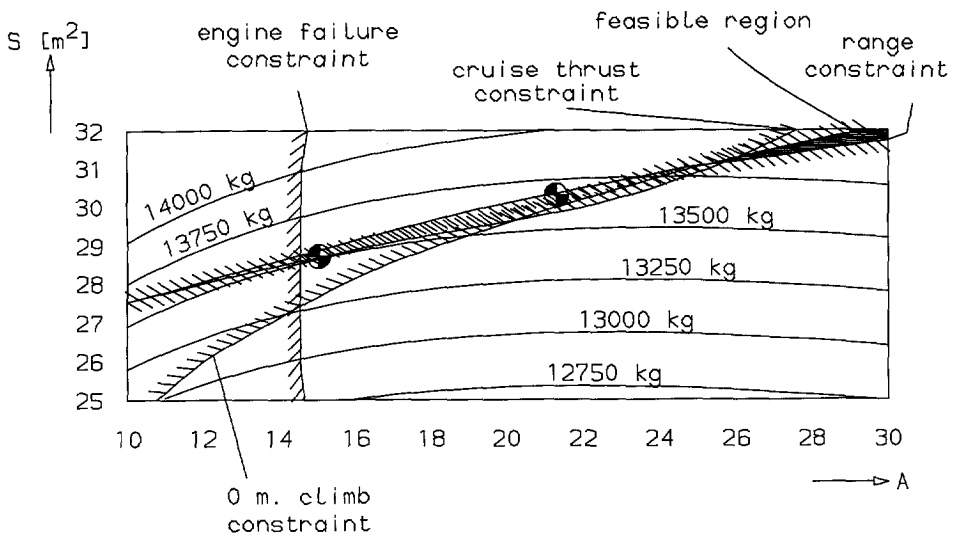
On comparing the results with the exact optimum, it is clear that, apart from the wing area and aspect ratio, the design variables tally very well. In order to investigate the reason for the difference in wing area and aspect ratio, the objective function and constraints for both methods were drawn in the S,A-hyperplane, Figure 10 and Figure 11. The other variables were fixed on their values at the corresponding optima. Therefore, it should be kept in mind that both figures represent different hyperplanes, although they only differ very slightly.

Apparently, only four constraints are active in the examined part of the design space; range, cruise thrust, zero altitude climb and one-engine-out ceiling. In the hyperplane of Figure 10, these constraints shut off the design space, apart from a small strip bordered by the three constraint boundaries active at the optimum.

The same figure is drawn for the exact analysis (Figure 11). For that purpose, all possible integer combinations of S and A were analyzed. Figure 11 shows, that the design space is shut off almost completely, but for a very small strip. However, in the optimization program, a certain tolerance is assigned to the constraints, or, in other words, a finite thickness is assigned to the constraint boundary curves. The purpose of this is to be able to "catch" an active constraint more easily and keep it active. A consequence is, that very small constraint violations are possible. In this example, the tolerance was set to  $1.10^{-4}$ . Apparently, the location of the optimum is such, that the three active constraints are violated as much as their tolerances permit.



**Figure 10** S,A hyperplane (approximate analysis)



**Figure 11** S,A hyperplane (exact analysis)

At first glance the two figures show very good agreement. Especially the isometric curves for the objective function tally very well. However, a closer examination shows



a few discrepancies between the constraint boundaries:

1. the curve for the one-engine-out ceiling constraint is located further to the right in Figure 11.
2. the line for the range constraint is located higher in Figure 11.
3. the curve for the zero altitude climb constraint lies considerably lower in Figure 11.

The latter apparently has the largest impact on the location of the optimum. As the constraint boundaries for the three active constraints run almost parallel, small changes in their relative positions have large consequences for the location of their intersection point. On the other hand, as the isometric curves for the objective function locally are orientated approximately in the same direction as the active constraint boundaries, the difference between the optimal objective function values is rather small: 13534 kg for the approximate optimum and 13682 kg for the exact optimum.

The question that arises is, what causes the difference between the two figures. There are two possibilities:

1. The constraint boundaries are highly nonlinear and just cannot be approximated accurately using second-order polynomials.
2. The slight differences in the five fixed variables cause the constraint boundaries to shift relative to each other.

The first explanation is hardly likely. As the values of the  $MCC^2$  show, it is expected that the second-order polynomials are very accurate. In order to test this anyhow, the constraint boundaries in the hyperplane of the exact optimum were approximated by the polynomials. As can be seen from Figure 12, in which the most critical constraint is pictured, the agreement is very good. Therefore, the second explanation is the most likely one. This is supported by the derivatives of the two most critical constraints at the optimum found by the approximate method. The scaled values are obtained by multiplying the derivatives by the value of the variable with respect to which they were differentiated.

---

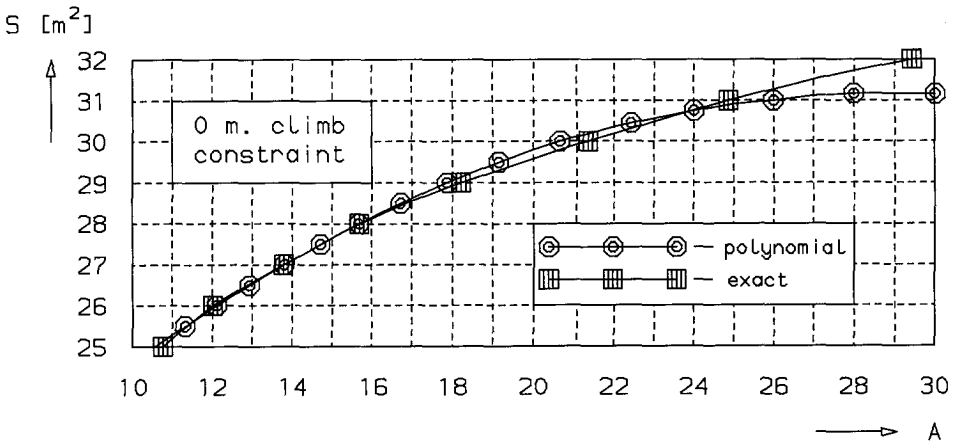


Figure 12 Sea level climb constraint

	finite-differencing	polynomial	scaled
S	0.00842	0.00865	0.24826
A	-0.00209	-0.00230	-0.03473
$\Lambda_{le}$	-0.00321	0.00004	0
$i_w$	0.12380	0.12907	0.00911
$M_{cr}$	0.10810	0.09456	0.05258
$\sigma_T$	-0.02431	-0.02464	0.01774
$\alpha_{cr}$	0.14787	0.13036	0.00649

Table 10 Derivatives of range constraint in approximate optimum

As expected,  $S$  has a large influence. The cruise Mach number, however, has an influence of the same order as the influence of  $A$ . The derivatives of the zero altitude climb constraint are:

	polynomial	scaled
S	-0.31390	-9.00893
A	0.16367	2.47142
$\Lambda_{le}$	-0.90688	-0.09069
$i_w$	-1.91599	-0.13527
$\sigma_T$	29.3305	21.1180

**Table 11** Derivatives of zero altitude climb constraint in approximate optimum

Apparently, the difference between the two optimal values for the engine thrust scale factor causes the large shift of this constraint. The fact that some derivatives completely differ from those obtained by finite-differencing occurs at many points and for every function but the objective function. In most cases it is the derivative with respect to the sweep angle that differs, although sometimes others too show differences. There are two possible explanations for this. The first one -as already explained earlier- is that a global approximation might locally render false results. The other possibility is that the results obtained by finite-differencing (a rather inaccurate method) are wrong. This point was not further investigated, as the global approximations turn out to be very well and sufficient to guide the design towards its optimum.

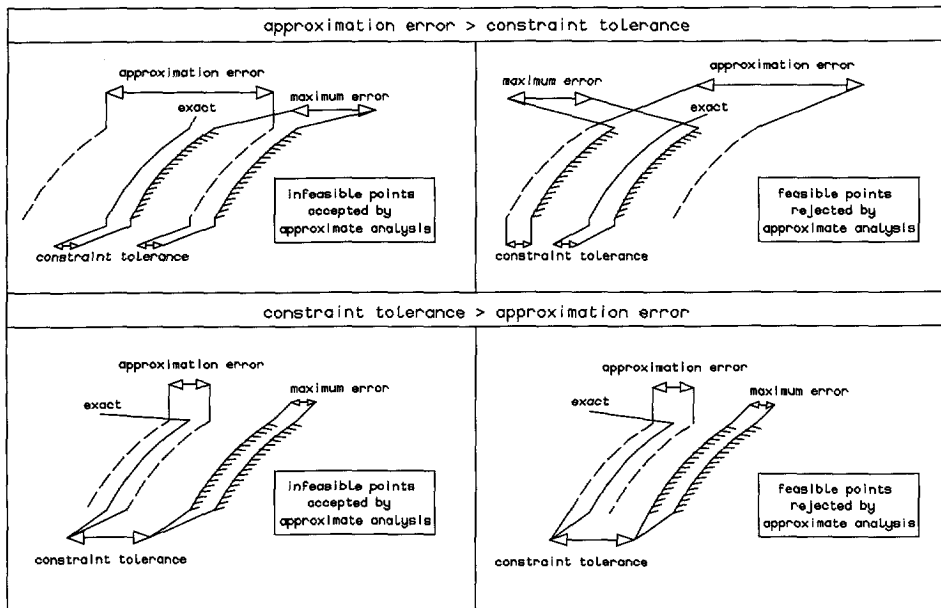
## 4.5 Conclusions

The present approximate analysis finds an optimum at the vertex of three constraints. In reality, however, this design has a violated range constraint of  $-1.3 \cdot 10^{-3}$ , clearly outside the specified constraint tolerance of the order  $1 \cdot 10^{-4}$ . This is the reason that this design was not chosen by the exact analysis, although its objective function value is lower than that of the exact optimum. The exact analysis locates the optimal design in a hyperspace that leaves no feasible region with respect to S and A, but nonetheless is within the specified tolerances. The approximation of this point, however, shows a range constraint with value  $-3.38 \cdot 10^{-2}$ , which is clearly infeasible. Apparently, two problems have occurred; firstly, due to approximation errors, a perfectly feasible design was rejected and secondly a design was chosen that in reality is infeasible. The optimizer moves between three almost parallel constraint boundaries until the maximum allowable violation is attained (the objective function still decreases). Both the approximation errors and the parallel orientation of the constraint boundaries cause the large differences between both methods for the optimal values for S and A.

The reason that these problems occur, can be explained with Figure 13. In both the exact optimization and the approximate optimization constraint tolerances of  $1 \cdot 10^{-4}$  are used. It is very likely that the accuracy of the approximations is not that high. Figure 13 shows that in this case both problems (rejection of feasible designs and

acceptation of infeasible ones) can occur. Making the tolerances larger than the estimated approximation error does not change this. The only solution is, to accept a constraint tolerance that is slightly in excess of the estimated approximation error. The approximate optimization should then be performed with a constraint tolerance equal to the difference between the exact tolerance and the estimated approximation error:

$\hat{\varepsilon}_t \leq (\varepsilon_t - \varepsilon_a)$ , in which the subscript t refers to the constraint tolerance and the a to the approximation error. Therefore, the maximum attainable constraint tolerance  $\hat{\varepsilon}_t$  dictates the order of possible exact constraint violations. Incidentally, the problem of possible rejection of feasible designs is not solved in this way, but the produced optimal designs will always turn out to be feasible ones, that is, within the tolerance  $\varepsilon_t$  ( $> \varepsilon_a$ ).



**Figure 13** Approximation error and constraint tolerance

Throughout this chapter, the methods used were referred to as either approximate or exact. Of course, the "exact" methods of analysis are themselves approximations to the physical reality. Therefore, problems like the ones described in this chapter, with active constraints that turn out to be violated after closer examination, are in any case bound to happen. In any constrained optimization problem, there will be at least one constraint active at the optimum. Due to the limited accuracy of the method used, there is great danger that this constraint will turn out to be violated, thus rendering

the whole design useless. To avoid this risk, that is apparently inherent in constrained optimization, it is again essential not to require constraint tolerances of a lower order than the expected accuracy of the methods used, and on the other hand, to use methods of analysis that are as accurate as possible.

It can therefore be concluded that the present method of approximation is very useful to simplify optimization problems that use complex methods of analysis. However, it must be realized, that the final amount of constraint violation might be of the same order as the accuracy of the approximation.

A very recent method of approximate analysis is the use of neural networks. In a neural network, the processes that are carried out inside the human brain are simulated. A neural network consists of a large amount of interconnected basic computing elements, the so-called neurons. A neuron transforms multiple inputs to a single output by means of a so-called activation function, which transforms a weighted sum of the inputs to a finite output, provided that the weighted sum of the inputs lies in a certain interval. This procedure imitates a biological neuron which only produces an output signal if the strengths of the incoming signals build up to a certain level.

The output of a neuron can act as the input for another neuron and in this way a network can be built that converts a number of input signals (design variables) to a number of output signals (objective function and constraints).

Initially, the weight factors on the input signals of each neuron are chosen at random. Subsequently, these values will be adapted in such a way that the network will predict the outputs belonging to a large set of different combinations of input variables with sufficient accuracy. This process is referred to as the learning process. Without getting into detail, it is just being mentioned that the way in which this learning process is generally carried out, is known as the back-propagation approach, which consists of comparing the actual output value produced by the network, to the known exact output value and correcting the weight factors of the different neurons backward-to-forward accordingly. This procedure is repeated for every sample in the learning data set. Since the output of a certain network is also determined by the different triggering intervals assigned to the neurons, these intervals might be adapted by a similar learning process too.

Once the learning process is completed, a network is obtained that features as an abstraction of the exact design analysis. This network can directly be coupled to any optimization algorithm since it is capable of performing design evaluations in a fast and efficient way. Just like the method of **response surfaces** employed in the present work, a neural network represents a global approximation to the design space. The exploration of the design space is carried out in advance and independent of the optimization process. The large advantage of neural networks over response surfaces is, that the order of the approximating analysis is not limited (the presently adopted method of response surfaces represents a second-order approximation to the design space, whereas a neural network can handle highly nonlinear relations with equal

---

accuracy). The disadvantage of the use of neural networks is of course the learning phase, which can be quite tedious and which might not produce the required accuracy. Furthermore, it is not known what it is that the network has learned. It is very sensitive for inadvertently present patterns or bias in the learning data set.

Another problem is, that there is total freedom in the choice of the number of neurons and the way in which they are interconnected, which leaves the user almost without a clue how to start. For computational efficiency, the number of neurons and connections should be minimized. However, the required accuracy of the network, together with the nature of the problem, might impose a minimum level of complication on the network. Therefore, a neural network might be subject of an optimization cycle itself, for instance with the use of a genetic algorithm (chapter 6), in which the individuals are different neural networks (with different performance), their genotype being a coded representation of the connections within the network. The determination of the phenotype of each individual network will consist of the learning phase (in which the individuals will develop themselves) followed by the assessment phase in which the performance of each network is quantified by testing it for accuracy and efficiency with a second set of data. The objective function might be a weighted sum of the average accuracy and the computational efficiency.

The use of neural networks in optimization is a rather new technique, but one which shows great promise. In the present work, it was not further investigated. However, since one of the conclusions of this dissertation is the importance of fast and simple, but accurate, substitutes for the complicated and tedious calculations inherently connected to multidisciplinary design and optimization, it might very well be that the neural network literally turns out to be the missing link between very sophisticated analysis on the one hand and multidisciplinary optimization on the other. Reasoning on in this manner, one could even imagine a neural network not providing the values for the objective function and constraints as a function of the design variables, but the other way around, thus solving the inverse design problem.

Summarizing, the neural network might be regarded as a substitute to the human designer, with far superior memory and learning capabilities. Once presented with a large set of examples, it will discern patterns and gather experience and knowledge that will enable it to predict the outcome of new problems presented to it. The only disadvantage is, that its knowledge (or perhaps intelligence) will be rather limited.

Applications of neural networks at first appeared in the military, where neural networks were used to recognize and identify sonar signatures, helicopter noise or reconnaissance photographs. Later, civil applications appeared as well. An application of a neural network in combination with multivariate optimization may be found in [ref.7] and [ref.12].

---

## 4.6 References

1. Barthelemy, J.M. and Hall, L.E.; *Automatic Differentiation as a Tool in Engineering Design*; AIAA 92-4743 CP
  2. Bischof, C. and Griewank, A.; *ADIFOR: a FORTRAN System for Portable Automatic Differentiation*; AIAA 92-4744 CP
  3. Chang, K.J.; Haftka, R.T.; Giles, J.L. and Kao, P.J.; *Sensitivity-Based Scaling for Approximating Structural Response*; Published as AIAA 91-0925
  4. Dovi, A.R.; Wrenn, G.A.; Barthelemy, J.-F.M.; Coen, P.G. and Hall, L.E.; *Multidisciplinary Design Integration Methodology for a Supersonic Transport Aircraft*; *Journal of Aircraft*, Vol.32, No.2, March-April 1995
  5. Fadel, G.M.; Riley, M.F. and Barthelemy, J.M.; *Two Point Exponential Approximation Method for Structural Optimization*; *Structural Optimization* 2, 117-124, Springer Verlag 1990
  6. Haftka, R.T.; *Combining Global and Local Approximations*; *AIAA journal*, Vol.29, No. 9, September 1991
  7. Hajela, P. and Berke, L.; *Neurobiological Computational Models in Structural Analysis and Design*; AIAA 90-1133-CP, April 1990
  8. Healy, M.J.; Kowalik, J.S. and Ramsay, J.W.; *Airplane Engine Selection by Optimization on Surface Fit Approximations*; *Journal of Aircraft* Vol.12, No.7, July 1975.
  9. Hutchison, M.; Unger, E.; Mason, W.; Grossman, B. and Haftka, R.; *Variable-Complexity Aerodynamic Optimization of an HSCT Wing Using Structural Wing-Weight Equations*; AIAA 92-0212
  10. Sobieszczanski-Sobieski, J.; *Optimization by Decomposition: A Step from Hierarchic to Non-Hierarchic Systems*; NASA CP-3031, part 1
  11. Sobieszczanski-Sobieski, J.; *Aircraft Optimization by a System Approach: Achievements and Trends*; ICAS 92-1.3.4.
  12. Swift, R.A. and Batill, S.M.; *Application of Neural Networks to Preliminary Structural Design*; AIAA 91-1038-CP, April 1991
-





# 5. The Applicability of Regression Surfaces for Approximating Purposes

## 5.1 Summary

The present chapter summarizes, by means of simple statistics, a few tests to determine whether a response surface obtained by a least-squares fit of a number of sample values is suitable for predicting the value of a certain parameter of interest as a function of its variables. These tests will be used to generate acceptable approximating functions which will be used as a substitute for the complicated and tedious exact calculations in a multivariate optimization study of a High-Speed Civil Transport plane (HSCT). Some previous experience with the use of approximating functions in optimization was reported in [ref.3], [ref.4], [ref.5] and [ref.6].

## 5.2 Notation

$a$	coefficient
$b$	approximate coefficient
$[b]$	vector of approximate coefficients
$E$	expectancy (mean)
$F(v_1, v_2)$	Snedecor-distribution with $v_1$ degrees of freedom of the numerator and $v_2$ degrees of freedom of the denominator
$H_0$	zero hypothesis
$i$	index
$j$	index
$k$	number of samples
$N(\mu, \sigma^2)$	Normal distribution with expectancy $\mu$ and variance $\sigma^2$
$n$	number of variables
$p$	number of polynomial coefficients
$R^2$	ratio of approximate response variance and true response variance ( $MCC^2$ )
$r$	residual
$S$	sum
$s^2$	mean square about regression (mean square of residuals)
$t(v)$	Student-distribution with $v$ degrees of freedom
$\underline{u}$	statistic
$V$	variance
$\underline{v}$	statistic
$[X]$	matrix
$x$	variable
$\bar{x}$	mean value of $x$
$y$	function response value
$\hat{y}$	function approximate value
$\bar{y}$	mean value of $y$
$[y]$	vector of function response values

---

$\beta$	coefficient
$\varepsilon$	residual (error)
$\mu$	mean (expectancy)
$\sigma^2$	variance
$\chi^2(v)$	chi-square distribution with $v$ degrees of freedom

### abbreviations

HSCT	High Speed Civil Transport
MCC <sup>2</sup>	Multiple Correlation Coefficient Squared ( $R^2$ )
R	Regression
SS	sum of squares
SS(R  $b_0$ )	sum of squares of regression after allowance has been made for $b_0$

## 5.3 Response Surfaces

In order to obtain the previously discussed response surfaces, it is necessary to generate a large amount of designs, thus providing a population of sample points on which the approximating function will be based. Such a relation will be called a regression line or curve (or, in more dimensions, a regression surface). For simplicity, first the case of a linear fit depending on one variable is considered. The following is largely based on [ref.2], whereas further information on the subject of regression equations can be found in [ref.1] and [ref.7]. For a more comprehensive and detailed treatment of various aspects of statistics and the use of regression equations in science and engineering, the works of Dr. G.E.P. Box are suggested. The subjects treated hereafter only require a basic knowledge of statistics as can readily be obtained after consulting any elementary textbook on statistics.

### 5.3.1 Linear regression in one variable

In this case, the true relation between the response  $y$  and its variable  $x$  is assumed to be:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (87)$$

and the approximating relation will be:

$$\hat{y} = b_0 + b_1 x \quad (88)$$

Using the least-squares method, the approximate coefficients,  $b_0$  and  $b_1$  will be determined such that the sum of squares of the approximating errors,  $S = \sum_{i=1}^k \varepsilon_i^2$ , is

minimized ( $k$  is the number of samples). This yields:

$$S = \sum_{i=1}^k (y_i - \beta_0 - \beta_1 x_i)^2 \quad (89)$$

Differentiating and replacing the coefficients by their estimates  $b_0$  and  $b_1$ , yields the following equations:

$$\sum_{i=1}^k (y_i - b_0 - b_1 x_i) = 0 \quad (90)$$

$$\sum_{i=1}^k x_i (y_i - b_0 - b_1 x_i) = 0 \quad (91)$$

After solving these equations and some rewriting, using  $\bar{x} = \frac{1}{k} \sum_{i=1}^k x_i$  and  $\bar{y} = \frac{1}{k} \sum_{i=1}^k y_i$ , it follows:

$$b_1 = \frac{\sum_{i=1}^k (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^k (x_i - \bar{x})^2} = \frac{\sum_{i=1}^k (x_i - \bar{x})y_i}{\sum_{i=1}^k (x_i - \bar{x})^2} \quad (92)$$

$$b_0 = \bar{y} - b_1 \bar{x} \quad (93)$$

which implies:

$$\hat{y}_i = \bar{y} + b_1 (x_i - \bar{x}) \quad (94)$$

Now writing for the difference between the approximate and the true value

$$y_i - \hat{y}_i = y_i - \bar{y} - (b_1 (x_i - \bar{x})) \quad (95)$$

and squaring both sides yields:

$$\sum_{i=1}^k (y_i - \hat{y}_i)^2 = \sum_{i=1}^k (y_i - \bar{y})^2 + \sum_{i=1}^k (\hat{y}_i - \bar{y})^2 - 2 \sum_{i=1}^k (y_i - \bar{y})(\hat{y}_i - \bar{y}) \quad (96)$$

The last term can be rewritten with the previously derived expression for  $b_1$  as:

$$-2 \sum_{i=1}^k (y_i - \bar{y})(\hat{y}_i - \bar{y}) = -2 \sum_{i=1}^k (y_i - \bar{y}) b_1 (x_i - \bar{x}) = -2 \sum_{i=1}^k (\hat{y}_i - \bar{y})^2 \quad (97)$$

Then it follows:

$$\sum_{i=1}^k (y_i - \bar{y})^2 = \sum_{i=1}^k (y_i - \hat{y}_i)^2 + \sum_{i=1}^k (\hat{y}_i - \bar{y})^2 \quad (98)$$

Or in words:

$$\begin{array}{ccccc} \text{sum of squares} & = & \text{sum of squares} & + & \text{sum of squares} \\ \text{about mean} & & \text{about regression} & & \text{due to regression} \end{array}$$

If the regression line constitutes a usable approximation, the sum of squares due to the regression should be as close as possible to the sum of squares about the mean, or in other words, the ratio:

$$R^2 = \frac{\sum_{i=1}^k (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^k (y_i - \bar{y})^2} \quad (99)$$

should be as close to one as possible. Since the expectancy or mean of the true responses and the approximate responses are the same (this is a direct result of the least-squares method), the ratio  $R^2$  is the ratio of the variance of the approximate responses to the variance of the true responses. Therefore,  $R^2$  (also known as  $MCC^2$  or Multiple Correlation Coefficient Squared) indicates the percentage of the true response variance which has been "explained" by the regression line.

In case the following relation is fit:  $y_i = \beta_0 + \varepsilon_i \Leftrightarrow \hat{y}_i = \bar{y}$  or  $b_1$  equals zero, then the Multiple Correlation Coefficient equals zero too. Therefore,  $R^2$  can also be interpreted as a measure of the usefulness of terms other than the constant. Because of this, the sum of squares due to regression is often divided in a part as a result of  $b_0$  ( $SS(b_0)$ ) and a part as a result of the other terms "after allowance has been made for  $b_0$ "

( $SS(R|b_0)$ , [ref.2], the capital R is short for Regression).

The following scheme then applies [ref.2]:

sum of squares		degrees of freedom	mean square
$SS(b_0)$	$k\bar{y}^2$	1	
$SS(R b_0)$	$\sum_{i=1}^k (\hat{y}_i - \bar{y})^2$	1	
residual SS	$\sum_{i=1}^k (y_i - \hat{y}_i)^2$	k-2	$s^2=SS/(k-2)$
total	$\sum_{i=1}^k y_i^2$	k	

Table 12    Sums of squares: linear regression in one variable

It should be noticed, that using the Multiple Correlation Coefficient as the only means of checking the adequacy of the approximating function is dangerous, since this factor can always be made one by implementing sufficient terms in the approximating function, or otherwise by adapting the number of samples used. In case the number of coefficients to be estimated equals the number of samples on which the least-squares fit is based, the resulting fit will be exact. It is, however, by no means certain that the relation thus obtained will constitute a usable approximation. In other words, it is necessary to investigate whether the obtained relation is statistically significant. This is done, by testing the hypothesis that the obtained coefficients should in reality all be zero (except for the constant).

In order to be able to introduce statistics to the regression analysis, it is necessary to make an assumption about the distribution of the errors or residuals. It is assumed that these errors are Normally distributed:  $\epsilon_i \sim N(0, \sigma^2)$  (this assumption will be tested later on).

$b_1$  is an unbiased estimator of the linear coefficient since  $E(b_1)=\beta_1$ . This follows easily from the fact that  $E(\epsilon_i)=0$ .

For the variance of  $b_1$  the following relation can be derived:

$$b_1 = \frac{\sum_{i=1}^k (x_i - \bar{x}) y_i}{\sum_{i=1}^k (x_i - \bar{x})^2} = a_1 y_1 + a_2 y_2 + \dots + a_k y_k \quad (100)$$

$$\begin{aligned} V(b_1) &= a_1^2 V(y_1) + a_2^2 V(y_2) + \dots + a_k^2 V(y_k) \\ &= (a_1^2 + a_2^2 + \dots + a_k^2) \sigma^2 = \frac{\sigma^2}{\sum_{i=1}^k (x_i - \bar{x})^2} \end{aligned} \quad (101)$$

(since the  $y_i$ 's are independent).

From the assumption of the Normal distribution of the errors, it follows:

$$b_1 \sim N \left( \beta_1, \frac{\sigma^2}{\sum_{i=1}^k (x_i - \bar{x})^2} \right) \quad (102)$$

A test for  $b_1$  is then:

$$\underline{u} = \frac{(b_1 - \beta_1) \sqrt{\sum_{i=1}^k (x_i - \bar{x})^2}}{\sigma} \sim N(0,1) \quad (103)$$

The problem with this test is, that the variance of the true response,  $\sigma^2$ , is unknown. However, the mean square of the residuals (mean square about regression), which is defined as

$$s^2 = \frac{\sum_{i=1}^k (y_i - \hat{y}_i)^2}{k-2} \quad (104)$$

is an unbiased estimator for the variance of the true response:

$$E(s^2) = \sigma^2 \quad (105)$$

This can be proven as follows:

$$E(s^2) = \frac{1}{k-2} \sum_{i=1}^k E[(y_i - \bar{y})^2 - (\hat{y}_i - \bar{y})^2] = \frac{1}{k-2} \left[ k\sigma^2 - \sum_{i=1}^k V(\hat{y}_i) \right] \quad (106)$$

since  $(y_i - \bar{y})^2 - (\hat{y}_i - \bar{y})^2 = (y_i - \hat{y}_i)^2 - 2(\hat{y}_i - \bar{y})(\hat{y}_i - y_i)$ , where the expectancy of the last term is zero.

For the variance of a specific approximate value, the following relation can be deduced, using the variance of  $b_1$  as presented previously:

$$\hat{y}_i = \bar{y} + b_1(x_i - \bar{x}) \quad (107)$$

$$V(\hat{y}_i) = V(\bar{y}) + (x_i - \bar{x})^2 V(b_1) = \frac{\sigma^2}{k} + \frac{(x_i - \bar{x})^2 \sigma^2}{\sum_{i=1}^k (x_i - \bar{x})^2} \quad (108)$$

Substituting this in the expression for the expectancy of  $s^2$  yields the following relation:

$$E(s^2) = \frac{1}{k-2} (k\sigma^2 - 2\sigma^2) = \sigma^2 \quad (109)$$

which was to be proven.

A test for the variance of the true response is then:

$$\underline{v} = \frac{(k-2)s^2}{\sigma^2} \sim \chi^2(k-2) \quad (110)$$

Using these expressions, the hypothesis can be tested that the linear coefficient, as

determined by the least-squares fit, should in reality be zero. A pivoting function to test the hypothesis  $H_0: \beta_1=0$  is:

$$\frac{\frac{\underline{u}}{\sqrt{\frac{\underline{v}}{(k-2)}}}}{\frac{\sigma}{s}} = \frac{(b_1 - \beta_1) \sqrt{\sum_{i=1}^k (x_i - \bar{x})^2}}{\sigma} \sim t(k-2) \quad (111)$$

or after substituting  $\beta_1=0$  and squaring:

$$\frac{b_1^2 \sum_{i=1}^k (x_i - \bar{x})^2}{s^2} \sim t^2(k-2) = F(1, k-2) \quad (112)$$

Using a previously introduced notation, the statistical function  $\underline{u}$  can be written as follows:

$$\underline{u}^2 = \frac{b_1^2 \sum_{i=1}^k (x_i - \bar{x})^2}{\sigma^2} = \frac{\sum_{i=1}^k (\hat{y}_i - \bar{y})^2}{\sigma^2} = \frac{SS(R|b_0)}{\sigma^2} \sim \chi^2(1) \quad (113)$$

with  $\beta_1=0$  (zero-hypothesis), yielding:

$$\frac{\underline{u}^2/1}{\underline{v}/(k-2)} \sim F(1, k-2) \quad (114)$$

If the calculated value for the obtained regression line exceeds the table value of an F-distribution for a certain error, for instance 5%, then the hypothesis that the terms other than the constant should in fact be zero cannot be accepted on statistical grounds, with a 5% chance of being wrong (that is, the first-order error is 5%). On the other hand, if the calculated value does not exceed the table value, the zero-hypothesis is accepted and a better regression line must be found. Nonetheless, there is a chance that this is incorrect. This so-called second-order error depends on the actual value of the F-test. However, the definition of the zero-hypothesis is such that the unjust rejection of it is considered the most serious, and the chance of this happening is less than 5%.



### 5.3.2 Matrix notation

In most applications, a regression surface of higher order than one will be needed, in order to take curvature into account as well. Also, like in the present study, the number of variables will exceed one, and a matrix approach to the least-squares problem will be more suitable. In the present study, the various relations to be approximated will be represented by second-order polynomials in  $n$  variables, therefore:

$$y = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i=1}^n \sum_{j=1}^n b_{ij} x_i x_j \quad (115)$$

If the number of samples taken is denoted by  $k$ , the residual can be expressed as follows:

$$b_0 + \sum_{i=1}^n b_i x_i^{(k)} + \sum_{i=1}^n \sum_{j=1}^n b_{ij} x_i^{(k)} x_j^{(k)} - y^{(k)} = r^{(k)} \quad (116)$$

Written in a slightly different shape:

$$a_{k0} \cdot b_0 + a_{k1} b_1 + \dots + a_{kn} b_n + a_{kn+1} b_{n+1} + \dots + a_{k\frac{1}{2}n(n+3)} b_{\frac{1}{2}n(n+3)} - y_k = r_k \quad (117)$$

with:

$$a_{k0} = 1$$

$$a_{k1} \dots a_{kn} = x_1^{(k)} \dots x_n^{(k)}$$

$$a_{k(n+1)} \dots a_{k\frac{1}{2}n(n+3)} = x_1^{(k)2} \dots x_1^{(k)} x_j^{(k)} \dots x_n^{(k)2}$$

Thus, the desired relation can be written as:

$$[y] = [X][b] \quad (118)$$

with:

$$[X] = \begin{matrix} 1 & x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} & x_1^{(1)2} & x_1^{(1)} x_2^{(1)} & \dots & x_n^{(1)2} \\ 1 & x_1^{(2)} & x_2^{(2)} & \dots & x_n^{(2)} & x_1^{(2)2} & x_1^{(2)} x_2^{(2)} & \dots & x_n^{(2)2} \\ \text{etc.} \end{matrix}$$

The least-squares solution with respect to the residuals can be obtained by solving the following equation:

$$[X]^T [X] [b] = [X]^T [y] \quad (119)$$

or:

$$[b] = ([X]^T [X])^{-1} [X]^T [y] \quad (120)$$

(this is a very elementary relation, the proof of which can be found in many basic textbooks). The matrix  $[X]$  is a  $[k \times p]$  matrix, the vector  $[y]$  has dimension  $k$ , whereas the vector  $[b]$  has dimension  $p$ , where  $p$  is the total number of coefficients (including  $b_0$ ) in the polynomial:

$$p = (n+1)(n+2)/2.$$

Like in the case of a line fit in one variable, the different sums of squares can be shown in a table:

sum of squares		degrees of freedom	mean square
$SS(b_0)$	$k\bar{y}^2$	1	
$SS(R   b_0)$	$[b]^T [X]^T [y] - k\bar{y}^2$	$p-1$	
residual SS	$[y]^T [y] - [b]^T [X]^T [y]$	$k-p$	$s^2 = SS/(k-p)$
total	$[y]^T [y]$	$k$	

Table 13 Sums of squares: matrix notation

The previously introduced Multiple Correlation Coefficient Squared,  $MCC^2$ , or  $R^2$ , in this case becomes:

$$R^2 = \frac{SS(R | b_0)}{[y]^T [y] - SS(b_0)} = \frac{\sum_{i=1}^k (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^k (y_i - \bar{y})^2} \quad (121)$$

In this case also, statistics for the hypothesis that all terms other than the constant should be zero can be defined:

$$\underline{u}^2 = \frac{SS(R|b_0)}{\sigma^2} \sim \chi^2(p-1) \quad (122)$$

$$\underline{v} = \frac{(k-p)s^2}{\sigma^2} \sim \chi^2(k-p) \quad (123)$$

$$\Rightarrow \frac{\underline{u}^2/(p-1)}{\underline{v}/(k-p)} \sim F(p-1, k-p) \quad (124)$$

$$\Rightarrow \frac{SS(R|b_0)/(p-1)}{s^2} \sim F(p-1, k-p) \quad (125)$$

If this value exceeds the table value for 5%, for example, the hypothesis is rejected. This does not mean that all terms in the regression equation are significant; it merely states that the proportion of the variation in the true responses, which has been accounted for by the equation is greater than would be expected by chance in 95% similar sets of data.

In [ref.2] a method is presented in which the different terms of the regression equation are added in a stepwise fashion starting with the most significant term. Partial F-tests are performed for the statistical significance of each term "after allowance has been made" for the previously added terms. Also, the previously added terms are tested to check whether they would still be significant if each of them would have been added *after* the lastly added term, etcetera. This procedure makes sure that only significant terms feature in the regression equation.

In the present study, this procedure was not applied. Instead, to all terms of the second-order polynomials coefficients were assigned, without checking of each of them is really statistically significant. Partly, this is done to constrain the calculation effort and partly because omitting certain terms would deteriorate the  $R^2$  factor. The coefficients of the terms less significantly contributing to the equation are expected to be small, so that their influence on the function to be estimated is not large, whereas these terms do contribute to the explanation of the variance by the regression equation.

In order to check the usability of the obtained regression equations as approximations for the true responses, two different methods, the  $R^2$  test and the F-test for statistical

significance have been described. First, it is recalled that the statistical analyses of the regression equation were based on the assumption that the residuals are Normally distributed. This assumption should be checked. The assumption  $\varepsilon_i \sim N(0, \sigma^2)$  implies that  $\frac{\varepsilon_i}{\sigma} \sim N(0, 1)$ . Since  $s^2$  is an unbiased estimator of  $\sigma^2$ , the value  $\frac{\varepsilon_i}{\sigma}$  can be examined for each sample point to check whether the assumption of the residuals being Normally distributed is a realistic one. As 95% of a  $N(0, 1)$  distribution lies in the interval  $(-1.96, 1.96)$ , the assumption would be supported (at least could not statistically be rejected) if roughly 95% of the observed values of  $|\varepsilon_i / s|$  would be less than 2.

As indicated before, the  $R^2$  test and the F-test impose contradicting demands on the number of samples on which a regression equation should be based. An exact fit would be obtained in case the number of samples would be taken equal to the number of coefficients to be determined, resulting in an exactly determined system of equations. In this case, the  $R^2$  factor would be one. However, the statistical significance of such a fit would be remote, and the obtained equation would not be suitable for approximating purposes on a global scale within the region of interest (the design space). On the other hand, when the number of samples on which the regression equation is based is very large, the fit would probably be rather poor. Therefore, the number of samples should be chosen such, that both the  $R^2$  and the F-test attain acceptable values. In fact, according to [ref.2] (based on [ref.8]), in order for a regression equation to be suitable for predictive purposes, the observed F-test value should exceed at least four times the table value. Since in the present study, the number of variables considered is such that the values  $p-1$  and  $k-p$  will comfortably exceed 60, the application of the "four-times rule" dictates that the F-test should exceed  $4 \times 1.53 = 6.12$ , or at least 6, with 95% fidelity.

If both the  $R^2$  factor and the F-test yield satisfactory values for a certain regression equation and if the assumption that the residuals are Normally distributed is sufficiently supported, it is still not certain that the obtained equation is suitable for predictive purposes. It has, so far, only been established that the equation is statistically significant, that is, it cannot be rejected that at least one of the non-constant terms is significant in explaining the observed variance in the true responses and that a sufficient percentage of this variance is explained by the equation. However, the errors in the prediction (the residuals) can still be so large that the equation is of no use (the regression surface is in this case merely fitted to the errors).

Therefore, it is necessary to check as well the standard error of the regression,  $s$ . This value is obtained by taking the square root of the mean square about the regression (mean square of the residuals,  $s^2$ ). Usually this value is expressed as a percentage of the mean response.

A possible problem may occur in case the value to be approximated must be positive (like for instance the wave drag coefficient or the lift gradient). The regression

equation may occasionally yield negative values for a certain combination of variables which would make the equation useless in an analysis or optimization program. A remedy for this problem is to fit the regression surface through the logarithms or the square roots of the responses and to take the exponents c.q. squares of the approximate values obtained from the regression equation. It was found that the square root trick yields the best results. It should be realized, though, that the tests as presented for the regression equations refer to their ability to approximate the square roots of the wanted values instead of the values themselves. They should, however, give sufficient indication as to the accuracy of the obtained regression equations in predicting the functions of interest.

In the present work, regression surfaces will be used for two purposes. Firstly, the gradient-based optimization will be applied to regression surfaces approximating the objective function and constraints. Secondly, as will be explained in part II, a panel method will be used to calculate the aerodynamic characteristics of the aircraft to be optimized. Since the calculations of the panel method are very time consuming whereas it is difficult to incorporate the program within an analysis program which in turn will be called many times by the optimization algorithm, the results of the panel method are approximated by regression surfaces as well.

The procedure to obtain a regression equation (second-order polynomial) of a certain value as a function of its variables is as follows:

- generate a number of random sets of variables (sample designs) by varying each variable independently at random between its predefined lower and upper bounds. It should be noted that the accuracy of the regression equations outside these boundaries will be doubtful. Experience teaches, that the accuracy of least-squares polynomial fits will deteriorate dramatically outside the design space for which they were derived.
- calculate the responses for each of the samples.
- take the square root of the responses.
- fit the least-squares polynomial through the samples.
- calculate the  $R^2$  factor, the F-test, the percentage of residuals which are Normally distributed and the standard error of the regression as a percentage of the mean response.
- if necessary, adapt the number of samples until these values have attained acceptable values (F-test according to "four-times rule"). For the first optimization run the requirements to these values may be somewhat slackened, for instance:  $R^2 > 80\%$  and the standard error of regression  $< 15\%$ .

In the context of the present application, it should be realized that the regression equations for the aerodynamic coefficients in supersonic flow are likely to be less accurate than the same in subsonic flow, which is of course the result of the natural discontinuities that occur in supersonic flow because of abrupt changes from supersonic to subsonic leading-edge flow and vice versa. Numerical discontinuities occur as well, resulting in relatively much noise on the solutions. These discontinuities

---

are, however, a major obstacle for the use of gradient-based optimization routines, a problem which can be elegantly averted by the use of regression surfaces, be it at the cost of the accuracy (see also [ref.4]). However, after a few optimization runs (with different starting points or initial designs), the design space in which the optimum is suspected to be located can be reduced and new polynomials can be generated with enhanced accuracy.

## 5.4 References

1. Davies, O.L.; *Statistical Methods in Research and Production*; Hafner Publishing Co. New York, 1958
2. Draper, N.R. and Smith, H.; *Applied Regression Analysis*; John Wiley & Sons, Inc., New York, 1966
3. Englund, W.C.; Stanley, D.O.; Lepsch, R.A.; McMillian, M.M. and Unal, R.; *Aerodynamic Configuration Design Using Response Surface Methodology Analysis*; AIAA 93-3967
4. Giunta, A.A.; Dudley, J.M.; Narducci, R.; Grossman, B.; Haftka, R.T.; Mason, W.H. and Watson, L.T.; *Noisy Aerodynamic Response and Smooth Approximations in HSCT Design*; AIAA 94-4367-CP.
5. Healy, M.J.; Kowalik, J.C. and Ramsay, J.W.; *Airplane Engine Selection by Optimization on Surface Fit Approximations*; Journal of Aircraft Vol.12, No.7, July 1975
6. Jensen, S.C.; Rettie, I.H. and Barber, E.A.; *Role of Figures of Merit in Design Optimization and Technology Assessment*; Journal of Aircraft Vol.18, No.2, 1981; AIAA 79-0234 January 1979
7. Myers, R.H.; *Response Surface Methodology*; Allyn and Bacon, Boston, Mass. 1971
8. Wetz, J.M.; *Criteria for judging Adequacy of Estimation by an Approximating Response Function*; Ph.D. dissertation, under direction of Dr. G.E.P. Box, 1964, University of Wisconsin.

# 6. Direct Search Methods with Special Reference to Genetic Algorithms

## 6.1 Summary

The main problem concerned with multidisciplinary optimization is, that calculating derivatives of the objective function and the constraints with respect to the design variables, is a very time-consuming task, while the accuracy of the obtained values in many cases is not very high. Therefore, it would be logical to consider optimization methods that do not need such derivatives. Such methods are known as direct search methods. The problem with these methods is, that the information about the character or shape of the design space which is normally provided by the derivatives, is now not available. This information, which is needed in order to guide the design towards the optimum, must therefore be obtained in a different manner. Unfortunately, this often implies that either an extremely large amount of designs is generated, or that not all the necessary information is available, leading to an improved, but not optimal design.

A short summary of the most important direct search optimization methods will be given in the first paragraph. Only the genetic algorithm, which is the subject of this chapter, will be treated more elaborately (paragraph 6.3.6).

## 6.2 Notation

A	wing aspect ratio
c	fitness ratio between two specified individuals
c'	fitness ratio between two specified individuals
E	energy function
f	objective function, fitness value
H	string schema
H'	subset of string schema
$h_i$	value of bit at $i^{\text{th}}$ position
i	index of bit position on string
	incidence
j	index
$k_B$	Boltzmann constant
l	string length (number of bits)
M	Mach number
m	number of copies of specified schema
n	number of individuals in population
	iteration number
o	order of schema
P	probability function
$p_c$	crossover probability
$p_m$	mutation probability

---

r	fitness ratio between two specified individuals minimum ranking
S	wing area
T	temperature parameter
t	time, generation number
x	design variable
y	design variable
z	design variable
$\alpha$	coefficient angle of attack
$\beta$	power
$\delta$	defining-length of schema
$\epsilon$	fitness coefficient
$\Lambda$	sweep angle
$\sigma$	coefficient
$\sigma_T$	thrust scaling coefficient

#### subscripts

0	initial
cr	cruise
le	leading-edge
n	iteration number
w	wing

#### abbreviations

SQP Sequential Quadratic Programming

### 6.3 A Summary of Direct Search Optimization Methods

In the following paragraphs a short summary of the most important direct search optimization methods will be given.

#### 6.3.1 Monte Carlo Methods

The most well-known methods belong to the category of random search, or Monte Carlo methods. These methods simply make use of the computer's ability to generate a large amount of designs in relatively short time. The method proceeds as follows: each independent variable is randomly chosen within the bounds assigned to it. In this way, many designs are generated, that are scattered randomly across the design space. The design which has the best value for the objective function is selected as the temporary optimal design and will be stored in memory. The other designs are not taken into consideration. In case, during the evaluation of a certain design point, a constraint turns out to be violated, the further evaluation of that design point is

---



immediately abandoned, thus saving computational time. This process is repeated several times, each time generating a large amount of random designs and keeping only the best one. If the optimal values of certain variables seem to converge to a specific region, the bounds on that variable may be tightened, thus reducing the design space. In this way the random search will take place in a gradually decreasing design space around the absolute optimum until the variation of the optimal values of the design variables stays within a specified termination interval.

A completely automated version of this optimization method including interactive graphical visualization of the convergence process is described in [ref.8].

An advantage of this method is, that there is no danger of the program converging to a local optimum, therefore non-convex problems can be handled. Furthermore, the method can handle discrete variables, which a gradient-based optimization method cannot, as it is impossible to calculate derivatives of discrete variables. This enables not only design variables like for instance the number of passengers or the number of engines to be optimized, but even variables that indicate if the airplane is equipped with a T-tail or a butterfly-tail, if a canard wing is applied or which type of fuel or propulsion system is used. These variables normally could only be determined by generating separate optimal solutions for different concepts. Most direct search optimization methods have these favorable characteristics.

A major disadvantage of the method, however, is the extremely large amount of designs that needs to be generated in order to reach the optimal solution. This amount can be as high as 20,000 generated designs for an optimization problem with less than 10 variables and constraints. Clearly, this makes the method completely unsuitable for large, time-consuming analyses that use rather complex computer programs like for instance finite element methods or panel methods.

### 6.3.2 Adaptive Grid Methods

The adaptive grid method is somewhat similar to the random search method described in the preceding paragraph, only it does not adapt the design space by reducing it, but by adapting the local density of the search grid in the neighborhood of the optimum. In [ref.14] such a method is described and relations are presented -as a function of the dimensionality of the problem- for the calculation effort to reach the optimum with a specified accuracy and probability that the optimum is a global one.

The method proceeds as follows: within the bounds imposed on the design variables, a grid is defined and the function values are evaluated at the grid points. In the neighborhood of the best point the grid density is increased and the procedure is repeated. The design space itself is not diminished since this would cumulatively decrease the probability that the investigated design space contains the global optimum.

The advantage of the method is that the relations between the calculation effort, the probability that the optimum found is a global optimum, the number of steps, the

---

accuracy and the dimensionality of the problem can be analyzed (and specified) in advance, which implies that the performance of the optimization is known a priori. This is a feature not available for other optimization methods. According to [ref.14], the method has a much higher probability of locating the global optimum than has the previously discussed Monte Carlo method for any number of steps.

The big disadvantage, like with most direct search methods, is the extremely large number of function evaluations needed, which makes the method only suitable for objective functions that can be analyzed real-time or nearly real-time. According to the relations presented in [ref.14], for the problem that is tackled in the present chapter (dimensionality=7), in order to obtain the same accuracy as was obtained using a gradient-based method and with an acceptable probability of locating the global optimum, about 10,000 designs have to be generated. For detailed information about this type of optimization method and the discussed relations, the reader is referred to [ref.14].

### 6.3.3 Pattern Search

The pattern search method globally guides the design towards the optimum in the direction of the line connecting the current with the previous base point. These base points are a series of improving designs generated by a local investigation of the design space. The method proceeds as follows: starting with an initial design point (base point), small steps are taken in the coordinate directions of each design variable. If the objective function improves along a specific direction, that direction is kept, otherwise the opposite direction is examined. In case a constraint boundary is encountered, the steplength is reduced accordingly. The result of this exploratory search is a second base point. The next step will be taken in the direction from the initial design point towards the newly generated base point. This direction is maintained until either a constraint boundary is encountered or until no further improvement occurs. From that point, a new exploratory search is started, yielding a third base point whereafter the search proceeds in the direction from the second to the third base point and so on.

Assuming that in large, multidisciplinary optimization the derivatives of the objective function and the constraints with respect to the design variables will have to be calculated using finite-differencing techniques, the pattern search method is no improvement at all. The exploratory search toward a new base point is practically the same as calculating the derivatives with forward differencing, which -as mentioned before- is a very cumbersome task for large and complex systems. More information is to be found in [ref.10], [ref.11] and [ref.16].

### 6.3.4 Taguchi Methods

These methods are named after the Japanese engineer Taguchi, who introduced the method as a means to rebuild Japanese post-war industry from the idea that optimization and quality control should be implemented in the design process as early as possible. The method, in fact, is a kind of intelligent parameter study, in

---

which a number of possible values is assigned to each design variable. These different values are known as the levels that a particular variable can attain. Instead of examining all possible combinations of the design variables (full factorial matrix), only a subset of designs is examined, based on so-called orthogonal matrices. In these matrices, the vectors of design variables appear as rows, in such a way that each pair of columns is orthogonal. This ensures that between each pair of columns all combinations of the two variables occur and this an equal number of times. In this way, a subset is taken from the design space that is unbiased towards any variable. Therefore, in this way, the influence of the various variables on the design should be well represented, but with much less computational effort than would be required in a full factorial parameter study. For instance, an orthogonal matrix with 27 rows could explore 13 variables at three levels by generating only 27 designs, whereas a full factorial parameter study of the same problem would require  $3^{13}$  or almost 1.6 million design evaluations.

When all the required design evaluations are completed, the mean value of the objective function is calculated for each variable at all its levels. In this way, the overall influence of a certain level can be determined for each variable and the level that yields the best design is chosen as the optimal one.

A major disadvantage of this method is, that interactions between two or more variables are not taken into account. If such interactions or couplings exist, the user should identify them himself and create an extra row in the orthogonal matrix that represents the interaction between the chosen two variables. Naturally, when applying the method to very large systems, only a limited amount of interactions can be studied. A further drawback of the method is, that no exact optimal solution is obtained, as the variables are only examined for a number of discrete values. The advantage, on the other hand, is that this kind of optimization is very well suitable for variables that are discrete by nature. Further information about Taguchi methods can be found in [ref.1] and [ref.15].

### 6.3.5 Simulated Annealing

Simulated Annealing Algorithms imitate the process of physical annealing, that is, the heating of a solid until it melts, followed by gradually cooling until it starts crystallizing. The final state that this process reaches, is a state of minimal free energy. This process is simulated by the Metropolis algorithm.

The basic assumption of the algorithm is, that the probability of a design attaining a certain state (combination of design variable values) is related to the energy (objective function value) associated with that state by the Boltzmann probability factor:

$$P(\bar{x}) = e^{\frac{-E(\bar{x})}{k_b T}} \quad (126)$$

in which  $E(\bar{x})$  is the energy associated with the vector of design variables  $\bar{x}$ ,  $T$  is the temperature parameter and  $k_B$  is the Boltzmann constant.

It can be easily established that, at equilibrium, the most probable state is that of lowest energy. The algorithm proceeds as follows: first the design variables are changed at random and the associated change in energy  $\Delta E$  is computed. If  $\Delta E$  is less than or equal to zero, the change is accepted, if  $\Delta E$  is positive, the change is accepted with a probability

$$P(\Delta E) = e^{\frac{-\Delta E}{k_B T}} \quad (127)$$

This procedure has the advantage that the optimization will not get stuck in a local optimum, since steps away from the optimum are permitted. With decreasing temperature, however, the probability that changes resulting in an increase of the objective function are accepted diminishes. After equilibrium is attained for a given temperature, the procedure is repeated for a lower temperature. The way in which the temperature is adapted is known as the cooling schedule, for instance exponentially:  $T_n = 0.9^n T_0$ . A cooling schedule for a certain optimization problem may be obtained by trial and error, or the system may simply be "heated" until it "melts" and subsequently "cooled down" in steps.

Simulated Annealing appears to be very well suited for optimization problems with very high dimensionality. It has been reported that good and accurate results can be obtained. The latter is a result of the inherent fine-tuning of the method, since at low temperatures only very small uphill changes are probable. The method is very robust (high ability to locate global optimum) and it has been reported to be relatively efficient. More information about Simulated Annealing can be found in [ref.9] and [ref.13].

### 6.3.6 Genetic Algorithms

Genetic search algorithms are based on Darwin's theory of survival of the fittest. According to this theory, an individual (design) with favorable genetic characteristics (values of design variables) will be most likely to survive and produce offspring (mating of the fittest). In a genetic algorithm, evolution is simulated by generating a population of designs, either at random (primeval soup) or by judgement or intuition, from which a subset is selected to produce offspring. It is expected that only the best designs will survive and mate again (produce new designs) thus increasing the fitness of the population on average and hopefully the highest fitness in the population as well. This immediately indicates the difference between genetic algorithms and neural networks, which is that in a genetic algorithm no individual develops (learns), only the generations develop.

Offspring is produced by means of three different processes:

- |                 |   |
|-----------------|---|
| 1) reproduction | (the selection of candidates for the crossover process) |
| 2) crossover    | (the exchange of genetic material between mates)        |
| 3) mutation     | (change of genetic material within an individual)       |

To facilitate this, the values of the design variables for each individual (i.e. design) are represented as binary numbers, arranged in long chains which act as the equivalents of DNA-strings. Reproduction takes place by picking members from the population at random, but with a bias towards those members whose fitness (measured by some augmented objective function) is relatively high in comparison to the average fitness of the population. The most commonly employed selection procedure is roulette wheel or proportional selection. With this procedure, to each individual a certain part of an imaginary roulette wheel is assigned, proportional to its fitness in relation to the average fitness of the population. When the roulette wheel is spun (by generating quasi-random numbers) the fittest individuals have the greatest probability of being picked. Another, less common selection procedure is tournament selection, which implies that the individuals are distributed in groups of tournament size  $n$ , and the best of each group is picked (each individual thus at least has one competitor).

Next, the individuals selected for crossover are grouped into pairs which will act as parents. For each pair it is established (again by generating quasi-random numbers) whether crossover will take place or whether the parents will go over into the next generation unchanged (the probability that crossover will take place,  $p_c$  is taken equal to 0.6 in this application). In case crossover will take place, two sites on their pseudo-DNA strings are selected at random. Subsequently, the sections between the sites are swapped between the two mates (two-point crossover). A different crossover operator is one-point crossover in which case only one site is selected at random and the tails of the pseudo-DNA strings are swapped between the parents. Also uniform crossover is sometimes employed. In this case it is established per gen (bit), whether the value of the father or the value of the mother is assigned to it. This reduces the risk that two highly-fit parents produce low-fitness offspring because of strong couplings between variables. Thus, crossover is a binary operator; two individuals create two new individuals. Recent research in the field of genetic optimization indicates that very good results are obtained if "orgies" are enabled; multiple parents exchanging genetic material.

After crossover has taken place, it is established per child, again by quasi-random number, whether mutation should take place. The probability of mutation is usually very small (in the order of 0.001) and the process should be regarded as a guard against premature loss of beneficial characteristics. If mutation takes place, a site on the DNA string of the child is selected and the bit under consideration is inverted. Mutation, therefore, is a unary operator; one individual creating one new individual. After all three stages of regeneration are completed, the new population is reduced to its original size by letting the least fit individuals die (survival of the fittest). In some genetic algorithms the parents always die. A special measure, called the elitist option, is included in this case to preserve the best individual in order to avoid deterioration

---

of the highest fitness in the next generation.

The process of creating a new generation will be illustrated hereafter (it is assumed that the parents die).

Imagine a population of four individuals, described by a single variable represented by a five-bit string, for instance:

0 1 1 0 1	(13)
1 1 0 0 0	(24)
0 1 0 0 0	(8)
1 0 0 1 1	(19)

The decoded values are given behind the string in parentheses. In biologic terminology, the binary (coded) representation is an individual's genotype, the decimal (decoded) representation its phenotype.

If the fitness value (objective function) would be defined as the decoded value squared, ( $f=x^2$ ), then the probability of reproduction for the four individuals would be respectively  $169/1170=0.14$ ;  $576/1170=0.49$ ;  $64/1170=0.06$  and  $361/1170=0.31$  (fitness relative to total fitness).

The expected number of copies in the next generation is given by the ratio of the individual fitness and the average fitness. This yields 0.58 ( $\Rightarrow 1$ ), 1.97 ( $\Rightarrow 2$ ), 0.22 ( $\Rightarrow 0$ ) and 1.23 ( $\Rightarrow 1$ ) respectively. After reproduction, the population will therefore look as follows:

0 1 1 0 1
1 1 0 0 0
1 1 0 0 0
1 0 0 1 1

Note that the third individual with the poorest fitness was not picked for reproduction; instead, the second one with the highest fitness was picked twice.

---

Now, assuming that the crossover location for the first two mates was randomly selected at location 4 and for the second two mates at location 2, the result of crossover will be as follows:

initial population	after reproduction	new population
0 1 1 0 1	0 1 1 0   1	0 1 1 0 0
1 1 0 0 0	1 1 0 0   0	1 1 0 0 1
0 1 0 0 0	1 1   0 0 0	1 1 0 1 1
1 0 0 1 1	1 0   0 1 1	1 0 0 0 0

The symbol | marks the crossover site.

Apparently, the fitness of the best individual has increased from 576 in the first generation to 729 in the second one.

It is clear that the great advantage of a genetic algorithm over a gradient-based method is, that a genetic algorithm works with a population of points instead of a single point, which largely reduces the possibility of convergence towards a local optimum (and therefore increases the robustness, i.e. the probability of finding the global optimum). This is indicated in Figure 14 which shows the robustness of three different classes of optimization algorithms versus their applicability.

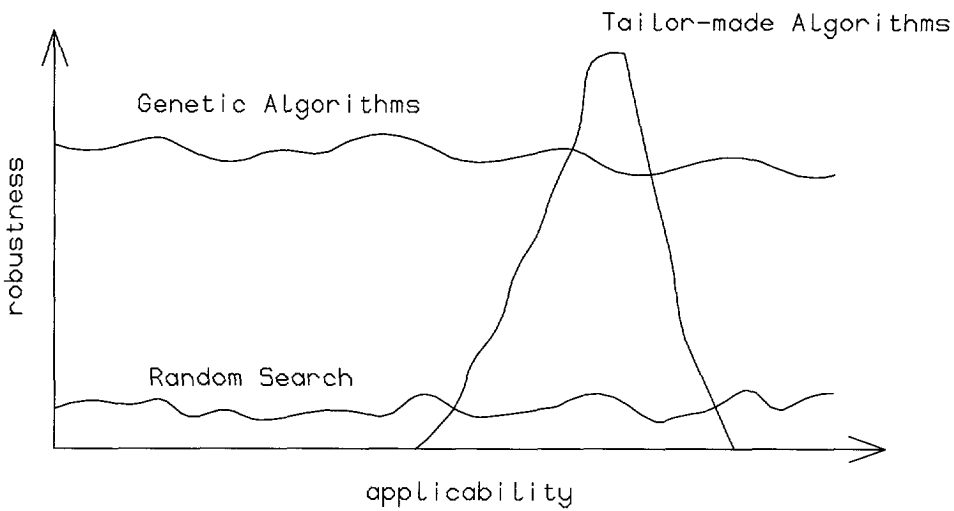


Figure 14 Robustness versus applicability for different classes of optimization methods

Although it would seem that the frequent use of quasi-random generators in the

algorithm does not distinct it from a classical Monte Carlo method, it is important to realize that quasi-random generators are only tools to carry out certain decisions in an otherwise perfectly intelligent search. On the other hand, of course, since the processes of crossover and mutation are not intelligent but performed at random, there is no guarantee that the method will actually find the optimum (or even show improvement in the next generation, which is why the elitist option is included in case the parents will not survive in the next generation). As will be explained later on, however, convergence does occur most of the time.

The powerful (though very simple) mechanism that guides the search toward the optimum is the combination of highly fit, short defining-length schemata, the so-called building blocks. A schema is a representation of a certain section of a string, for instance, the string 0111000 can be represented by different schemata such as:

H1: \* 1 \* \* \* 0

H2: \* \* \* 1 0 \* \*

in which the asterisk (\*) can be either a 0 or a 1. Thus, a schema represents a family of pseudo-DNA strings and can therefore be interpreted as a hyperplane in the design space.

The order of a schema,  $o(H)$ , is defined as the number of fixed positions and both schemata therefore are of order 2. The defining-length of a schema,  $\delta(H)$ , is defined as the difference between the first and the last fixed position and it can be easily verified that  $\delta(H1)=5$  and  $\delta(H2)=1$ .

Two examples of schemata representing the previously used five-bit representation of  $x$  for the optimization of  $f=x^2$  are illustrated in Figure 15.

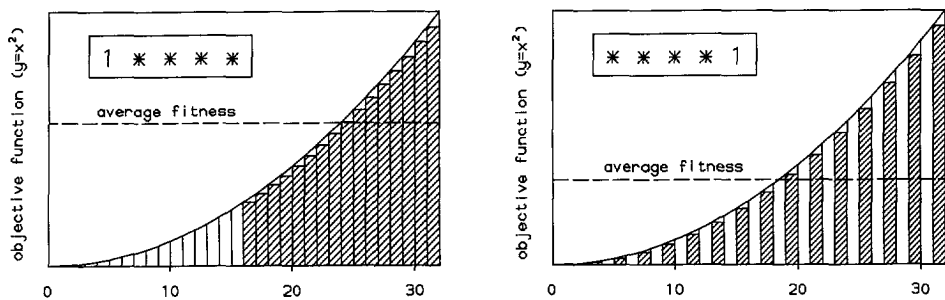


Figure 15 Difference between the schemata 1\*\*\*\* and \*\*\*\*1

Clearly, the two schemata represent different sections of the design space with different average fitness.



The advantage of short defining-length schemata is obvious; they have the smallest chance of getting disrupted by crossover (schemata are relatively unlikely to be disrupted by mutation) and the best chance of being picked for reproduction. If the average fitness of a certain schema is relatively high then the expected number of copies of that schema in the next generation will increase exponentially according to [ref.4]:

$$m(H, t+1) \geq m(H, t) \frac{f(H)}{\bar{f}} \left[ 1 - p_c \frac{\delta(H)}{l-1} - o(H) p_m \right] \quad (128)$$

in which  $m(H, t)$  is the number of copies of schema  $H$  in generation  $t$  and  $p_c$  and  $p_m$  are the probabilities of crossover and mutation respectively. The average fitness of a certain schema  $f(H)$  is obtained by averaging the fitnesses of all possible individuals represented by that schema. Unfortunately, the average fitness of a certain schema can in practice only be estimated from the individuals in the current population. It can be easily verified that the first term on the right-hand-side represents the probability of being selected for reproduction, whereas the second and third term represent the probabilities of disruption of the schema by crossover, respectively mutation.

Proceeding in the manner explained, generating new populations by reproduction, crossover and mutation, involves the risk of premature convergence. In the initial population there will be a few good individuals amongst many mediocre ones. In the next generations the good individuals will start to dominate the population, which is the cause for premature convergence toward an improved, but not optimal solution. Once almost all individuals of the population have about the same fitness, the method will degenerate into a random search amongst largely equally fit individuals.

To prevent the occurrence of these effects, the fitness value is often scaled with respect to the current average fitness or the best fitness over the last  $n$  generations. This will reduce the difference in fitness of the individuals in the first generations and increase it in the final ones.

As indicated before, there is no guarantee that the algorithm will work. When parts of the strings of two excellent designs are exchanged, the result will not necessarily be an even better design, especially when strong couplings exist.

Problems with such characteristics as to cause the algorithm to diverge are called deceptive problems. As will be shown, deceptive problems are in many cases not *hard*, that is, despite its negative characteristics the problem will converge toward the right solution.

If a certain problem is deceptive, the average fitness of the building blocks (highly fit, short defining-length schemata) of suboptimal individuals is higher than that of the

building blocks of the optimal individual. As an example consider a population defined by the following second-order schemata [ref.4]:

*** 0 * * * * 0 *	$f_{00}$
*** 0 * * * * 1 *	$f_{01}$
*** 1 * * * * 0 *	$f_{10}$
*** 1 * * * * 1 *	$f_{11}$

The average fitness of each schema is represented by the symbol  $f$  behind the string representation.

If the optimum is represented by the fourth schema, then the following relations must apply:

$$f_{11} > f_{00} ; f_{11} > f_{01} ; f_{11} > f_{10}$$

For a deceptive problem, the building blocks of the nonoptimal schemata should be better (have higher average fitness) than those of the optimal one, or:

$$f(0^*) > f(1^*)$$

$$f(*0) > f(*1)$$

Writing these conditions in a different shape yields:

$$\frac{f(00) + f(01)}{2} > \frac{f(10) + f(11)}{2}$$

$$\frac{f(00) + f(10)}{2} > \frac{f(01) + f(11)}{2}$$

Both conditions cannot be satisfied simultaneously or else the fourth individual will not be the optimum. In this example, the first condition is retained. With the following substitutions, the optimality condition and the deceptive condition can be written in a different shape:

$$r = \frac{f_{11}}{f_{00}} \quad c = \frac{f_{01}}{f_{00}} \quad c' = \frac{f_{10}}{f_{00}}$$

which yields:

$$r > c; \quad r > 1; \quad r > c'$$

$$r < 1+c+c'$$

From this it follows:

$$c' < 1; \quad c' < c$$

It appears however, that even if a deceptive problem is constructed, it is often not *hard* because the expected loss of genetic material as predicted by equation (128) does not occur. This is caused by the fact that in this equation the creation of schemata is not included and the formula is therefore overly conservative. In case two noncomplementary schemata are crossed, none is lost, like for instance in the case of mating a 00 with a 01 yielding a 01 and a 00. Therefore, although still no hard guarantees can be provided, usually convergence occurs.

A different way to look at deceptive problems is given in [ref.4]. In this case the fitness of a schema is represented as a series of coefficients as follows:

$$f(H) = \sum_{H' \subseteq H} \sigma(H') \varepsilon_{j(H')} \quad (129)$$

in which  $H'$  are subsets that are contained in the schema  $H$ . The function  $\sigma$  assumes the value 1 when a schema contains an even number of 0's and a -1 otherwise:

$$\sigma(H) = \prod_{i=1}^l (-1)^{\beta(h_i)} \quad (130)$$

in which  $h_i$  is the value (0,1 or \*) at the  $i^{\text{th}}$  position of the string. The power  $\beta$  assumes a value of 1 when  $h_i = 0$  and 0 otherwise.

The index  $j$  is defined as:

$$j(H) = \sum_{i=1}^l \alpha(h_i) 2^{i-1} \quad (131)$$

where  $\alpha$  assumes a value of 0 when  $h_i$  equals \* and a 1 otherwise.

As an example [ref.4] considers the problem of minimizing  $f=x^2$  coded as a three bit string. For each schema, the following expressions can be derived for the fitnesses in terms of the coefficients  $\epsilon$ :

$$\begin{array}{ll}
 f(***) = \epsilon_0 & f(***) = \epsilon_0 \\
 f(**1) = \epsilon_0 + \epsilon_1 & f(**0) = \epsilon_0 - \epsilon_1 \\
 f(*1*) = \epsilon_0 + \epsilon_2 & f(*0*) = \epsilon_0 - \epsilon_2 \\
 f(*11) = \epsilon_0 + \epsilon_1 + \epsilon_2 + \epsilon_3 & f(*00) = \epsilon_0 - \epsilon_1 - \epsilon_2 + \epsilon_3 \\
 f(1**) = \epsilon_0 + \epsilon_4 & f(0**) = \epsilon_0 - \epsilon_4 \\
 f(1*1) = \epsilon_0 + \epsilon_1 + \epsilon_4 + \epsilon_5 & f(0*0) = \epsilon_0 - \epsilon_1 - \epsilon_4 + \epsilon_5 \\
 f(11*) = \epsilon_0 + \epsilon_2 + \epsilon_4 + \epsilon_6 & f(00*) = \epsilon_0 - \epsilon_2 - \epsilon_4 + \epsilon_6 \\
 f(111) = \epsilon_0 + \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4 + \epsilon_5 + \epsilon_6 + \epsilon_7 & f(000) = \epsilon_0 - \epsilon_1 - \epsilon_2 + \epsilon_3 - \epsilon_4 + \epsilon_5 + \epsilon_6 - \epsilon_7
 \end{array}$$

The average fitness of each schema can easily be computed and from these values follow the coefficients:

	schema	average fitness $\epsilon(j)$	
0	***	17.5	17.5
1	**1	21.0	3.5
2	*1*	24.5	7.0
3	*11	29.0	1.0
4	1**	31.5	14.0
5	1*1	37.0	2.0
6	11*	42.5	4.0
7	111	49.0	0.0

To check if the present problem is deceptive, it must be verified whether the schemata containing a 0 have higher fitness than the corresponding schemata containing 1's (since the optimum is 111):

$$\begin{array}{l}
 f(**1) < f(**0) \\
 f(*1*) < f(*0*) \\
 f(1**) < f(0**)
 \end{array}$$

With the previously deduced coefficient representation this implies:

$$\begin{array}{l}
 \epsilon_1 < 0 \\
 \epsilon_2 < 0 \\
 \epsilon_4 < 0
 \end{array}$$

and apparently the problem at hand is not one-bit deceptive.

The disadvantage of genetic algorithms, besides the fact that convergence is not guaranteed and the possibility of premature convergence, is the large number of function evaluations needed. The inefficiency of genetic algorithms is well illustrated by the example in [ref.5]. In this example it is endeavored to find the minimum of the function  $f=x^2+y^2+z^2$  on the interval  $[-5.12, 5.12]$  for each variable. Obviously, the

optimum is zero at the origin. The printed output of [ref.5] shows that after 1000 function evaluations (19 generations) using a population of 50 individuals defined by strings of 30 bits, the best value found was 0.1202 for  $(x,y,z)=(-0.29,0.19,0)$ .

It is very unlikely that a genetic algorithm will find the exact optimum. Instead, it will come up with a number of designs that are close to the optimum. This is partly caused by a very illustrative phenomenon, known as hitchhiking. In the following population:

```
1 0 0 1 1 (19)
1 0 1 0 1 (21)
1 0 1 1 0 (22)
0 1 0 0 0 (8)
0 1 1 1 1 (15)
```

the first three individuals have the highest probability of being selected for reproduction. Since all these individuals carry a 1 on the first position and a 0 on the second position that "takes a hike" from the beneficial 1 on the first position, it is most likely that the best result that is going to be obtained eventually is 10111. The probability that the building block 10\*\*\* will be disrupted is very small because of its high fitness and short defining-length.

The algorithm used in this application is the GENESIS algorithm by J.J. Grefenstette [ref.5]. One of the features of this program is that the user may define the design variables as floating point numbers instead of binary strings. The string representation (number of bits per variable) as well as the population size (number of designs per population) can be user-defined. For each design variable, the user sets the upper and lower bounds. The number of function evaluations can be set, as well as the probabilities for crossover and mutation (default values are 0.6 and 0.001 respectively). An option for a restart after the maximum allowable number of function evaluations has been performed is included.

The risk of premature convergence (a couple of excellent individuals rapidly taking over the entire population) can be reduced by selecting the individuals for reproduction by the so-called ranking procedure instead of the proportional selection procedure.

When using proportional selection, the probability to produce offspring is proportional to an individual's fitness relative to the sum of all individuals' fitnesses in the current population. This implies that an individual whose fitness is very much above average will have a very high probability to produce offspring and consequently will have many copies in the next generation. Depending on the measure of unbalance between the different fitnesses, these above average individuals can quickly start to dominate the subsequent generations.

In the example of page 98, the probabilities to produce offspring, according to proportional selection, were respectively: 0.14; 0.49; 0.06 and 0.31, yielding an

---

expected number of offspring of 0.56 ( $\Rightarrow$  1); 1.96 ( $\Rightarrow$  2); 0.24 ( $\Rightarrow$  0) and 1.24 ( $\Rightarrow$  1). For a population somewhat more unbalanced, like for instance:

0 1 1 0 0 (12)  
 1 1 1 0 0 (28)  
 0 1 0 1 1 (11)  
 0 1 1 0 1 (13)

the probabilities to produce offspring become respectively: 0.12; 0.64; 0.10 and 0.14 yielding an expected number of offspring of 0.48 ( $\Rightarrow$  0); 2.56 ( $\Rightarrow$  3); 0.40 ( $\Rightarrow$  0) and 0.56 ( $\Rightarrow$  1). It is clear that the second individual will dominate the next generation.

A different selection procedure is known as the ranking procedure which is included in the GENESIS program as an option. According to the ranking procedure the probability to produce offspring is proportional to an individual's ranking in the population, the best individual being ranked  $n-1$  ( $n$  being the number of individuals in the population) and the worst individual zero. With this selection procedure, the probability to produce offspring will be for both examples: 1/6; 3/6; 0/6; 2/6. The expected number of offspring will be 1, 2, 0 and 1. For the best individual, the expected number of offspring is equal to  $\frac{(n-1)}{\frac{1}{2}(n-1)}$ , which is therefore always equal to

2. The worst individual will have no offspring. The expected number of offspring for the other individuals will be linearly scaled between the best and the worst individual. This scaling can still be adapted somewhat by means of a minimum ranking option. Instead of ranking the worst individual zero, it can be ranked at a certain value, say  $r$ , whereas the best individual will be ranked  $2-r$ . Increasing  $r$  will increase the number of individuals which will contribute 1 offspring to the next generation, the upper limit being that each individual will contribute exactly 1 offspring.

The program contains some special features like the possibility to define a generation gap (percentage of old population to be replaced, in this case 100%), a scaling window (interval of generations after which the scaling reference fitness is updated), the number of consecutive generations that are permitted without evaluation of the fitnesses and the number of best individuals that should be carried on to the next generation unchanged (elitist option). This last option is advisable in order to prevent the destruction of good individuals by crossover.

Summarizing, the same sort of problems that occur with other direct search algorithms also appear in genetic algorithms. As the method proceeds iteratively, for each new generation the members of the population have to be evaluated. If the analysis of the designs incorporates complex computer programs, this will make convergence very slow. Furthermore, it has been reported that the method is not very accurate (it is not suitable to find the exact location of the optimum). Also, the method is very sensitive to coupled characteristics. In case two designs are selected

for crossover on the account of their having above average fitness, combining their properties in the offspring will not necessarily result in a better design and might even result in very bad characteristics.

On the other hand this method, like other direct search methods, is capable of handling non-convex problems and discrete variables as well as highly discontinuous or noisy functions.

It is these properties and the rapid increase in computer speed that has led to the fast rise in popularity of genetic search algorithms over the last few years. Although the extremely large number of function evaluations necessary to approach the optimum is still a large drawback, the use of genetic algorithms may be helpful in locating those regions in the design space where local optima may be found, thus enabling an effective reduction of the design space. In industry today much attention is paid to possible reductions of the time required for a single function evaluation (see also [ref.17]).

Using the presently adopted method of a gradient-based algorithm using approximate analysis (chapter 4), this implies that a more accurate approximate analysis can be obtained because of the reduction of the design space. Furthermore, the combination of the genetic optimization method with a gradient-based method and approximate analysis will enable the location of the global optimum without the risk of premature convergence or divergence. The use of a hybrid method as suggested combines the good characteristics of the genetic algorithm (robust) with that of the gradient method (efficient, exact). The genetic search can be constrained by the maximum allowed number of function evaluations.

Further information about genetic search algorithms can be found in [ref.2], [ref.3], [ref.6] and [ref.7].

## 6.4 Results

The genetic algorithm GENESIS [ref.5] was applied to the case of a twin-engined subsonic airplane design which was optimized in chapter 4. Two different methods were applied in that chapter, that is, a sequential quadratic programming method (E04VCF, available from the NAG Library [ref.12]) applied directly on the analysis model, and the same optimization method applied on second-order response surfaces based on the same analysis method (approximate analysis).

The same problem was tackled with the GENESIS algorithm ([ref.5]). Since genetic algorithms do not have special features to handle constrained problems, it is necessary to define a penalty function. In this case the squared value of each violated constraint times a penalty coefficient of  $10^9$  was added to the objective function. In case a design evaluation could not be completed because for example an angle of attack exceeded the stall angle of attack or because the available thrust during take-off was insufficient to accelerate the airplane, the calculations were broken off and the

---

objective function was set at a large number. The population size was put at 50 and each variable was represented by a ten bit binary number, leading to a length of 70 bits for the pseudo-DNA string of each individual.

The designs in the latest generated population (generation 84; after 2512 function evaluations) are gathered in the following table (only designs of the same order of magnitude as the current optimum are shown):

design variables (angles in rad)							
S [m <sup>2</sup> ]	A	$\Lambda_{le}$	$i_w$	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$	f
29.24	13.66	0.28	0.0948	0.61	0.88	0.0169	13999.4
29.24	13.66	0.28	0.0948	0.61	0.88	0.0169	13999.6
30.76	13.66	0.28	0.0948	0.61	1.33	0.0169	15085.5
30.53	12.27	0.38	0.0992	0.61	1.40	0.0161	15275.1
30.53	12.27	0.38	0.0992	0.59	1.40	0.0161	15275.1
34.55	12.51	0.28	0.0933	0.59	0.91	0.0165	15335.5
37.96	12.56	0.28	0.0948	0.58	0.88	0.0162	16106.9
30.45	13.66	0.28	0.0992	0.53	1.38	0.0615	16722.4
39.09	17.41	0.33	0.0948	0.59	1.45	0.0604	16951.0
28.57	12.59	0.28	0.0960	0.60	0.88	0.0169	17988.8
39.40	10.08	0.10	0.0938	0.60	1.58	0.0614	18023.5
48.01	13.60	0.55	0.0930	0.60	1.07	0.0205	19069.8
29.55	13.44	0.29	0.0678	0.61	0.89	0.0169	24210.6
29.20	12.39	0.31	0.0557	0.61	1.21	0.0306	72542.8

Table 14 Best designs of last population



A different way of presenting the results of the method, is given in the next table. Here, the ten best designs of all generations are gathered:

design variables (angles in rad)								
S [m <sup>2</sup> ]	A	$\Lambda_{je}$	$i_w$	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$	f	gen.
29.24	20.70	0.28	0.0928	0.59	0.89	0.0169	14022	66
29.24	20.70	0.28	0.0948	0.61	0.89	0.0169	14022	60
29.24	20.70	0.28	0.0948	0.61	0.89	0.0163	14022	67
29.24	20.70	0.28	0.0948	0.61	0.88	0.0169	14000	66
29.55	13.66	0.28	0.0948	0.61	0.88	0.0168	14066	80
29.24	13.66	0.28	0.0948	0.61	0.88	0.0169	14000	83
29.24	13.66	0.28	0.0948	0.61	0.88	0.0169	13999	70
29.24	20.70	0.28	0.0948	0.61	0.90	0.0619	14048	50
29.24	20.70	0.28	0.0928	0.61	0.89	0.0609	14022	64
29.26	11.41	0.45	0.0948	0.61	0.90	0.0165	14058	75

gen. = generation number

**Table 15** Best 10 designs of all generations

As can be deduced from these tables, the algorithm converges to one single optimum except for the aspect ratio. The same large difference that was already encountered in chapter 4 shows here as well.

Although the algorithm has found some good designs, there is no real convergence in the sense that all designs in the final population are almost the same. Although the wing area seems to converge to values around 29 m<sup>2</sup>, the aspect ratio to values around either 13 or 20, the incidence to values around 0.095 radians and the cruise angle of attack to values around 0.017 radians, there is still much scatter in the values for the sweep angle, the cruise Mach number and the engine scale factor. Furthermore, the best design differs substantially from the one found by the gradient method.

The latter is caused by the fact that the optimal values found by the gradient method are slightly infeasible (because the SQP optimizer uses a constraint tolerance in order to be able to capture active constraints and keep them in the active set more easily). The infeasibility of the approximate optimum is of course due to the approximation error. According to the penalty function used in the present method, the actual objective function for the two gradient methods would be 14383 and 15224

respectively. This is obviously much worse than the values found by GENESIS.

To investigate this matter further, a constraint tolerance of  $10^{-3}$  (same as the gradient method) was imposed and the program was executed once more. In order to establish whether better convergence could be obtained, the maximum number of function evaluations was this time put at 5000.

The results, after 5002 function evaluations (173 generations), are presented in the following tables. The ten best results are:

design variables (angles in rad)								
S [m <sup>2</sup> ]	A	$\Lambda_{le}$	$i_w$	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$	f	gen.
30.61	18.69	0.13	0.0881	0.57	0.74	0.0144	13830	170
30.91	17.80	0.11	0.0881	0.58	0.74	0.0144	13954	152
30.60	17.80	0.14	0.0881	0.57	0.74	0.0144	13854	168
30.60	17.80	0.11	0.0881	0.57	0.74	0.0144	13831	167
31.58	23.82	0.10	0.0995	0.57	0.72	0.0137	13931	151
30.60	18.68	0.13	0.0881	0.57	0.74	0.0143	13830	171
31.58	23.45	0.11	0.0995	0.57	0.72	0.0137	13931	135
31.58	24.06	0.11	0.0995	0.57	0.72	0.0137	13930	137
31.58	23.76	0.11	0.0995	0.57	0.72	0.0137	13930	141
31.58	24.00	0.11	0.0995	0.57	0.72	0.0137	13930	149

gen. = generation number

Table 16 Best 10 designs of last generation

The current optimum is:

design variables (angles in rad)						
S [m <sup>2</sup> ]	A	$\Lambda_{le}$	$i_w$	$M_{cr}$	$\sigma_T$	$\alpha_{cr}$
30.61	18.69	0.133	0.0881	0.568	0.742	0.0143
objective function and constraints						
objective function [kg]	wing span [m]	range	cruise thrust	zero altitude climb	one-engine-out ceiling	take-off field length
13830	23.91	-0.0006	0.0063	0.3877	1.0228	0.2239

**Table 17** Current optimum

Finally, 1000 more function evaluations were performed, but the result did not differ much from those obtained with less function evaluations. Although some variables show convergence toward one or more specific sections of the design space, others are still much scattered. Also, the objective function values differ much and the number of infeasible or incomputable designs is still remarkably high.

Looking at the bit-representation of the pseudo-DNA strings the evolution according to the building-block theorem is clearly visible. In Figure 16 the population at the second generation (after 100 function evaluations) is shown next to the population at the 208<sup>th</sup> generation (after 6000 function evaluations). For clarity, the 0's in the strings have been replaced by blanks. The change from a population with a purely chaotic distribution of genetic material into a population in which a very distinct pattern among the strings can be discerned is very striking.

## 6.5 Conclusions

The conclusions that can be drawn from the presented results are as was to be expected. The number of function evaluations needed for a successful genetic optimization as compared to gradient methods is extremely large. Even after 6000 function evaluations the method has not found a better optimum than the gradient method, worse, it has not even reached the optimum found by the SQP optimizer. Furthermore, the convergence rate in terms of the variance in the design variables in the last population is not great. This is also indicated by the average fitness of the final population which is  $2.243 \cdot 10^6$  (although it must be mentioned that the fitness value for incomputable designs -as indicated before- was put at  $1 \cdot 10^6$ ). The number of incomputable designs in the final population of 50 individuals is still quite high: 11, or 22% of the population. Furthermore, many designs violate the constraints

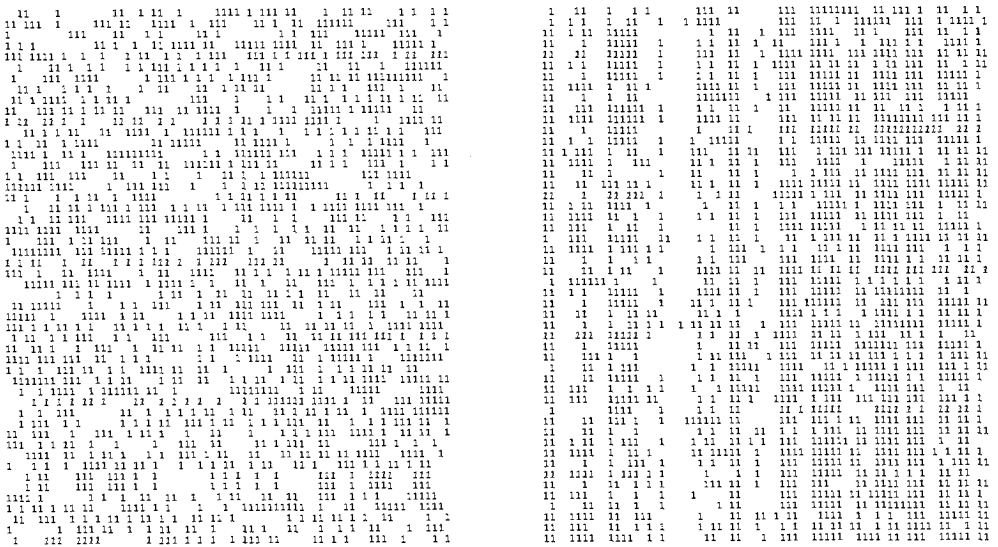


Figure 16 Comparison of populations of generations 2 and 208

substantially, leading to very high values for the augmented objective function.

Also, the final optimum often does not differ dramatically from the best designs that were found in very early generations. From the best design evaluated at generation 0, which was created purely at random, (objective function value: 14770) to the final best result (objective function value: 13789) an improvement of only 6.6% was achieved at the cost of more than 6000 function evaluations. The true optimum of 13683 was not found.

On the other hand, however, the method seems to be very well suited to constrain the design space for a gradient method. The best designs all seem to converge on certain clearly recognizable values and the presence of local optima is easily established (see for instance the two different values to which the aspect ratio converges).

It can further be concluded, that the use of the program GENESIS is a very simple and user-friendly means to get a first impression of the design space, because it requires only function values and no gradients and because it requires no knowledge of optimization theory whatsoever.

It seems therefore advisable to use a hybrid optimization method, consisting of a genetic algorithm to scan and reduce the design space in combination with a gradient method to carry out the final fine-tuning. Apparently, there is not much use in carrying out too many function evaluations during the genetic search, since the benefit does not balance the cost.

## 6.6 References

1. Bendell, T.; *Taguchi Methods*; Proceedings of the 1988 European Conference, Elsevier Applied Science
  2. Gage, P. and Kroo, I.; *A Role for Genetic Algorithms in a Preliminary Design Environment*; AIAA 93-3933, August 1993
  3. Gage, P. and Kroo, I.; *Representation Issues for Design Topological Optimization by Genetic Algorithms*; Department of Aeronautics and Astronautics, Stanford University, Stanford CA 94305, 1995
  4. Goldberg, D.E.; *Genetic Algorithms in Search, Optimization and Machine Learning*; Addison-Wesley Publishing Company Inc., 1989
  5. Grefenstette, J.J.; *A User's Guide to GENESIS*; Version 5.0, October 1990
  6. Hajela, P.; *Genetic Search-An Approach to the Nonconvex Optimization Problem*; AIAA Journal Vol.28, No.7, July 1990
  7. Hajela, P. and Lin, C.-Y.; *Genetic Search Strategies in Multicriterion Optimal Design*; AIAA 91-1040-CP
  8. Kalker-Kalkman, C.M.; *Optimal Design with the Aid of Randomization Methods*; Engineering with Computers 7, 173-183; Springer Verlag 1991
  9. Kirkpatrick, S.; Gelatt, C.D. Jr. and Vecchi, M.P.; *Optimization by Simulated Annealing*; Science, Vol.220, May 1983
  10. Lootsma, F.A.; *Algorithms for Unconstrained Optimization*; Delft University of Technology, June 1986
  11. Lootsma, F.A.; *Duality in Non-Linear Optimization*; Delft University of Technology, November/December 1991
  12. Numerical Algorithms Group; *Fortran Libraries, Mk.12, Vol.4*
  13. Ping Lu and Asif-Khan, M.; *Nonsmooth Trajectory Optimization: An Approach Using Continuous Simulated Annealing*; Journal of Guidance, Control and Dynamics, Vol.17, No.4, July-August 1994
  14. Shahrudi, K.E.; *Non-Random Adaptive Grid Method for High-Speed Optimization of Highly Dimensional, Badly Behaving Real-Time Functions*; ASME paper, 94-GT-487, June 1994
  15. Stanley, D.; Unal, R. and Joyner, R.; *Application of Taguchi Methods to Dual Mixture Ratio Propulsion System Optimization for SSTD Vehicles*; AIAA 92-0213
  16. Swann, W.H.; *Direct Search Methods*; published in *Numerical Methods for Unconstrained Optimization* (W. Murray ed.), Academic Press 1972
  17. van der Velden, A.J.M.; *Tools for Applied Engineering Optimization*; AGARD R-803, part 1, March 1994
-



---

## **Part II**

# **The Design and Analysis of a Second- Generation Supersonic Transport Aircraft**

---





# 7. Introduction

## 7.1 Summary

The second part of this dissertation is mainly concerned with the development of a design and analysis program to evaluate supersonic transport planes. The choice to apply the multivariate optimization techniques employed in this work to a supersonic transport aircraft is based on a number of reasons. First of all, since the couplings between the different disciplines are particularly strong for this type of airplane, a supersonic design can serve as an excellent test case. Secondly, interesting and perhaps unexpected results regarding the lay-out of a supersonic transport design might come up. Finally, since the design space for a second-generation supersonic transport is extremely small, because of the mentioned strong interdisciplinary couplings and the conflicting requirements imposed on the design with respect to subsonic and supersonic operation as well as environmental issues, it is hardly possible to generate a feasible design without the use of multidisciplinary optimization techniques.

The type of airplane treated in this work, in agreement with similar studies, will be required to carry 250 passengers over a transpacific distance (range at least 9000 km), while simultaneously being able to take-off from existing runways and complying with basic handling, safety and performance requirements (see page 306). Only a single configuration will be considered, namely a "classic" wing-body-tail configuration, equipped with four wing-mounted low-bypass engines.

## 7.2 The Development of a Second-Generation Supersonic Transport

As a consequence of the expected rise in air travel over the next decade, industry and airlines (notably the current operators of Concorde, British Airways and Air France) have shown a renewed interest in the development of a second-generation supersonic transport plane. According to [ref.21] global air traffic is expected to double until 2005 and again between 2005 and 2025. It is estimated that the number of passenger-miles on international lines suitable for HSCT (High-Speed Civil Transport) operation will grow from 795 billion in 2005 to over 2100 billion in 2025 ([ref.21]). Boeing estimate that the potential market for a supersonic transport will be 55% of a total 1700 billion passenger-kilometers or that 25% of the long-range market will be served by supersonic transports ([ref.25]). British Aerospace and Boeing see a demand for between 500 and 1000 supersonic cruising transport aircraft ([ref.12] and [ref.25]). Aerospatiale use the figure of 500 aircraft needed by 2025 ([ref.16]). On the other hand, German estimates only forecast a need of between 110 and 370 aircraft in 2010 ([ref.25]). In [ref.16] it is suggested that DASA tends to the Very Large Commercial Transport solution (Airbus A3XX) to the increased capacity problem.

In the currently undertaken efforts particular attention is being paid to the lessons learned from Concorde. First flown in 1969, Concorde is still the only commercially operated supersonic cruising airplane in history. Even few military airplanes are

---

capable of cruising at supersonic speeds (exclusively strategic reconnaissance planes like the Lockheed SR-71 *Blackbird*, or strategic bombers like for instance the Dassault-Bréguet *Mirage IV*, the Convair B-58 *Hustler*, the General Dynamics F-111 *Aardvark* and the Rockwell B1B *Lancer*, as well as some Soviet/Russian designs (notably the Tu-22M *Backfire* and Tu-160 *Blackjack*)). Currently, the first fighter with supercruise capabilities, the Lockheed-Boeing-General Dynamics F-22 has made its maiden flight. In most of these examples, the supersonic cruise capability imposed a heavy range penalty, leading to a reconsideration of the mission (most Russian supersonic bombers as well as the F-111 are more accurately described as tactical bombers, since their range does not allow transatlantic operations). Other designs are only capable to perform short supersonic dashes using reheat. It has often been suggested that the total supersonic flight hours of Concorde exceed all accumulated military supersonic flight hours ever. However, an airplane that is capable of cruising for longer times at supersonic speeds in an economical way still has to be designed.

Figure 17 shows the well known cruise speed gap of about 10 years between civil and military aircraft designs (taken from [ref.22]).

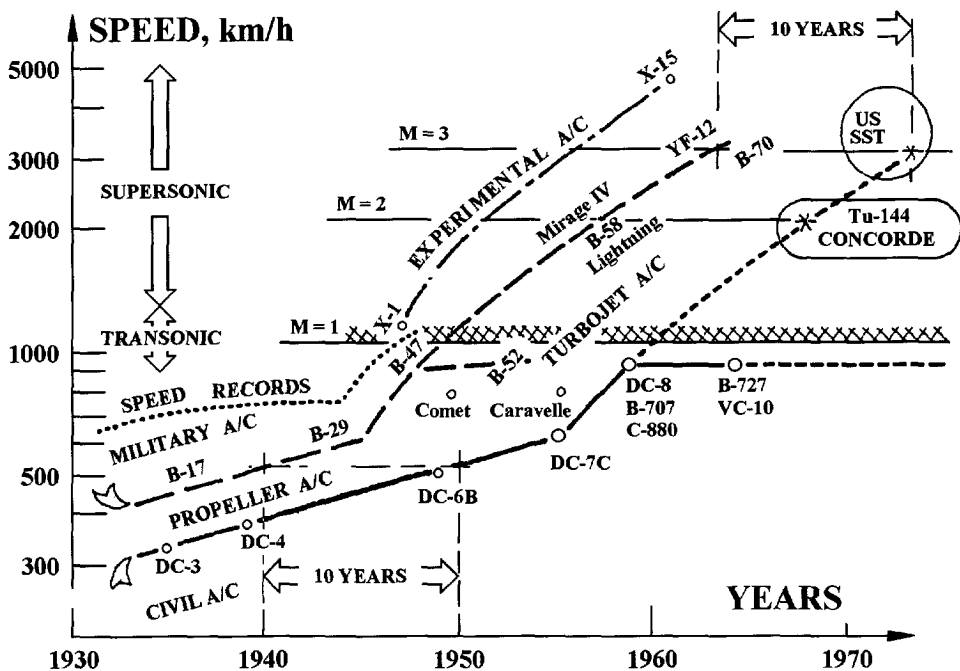


Figure 17 Cruise speed gap between civil and military aircraft ([ref.22])

Several causes have been suggested for the commercial failure of Concorde. Most important of these has been the gross underestimation of the task, from a technical point of view, leading to development time and cost that by far exceeded initial

estimates. Accordingly, the airplane weight increased rapidly, leading eventually to a design that was too expensive with too small capacity and range (the final maximum take-off weight was almost three times as high as the first estimates indicated, [ref.27], originally from [ref.19]). Furthermore, when Concorde finally entered service, it was in the middle of the first oil crisis. In the meantime, subsonic engine technology had made major progress with the development of the bypass engine, leading to important reductions in fuel consumption and engine noise for subsonic aircraft. The fact that Concorde, because of the use of reheat during take-off, did not comply with noise regulations and the fact that it had a special status in air traffic control caused it to be merely "tolerated" on some airports, instead of it being fully serviceable on a large number of air routes. With respect to safety regulations too, exceptions were made for Concorde (despite this, Concorde has a remarkable safety record, see Appendix). The final blow came when a public rage developed in the United States over the sonic boom, initially causing the cancellation of British Airways' Middle East line and in the end resulting effectively in a global ban on overland supersonic flight. Finally, only 13 airplanes would be operated by Air France and British Airways (the last ones sold at a symbolic price).

Recently, plans were revealed to carry out minor modifications to Concorde airframes in order to extend their life beyond 2005 ([ref.17]). Apparently, the airframes are in a better condition than expected since Concorde does not suffer from corrosion because of aerodynamic heating of the structure and its cruise altitude in the dry stratosphere.

The failure of the American SST (Supersonic Transport) development is generally ascribed to the fact that the Americans wanted a cruise Mach number of 2.7 or more as opposed to the European studies which put an upper limit of 2.2 on the Mach number from the outset, with respect to aerodynamic heating. In 1964, three companies responded to the invitation of the FAA (Federal Aviation Administration) to present conceptual designs: Boeing, Lockheed and North American. The Boeing studies showed great interest in the swing-wing concept as a compromise between the high leading-edge sweep angles necessary for the specified cruise Mach number and the required span for acceptable take-off and low-speed performance. The swing-wing concept, however, involved a heavy pivot mechanism whereas the aircraft structure already became relatively heavy because of the excessive use of Titanium alloys necessary to provide the required strength and fatigue characteristics for a structure heated by an airflow at almost three times the speed of sound. A well known quote is that "one could either carry the pivot or the passengers, but not both at the same time".

The Lockheed design consisted of a double-delta concept whereas the North American design was equipped with a canard wing to counteract the increase in cruise trim drag caused by the rearward shift of the aerodynamic center during transition to supersonic speeds. When finally only the Boeing company remained in the race, it reopened the configuration issue and came up with three different concepts, variable geometry, canard and double-delta (1966-1969). Finally Boeing favored the "conventional" delta wing plus horizontal stabilizer concept, but the weight problem continued to be a burden. The project was abandoned in 1971.

---

The efforts in the former USSR started with the development of the Myasishev supersonic bombers M-50 and M-52 (*Bounder*) in the 1950's. These designs were equipped with stubby, low-slenderness delta wings and were probably only capable to perform short supersonic bursts. A pragmatic program, mainly based on published (and allegedly some non-published) Western data eventually resulted in the Tupolev Tu-144 *Charger* which first flew in 1968. Although nicknamed "*Concordski*" for its resemblance with Concorde, the design differed in many aspects from its western competitor. The wing of the Tu-144 was not twisted as opposed to Concorde and the aircraft was equipped with four bypass engines with afterburners on both the hot and the cold flow (the Tu-144 produced less noise at take-off than Concorde, which is equipped with straight jets ([ref.18])). Just like the Concorde designers used the Handley-Page-115 and the BAC-221 (a modified version of the Fairey Delta 2), the Russians had a flying testbed to obtain valuable data on the low-speed behavior of delta wings. To facilitate this, a MiG-21 was equipped with a double-delta, designated MiG-21I *Analog* ([ref.7]). In an initial version of the Tu-144, the engines were placed in a group of four under the fuselage. The nose gear was retracted between the pods near the centerline but the main gear struts had to be retracted in the outer parts of the wings causing a large reduction in fuel volume. Much Titanium was used in the structure and it had a very complicated and heavy undercarriage.

In a second version of the design, the engines were placed in two pairs under the wing and each main landing gear strut was retracted between the two pods. Also, retractable canard wings were added to reduce the rotation speed of the airplane. Although this was the main function of the canards, they were also used to pass through the transonic regime more quickly and easily ([ref.3]). It was established that a substantial reduction in drag could be realized if the canards were deployed when the aircraft approached Mach 1 and were retracted after Mach 1.3. Passengers reported that they felt no additional g-loading during transition (which apparently did occur otherwise). Due to the poor L/D ratio it was necessary to use reheat during cruising which rendered the design useless as a commercial airliner. It has been suggested, that the aerodynamic inefficiency of the Tu-144 was partly overcome by the excessive use of Titanium, which was deemed too innovative a material for Concorde, but with which the Russians had much experience ([ref.10]). Two aircraft crashed, one of which during the Bourget airshow of '73 ([ref.2]), probably due to structural failure after a negative g overload ([ref.10]), possibly due to stalling of the canard surfaces after a steep climb-out ([ref.2]).

After the two fatal accidents, the Russians repeatedly asked Concorde officials for help with regard to the major problems they had encountered. Among these were oil leaks, fuel system failures and landing gear axles running hot. Because of the excessive fuel consumption displayed by the Tu-144, the Russians were particularly interested in obtaining Western fuel control units. A \$ 17 million contract with Lucas Aerospace was vetoed in 1977 by the British Government, because the Kuznetsov KN-144 engines also powered the Tu-22M *Backfire* bomber. Among other items of interest on the list submitted by the Russians to the west were anti-icing systems for the engine intakes, fuel system pipes, speed stability control units, warning devices for lightning protection and acoustic loading of airframe and controls and

---

improvements in residual strength of the fuselage after damage and undercarriage life ([ref.4]).

Only operated experimentally on a commercial line in Siberia (Moscow- Alma Ata; 3200 km), the Tu-144 was withdrawn from service in 1983 ([ref.5]).

Recently, NASA revealed plans to use a modified Tu-144 (Tu-144LL) as a research testbed for its High-Speed Research (HSR) program ([ref.13] and [ref.14]). The modification mainly consists of new airborne data sensors and new engines, normally used on the Tu-22M *Backfire* bomber. Roll-out has taken place March 17, 1996 at the Zhukovsky airfield near Moscow ([NASA press release]).

More information about the first supersonic transport design programs in the 60's can be found in [ref.18].

Following the cancellation of the US SST, NASA continued to receive government funding for research on major topics concerning supersonic transport development in order to keep the option open for the future. Main fields of interest were materials, computational fluid dynamics and propulsion technology. These activities resulted in 1972 in the start of the Advanced Supersonic Transport (AST) program, which evolved into the Supersonic Cruise Aircraft Research (SCAR) program in 1973 ([ref.1]). From 1986 NASA and industry were involved in the High Speed Civil Transport (HSCT) study. The results from studies by Boeing and Douglas carried out under the HSCT program are described in [ref.8] and [ref.9]. An interesting feature of these studies is that Boeing consider cruise Mach numbers of 2.4 and 3.2 whereas Douglas in the final phases of their study focus on Mach numbers of 3.2 and even 5.0, despite the fact that the choice of such high Mach numbers is generally regarded as having caused the failure of the American SST development. It is further of interest that these studies were taking the effect of a future SST on the environment into account.

As soon as the renewed interest in a supersonic transport became apparent, it was recognized that in order to reduce the risk for the companies involved, cooperation would be of great importance. From 1986-1989, initiatives to arrange such cooperation between American and European companies were undertaken by the Batelle Institute of Columbus, Ohio, but they eventually failed as a result of the American contractors and NASA being unprepared to be coordinated from the outside ([ref.24]). Other concerns were the protection of company data or anticipated difficulties in working together with other companies at all. Furthermore, some companies expressed concerns that international discussion on HSCT design could trigger premature public reaction and jeopardize NASA efforts to obtain congressional support and funding. Finally it was feared that subsonic sales might suffer if airlines anticipated too much on the availability of a supersonic transport in the near future (proceedings of the European Symposium on Future Supersonic Hypersonic Transportation Systems, held in Strasbourg, 6-8 November, 1989).

From 1990 onwards, NASA has been involved in the High Speed Research (HSR)

---

program, together with the major American airframe and engine manufacturers. The program is split in two phases, the first of which is to take 6 years and cost about \$ 450 million and the second lasting about 8 years spending about \$ 1.5 billion. During this second phase, which has started in 1993 ([ref.11]), overlapping the first phase, a final configuration is to be obtained and an engine contract assigned. At the end of the second phase, that is, in 2001, a decision will be made on launching the aircraft, for which the first flight is envisaged in 2004 and certification in 2005/6. As mentioned previously, during the second phase flying testbeds like the F-16 XL and the Tu-144 will be used.

Also from 1990 on, international cooperation has taken off in the shape of the Supersonic Study Group, initiated by the so-called group of five, consisting of Aerospatiale, British Aerospace, Deutsche Aerospace, Boeing and McDonnell-Douglas, with the intention to cooperate on feasibility studies for a supersonic transport design that is to enter service in 2005. Also associated with this initiative are Fuji, Mitsubishi and Kawasaki Heavy Industries (joined in the Japan Aircraft Development Corporation), Alenia and Tupolev (group of eight). In 1994 a Memorandum of Understanding was signed by DASA (which until then had a separate Memorandum of Understanding with Boeing and McDonnell-Douglas), Aerospatiale and British Aerospace ([ref.6]). This initiative will not compete with the Supersonic Study Group, however.

Despite the intention on paper to cooperate, there is still a wide gap between American and European points of view. This has already caused concern with British Airways and Air France that the project will become completely US dominated, because of the large difference in funding between the United States and Europe. The US Government committed almost \$ 2 billion, [ref.12], NASA financed research for more than \$ 260 million since 1990 and \$ 187 million for 1994 alone ([ref.6] and [ref.21]). For the next eight years investments of more than \$ 650 million are envisaged by NASA. In [ref.6] it is suggested that development costs of a 250 passenger, Mach 2 transport would be between \$ 10 and \$ 15 billion (Concorde total development costs exceeded \$ 4 billion (present-day dollars), [ref.24]). As a contrast, French and British investments have not exceeded \$ 15 million per year on average. The Japanese investment in their supersonic/hypersonic project amounts ¥ 6 billion ([ref.15]).

Based on Concorde experience, all currently undertaken projects depart from the point of view that a second-generation supersonic transport should be able to compete with contemporary subsonic designs. Such competing designs should then include those that achieve increased capacity by enhanced volume instead of cruise speed and those that achieve increased passenger comfort by offering larger seat pitch instead of reduced flight time.

In view of the problem of airport congestion, the alternative of the Very Large Commercial Transport might be favored since it increases transportation capacity at the same number of flights, whereas a Supersonic Transport realizes increased transportation capacity by an increased number of flights. It is also suggested by some

---

sources, that the Very Large Commercial Transport represents a smaller technology jump than the Supersonic Transport, although other sources contradict this. However, in view of the environmental issues and the small design space available for a supersonic transport in combination with the larger uncertainties in the prediction of the critical constraints, the Very Large Commercial Transport option might involve a smaller risk for the companies undertaking the project. It is clear that no company will be able to develop either project completely by itself (an international study group, similar to the Supersonic Study Group, has been founded involving Boeing and Airbus to study the Very Large Commercial Transport), let alone both the SST and the Very Large Commercial Transport (the wisdom of developing two competing concepts is questionable anyway). In [ref.20], development cost of the Very Large Commercial Transport are estimated to amount to \$ 12 billion.

In the summer of 1995, the Very Large Commercial Transport project was put on ice, with the industry's attention focused on stretched versions of the B747 and A340 ([ref.16] and [ref.20]) to fill the capacity gap, pending a final decision on the new concepts. Bearing in mind, that a stretched version of either the Supersonic Transport or the Very Large Commercial Transport is virtually impossible, leaving no margins to respond to changing market predictions, the final decision might be postponed for a long time indeed. With Aerospatiale seemingly favoring the supersonic concept and DASA tending to the Very Large Commercial Transport, European consensus may be even more difficult to accomplish than usual. It is clear, however, that despite huge uncertainties in the market predictions, air traffic will increase, whereas the extent to which the capacity gap can be bridged by stretching existing designs is limited, which eventually will make the choice between the subsonic and the supersonic concept unavoidable.

Most importantly, a future supersonic airplane shall have to comply with the same safety and environmental regulations as contemporary subsonic aircraft, whereas designs requiring special status in air traffic control or major adaptations to existing infrastructure cannot be considered as serious options. This will in advance rule out the use of alternative propellants (hydrogen, methane) and reheat, whereas limits must be imposed on the length and span of the aircraft as well as its take-off field length. All currently undertaken studies further agree that a future supersonic transport should have a capacity of at least 250 passengers and a range of at least 9000 km (that is, transpacific instead of transatlantic), see also [ref.6], [ref.21], [ref.23], [ref.24] and [ref.28]. The pacific routes are considered to be of greatest importance because in this area the rise in air travel is expected to be largest, whereas the pacific region is especially well suited for supersonic flight (almost no overland stretches). The maximum take-off weight is expected to be about 320-360 metric ton.

The major difference in point of view between the contractors (and within companies as well) concerns the cruise Mach number. The choice made by the designers of Concorde, to limit the Mach number to a value of 2.2 was purely based on the maximum acceptable temperature for an Aluminum structure. Today, with the use of modern composites (like carbon reinforced materials) higher Mach numbers might be possible without the weight penalties resulting from the use of Titanium or steel. On

---

the other hand, gains in block-time become very marginal once the cruise Mach number is increased beyond about 3.

Unfortunately, as a result of the increased range and flight time of a future SST in comparison to Concorde, a couple of serious problems occur that force the cruise Mach number to lower values (possibly below 2). The need to supply the passengers with a fresh flow of (cool) air during the entire flight imposes extreme demands on the air conditioning system. It seems unavoidable to use the fuel as a heat sink resulting in a continuously increasing fuel temperature. Furthermore, the fuel is heated because of aerodynamic heating of the skin. Recent calculations by Deutsche Aerospace indicate that after 3-3.5 hours cruising (depending on the Mach number) the fuel temperature has reached a critical value at which a dramatic deterioration of the heat of combustion [J/kg] takes place. According to [ref.26], this limits the cruise Mach number of flights exceeding three hours to about 2.0. The use of an alternative fuel like JP-7 (used in the Lockheed SR-71 *Blackbird*), instead of JP-4 would require larger fuel tanks since the ratio of density and combustion value is less favorable than for JP-4. Besides, the use of a rare type of fuel would require adaptations to airport facilities, which is in conflict with the prerequisite that for a commercially viable supersonic airplane, no exceptions should be made that are not required for competing subsonic designs.

Another reason for the tendency to reduce the cruise Mach number is the sonic boom problem. It is unlikely that supersonic overland flights will ever be permitted whereas efforts to minimize the bow shock overpressure for a given airplane weight, altitude and Mach number by choosing an appropriate cross-section distribution are today considered unrealistical. Although this would indicate that the sonic boom problem should be translated into a subsonic range constraint since it appears to be unavoidable that a certain part of the mission will be flown at subsonic speeds ([ref.24]), it is likely that a limit will be imposed on the maximum bow shock overpressure in future nonetheless, because of expected future regulations concerning the effect of overflying SST's on ships and ocean wildlife.

It is beyond doubt that a future supersonic transport must comply with the same environmental regulations as a competing subsonic design. Especially the noise requirements will be very difficult to fulfil. The feasible design space of a supersonic airplane, which is already inherently very small because of the conflicting demands imposed on the design with respect to subsonic and supersonic operation requirements, is further decreased by the environmental constraints. The small design space available to Concorde (about 128 passengers, range 6000 km) may not be present for a future SST (250 passengers, range 9000 km) once the same requirements are applied as on a subsonic design. Furthermore, since many environmental constraints are not only dictated by science but by politics, they are very susceptible to changes in political climate and public opinion.

The higher the cruise Mach number, the smaller the design space will be and once the margins become smaller than the accuracy of the calculation methods used, the feasibility of the design will become very uncertain, even if the environmental

---



regulations do not change.

It has already been signalled in industry that, in order to have sufficient margin, the cruise Mach number should be chosen between 1.6 and 1.8. Many designers however, have always been opposed to the compromise of flying "a little supersonic", because it will jeopardize the competitiveness of the supersonic transport. It must be acknowledged that a supersonic transport will always have higher Direct Operating Costs than a contemporary subsonic airplane designed according to the same specifications. Furthermore, regulations most likely will follow trends in subsonic technology, for instance in the field of take-off noise. In the end, the economical viability of a supersonic transport plane will be determined by the added ticket price as compared to a subsonic plane and the extent to which the passenger is prepared to pay this price in relation to the reduction in flight time offered. According to officials quoted in [ref.6], the Direct Operating Costs of a second-generation supersonic transport will be 40-50% higher than that of current long-range subsonic airplanes. In [ref.12] a ticket surcharge of only about \$ 150 over a ticket for a flight on a conventional aircraft is expected. Other sources, however estimate a ticket price increase of about 20%. The willingness to pay this extra price may become questionable once cruise Mach numbers are reduced to values as low as 1.6.

From these considerations it becomes clear that the design of a second-generation supersonic transport should be a multivariable and multidisciplinary effort, taking as many interdisciplinary couplings into consideration as possible. As few parameters as possible should be fixed in advance, in order to increase the degrees of freedom of the design and optimization problem (this implies for instance that the powerplant design should be incorporated in the airplane design, instead of choosing an existing engine as is often practice in subsonic airplane design). Only in this way is it possible to obtain not just a feasible design, but the best feasible design.

For more information about past and present supersonic transport projects, including environmental issues, propulsion systems and market assessments, the reader is referred to the proceedings of the European Symposium on Future Supersonic Hypersonic Transportation Systems, held in Strasbourg, 6-8 November, 1989.

### 7.3 References

1. anon.; Proceedings of the SCAR Conference; NASA CP-001, November 1976
  2. Aviation Week & Space Technology; June 11, 1973, pp. 12-23
  3. Aviation Week & Space Technology; May 13, 1974, p. 26
  4. Aviation Week & Space Technology; December 4, 1978, pp. 26,27
  5. Aviation Week & Space Technology; February 14, 1983
  6. Aviation Week & Space Technology; April 11, 1994, pp. 20,21
  7. Belyakov, R.A. and Marmain, J.; *MiG 1939-1989*; Docavia / Éditions Larivière, 1989
  8. Boeing Commercial Airplanes / New Airplane Development; *High Speed Civil Transport Study*; NASA CR 4233/4234, September 1989
-

9. Douglas Aircraft Company; *Study of High-Speed Civil Transports*; NASA CR 4235, December 1989
10. Flight International; September 30, 1978
11. Flight International; 17-23 November, 1993, p. 5
12. Flight International; 15-21 December 1993, pp. 3,10
13. Flight International; 2-8 March, 1994, p. 10
14. Flight International; 7-13 September, 1994, p. 31
15. Flight International; 8-14 December, 1994
16. Flight International; 19-25 July, 1995
17. Flight International; 4-10 October 1995, p. 10
18. Forestier, J.; Lecomte, P. and Poisson-Quinton, P.; *Les Programmes de Transport Supersonique dans les Années Soixante*; Proceedings of the European Symposium on Future Supersonic Hypersonic Transportation Systems, Strasbourg, 6-8 November, 1989
19. Maurin, E. et al.; *Struktureller Aufbau des Überschallverkehrsflugzeuges Concorde*; Luftfahrttechnik-Raumfahrttechnik 12, 1966
20. NRC Handelsblad; Dinsdag 14 November 1995
21. Piellisch, R.; *Mach 2 et plus*; Nouvelle Revue d'Aeronautique et d'Astronautique, No.1, 1994
22. Poisson-Quinton, P.; *First Generation Supersonic Transports*; ONERA T.P. No. 1976-113, 1976
23. Poisson-Quinton, P.; *Future SST's - A European Approach*; Aerospace America, September 1994
24. Reimers, H.D.; *Das Überschallverkehrsflugzeug der 2. Generation-eine zweite Chance ?!*; DGLR Jahrbuch 93-03-029, 1993, pp. 1239-1250.
25. Tardif, T. and Casamayou, J.-P.; *Supersonique du Futur: l'Europe face aux Etats-Unis*; Air & Cosmos/Aviation International No.1492, Vendredi 4 Novembre 1994
26. van der Velden, A.J.M.; *Multi-disciplinary SCT Design Optimization*; AIAA 93-9391, August 1993
27. van der Velden, A.J.M.; *Aerodynamic Design and Synthesis of the Oblique Flying Wing Supersonic Transport*; Ph.D. dissertation, Stanford University SUDAAR 621, June 1992
28. Welge, H.R.; Sutton, J.O. and Metwally, M.; *Engineering Status of the McDonnell Douglas HSCT Program*; SAE Technical Paper 912215, September 1991

## Appendix      A Summary Of Concorde Accidents

A summary of all accidents with Concorde until 1994 is given, based on the information from [ref.1]. In this time six accidents occurred, two with aircraft operated by Air France and four with aircraft operated by British Airways. In two of the latter accidents the same aircraft was involved. Although the damage to the aircraft was in four cases categorized as substantial, in no case were there any casualties and the aircraft in all cases landed safely. It is further noteworthy that, apart from the first case, the accidents either involved tire puncture on the main landing gear or the loss of parts of the rudder during supersonic flight. The six accidents are summarized in

---

the following table:

date	registr.	operator	location	damage to aircraft	cas.	remarks
28-11-'77	-	Air France	Dakar	unknown	none	hard landing
14-6-'79	F-BVFC	Air France	Washington	substantial	none	tire puncture
12-4-'89	G-BOAF	British Airways	Tasman Sea	substantial	none	loss of rudder
4-1-'91	G-BOAE	British Airways	Atlantic Ocean	substantial	none	loss of rudder
21-3-'92	G-BOAB	British Airways	Atlantic Ocean	minor	none	loss of rudder
15-7-'93	G-BOAF	British Airways	Heathrow	substantial	none	tire puncture

registr. : registration

cas. : casualties

**Table 18** Overview of Concorde accidents

**28-11-'77** The aircraft was reported to have made a bad landing at Dakar. The tailwheel crushed and the rear of the engines scraped along the runway over several hundred meters.

**14-6-'79** During the run-up to take-off, after V1, two tires on the left-hand landing gear were destroyed causing the aircraft to suffer sundry damage. After fruitless attempts to retract the landing gear and make a diversion, the plane touched down at Washington without further damage 24 minutes later. The cause of the deflation of tire no. 6 was not established. Tire no. 6 was destroyed by running when flat. The bursting of tire no. 5 was due to running under overload after the destruction of tire no. 6. The sundry structural damage and damage to the circuitry were direct consequences of the projection of debris from the tires and wheels.

**12-4-'89** En route from Christchurch to Sydney, during acceleration to Mach 1.7 and climbing through Flight Level 440 a "thud" was heard by the crew and passengers. Subsequent analysis suggests this was a minor engine surge not connected with the accident sequence. As the aircraft was descending through Flight Level 400 at Mach 1.3, moderate vibration occurred, lasting two to three minutes. This was almost certainly caused by portions of the upper rudder separating from the aircraft. Eventually almost all of the top

wedge of the upper rudder aft of the main spar and above the Power Flight Control Unit attachment structure separated. The aircraft handling was unaffected and an uneventful approach and landing were carried out at Sydney Airport.

The following causal factors were identified:

- 1) The inflight break-up of the bonded honeycomb structure of the upper wedge of the upper rudder occurred as a result of extensive prior delamination of the skin from the honeycomb core.
- 2) Moisture ingress past the rivets in the trailing-edge lead to corrosion between the honeycomb structure and the skin of the upper wedge, and to deterioration of the cohesive bond strength.
- 3) There was incomplete compliance by production staff with the drawing requirements relating to the modification of the rudder trailing-edge.
- 4) The rudder trailing-edge extension modification, as specified in the relevant drawings, placed requirements on production and inspection personnel which were difficult to attain.

4-1-'91 The aircraft departed at 1055 hours from Heathrow to John F. Kennedy Airport, New York. At 1207 hours, while flying at Mach 2, Flight Level 560, the crew felt what they thought was an engine surge. The effect lasted one half to one second and all engine parameters appeared normal. The event marker which signals the Flight Data Recorder was manually pressed by the crew. More unusual vibration was felt during descent and deceleration at 1403 hours at Mach 1.1, Flight Level 410. This lasted for 10 seconds and again the event marker was pressed. No anomalies in any engine or other parameter could later be detected in the Flight Data Recorder traces. However, on landing at New York a portion of the lower rudder and most of the skin on the right side above the Power Flight Control Unit attachment structure was missing. There was also some indentation, scoring and paint smearing on the tailcone almost immediately below the rudder and some damage near the bottom aft corner of the lower rudder where it had been holed.

21-3-'92 A large portion of the rudder separated in flight.

15-7-'93 During the landing roll tire no. 4 (on the right-hand main landing gear) burst due to break seizure. The resulting debris damaged the right wing underside and caused a loss of the "green" hydraulic system. The number 3 engine sustained ingestion damage causing it to stick in reverse.

## Reference

1. Civil Aviation Authority; *World Airline Accident Summary*; Cheltenham, United Kingdom, October 1993
-

# 8. The Mass Prediction Model

## 8.1 Summary

The prediction of the empty weight of an aircraft is always a cumbersome task. It is very important that an accurate value for this weight is available at an early stage of the design process, as it affects all performance criteria. Unfortunately, weight estimations in the conceptual design often have inaccuracies of up to 20%. In engineering practice, this results not seldom in big problems at the final stages of the design. A well known example is Concorde, whose take-off weight increased from about 60 metric ton in the earliest stages of the project to more than 180 metric ton for the final design. A more recent example of bad weight prediction leading to severe performance penalties is the Airbus A320.

The use of more sophisticated methods, like finite element methods does not reduce the problem, as these methods often need details of the structural design that are not yet available at the conceptual stage. Furthermore, these methods would only yield the construction weight, which is only part of the structural weight. Invariably, the finite element methods require the aerodynamical loading of the aircraft components as input, information that cannot be given in detail at the conceptual design stage without unproportional calculation effort.

Therefore, weight prediction methods are statistical or semi-empirical with the mentioned small accuracy. For supersonic cruising airplanes the problem is even bigger as no statistical information is available since only one commercial supersonic airplane is operational at present. Of military airplanes, only bombers and strategic reconnaissance planes would give comparable information, but military sources are usually not freely accessible.

Since one of the targets of multidisciplinary optimization is to reduce costly feed-backs in the design process like those resulting from bad weight predictions, it is important to be at least able to model sensitivities of the empty weight to the design variables accurately. As most design variables in this study concern the wing design, it was decided to use a more sophisticated weight prediction method for the wing weight [ref.4]. The other weight contributions are estimated with the use of the statistical and semi-empirical methods of [ref.3] which are derived for subsonic aircraft. Since the cruise Mach number in this study was limited to 2.5, it is assumed that the construction material will be Aluminum, which is what the subsonic prediction methods are based on. The author is aware that this represents technology of the 70's whereas the use of modern composite materials will enable important weight reductions for a second-generation supersonic transport. However, no statistical information is available with respect to such materials. Because of this, the present mass prediction method will be rather conservative.

The weight of many components depends on the final maximum take-off weight which necessitates the use of an iterative procedure. In this application a method based on the regula-falsi method was used, because of its guaranteed convergence.

---

The weight at the end of a calculation cycle is substituted in the next cycle until two values are obtained between which the solution is located. From that point on the regula-falsi procedure is applied. A disadvantage of the regula-falsi method is the fact that it proceeds rather slow under some circumstances. This problem is partly overcome by an additional requirement that the procedure should make a progress of at least 10% (measured relative to the distance between the two bordering values).

## 8.2 Notation

A	coefficient for landing gear weight estimation
B	coefficient for landing gear weight estimation
C	coefficient for landing gear weight estimation
D	coefficient for landing gear weight estimation
d	diameter
f	function
$K_{pg}$	factor for propulsion group weight estimation
$K_{sc}$	factor for surface control group weight estimation
$K_{uc}$	factor for undercarriage weight estimation
$K_{wf}$	factor for fuselage weight estimation
l	length
$\dot{m}$	massflow
n	number
S	area
T	thrust
V	speed
W	weight (for the sake of tradition, weights are given in kg)
$W_0$	maximum take-off weight
$\epsilon$	compressor pressure ratio
$\Lambda$	sweep angle
$\lambda$	bypass ratio

### subscripts

ase	airframe services and equipment group
cr	cruise
eng	engines
f	fuselage
pax	passengers (payload)
pp	powerplant
sc	surface control group
t	tail (plane)
to	take-off
uc	undercarriage
w	wing

wet   wetted

In accordance with tradition, weights are given in kilograms, although -according to the SI system- the unit of weight is Newton and the unit of mass is kilogram.

### 8.3 Wing Weight Estimation

The wing structural weight according to [ref.4], is divided into two components: the primary weight, which includes the weight of all material between and including the forward and the rear wing spar and the secondary weight, which comprises the high-lift devices, spoilers, ailerons and miscellaneous weight contributions. The primary weight consists of the basic construction weight needed to resist shear and bending loads plus weight penalties for the nonoptimality of the structure, which includes cut-outs, engine and undercarriage mountings etc. To be able to resist flutter, an extra torsional stiffness requirement is added to the primary weight.

For the calculation of the basic weight, a weight relief factor is included, which accounts for the bending load relief due to the wing weight itself, wing-mounted loads such as engines and fuel. The fuel bending moment reduction is only applied in case gust loads are more critical for bending than manoeuvre loads. The critical condition for the gust case is assumed to be the cruise condition. To calculate the weight relief due to the engines, it is necessary to know the spanwise position of the engines. In the present study, it is assumed that this position depends on the length of the main gear struts. This length is determined by assuming a main gear position just aft of the most rearward center of gravity during ground operation. The main gear strut length can then be determined by the requirement that the airplane can be rotated over at least 12 degrees without the risk of tailsrape. Since the most rearward center of gravity position during ground operation depends on the results of the mass calculations, it is necessary to perform these calculations twice (the influence of the spanwise engine position on the calculations is small enough to make further iteration unnecessary).

The weight relief due to the wing weight itself depends on the wing weight ratio and would therefore require an additional iteration.

The most important progress in multivariate optimization still can be achieved in reducing the CPU time required for a single design analysis. This is especially important in case global/non-gradient-guided search algorithms are used. This implies, that a sensible trade-off should be made between the accuracy of the analysis and the time required to perform it. It was established that the mass prediction method as outlined previously was extremely time consuming. The influence of the wing weight ratio iteration on the weight prediction is very small. Since for each wing weight ratio iteration cycle a complete maximum take-off weight iteration has to be performed the influence on the total required CPU time is relatively large. Therefore, this iteration was removed from the procedure. Instead, the wing weight ratio was estimated by a simple empirical relation as suggested in [ref.4].

---

Also, the resubstitution of the spanwise location of the engines into the mass prediction procedure was not retained. The accuracy of the value for the weight relief due to the engines thus obtained is doubtful, whereas the procedure doubles the required CPU time for the mass calculations. Instead, a typical value for the weight relief for four-engined airplanes, as suggested in [ref.4] was used.

Finally, the mass prediction method depends on the cruise altitude. An initial estimate for this value is made using the untrimmed lift coefficient, since the center of gravity locations are first known after the mass prediction calculations are completed. Instead of resubstituting the now known center of gravity locations and solve the cruise altitude iteratively, it was established that sufficient accuracy was obtained without a time consuming iterative procedure.

The major drawback in the application of the method, is the fact that the bending moments are based on an integrated lift distribution and therefore require the lateral center of pressure position. For a supersonic cruising airplane, the critical value of this parameter is very difficult to assess, especially when partial vortex flow occurs. Therefore, a simple approximation was used, valid for straight tapered wings at low speeds, which is the mean of a chord proportional and an elliptic pressure distribution. This assumption could, however, result in relatively large errors in the prediction of the basic wing weight.

Another point of concern is the possibility that the critical load case may change during the iterative procedure to determine the take-off weight. Beyond a certain weight, the manoeuvre case becomes critical, resulting in an abrupt change of weight. In this case it is impossible to obtain a solution. The regula-falsi method is capable, however, to recognize the occurrence of this problem since it will "jump over the solution". The iteration will be cut off at the moment that the wing weight is overestimated (see Figure 18, right-hand-side of the solution, overestimation of maximum take-off weight). As can be seen in Figure 18, the difference between the converged solution for critical gust loads and that for critical manoeuvre loads can be quite substantial, as a result of the small slope of the gust-critical line and the large increase in weight once the manoeuvre case becomes critical. This shape of Figure 18 also increases the earlier mentioned problem of the slow convergence of the regula-falsi method, which -as indicated before- is compensated by requiring at least 10% progress. If the progress is smaller, than the next value for the take-off weight is chosen in the middle of the two bordering values.

## 8.4 Maximum Take-off Weight

The other weight components are determined using the statistical and semi-empirical formulas of [ref.3].

---



The weight of the powerplant installation is related to the take-off performance of the engine by the following relation:

$$W_{pp} = n_{eng} K_{pg} T_{to} \left( \frac{10 \epsilon_{to}^{0.25} \dot{m}_{to}}{T_{to} (1 + \lambda_{to})} + 0.12 \left( 1 - \frac{1}{\sqrt{1 + 0.75 \lambda_{to}}} \right) \right) \quad (132)$$

in which  $K_{pg}$  equals 1.15 for podded jet engines.

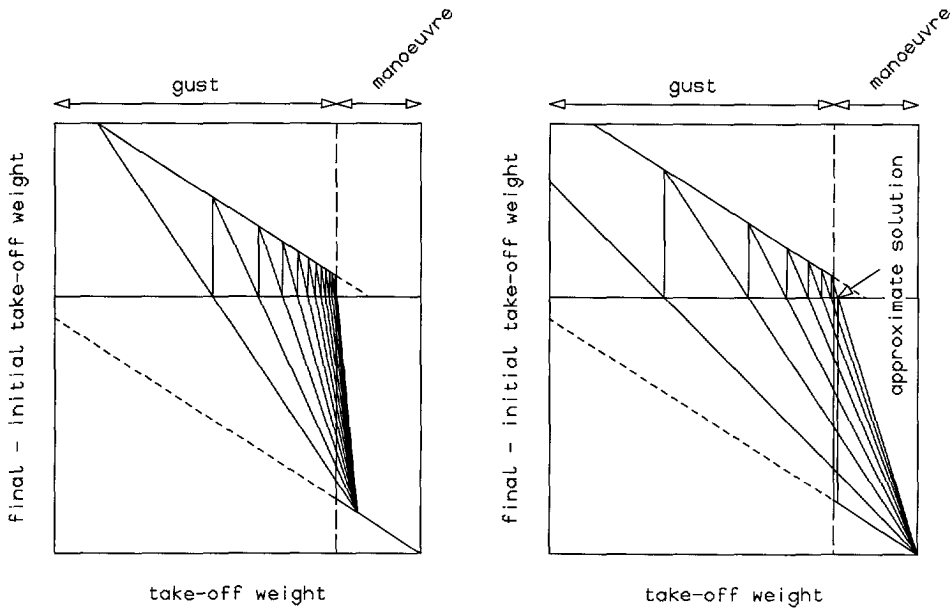


Figure 18 Regula-falsi procedure for Maximum Take-off Weight solution

The fuel weight is directly calculated from the available tank volume, whereas the payload weight is calculated by the following simple relation [kg]:

$$W_{pax} = 700 + 95 n_{pax} \quad (133)$$

These two weight contributions, payload and fuel, can therefore be calculated directly from the geometric variables since they are not dependent on the final take-off weight (the powerplant weight only depends on the final take-off weight through the take-off Mach number. In case a representative Mach number would be used instead, for instance Mach 0.2, this weight contribution can be calculated directly from the design

variables as well). All other components must be computed iteratively as explained before.

The fuselage weight is computed using the following equation [kg]:

$$W_f = 1.04 K_{wf} \sqrt{\left( \frac{V_{cr} l_t}{2d_f} \right)} S_{wet_f}^{1.2} \quad (134)$$

in which  $K_{wf} = 0.23$ . The cruise speed in this equation should be given as Equivalent Airspeed and is therefore dependent of the maximum take-off weight (cruise altitude). The presented equation is valid for fuselages with pressurized cabins, with no engine or undercarriage attached to the fuselage.

The weight of the tailplanes [kg] is calculated as a function of the parameter

$$\frac{S^{0.2} V_{cr}}{\sqrt{\cos \Lambda_{1/2}}} :$$

$$W_t = S_t f \left( \frac{S_t^{0.2} V_{cr}}{\sqrt{\cos \Lambda_{1/2_t}}} \right) \quad (135)$$

Again, the cruise speed should be given as Equivalent Airspeed.

The weight of the undercarriage [kg] is computed by the following relation:

$$W_{uc} = K_{uc} (A + B W_0^{3/4} + C W_0 + D W_0^{3/2}) \quad (136)$$

in which  $K_{uc} = 1.0$  for low-wing configurations. The coefficients A, B, C and D attain different values for the nose and the main gear, as indicated in the following table:

	A	B	C	D
main gear	18.1	0.131	0.019	$2.23 \cdot 10^{-5}$
nose gear	9.1	0.082	-	$2.97 \cdot 10^{-6}$

Table 19 Coefficients for landing gear weight prediction

For simplicity, the weight group referred to as airframe services and equipment in [ref.3], is not broken down into the different components but taken as 10% of the maximum take-off weight instead.

Finally the weight of the surface control systems is taken as:

$$W_{sc} = 0.59 W_0^{2/3}$$

(137)

The method was applied to Concorde, taking the actual values for the fuel weight and the payload weight. The actual values for Concorde as presented in the following table were deduced from data published in [ref.1] and [ref.2]. The method to calculate the fuselage weight depends on the tail length (tail load). Since Concorde is not equipped with a horizontal stabilizer, the distance between the elevon semi-chord and the wing semi-chord points was taken for this value. The weight of the fin is assumed to be 2% of the maximum take-off weight, whereas the weight of the powerplant installation is 12.6% of the maximum take-off weight (percentage based on Concorde data). As explained previously, the weight of the airframe services and equipment is taken as 10% of the maximum take-off weight.

Component	Computed [kg]	Actual [kg]	Error (%)
wing	15808	13733	+15.1
powerplant	23805	23496	+1.3
fuselage	11401	11748	-3.0
payload	12700	12700	-
fuel	92985	92985	-
tail	3779	3255	+16.1
landing gear	7622	7224	+5.5
controls	1942	19766	+5.4
equipment	18893		
total	188935	184907	+2.2
The equipment contribution was fixed at 10%, the powerplant weight at 12.6% and the fin weight at 2% of the maximum take-off weight.			

Table 20    Comparison between computed and actual Concorde weight breakdown

This result seems to be very encouraging, although it was slightly manipulated by varying the percentage of the equipment contribution, which should be between 8% and 12% of the maximum take-off weight. As explained, there is hardly any statistical data available for validation of a supersonic mass prediction method, and this example does not pretend to be a validation. It is merely an illustration of the effect of errors in the estimation of component weights on the total error.

As indicated before, the method is rather conservative, being based on subsonic technology of the 1970's. This will reflect in the final result. It might be possible to compensate the method for some state of the art technology but it is questionable whether this will much enhance the accuracy, since too much insecurities still remain. Among these are the percentage of new materials that will be used in the construction, the question if full leading-edge flaps will be needed, the extra structural weight for the windows with respect to decompression safety regulations, undercarriage weight, wiring weight (fly-by-wire), the weight of a stability augmentation system and possibly the weight of noise suppression systems. Furthermore, a correction of certain mass contributions would be based on either a very rough guess or require too much detail in proportion to other components. Therefore, the presented method was accepted as a means to predict the influence of the design variables on the aircraft weight, if not as an accurate weight prediction method in terms of absolute values. It is quite likely that the very small design space as observed in chapter 15 is partly caused by this overly conservative weight prediction method.

## 8.5 References

1. anon.; *Jane's All the Worlds's Aircraft*; Jane's Information Group Limited, U.K.
  2. Mizuno, H.; Hagiwara, S.; Hanai, T. and Takami, H.; *Feasibility Study on the Second Generation SST*; AIAA-91-3104, 1991
  3. Torenbeek, E.; *Synthesis of Subsonic Airplane Design*; Delft University Press, Kluwer Academic Publishers, 1988
  4. Torenbeek, E.; *Development and Application of a Comprehensive, Design-sensitive Weight Prediction Method for Wing Structures of Transport Category Aircraft*; Report LR-693, September 1992, Delft University of Technology
-

# 9. The Aerodynamics Model

## 9.1 Summary

In this chapter, based on the Polhamus leading-edge suction analogy and the Carlson attainable thrust concept, closed-form relations are derived which provide a means to calculate the lift and drag coefficient of flat swept or delta wings. The partial loss of the leading-edge suction force caused by flow separation and the resulting occurrence of leading-edge vortices have been taken into account. The method is valid both in subsonic and supersonic flow. The major part of this chapter was published as [ref.2].

## 9.2 Notation

A	aspect ratio ( $b^2/S$ )
b	wing span
$C_D$	drag coefficient (3D)
$C_L$	lift coefficient (3D)
$C_{L_\alpha}$	lift gradient (3D)
$C_N$	normal force coefficient (3D)
$C_S$	suction force coefficient normal to leading-edge (3D)
$C_T$	thrust (suction force) coefficient (3D)
$C_X$	tangential force coefficient (3D)
c	chord
$c_d$	drag coefficient (2D)
$c_l$	lift coefficient (2D)
$c_{l_\alpha}$	lift gradient (2D)
$c_n$	normal force coefficient (2D)
$c_p$	pressure coefficient
$(\Delta c_p \sqrt{x})_0$	limiting value of leading-edge pressure coefficient
$c_s$	suction force coefficient perpendicular to leading-edge (2D)
$c_t$	local thrust (suction force) coefficient
D	drag
E	elliptical integral of the second kind
i	incidence
$K_A$	attainable thrust factor
$\bar{K}_A$	average attainable thrust factor over span
$K_T$	leading-edge thrust factor
$K_V$	vortex flow factor
$K_i$	induced downwash factor
$K_p$	lift gradient in potential flow
$K_w$	lift-induced supersonic drag factor
$k_1 - k_4$	coefficients in airfoil section definition

---

L	lift
l	lift per unit span
$l_i$	spanwise location of wing kink measured from root
M	Mach number
$M_e$	equivalent Mach number for vacuum limiting pressure
m	leading-edge parameter
N	normal force
n	normal force per unit span
R	Reynolds number based on chord
r	leading-edge nose radius
S	wing area
s	semi net span
T	thrust force
t	local thrust force
	tangential force per unit span
	wing thickness
U	free stream velocity
$w_i$	induced downwash velocity
x	chordwise airfoil section coordinate
y	spanwise coordinate
	airfoil section thickness coordinate
$\alpha$	angle of attack
$\alpha_i^*$	fuselage angle of attack for zero fuselage lift
$\Gamma$	circulation strength
$\gamma$	specific heat ratio of air
$\eta$	exponent in approximation of elliptic integral of second kind
$\Lambda$	sweep angle
$\mu$	Mach angle
$\rho$	density of air
$\phi$	angle in definition of elliptical integral
	wing span efficiency factor

### subscripts

0	at leading-edge of airfoil section
f	fuselage
i	inner
	induced
le	leading-edge
n	normal to wing leading-edge in wing plane
o	outer
p	potential
v	vortex
w	wing
wf	wing-fuselage

### 9.3 Introduction

Since the design of airplanes capable of cruising at supersonic speeds involves finding a compromise between the contradicting performance requirements at high-speed flight and low-speed flight, it is very important to be able to predict the aircraft's characteristics accurately throughout the operational envelope. Therefore, especially in case of wave drag calculations on configurations somewhat more complicated than the pure delta wing, more advanced aerodynamic models than handbook methods like the USAF Stability and Control DATCOM ([ref.8]) provide, are often used. Also, in the case of wings with a supersonic leading-edge, in those regions where Evvard's principle applies (the area between the Mach line from the tip leading-edge, the tip chord and the trailing-edge), the pressure distribution differs from the two-dimensional case (constant loading) which applies elsewhere on the section with the supersonic leading-edge. In order to account for this effect in the lift and drag calculations, a method based on the near-field integration of the perturbation potential, for instance a panel method, is needed. Furthermore, as far as the subsonic case is concerned, many handbook methods have a limited applicability, usually with respect to moderate sweep angles and not too small aspect ratios.

Unfortunately, even the more advanced computational aerodynamic models do not fully account for an effect that is very important in the analysis of wings that are especially designed for high speeds. In subsonic flow a leading-edge suction force occurs on leading-edges that are sufficiently blunt. Depending on the leading-edge sweep angle of the wing, this force can be retained in supersonic flow, provided that the leading-edge flow remains subsonic. As, in supersonic cruise, lift-to-drag ratio's tend to be relatively low because of the presence of wave drag, this can be very much worthwhile for vehicles designed for long range.

If the leading-edge radius of the wing is small, however, the suction force cannot be generated as the flow around the nose will separate, resulting in strong stable vortices over the wing. This will especially occur at high angles of attack, for instance during the landing approach, where the vortex flow will provide a large amount of extra lift, thus enabling a lower approach speed and accordingly shorter landing field length. As the aspect ratio of the wing for supersonic cruising airplanes is usually relatively small, leading to high approach angles of attack, this is an important effect. The disadvantage of the development of vortex flow is, however, that it results in a large increase in drag, as on the one hand the suction force disappears, while on the other hand an extra component of induced drag is generated.

At present, the only methods that are capable of taking these effects into account are specially adapted panel methods or Euler methods. Since these methods require far too much computational time to be suitable in the conceptual stage of the design process, other means have to be employed to take this phenomenon into account, which is very important for supersonic airplanes. The purpose of the present chapter is to summarize available methods from literature, especially from [ref.4], [ref.5], [ref.14] and [ref.15] and to use these in combination with a panel method used at the

---

Delft University of Technology, the NLRAERO program ([ref.9], [ref.10] and [ref.13]).

## 9.4 Leading-edge Suction Analogy

A very useful -though not entirely physically supported- theory to predict drag and lift due to leading-edge vortices is the so-called leading-edge suction analogy, introduced by Polhamus ([ref.14] and [ref.15]). The basis for this theory is the assumption, that the force required to maintain the reattached flow behind the vortex is equal to the force required to maintain the attached potential flow around the leading-edge. If this is the case, then the lift and drag due to the vortex flow can be found by calculating the leading-edge suction force as it occurs in potential flow and rotating it 90 degrees upward. It has been established that this theory accounts very well for the strong non-linear behavior of the lift coefficient as observed in wind-tunnel tests with highly swept, sharp leading-edge wings.

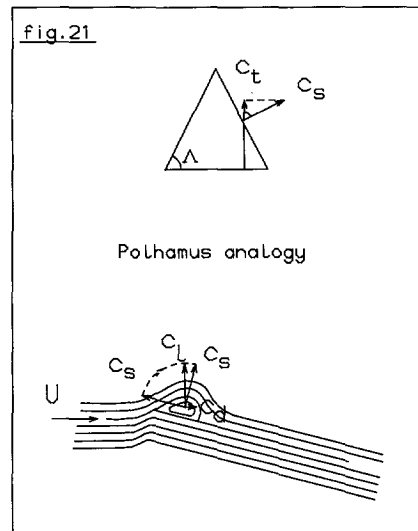
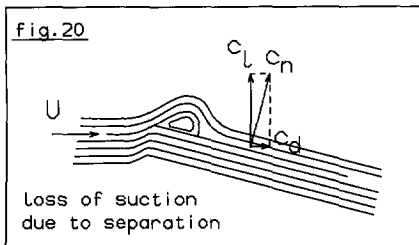
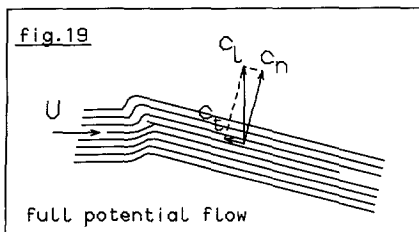


Figure 19, 20 and 21 Polhamus leading edge suction analogy

## Subsonic Flow

In incompressible two-dimensional potential flow, using Blasius' theorem, the following two expressions can be derived for the tangential and the normal force per



unit span, acting on a flat plate:

$$t = \rho U \Gamma \sin \alpha \quad (138)$$

$$n = \rho U \Gamma \cos \alpha \quad (139)$$

The circulation  $\Gamma$  follows from the Kutta-condition, which ensures a smooth, attached trailing-edge flow:

$$\Gamma = \pi U c \sin \alpha \quad (140)$$

Thus, the lift per unit span equals:

$$l = \rho U \Gamma \quad (141)$$

or:

$$c_l = 2\pi \sin \alpha \quad (142)$$

In incompressible two-dimensional potential flow, the drag force equals zero (d'Alembert's paradox, Figure 19).

If, due to separation at a sharp leading-edge the suction force disappears, the following results are obtained (Figure 20):

$$c_n = \frac{\rho U \Gamma \cos \alpha}{\frac{1}{2} \rho U^2 c} = 2\pi \sin \alpha \cos \alpha \quad (143)$$

$$c_l = c_n \cos \alpha = 2\pi \sin \alpha \cos^2 \alpha \quad (144)$$

$$c_d = c_n \sin \alpha = 2\pi \sin^2 \alpha \cos \alpha \quad (145)$$

Hereafter, these expressions will be given the subscript p, as they represent the potential-flow contribution.

According to the Polhamus leading-edge suction analogy, the contribution to drag and lift due to the vortex flow, can be derived by taking the theoretical potential-flow suction force that acts perpendicular to the leading-edge and rotate it 90 degrees upward (Figure 21):

$$c_{l_v} = c_s \cos \alpha = \frac{c_t}{\cos \Lambda} \cos \alpha = \frac{2\pi}{\cos \Lambda} \sin^2 \alpha \cos \alpha \quad (146)$$

$$c_{d_v} = c_{l_v} \tan \alpha \quad (147)$$

In three-dimensional flow, the drag is not zero. Due to the trailing vortex wake, a downwash velocity is induced, normal to the chord. This will cause a reduction of the magnitude of the suction force, according to:

$$T = \rho \Gamma b (U \sin \alpha - w_i) \quad (148)$$

The circulation strength is still determined by the Kutta-condition:

$$\Gamma = \frac{K_p}{2} U \frac{S}{b} \sin \alpha \quad (149)$$

Comparison with the expressions for two-dimensional potential flow shows that  $K_p$  is equal to the lift gradient  $C_{l_\alpha}$ .

Substitution yields for the suction force coefficient:

$$C_T = \left( 1 - \frac{w_i}{U \sin \alpha} \right) K_p \sin^2 \alpha \quad (150)$$

As the downwash depends on the circulation distribution only, which according to the Kutta-condition is determined by  $U \sin \alpha$ , the following relation may be introduced:

$$\frac{w_i}{U \sin \alpha} = K_p K_i \quad (151)$$

Then the following expression is found for the suction force coefficient in three-dimensional potential flow:

$$C_T = (K_p - K_p^2 K_i) \sin^2 \alpha \approx \left( \frac{\partial C_T}{\partial \alpha^2} \right) \alpha^2 \quad (152)$$

The three-dimensional vortex contribution to the lift coefficient can then be expressed as:

$$C_{L_v} = K_v \sin^2 \alpha \cos \alpha \quad (153)$$

with:

$$K_v = \left( \frac{\partial C_T}{\partial \alpha^2} \right) \frac{1}{\cos \Lambda_{le}} \quad (154)$$

The induced drag in three-dimensional potential flow is equal to:

$$C_{D_i} = C_N \sin \alpha - C_T \cos \alpha \quad (155)$$

The ratio of these two terms is shown in Figure 22 which stresses the importance of retaining the leading-edge suction force. It is obvious that in subsonic flow, the suction force can be quite substantial and cause a large reduction of the induced drag. In supersonic flow, on the other hand, although the suction force in comparison to the potential flow contribution to the drag is not as large, very significant reductions of the induced drag can be realized through geometrical adaptations.

With the previously deduced expressions, equation (155) may be written as:

$$C_{D_i} = C_{L_\alpha} \sin^2 \alpha \cos \alpha - C_T \cos \alpha \quad (156)$$

or:

$$C_{D_i} \approx \alpha^2 C_{L_\alpha} - C_T = C_L^2 \left( \frac{1}{C_{L_\alpha}} - \frac{C_T}{C_L^2} \right) \quad (157)$$

using  $C_L \approx C_{L_\alpha} \alpha$ , which applies in linearized flow. As, in subsonic flow, the induced

drag of flat wings can be expressed as  $C_{D_i} = \frac{C_L^2}{\pi A \phi}$ , the following results can be obtained ([ref.11]):

$$C_T = \alpha^2 \left( C_{L_\alpha} - \frac{C_{L_\alpha}^2}{\pi A \phi} \right) \quad (158)$$

$$\Rightarrow \frac{\partial C_T}{\partial \alpha^2} = C_{L_\alpha} - \frac{C_{L_\alpha}^2}{\pi A \phi} \quad (159)$$

$$\Rightarrow K_v = \left( C_{L_\alpha} - \frac{C_{L_\alpha}^2}{\pi A \phi} \right) \frac{1}{\cos \Lambda_{le}} \quad (160)$$

If the subsonic lift gradient and the induced drag or the span efficiency factor  $\phi$  have been determined either from a panel method or from a simple approximating expression, the vortex coefficient can be calculated.

### Supersonic Flow

In [ref.7] it is concluded that although leading-edge forces are not large in supersonic flow, they can have a significant influence on the design. Especially the severely twisted and cambered designs resulting from methods that do not take the leading-edge thrust into account could be replaced by more moderate surfaces once it is acknowledged that the leading-edge thrust does not equal zero.

For supersonic flow, an expression for the local suction force coefficient is given in [ref.4] (originally from [ref.3]) for flat delta wings:

$$c_i = \frac{\pi b}{8 S} \tan \Lambda_{le} \sqrt{1 - m^2} (\Delta c_p \sqrt{x})_0^2 \quad (161)$$

based on the mean geometric chord, in which the last term represents the limiting value of the pressure coefficient on the leading-edge. An empirical relation to estimate

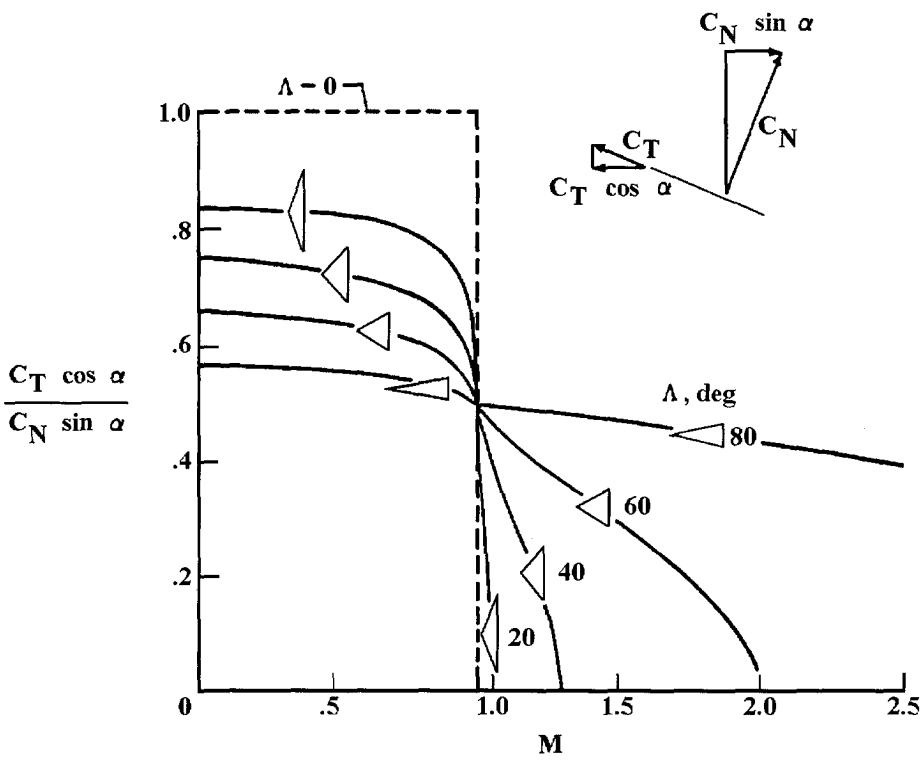


Figure 22 Normal force and suction force contribution to induced drag ([ref.4])

this value was given as:

$$(\Delta c_p \sqrt{x})_0 = \frac{4 \sin \alpha \sqrt{y \cot \Lambda_{le}}}{\sqrt{2} E} \tag{162}$$

in which E is the elliptical integral of the second kind defined as:

$$E(m) = \int_0^{\pi/2} \sqrt{1 - (1 - m^2) \sin^2 \varphi} \, d\varphi \tag{163}$$

According to [ref.4], this integral can be approximated as:

$$E = 1 + \left( \frac{\pi}{2} - 1 \right) m^\eta, \quad \eta = 1.226 + 0.15\pi(1 - \sqrt{m}). \quad (164)$$

Substituting equation (162) into equation (161) yields for the local suction force coefficient:

$$c_t = \frac{\pi A \sqrt{1-m^2}}{E^2} \frac{y}{b} \sin^2 \alpha \quad (165)$$

Integration over the wing span yields the following expression for the supersonic suction force on delta wings:

$$C_r = \frac{\pi A}{4E^2} \sin^2 \alpha \sqrt{1-m^2} \quad (166)$$

The relation for the limiting value of the pressure coefficient is based on theoretical results. In reality, it was found that these theoretical values often overestimate the suction forces. In order to cope with this problem, Carlson in [ref.5] introduced a method to correlate theoretical values of subsonic and supersonic suction force coefficients with results of wind-tunnel experiments. This method has become known as the Carlson attainable thrust concept, which will be explained later.

An expression for the vortex coefficient in supersonic flow can be derived from equations (154) and (166):

$$K_v = \frac{\pi A}{4E^2} \frac{\sqrt{1-m^2}}{\cos \Lambda_{le}} \quad (167)$$

For a delta wing this becomes:

$$K_v = \frac{\pi A}{4E^2} \sqrt{1-m^2} \sqrt{1 + \frac{16}{A^2}} \quad (168)$$

The parameter  $m = \frac{\cot \Lambda_{le}}{\tan \mu}$ , with  $\mu$  as Mach angle. For sonic leading-edges, the

parameter  $m$  becomes 1.

As stated before, the potential flow coefficient  $K_p$  is equal to the lift gradient. In supersonic flow, according to [ref.16], this coefficient can be taken as:

$$K_p = \frac{2\pi m}{E\sqrt{M^2-1}} \quad (169)$$

in which  $E$  is the previously defined elliptical integral of the second kind.

For a sonic leading-edge (or supersonic, as long as  $m$  is kept equal to 1), this relation changes into:

$$K_p = \frac{4}{\sqrt{M^2-1}} \quad (170)$$

which is in accordance with Ackeret's linearized supersonic theory. For a delta wing the relation for  $K_p$  becomes:

$$K_p = \frac{\pi A}{2E} \quad (171)$$

In order to obtain the lift and drag due to lift coefficient in a flow with leading-edge vortices, the contributions from the potential flow without leading-edge suction and the Polhamus leading-edge suction analogy are to be added:

$$C_L = C_{L_p} + C_{L_v} = K_p \sin \alpha \cos^2 \alpha + K_v \sin^2 \alpha \cos \alpha \quad (172)$$

$$C_D = C_L \tan \alpha = K_p \sin^2 \alpha \cos \alpha + K_v \sin^3 \alpha \quad (173)$$

## 9.5 Attainable Thrust Concept

According to Carlson ([ref.4], [ref.5], [ref.6] and [ref.7]), for a family of symmetric airfoils defined by  $y = \pm(k_1\sqrt{x} + k_2x + k_3x^{3/2} + k_4x^2)$  an expression can be derived for a

factor  $K_A$  which is the ratio between the attainable and the theoretical local suction coefficient:

$$C_{t_A} = K_A \cdot C_t \quad (174)$$

In Figure 23 a typical pressure distribution is sketched on the left-hand-side for an airfoil at large angle of attack. The theoretical pressure peak can become quite large, often in excess of the vacuum pressure coefficient, which is obviously unrealistic. In case the pressure coefficient is limited, the picture as sketched at the right-hand-side of Figure 23 remains.

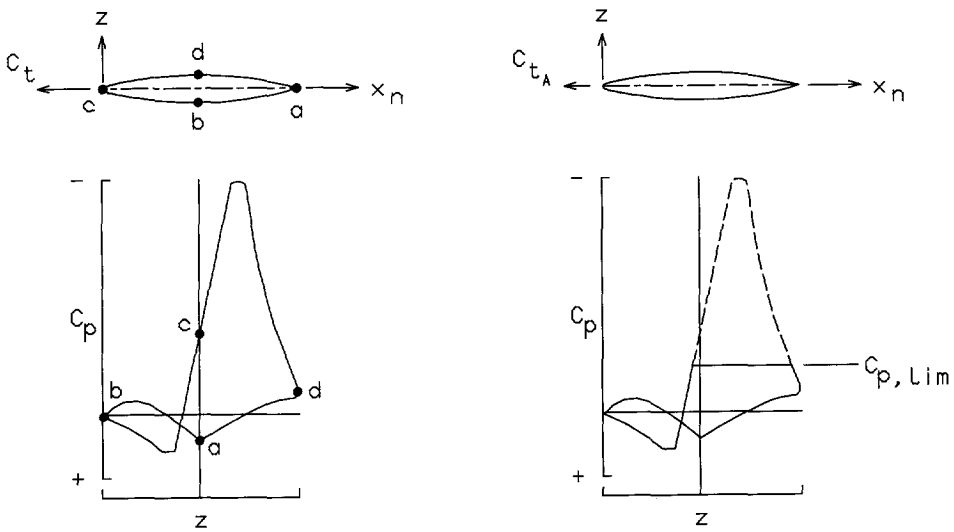


Figure 23 Typical pressure distribution of airfoil at large angle of attack ([ref.5])

According to [ref.6], limiting the pressure peak to practically achievable pressures has a relatively insignificant effect on the normal force but it severely limits the value of the theoretical suction force, which will therefore be corrected according to equation



(174). Carlson and coworkers found that this attainable thrust coefficient correlated well with the parameter:

$$\frac{c_{t_n} \sqrt{1 - M_n^2}}{\left( \frac{t}{c} \right)_n \left( \frac{r}{c} \right)_n^{0.4}} \quad (175)$$

For the attainable thrust coefficient the following relation was found ([ref.5]):

$$K_A = \frac{2(1 - M_n^2)}{M_n} \left[ \frac{\left( \frac{t}{c} \right)_n \left( \frac{r}{c} \right)_n^{0.4}}{c_{t_n} \sqrt{1 - M_n^2}} \right]^{0.6} \quad (176)$$

This relation is depicted in Figure 24.

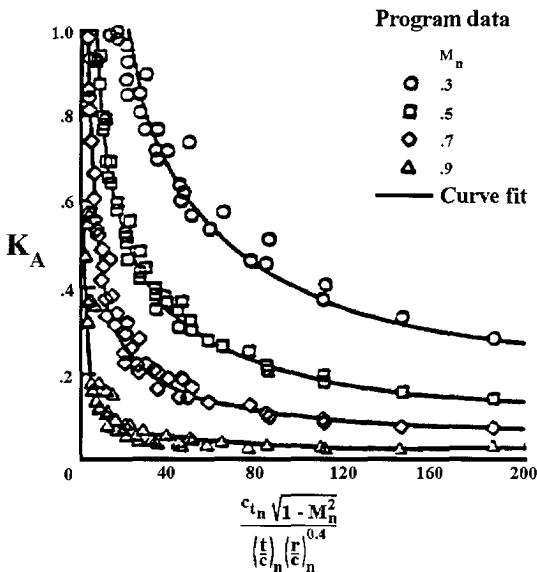


Figure 24 Attainable thrust factor ([ref.5])

The subscripts  $n$  refer to parameters normal to the leading-edge at a selected spanwise station of the original wing. The reasoning behind this is that the flow

conditions normal to the wing leading-edge are considered to determine the leading-edge flow behavior. To obtain these parameters, a phantom wing is superimposed over the original wing in such a way that the leading-edge and trailing-edge sweep angles are equal to those of the original wing (see Figure 25, taken from [ref.5]). The maximum thickness line of the phantom wing is chosen such, that it runs through the maximum thickness location of the original wing at the selected spanwise location, while the maximum thickness of the normal section of the phantom wing is placed at midchord. The required airfoil parameters follow by means of simple geometrical relations, which can be found in [ref.5].

Equation (176) was derived for a vacuum limiting pressure. It can be extended to more realistic limiting pressures, as experienced in wind-tunnel tests, by means of the Prandtl-Glauert rule: the vacuum pressure which occurs at a certain equivalent Mach number  $M_e$ , can be related to an arbitrary limiting pressure at the normal Mach number under consideration, subject to the following condition:

$$c_{p_{lim}}(M_e) = c_{p_{lim}}(M_n) \frac{\sqrt{1-M_n^2}}{\sqrt{1-M_e^2}} \quad (177)$$

If the limiting pressure coefficient at the equivalent Mach number is set equal to the vacuum pressure coefficient, the following relations hold:

$$K_A = \frac{2(1-M_e^2)}{M_e} \left[ \frac{\left(\frac{t}{c}\right)_n \left(\frac{r}{c}\right)_n^{0.4}}{c_{t_n} \sqrt{1-M_n^2}} \right]^{0.6} \quad (178)$$

with:

$$M_e = \frac{-\sqrt{2}}{\gamma c_{p_{lim}} \sqrt{1-M_n^2}} \sqrt{\sqrt{1 + \left(\gamma c_{p_{lim}} \sqrt{1-M_n^2}\right)^2} - 1} \quad (179)$$

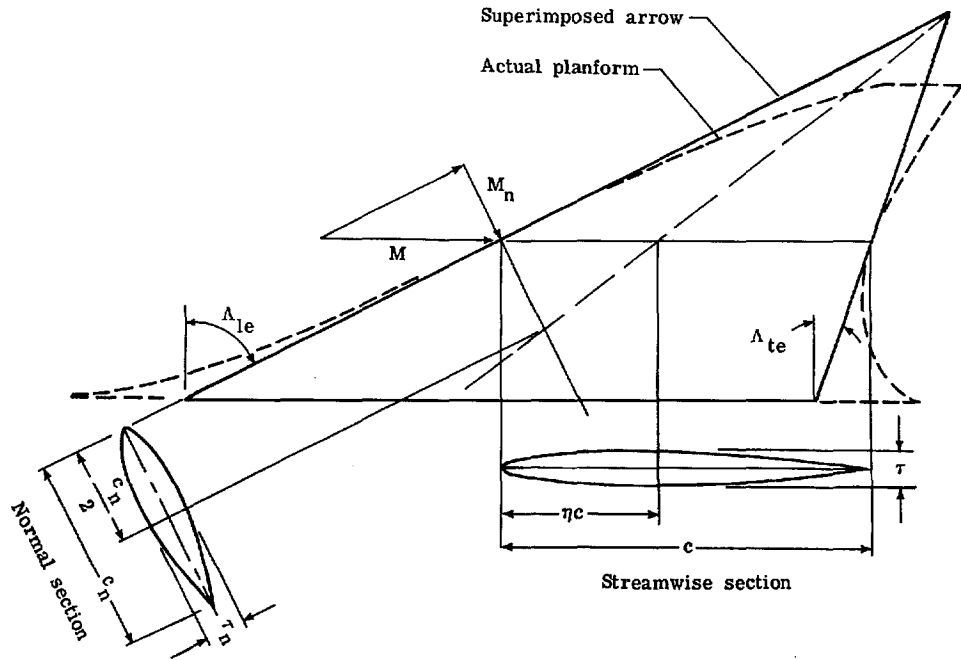


Figure 25 Definition of normal chord section ([ref.5])

By examining various different airfoil sections at a wide range of Mach numbers and Reynolds numbers in the wind-tunnel, Carlson derived the following empirical relation for the limiting pressure coefficient:

$$C_{Plim} = \frac{-2}{\gamma M_n^2} \left( \frac{R_n \cdot 10^{-6}}{R_n \cdot 10^{-6} + 10^{4-3M_n}} \right)^{0.05 + 0.35(1-M_n)^2} \quad (180)$$

The Reynolds numbers in this expression are based on the mean aerodynamic chord and transformed to a chord normal to the leading-edge, again using simple geometrical relations. The correlation of the limiting pressure coefficient with the normal Mach number and Reynolds number is shown in Figure 26, which was taken from [ref.5].

In case the factor  $K_A$  becomes greater than 1, it is kept equal to 1 (the attainable suction force is never allowed to exceed the theoretical value).

Instead of calculating the attainable thrust factor for a large number of spanwise locations and integrating over the span, in this application one representative arrow

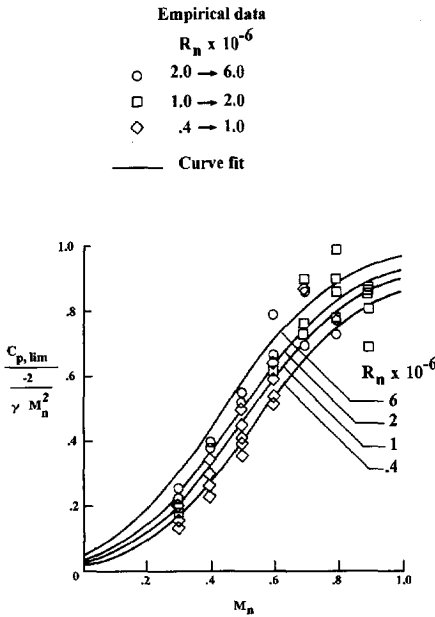


Figure 26 Limiting pressure coefficient ([ref.5])

wing is selected, leading to an average value  $K_A$ . Consequently, for double-delta wings, this could imply that in some cases where the average normal Mach number is supersonic, no suction force is taken into account, although locally some wing sections might be sufficiently swept to generate leading-edge suction. Therefore, in the next section, alternative relations will be presented for double-delta wings.

In case the theoretical suction force is not fully developed ( $K_A < 1$ ), for instance because the leading-edge radius is too small, it is assumed that the theoretical suction force vector is rotated upwards in such a way that the tangential force is equal to (see Figure 27):

$$\Delta C_x = -K_A C_T \quad (181)$$

This causes an extra normal force (compare with Polhamus' leading-edge suction analogy) that is equal to:

$$\Delta C_N = \frac{C_T}{\cos \Lambda_{le}} \sin(\arccos K_A) \quad (182)$$

or:

$$\Delta C_N = \left( \frac{\partial C_T}{\partial \alpha^2} \right) \sin^2 \alpha \sqrt{1 - K_A^2} \quad (183)$$

For a representative arrow or delta wing this can be written as:

$$\Delta C_L = K_v \sin^2 \alpha \cos \alpha \sqrt{1 - K_A^2} \quad (184)$$

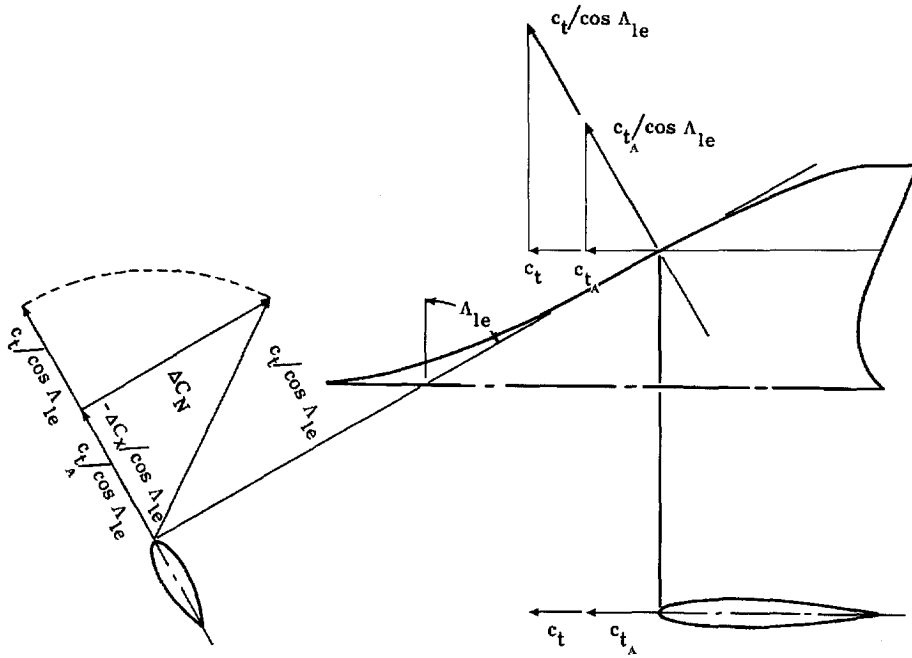


Figure 27 Attainable thrust concept ([ref.5])

Using in this fashion Polhamus' leading-edge suction analogy and Carlson's attainable leading-edge thrust concept it is possible to combine both effects, suction force and vortex flow, into simple formulas for lift and lift-dependent drag of flat wings:

$$C_L = K_p \sin \alpha \cos^2 \alpha + K_T K_A \sin^3 \alpha + K_v \sqrt{1 - K_A^2} \sin^2 \alpha \cos \alpha \quad (185)$$

$$C_D = K_p \sin^2 \alpha \cos \alpha - K_T K_A \sin^2 \alpha \cos \alpha + K_v \sqrt{1 - K_A^2} \sin^3 \alpha \quad (186)$$

In these expressions, the thrust force coefficient was replaced by:

$$C_T = K_T \sin^2 \alpha, \text{ with:}$$

$$K_T = \frac{\pi A}{4E^2} \sqrt{1 - m^2} \text{ for supersonic flow, and}$$

$$K_T = (K_p - K_p^2 K_i), \text{ or } K_T = (K_p - \frac{K_p^2}{\pi A \phi}) \text{ in subsonic flow.}$$

As, for supersonic flow, both the local suction force coefficient and the attainable thrust coefficient depend on the spanwise coordinate  $y$ , an average value should be deduced in order to justify the previous notation. The attainable suction coefficient is defined as:

$$C_{T_A} = \frac{2}{S} \int_0^{b/2} K_A(y) c_t \frac{S}{b} dy \quad (187)$$

Substituting equation (165) then yields:

$$C_{T_A} = \frac{2\pi}{SE^2} \sqrt{1 - m^2} \sin^2 \alpha \int_0^{b/2} y K_A(y) dy = K_T K_A \sin^2 \alpha \quad (188)$$

In order to keep the supersonic suction force coefficient constant at its value at full suction ( $K_A = 1$  ;  $K_T = \frac{\pi A}{4E^2} \sqrt{1 - m^2}$ ), the average attainable thrust coefficient should be based on  $y = 0.552s_w$  (see Appendix).

The relation for the local suction force coefficient (based on the mean chord) as it

features in the formula for the attainable suction coefficient is:

$$c_t = \frac{\pi A}{E^2} \sqrt{1-m^2} \left( \frac{\bar{y}}{2s_w} \right) \sin^2 \alpha \quad (189)$$

It should be noted that  $K_A$  also depends on the angle of attack. In order to treat it purely as a factor on the total suction force, this dependency was not elaborated in the above expressions, but kept within the definition of  $K_A$ . Should this definition of  $K_A$  be substituted, however, then it would turn out that the suction force term is proportional to  $\sin^{4/5} \alpha$ .

For supersonic flow, the two extreme cases will be compared hereafter:

**Full leading-edge suction (no vortex lift  $K_A = 1$ )**

$$C_D = K_p \sin^2 \alpha \cos \alpha - K_T \sin^2 \alpha \cos \alpha \quad (190)$$

$$K_T = \frac{\pi A}{4E^2} \sqrt{1-m^2} \quad (191)$$

$$K_p = \frac{\pi A}{2E} \quad (192)$$

Substituting  $\alpha = \frac{C_L}{K_p}$  (which is only valid in linearized potential flow, but may be used here as no non-linear vortex flow occurs) the following expression is obtained:

$$C_D = \frac{C_L^2}{\pi A} (2E - \sqrt{1-m^2}) \quad (193)$$

**No leading-edge suction (full vortex lift  $K_A = 0$ )**

This situation may occur when sharp airfoils (zero leading-edge radius) are applied.

$$C_D = K_p \sin^2 \alpha \cos \alpha + K_v \sin^3 \alpha \approx K_p \alpha^2 + K_v \alpha^3 \quad (194)$$

Substituting  $\alpha = \frac{C_{L_p}}{K_p}$  yields, with equation (167) and  $K_p = \frac{\pi A}{2E}$ :

$$C_D = \frac{C_{L_p}^2}{\pi A} 2E \left( 1 + \frac{\sqrt{1-m^2}}{\pi A \cos \Lambda_{le}} C_{L_p} \right) \quad (195)$$

Since

$$C_L = C_{L_p} \left( 1 + \frac{C_{L_v}}{C_{L_p}} \right) = C_{L_p} + C_{L_p}^2 \frac{K_v}{K_p^2} \quad (196)$$

the following expression can be obtained after solving for  $C_{L_p}$ :

$$C_{L_p} = \frac{2C_L}{1 + \sqrt{1 + 4 \frac{\sqrt{1-m^2}}{\pi A \cos \Lambda_{le}} C_L}} \quad (197)$$

Substituting this, after some rewriting, the following formula for the total induced drag of a wing developing non-linear lift according to the Polhamus leading-edge suction analogy is derived:

$$C_D = \frac{C_L^2}{\pi A} \frac{4E}{1 + \sqrt{1 + 4 \frac{\sqrt{1-m^2}}{\pi A \cos \Lambda_{le}} C_L}} \quad (198)$$

Notice that the lift coefficient now features in the denominator as well. It can be verified easily that if the vortex coefficient becomes zero (for instance because the leading-edge of the wing becomes sonic, resulting in  $m$  becoming equal to one), the

drag equals  $C_D = \frac{C_L^2}{K_p}$ .

It should be noted, that even for wings with sharp leading-edges, the calculated vortex coefficient may not be fully attainable, without this resulting in partial recovery of the suction force ([ref.12]). In such cases, the vortex does not develop and the



vortex lift is not fully achieved. A figure in [ref.12] shows the ratio of the measured and the calculated value of the vortex coefficient of delta wings in subsonic flow as a function of the aspect ratio, according to which the vortex lift will be zero for wings with aspect ratio exceeding four. Since it is obvious that the leading-edge suction analogy has limited applicability (only wings with sufficient sweep) but Polhamus does not give any specification of such limitations, the mentioned figure might be used as an indication. In the present work it was not taken into account, however. The validity of the vortex coefficient with respect to vortex burst is illustrated too in [ref.12]. Roughly, the present analysis is valid up to angles of attack of 20 degrees for wings with leading-edge sweep angles below 60 degrees; for wings with higher leading-edge sweep angles the angles of attack may be as high as 45 degrees before vortex breakdown occurs. For a supersonic transport these limitations therefore do not seem to be of concern.

## 9.6 Application to Double-delta Wings

For double-delta wings, the integration of the product of the local supersonic suction force coefficient and the attainable thrust coefficient should be carried out over the panel inboard of the kink and over the panel outboard of the kink. The concept of augmented vortex lift in which the inboard vortex amplifies the outboard one was not applied in this study. Substituting equation (165) in both parts of the integral yields:

$$C_T = (K_{T_i} + K_{T_o}) \sin^2 \alpha \quad (199)$$

$$K_{T_i} = \frac{\pi A}{4E_i^2} \sqrt{1 - m_i^2} \left( \frac{l_i}{s_w} \right)^2 \quad (200)$$

$$K_{T_o} = \frac{\pi A}{4E_o^2} \sqrt{1 - m_o^2} \left( 1 - \left( \frac{l_i}{s_w} \right)^2 \right) \quad (201)$$

For each panel an average value for  $K_A$  as well as for  $K_T$  must be selected. The value for the average inner attainable thrust coefficient should be based on (see Appendix):

$$\bar{y} = 0.552 l_i \quad (202)$$

whereas the average outer attainable thrust coefficient should be based on:

$$\bar{y} = 0.552 s_w \left( \frac{1 - \left( \frac{l_i}{s_w} \right)^2}{1 - \left( \frac{l_i}{s_w} \right)^{7/5}} \right)^{5/3} \quad (203)$$

in which  $l_i$  is the span of one inboard wing panel, measured from the wing root to the spanwise location of the kink (see Figure 28).

In case of the presence of a fuselage, it should be noted that  $K_{Ti}$  and  $K_{To}$  are based on the net wing area (integration was performed over the wetted wing span).

In the present context, it is assumed that the thickness-to-chord ratios and the nose radii for the airfoil sections are constant for the entire wing. If this is not the case, either average values can be taken, or the spanwise variation can be implemented in the integration.

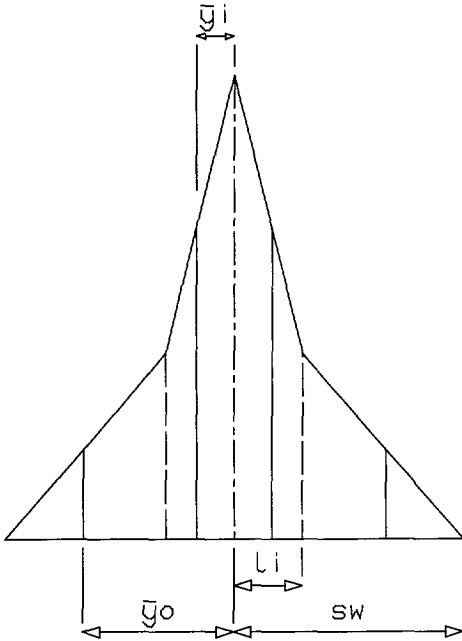


Figure 28 Double delta wing geometry definition

For subsonic flow, the relations for  $K_T$  (based on the gross wing area) become:

$$K_{T_i} = (K_p - K_i K_p^2) \left( \frac{l_i}{s_w} \right) \quad (204)$$

$$K_{T_o} = (K_p - K_i K_p^2) \left( 1 - \frac{l_i}{s_w} \right) \quad (205)$$

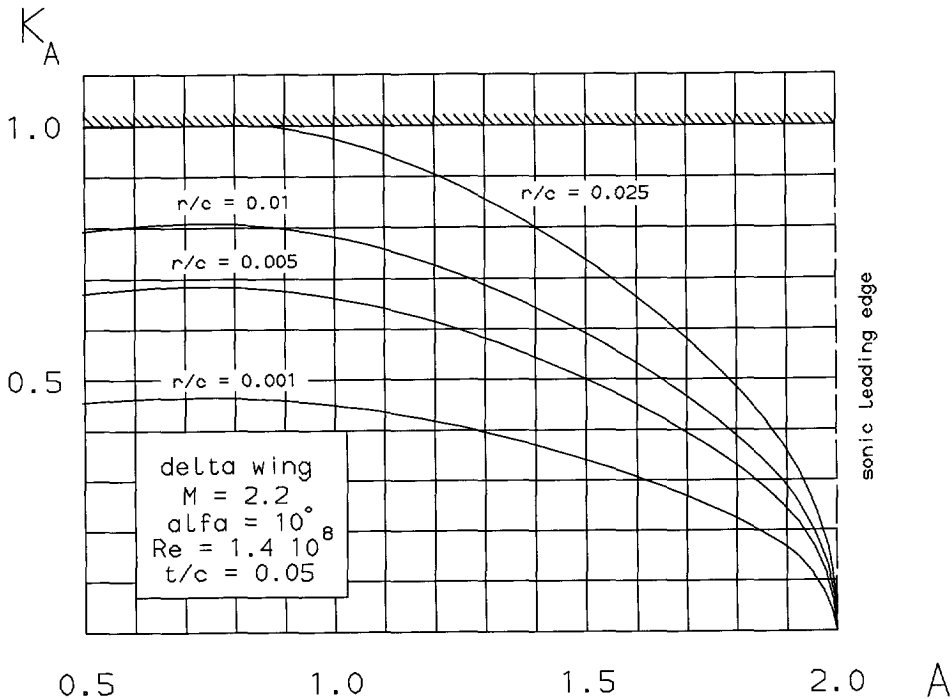
In case the leading-edge of the outer wing panel is supersonic, the lift gradient should be split as well. Using the previously introduced expression for  $K_p$  this becomes:

$$K_p = \frac{2\pi}{\sqrt{M^2 - 1} S} \left( \frac{m_i S_i}{E_i} + \frac{m_o S_o}{E_o} \right) \quad (206)$$

If a fuselage is present, the lift gradient is based on the gross wing area. If supersonic tip effects are to be taken into account as well, the lift gradient should be obtained from a panel method, as explained in the introduction.

## 9.7 Results

For a pure delta wing values of the attainable thrust coefficient as a function of the nose radius are presented in Figure 29. The attainable suction force coefficient -as discussed earlier- is not permitted to exceed the theoretical value. The influence of the nose-radius is obvious.



**Figure 29** Attainable thrust coefficient for a family of pure delta wings at  $M=2.2$

The ratio of the lift-to-drag ratio's for both extreme values of the attainable thrust coefficient in supersonic flow is shown in Figure 30, for a pure delta wing with unit aspect ratio.

As might be expected, for low values of the lift coefficient, the presence of leading-edge vortices will lead to poor lift-to-drag ratio's due to the increased amount of drag. However, at large values of the lift coefficient, a region exists in which the vortex flow seems to be more efficient. Airfoil sections with small leading-edge radii do not seem able to compete with airfoil sections equipped with round leading-edges, as far as maximum lift-to-drag ratio is concerned, as the favorable region is that of high angles of attack, at values which will in most cases lie beyond the values for maximum lift-to-drag ratio. However, thin wings are favorable in order to keep supersonic wave drag due to volume as low as possible. This implies that, in case a range constraint is imposed, an optimum value exists for the thickness-to-chord ratio and the leading-edge radius. In that case, the optimal wing incidence might be taken relatively high if the optimal leading-edges are taken relatively sharp, in order to remain in the favorable region of high lift coefficients. For a given wing area and cruise weight, this might lead to higher cruise altitudes, an effect that in turn will

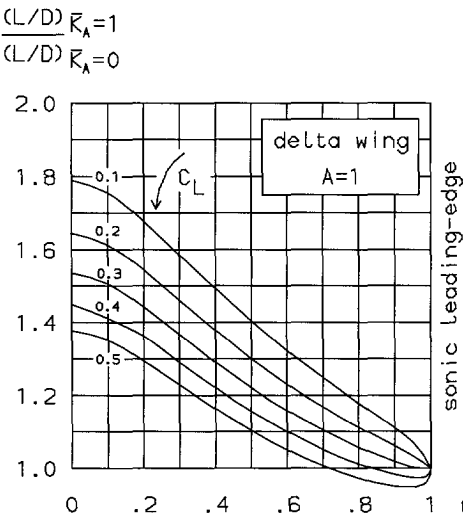
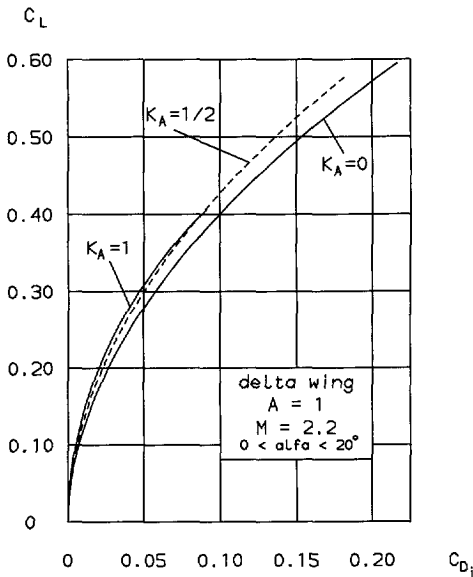


Figure 30 Ratio of full suction and full vortex  $L/D$  for a pure delta wing of unit aspect ratio

influence engine performance. It might therefore be concluded, that by taking leading-edge suction and vortex flow into account, a very important variable, that is, the leading-edge radius, can be introduced in the process of high-speed civil transport optimization.

For a pure delta wing of unit aspect ratio, the induced drag polars are drawn in Figure 31, for the cases of full leading-edge suction, full leading-edge vortex flow and for vortex flow with fifty percent of the theoretical suction force still present. As can be clearly seen, with full leading-edge suction, the drag force is much lower than in case separation occurs. On the other hand, in case full vortex flow occurs, a dramatic increase in lift is possible: in Figure 31 and Figure 32 at an angle of attack of 20 degrees, an increase of 50% is achieved. Furthermore, it can be concluded, that the presence of leading-edge vortices can be very advantageous for high wing loading/low aspect ratio vehicles in the low speed regime, especially during take-off and landing. The increase of lift due to leading-edge vortices enables Concorde to land without the use of flaps. This suggests the use of a high lift device with zero leading-edge nose radius, which can be deployed during approach, thus causing a large increase in lift and drag. In fact, the use of similar "vortex flaps" has been studied on the F-16 XL aircraft at the NASA Langley Research Center for possible use on a future High Speed Civil Transport ([ref.1]). On military aircraft, there is an additional application during high wing loading manoeuvres.

Since the attainable thrust coefficient depends on the angle of attack, it will normally decrease as the wing is rotated toward higher angles of attack unless full leading-edge suction can be retained under the circumstances. This implies that the leading-edge suction vector rotates backwards with the wing. In order to illustrate this, the induced



**Figure 31 Drag polars for different values of the attainable thrust coefficient**

drag polar and the lift versus angle of attack curve according to the attainable thrust concept, are plotted in Figure 32 for the same wing as used previously. The limits for full leading-edge suction and full vortex flow are shown in this figure as well. According to the right-hand figure, the full leading-edge suction force is retained up to an angle of attack of approximately eight degrees. At this point, the curve leaves the line for  $K_A=1$  and bends toward the full vortex flow line which it approaches for high angles of attack.

This behavior is depicted in a somewhat different way in Figure 33, in which, again for the same wing, the lift-dependent drag factor is plotted against the lift coefficient. Again, the two extreme cases for  $K_A=0$  and  $K_A=1$  are plotted as well, for both cases the linear as well as the non-linear relation. The linear relation for 100% leading-edge suction is:

$$\frac{C_{D_i}}{C_L^2} = \frac{2E - \sqrt{1 - m^2}}{\pi A} \quad (207)$$

which is independent of the lift coefficient. The linearized relation for the full vortex case depends on the lift coefficient, and so do (trivially) the non-linearized relations. Again, it is clear that for a lift coefficient of approximately 0.20, the curve diverges from the full leading-edge suction line.

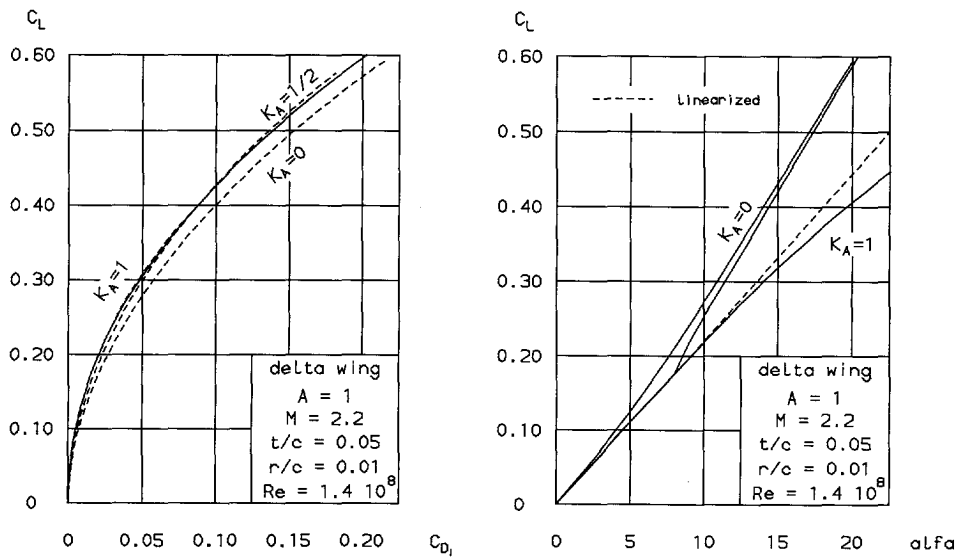


Figure 32 Drag polar and  $C_L$ - $\alpha$  line according to attainable thrust concept

A more clarifying picture of the influence of the leading-edge flow on the lift-dependent drag is found in Figure 34. As far as the linearized equations are concerned, this is an often published figure, in which the lift-dependent drag factor divided by  $\pi A$  is plotted against the leading-edge parameter  $m$ . The linearized relation for the full leading-edge suction case is simply:

$$\frac{C_{D_i}}{C_L^2/\pi A} = 2E - \sqrt{1 - m^2} \quad (208)$$

The case for no suction (which is only a theoretical one) then becomes:

$$\frac{C_{D_i}}{C_L^2/\pi A} = 2E \quad (209)$$

As soon as the leading-edge becomes sonic, the flow becomes two-dimensional (constant wing loading) and leading-edge suction or vortex flow can no longer occur.

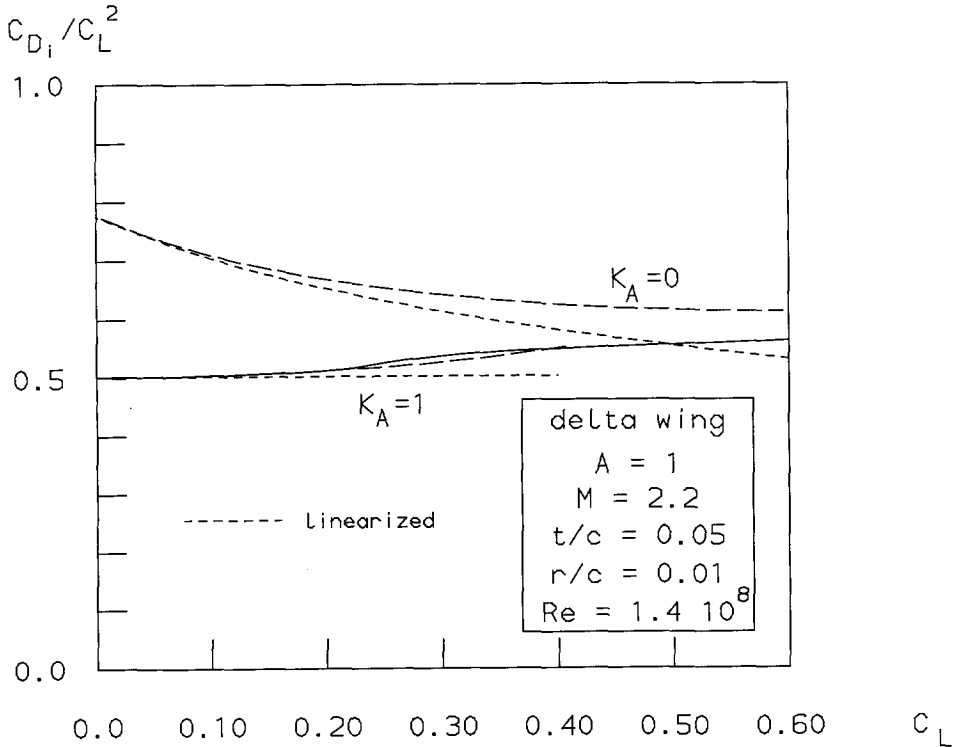


Figure 33 Induced drag factor as a function of the lift coefficient

At that point the well known linear relations of Ackeret apply:

$$C_L = \frac{4\alpha}{\sqrt{M^2-1}} \text{ and } C_D = \frac{4\alpha^2}{\sqrt{M^2-1}}, \text{ yielding:}$$

$$C_D = \frac{1}{4} \sqrt{M^2-1} C_L^2 \quad (210)$$

This figure clearly illustrates the importance of sufficiently swept wings with sufficiently rounded leading-edges. To reduce the lift-dependent drag, high relative wing thicknesses are advantageous too, however, this will negatively affect the wing zero-lift wave drag. With the equations derived in this chapter, curves can be drawn in the figure for the nonlinear full leading-edge suction case and the full vortex case. The increase in drag caused by the vortices is obvious. Since the nonlinear relations are dependent of the angle of attack, they apply for a specified angle of attack only; in



this example, for clarity, a relative large value of ten degrees was chosen. Also, in this figure, the line according to the attainable thrust concept was drawn, which closely follows the line for full leading-edge suction.

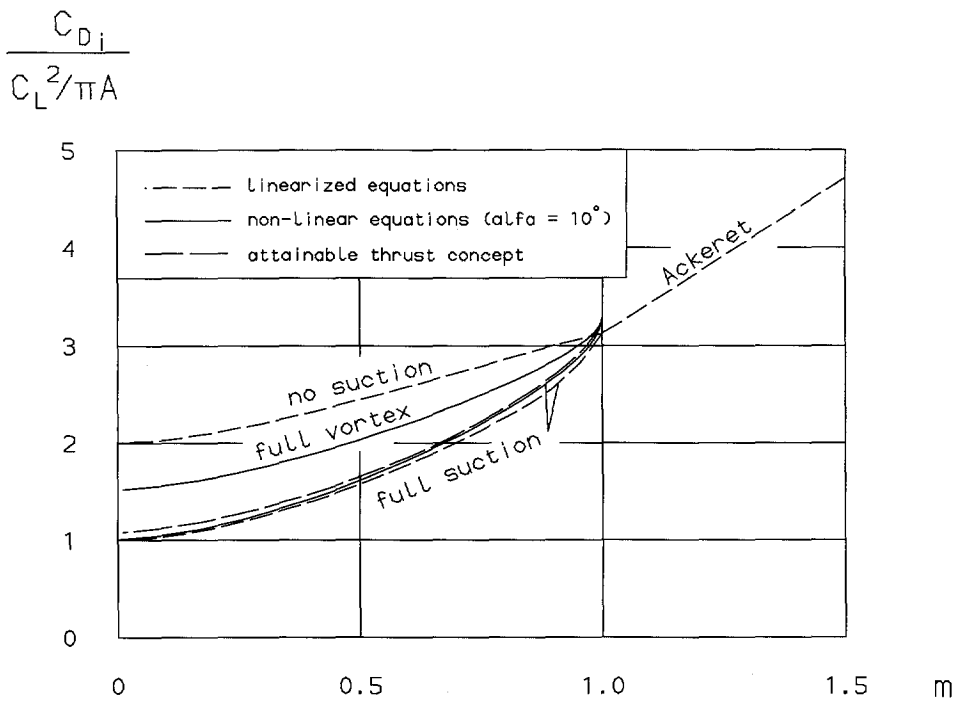


Figure 34 Induced drag factor for full suction, no suction, full vortex and attainable thrust

The impact of the leading-edge parameter (or aspect ratio for a delta wing at constant Mach number) and the wing nose radius and relative thickness is illustrated in Figure 35, where a fuselage (Sears-Haack body of length 60 m and diameter 5 m) was added to a wing of constant area (400 m<sup>2</sup>) and varying aspect ratio. In this simple example, it is assumed that all volume is carried inside the fuselage (wing zero-lift wave drag negligible), whereas all lift is generated by the wing (no fuselage lift induced drag). In the figure the maximum cruise lift-to-drag ratio is plotted against the aspect ratio for the linearized case of full leading-edge suction as well as for the case of full vortex flow. In case of full leading-edge suction, the decrease of the induced drag with increasing aspect ratio and the simultaneous decrease of the leading-edge suction forces leads to an optimal value for the aspect ratio. In case, however, the leading-edge suction force cannot be generated because the wing is too sharp or too thin (in order to reduce volume drag in a practical design), the L/D ratio

is drastically reduced due to the occurrence of vortex flow, and no optimum exists for a subsonic leading-edge.

Since the design of a supersonic airplane wing is always a compromise between high-speed and low-speed performance requirements, it is important to be able to analyze the low-speed characteristics of typical supersonic wing designs. With the use of the equations presented in this chapter, it is possible to analyze the influence of certain parameters on the approach lift coefficient. Delta wings and highly swept wings, if at all, only stall at very high angles of attack (because of vortex breakdown). Therefore, usually no maximum lift coefficient is defined, but an approach lift coefficient which is based on a maximum allowable approach angle of attack with respect to runway visibility. In this example it is put at 12 degrees. In Figure 36 the approach lift coefficient is plotted against the aspect ratio of a family of pure delta wings.

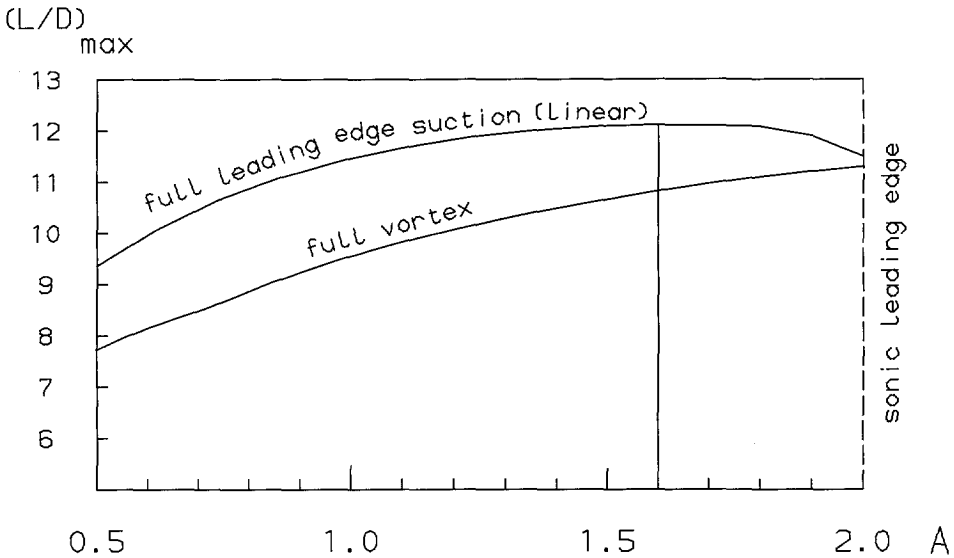


Figure 35 Effect of aspect ratio on maximum lift-to-drag ratio

The results have been calculated for a Mach number of 0.5 for wings with a relative nose radius of  $r/c=0.01$  and a relative thickness of  $t/c=0.025$ . The subsonic lift gradient was calculated using the following relation from [ref.8]:

$$C_{L_\alpha} = \frac{2\pi}{\frac{2}{A} + \left( \frac{(1-M^2)}{\cos^2 \Lambda_{1/2}} + \left( \frac{2}{A} \right)^2 \right)^{1/2}} \quad (211)$$

A constant elliptical span efficiency factor  $\phi=0.8$  was used in order to obtain the theoretical subsonic suction force. From the figure it follows, that up to an aspect ratio of 1.5, full leading-edge suction can be retained. If the attainable suction coefficient is artificially set equal to zero (by deploying the vortex flaps) the approach lift coefficient increases, as expected. It should be realized, however, that this is by no means trivial. If for a certain wing at angle of attack, the attainable thrust coefficient happens to be such that the upward rotated suction force vector points exactly vertical, then setting the attainable suction coefficient equal to zero will cause the vector to rotate backwards until it acts perpendicular to the wing chord, leading to a decrease in the vertical resultant. Using the relations from this chapter, this effect can be investigated by examining the ratio of the increase in lift coefficient with respect to the potential flow contribution, with and without vortex flaps deployed:

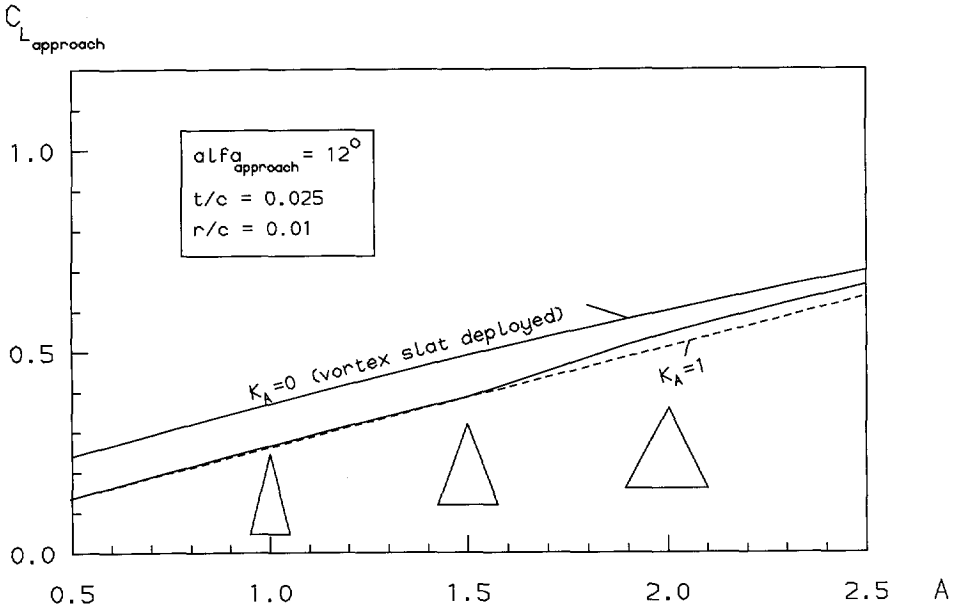
$$\frac{\Delta C_{L_{K_A > 0}}}{\Delta C_{L_{K_A = 0}}} = \frac{K_A K_T}{K_V} \tan \alpha + \sqrt{1 - K_A^2} \quad (212)$$

## 9.8 Implementation of the Method

The method described so far in this chapter accounts for the total wing-lift-dependent drag, that is, (trailing-edge vortex) induced drag, leading-edge vortex induced drag and in supersonic flow, lift-dependent wave drag. In case a fuselage is present as well, a more extensive analysis is necessary, since the presented method does not take into account the zero-lift wave drag (of neither the fuselage nor the wing) and the fuselage lift.

Using a panel method (NLRAERO, [ref.9], [ref.10] and [ref.13]) which is based on linear theory, for a given configuration only two runs at zero angle of attack and at a small positive angle of attack have to be performed to produce all required results. After these two runs the zero-lift drag is known (except for the viscous contributions)

as well as the lift gradient and  $\frac{C_{D_i}}{C_L^2}$  of the complete configuration. Some details of the



**Figure 36** Effect of deploying vortex slat on approach lift coefficient

NLRAERO program will be treated in the next chapter.

In case of the presence of a fuselage, the total lift follows from:

$$C_L = C_{L_w} + C_{L_f} = K_{P_w} \alpha_w + (K_{P_{wf}} - K_{P_w}) \alpha_f = K_{P_{wf}} (\alpha_f - \alpha_{f_{0L}}) \quad (213)$$

$$\text{with } \alpha_{f_{0L}} = \frac{-C_{L_{\alpha_f=0}}}{K_{P_{wf}}}.$$

In this case, the lift gradient as obtained from the panel method is that of the wing-

fuselage combination,  $K_{P_{wf}}$  whereas  $K_{P_w} = \frac{C_{L_{\alpha_f=0}}}{i_w}$ . In case the wing incidence equals

zero, the wing lift gradient is equal to the wing-fuselage lift gradient. This implies, that all lift is 'battered' over the wing. This is allowed, in this case, since the normal forces of both the wing and the fuselage act at the same angle and may therefore be added.

The assumption that  $K_{P_w} = \frac{C_{L_{\alpha_i=0}}}{i_w}$  is only acceptable in subsonic flow, since it implies that the fuselage does not contribute to the lift at zero angle of attack. It should be noted that this does not mean that the fuselage lift is completely ignored, it only fixes the fuselage-lift versus angle of attack line, since it must run through the value of the total lift at  $\alpha_w=0$  and through zero at  $\alpha_f=0$ . Since the line for the total lift versus angle of attack is fixed after two runs of the panel method at different angles of attack, the wing lift versus angle of attack line is fixed as well.

In supersonic flow, the assumption of a zero fuselage lift contribution at zero fuselage angle of attack is not acceptable. Furthermore, in case the panel method calculates positive lift at zero wing angle of attack, this assumption would lead to a negatively sloped fuselage lift line.

The actual situation is as follows:

$$C_{L_f} = K_{P_f} (\alpha_f - \alpha_f^*) \quad (214)$$

and, referring to Figure 37, it can easily be deduced that:

$$K_{P_f} = \frac{K_{P_{wf}} i_w - C_{L_{\alpha_i=0}}}{i_w + \alpha_f^*} \quad (215)$$

$$K_{P_w} = \frac{K_{P_{wf}} \alpha_f^* + C_{L_{\alpha_i=0}}}{i_w + \alpha_f^*} \quad (216)$$

in which  $K_{P_{wf}}$ ,  $i_w$  and  $C_{L_{\alpha_i=0}}$  are known.

The actual value of  $\alpha_f^*$  is very hard to determine, since the slope of the fuselage lift line is rather small and cannot be determined with sufficient accuracy. However, the value of  $\alpha_f^*$  can have a quite substantial influence on the fuselage and wing lift slopes as indicated in Figure 38.

As the wing lift slope can be determined with a higher accuracy, it was decided to use this value to complete the model. This is possible in supersonic flow, since in this case the results of the panel method are obtained by pressure integration, hence the fuselage and wing contributions can be separated. In the subsonic case a Trefftz-plane analysis is performed, which yields a greater accuracy but loses the possibility to discern the fuselage from the wing contribution.

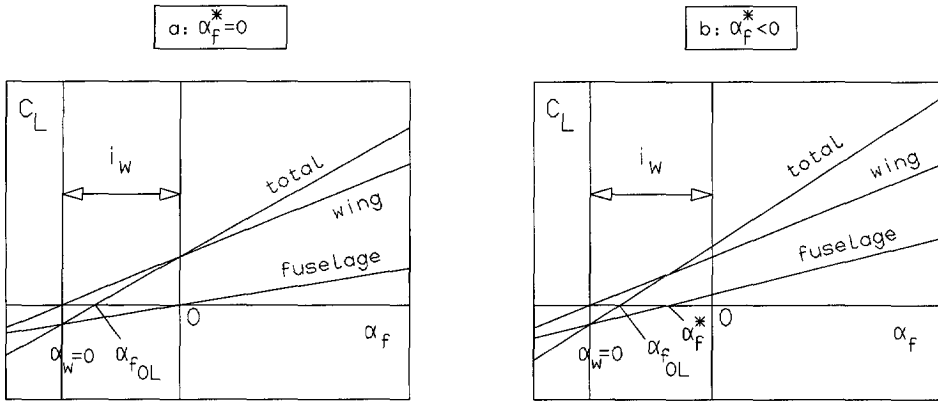


Figure 37 Wing-, fuselage- and total lift versus angle of attack

It should be realized that the obtained wing lift slope includes interference, implying that it will not run exactly through zero at zero wing angle of attack. The difference, however, is very small.

Since, in the supersonic case, the leading-edge suction force coefficient is known from an approximating formula, the wing lift and drag can be easily computed by the equations presented in this chapter. The fuselage contribution to the lift-dependent drag (which includes the wave drag due to lift) can readily be added using:

$$\Delta C_{L_f} = K_{p_f} (\alpha_f - \alpha_f^*) \quad (217)$$

$$\Delta C_{D_f} = C_{L_f} \alpha_f = \frac{C_{L_f}^2}{K_{p_f}} + C_{L_f} \alpha_f^* \quad (218)$$

$$\text{with } K_{p_f} = (K_{p_{wf}} - K_{p_w}).$$

Now, only the zero-lift wave drag as determined by the panel method is to be added in order to obtain the total configuration lift and drag.

For the subsonic case, the situation is somewhat more complicated, since the leading-edge suction force is not known and must be obtained from the panel method results. This is done by subtracting the potential flow contribution of both the wing and the fuselage from the induced drag as calculated by the panel method. This results in the

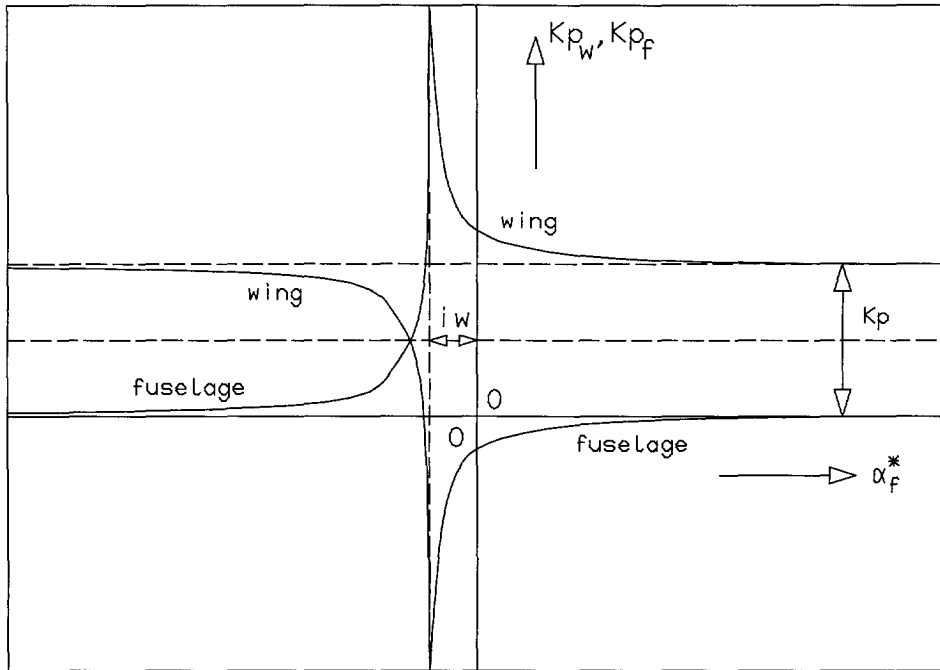


Figure 38 Influence of  $\alpha_f^*$  on wing and fuselage lift slope

following expression for the suction force factor:

$$K_T = \left( K_{P_w} - K_i K_{P_w}^2 \right) + \left( K_{P_f} - K_i K_{P_f}^2 \right) \left( \frac{\alpha_f}{\alpha_w} \right)^2 + \left( K_{P_f} - K_i K_{P_f}^2 \right) \left( \frac{\alpha_f^*}{\alpha_w} \right)^2$$

$$- 2K_i K_{P_w} K_{P_f} \left( \frac{\alpha_f - \alpha_f^*}{\alpha_w} \right) - 2 \left( K_{P_f} - K_i K_{P_f}^2 \right) \left( \frac{\alpha_f \alpha_f^*}{\alpha_w^2} \right) \quad (219)$$

in which  $K_i = \frac{C_{Di}}{C_{L_{tot}}^2}$ .

This gives the thrust coefficient as calculated by the NLRAERO program. As explained in the previous sections, this value will be compensated with the use of the Polhamus leading-edge suction analogy and the Carlson attainable thrust concept. In

case the wing incidence is zero, the wing lift gradient becomes equal to the total configuration lift gradient -as explained before- and the fuselage lift gradient becomes zero. It can be easily verified that this reduces the above formula to the first term only, a result that was previously presented for the subsonic wing alone. However, now the fuselage lift as well as mutual interference effects are included.

As explained previously, the value of  $\alpha_i^*$  can be taken as zero, since equation (219) is used for subsonic flow.

The vortex coefficient (based on full theoretical suction) follows from:  $K_v = \frac{K_T}{\cos \Lambda_{le}}$ .

Just like in the supersonic case, the lift-dependent drag of the fuselage must be added (in subsonic flow, the fuselage generates a small amount of lift due to it being placed in the flowfield of the wing). The zero-lift drag as calculated by a panel method will be zero for flat wings.

Viscous drag contributions have to be calculated independently, since viscous effects are not modeled in panel methods. A modification to NLRAERO was implemented by Middel ([ref.13]) in order to account for friction drag, form drag and separation drag, but the method is only expected to yield accurate results for the subsonic cruise condition and not for off-design conditions or supersonic flow, where temperature effects can become important. It seems therefore advisable to calculate viscous effects independently from the panel program. This eliminates the altitude from the set of input variables.

## 9.9 Conclusions

The methods presented in this chapter enable the designer to account for the influence of leading-edge vortices on the lift and drag coefficient of flat delta wings and derived planforms, both at subsonic and supersonic speeds. Without this method, this effect which is very important for any supersonic airplane design or optimization study, could not as easy (and as cheap) be taken into account at the earliest stage of conceptual design.

No validation or verification of the method has been carried out in this chapter, since the presented method is a combination of two established methods integrated over the wing span, methods which have demonstrated the capability to capture the non-linear effects very well.



## 9.10      References

1. Aviation Week & Space Technology; September 13, 1993, pp. 50-52
  2. Bos, A.H.W.; *A Method to account for Partial Leading-Edge Suction and Vortex Flow for Swept and Delta Wings at Subsonic and Supersonic Speeds*; Report LR-779, ISBN 90-5623-004-2, Delft University of Technology, December 1994
  3. Brown, C.E.; *Theoretical Lift and Drag of Thin Triangular Wings at Supersonic Speeds*; NACA Rept. No. 839, November, 1946
  4. Carlson, H.W. and Mack, R.J.; *Estimation of Leading-Edge Thrust for Supersonic Wings of Arbitrary Planform*; NASA TP 1270, October 1978
  5. Carlson, H.W.; Mack, R.J. and Barger, R.L.; *Estimation of Attainable Leading-Edge Thrust for Wings at Subsonic and Supersonic Speeds*; NASA TP 1500, October 1979
  6. Carlson, H.W. and Mack, R.J.; *Studies of Leading-Edge Thrust Phenomena*; Journal of Aircraft Vol.17, No.12, December 1980
  7. Carlson, H.W. and Miller, D.S.; *Influence of Leading-Edge Thrust on Twisted and Cambered Wing Design for Supersonic Cruise*; Journal of Aircraft Vol.20, No.5, May 1983
  8. Hoak, D.E.; *USAF Stability and Control DATCOM*; 2<sup>nd</sup> ed., 1968
  9. Hoeijmakers, H.W.M.; *A Panel Method for the Determination of the Aerodynamic Characteristics of Complex Configurations in Linearized Subsonic and Supersonic Flow*; National Aerospace Laboratory NLR, NLR TR 80124 U, part 1,2,3, December 1980
  10. Hoeijmakers, H.W.M.; National Aerospace Laboratory NLR, Memorandum AT-83-005 U, November 1983
  11. Hutchison, M.G.; Huang, X.; Mason, W.H.; Haftka, R.T. and Grossman, B.; *Variable-Complexity Aerodynamic-Structural Design of a High-Speed Civil Transport Wing*; AIAA 92-4695, September 1992
  12. Mendenhall, M.R. and Nielsen, J.N.; *Effect of Symmetrical Vortex Shedding on the Longitudinal Aerodynamic Characteristics of Wing-Body-Tail Combinations*; NASA CR-2473, 1975
  13. Middel, J.; *Development of a Computer Assisted Toolbox for Aerodynamic Design of Aircraft at Subcritical Conditions with Application to Three-Surface Aircraft*; Ph.D. dissertation, ISBN 90-6275-768-5/CIP, Delft University Press, 1992
  14. Polhamus, E.C.; *A Concept of the Vortex Lift of Sharp-Edge Delta Wings Based on a Leading-Edge Suction Analogy*; NASA TN D-3767, October 1966
  15. Polhamus, E.C.; *Prediction of Vortex Lift Characteristics by a Leading-Edge Suction Analogy*; Journal of Aircraft Vol.8, No.4, April 1971
  16. Schlichting, H. and Truckenbrodt, E.; *Aerodynamik des Flugzeuges*; Springer Verlag, Berlin 1960
-

## Appendix Calculation of Spanwise Coordinates for Mean Attainable Thrust Coefficient

### pure delta wings

According to equation (188), in order to use the mean attainable thrust coefficient as a factor on the suction force coefficient for full suction (equation (191)), it must be equal to:

$$K_A = \frac{8}{b^2} C \int_0^{b/2} y^{2/5} dy = \frac{C}{\bar{y}^{3/5}} \quad (220)$$

in which  $K_A$  was written in accordance with equation (176) as a constant  $C$  times  $y^{3/5}$ .

$$\text{This yields after solving the integral: } \bar{y} = \left( \frac{7}{40} 2^{7/5} \right)^{5/3} b \approx 0.276b = 0.552s_w$$

### double delta wings

For double delta wings, the integration of the local suction force coefficient is carried out in two parts, over the inner and over the outer wing panel. Using again equation (165) for the local suction force coefficient and writing  $K_A = C.y^{3/5}$  the following expression for the attainable thrust coefficient on the inner wing panel is obtained:

$$C_{T_{A_i}} = \frac{\pi A}{4E_i^2} \sqrt{1-m_i^2} \sin^2 \alpha \frac{8C}{b^2} \int_0^{l_i} y^{2/5} dy = K_{T_i} K_A \sin^2 \alpha \quad (221)$$

After substituting equation (200), this yields for the mean value of the attainable thrust coefficient of the inner wing panel:

$$K_{A_i} = \frac{2}{l_i^2} C \int_0^{l_i} y^{2/5} dy = \frac{C}{\bar{y}_i^{3/5}} \quad (222)$$

After solving the integral, this yields:  $\bar{y}_i = 0.552l_i$ .

For the outer panel follows in the same way, using equation (201):

$$K_{A_o} = \frac{2}{s_w^2 \left( 1 - \left( \frac{l_i}{s_w} \right)^2 \right)} C \int_0^{s_w} y^{2/5} dy = \frac{C}{\bar{y}_o^{3/5}} \quad (223)$$

After solving the integral, this yields for the spanwise coordinate on which the mean attainable thrust coefficient of the outer panel must be based:

$$\bar{y}_o = 0.552 s_w \frac{\left( 1 - \left( \frac{l_i}{s_w} \right)^2 \right)^{5/3}}{\left( 1 - \left( \frac{l_i}{s_w} \right)^2 \right)^{7/5}} \quad (224)$$



# 10. Using the NLRAERO Panel Method in Combination with the Attainable Thrust Concept in Multivariate Optimization

## 10.1 Summary

In chapter 9 a method was presented to calculate lift and induced drag of flat wings developing full or partial leading-edge vortex flow. In order to implement this method, several parameters are needed as input, being for subsonic flow the total configuration lift gradient, the lift coefficient at zero degrees angle of attack and the induced drag coefficient  $k$ . For supersonic flow the required parameters are the total configuration lift gradient, the wing lift gradient, the lift gradient at zero degrees angle of attack and the wave drag coefficient. It was suggested that these parameters should best be supplied by a panel method because subsonic (semi-)empirical relations are usually not applicable to configurations designed for supersonic cruise, whereas lift gradient and wave drag values in supersonic flow are expected to be more accurately predicted by a panel method, since interference effects and Evvard's principle are thus taken into account. To carry out stability and control calculations, the pitching moment coefficient at zero fuselage angle of attack and the pitching moment gradient are needed too (since the centers of gravity of the configurations to be analyzed are not known, these parameters are taken relative to the fuselage nose).

Since the total aerodynamic analysis procedure is to be part of a multivariate optimization, it was decided not to implement the NLRAERO method directly into the optimization procedure because CPU time and cost would become prohibitively large. Instead, as reported in chapter 5, a global approximation by means of regression surfaces was used for the data generated by the NLRAERO program and used in the aerodynamic analysis procedure.

In this chapter, the panel method as employed in the present work, will be outlined and some difficulties that were encountered will be discussed. Following this, the accuracy of the aerodynamic approximations will be evaluated.

Furthermore, the fact that the present analysis does not include cambered and warped wings to optimize the pressure distribution, but for simplicity is constrained to flat wings, might imply that the level of aerodynamic efficiency required for the type of airplane considered cannot be attained. In this chapter, a quantification of the influence of optimal wing camber and twist on the induced drag will be carried out.

## 10.2 Notation

A aspect ratio

---

---

	matrix of aerodynamic influence coefficients
$a$	notch ratio
$a_1$	notch ratio
$a_2$	leading-edge kink coordinate parameter
$a_3$	leading-edge kink coordinate parameter
$b$	right-hand-side matrix
	wing span
$C_D$	drag coefficient
$C_{D_0}$	zero-lift drag coefficient
$C_{D_F}$	friction drag coefficient
$C_{D_L}$	drag due to lift coefficient (induced drag coefficient)
$C_{D_W}$	wave drag coefficient
$C_F$	friction coefficient
$C_L$	lift coefficient
$C_{L_{\alpha_i=0}}$	lift coefficient at zero fuselage angle of attack
$C_{M_\alpha}$	pitching moment gradient
$C_T$	leading-edge thrust (suction) coefficient
$c$	chord
$D_L$	drag due to lift (induced drag)
$d$	diameter
$E$	elliptic integral of the second kind
$h$	cruise altitude
$K_0$	correction factor on volume drag
$K_p$	lift gradient
$K_T$	leading-edge thrust (suction) parameter
$K_v$	vortex parameter
$K_w$	wave drag factor
$k$	induced drag factor
$L$	lift
$l$	length
$M$	Mach number
$m$	leading-edge parameter
$q$	dynamic pressure
$R$	range
$Re$	Reynolds number
$r$	planform parameter
$S$	gross wing area
$s$	semi span
$t$	wing thickness
$Vol$	volume
$W$	weight
$W_0$	maximum take-off weight
$x$	matrix of unknown sources and doublets

---

$\alpha$	angle of attack
$\alpha_f^*$	fuselage angle of attack for zero fuselage lift
$\beta$	$\sqrt{M^2 - 1}$
$\eta$	exponent in approximation of elliptic integral of second kind
$\Lambda$	sweep angle

**subscripts**

ab	afterbody
bw	basic wing
c	compressible
cyl	cylindrical section
e	exposed
f	fuselage
fs	full suction
i	incompressible
	inner
le	leading-edge
min	minimum
o	outer
opt	optimal
t	tip
w	wing
wf	wing-fuselage combination

**abbreviations**

ADAS	Aircraft Design and Analysis System
L	lower triangular matrix
MCC <sup>2</sup>	Multiple Correlation Coefficient Squared
NLR	Nationaal Lucht- en Ruimtevaart Laboratorium (National Aerospace Laboratory)
U	upper triangular matrix

In accordance with tradition, weights are given in kilograms, although -according to the SI system- the unit of weight is Newton and the unit of mass is kilogram.

### 10.3 The NLRAERO Program

Panel methods are based on the discretization of the integral representation of the solution of the Prandtl-Glauert equation valid in linearized potential flow. By using Green's theorem, this solution (perturbation velocity potential) can be written as a function of elementary solutions of the Prandtl-Glauert equation. These elementary solutions are source and doublet distributions with as yet unknown strengths. In subsonic flow, the velocity potential at a certain point is affected by all singularity

distributions on the geometry and the wake (explained further on), whereas in supersonic flow, the solution is only affected by singularity distributions that are situated within the forward (projected) Mach cone, protruding from the point under consideration.

In a panel method, the geometry of the aircraft is discretized into panels, whereas the source and doublet distributions are discretized as polynomials as a function of the local panel axis variables. In many panel methods however, the values of the source and doublet strength are kept constant over each panel.

In order to model lift, a potential jump over the wake, equal to the circulation, is defined in such a way as to ensure smooth trailing-edge flow (Kutta condition, only in case of subsonic trailing-edge). To facilitate this, a wake panel is defined behind each streamwise strip of wing panels, extending to infinity. On this wake panel, the mass conservation requirement is enforced, implying that the source strength on the wake is zero, whereas the doublet strength follows from the Kutta condition (potential jump over the wake). On the wetted panels the boundary condition is enforced that the normal velocities (which are equal to the derivatives of the velocity potential with respect to the normal vector of the panel under consideration) should be zero, since they are impermeable. Therefore, collocation points are defined at which the boundary conditions are enforced, leading to a square system of equations of shape  $Ax=b$ , in which  $A$  is the matrix of normal velocities induced by one panel's singularity distribution on another panel's collocation point (matrix of aerodynamic influence coefficients). The singular value, of the velocity induced by a panel on its own collocation point is calculated by shifting that collocation point by an infinitesimal distance. Only a source distribution is assigned to the body panels, whereas the strength of the wing source distribution follows directly from the thickness distribution by means of a thin airfoil approximation. Therefore, the unknowns in the system of equations are the body source strengths and the wing doublet strengths and the dimension of the system is therefore equal to the sum of the body and the wing panels. The solution of the system yields the values of the source and doublet strengths at each panel, from which the perturbation potentials, the induced velocities and the pressures at each panel can be calculated. By integration of these pressures over the complete configuration, the total lift and induced drag forces can be obtained. In supersonic flow, the wave drag due to lift and volume is included, in subsonic flow the zero-lift drag obtained by the panel method is zero for flat wings.

Most panel methods calculate the linearized perturbation potentials, in order to yield pressures which are subsequently integrated over the wetted surfaces of the model (in supersonic flow over that part of the surface which is inside the projected Mach-cone) to yield forces. This linearized model is valid only under the assumption that the directional cosines of the normal vectors of the wetted surfaces and the vertical coordinate axis are approximately one. Obviously, in the neighborhood of airfoil section noses, this is not the case. Therefore, the estimation of leading-edge suction using panel methods is not very accurate. A more accurate drag prediction would be obtained if a Trefftz-plane momentum analysis would be performed ([ref.8], this is

---



very complicated in supersonic flow). The effects of vortex flow or curled-up wakes, are not taken into account by most panel methods.

The panel method used in this application, NLRAERO, was developed by Hoeijmakers [ref.4] and [ref.5] at the Dutch National Aerospace Laboratory (NLR) and modified by Middel [ref.8]. This modification mainly consists of the replacement of the geometrical part of the original program, such that it is compatible with the TU Delft's ADAS/Medusa design and analysis system. This enables the user to draw a configuration in the Medusa CAD system, which is subsequently translated to an internal NLRAERO geometry file. In the present application, however, this facility was bypassed because of the large number of configurations that have to be analyzed in order to derive approximating regression surfaces. Instead, the generation of the NLRAERO geometry files was made independent of the Medusa system.

A different panelling scheme was implemented by Middel [ref.8]. According to this scheme, the fuselage and wing are divided into a fixed number of segments, that are subsequently panelled. The partition of the fuselage segments is such, that at each intersection of a lifting surface with the fuselage, two fuselage sections are defined: one above the lifting surface and one below it. Furthermore, each part of the fuselage that is not intersected by a lifting surface is a separate segment as well. Thus, a wing-body combination consists of four fuselage segments, whereas a wing-body-tail configuration consists of seven fuselage segments. The lifting surfaces naturally are divided into the wing segment, horizontal stabilizer segment, etc. Furthermore, each lifting surface is divided into one segment inside the fuselage and a number of wetted segments, depending on the number of kinks in the leading- or trailing-edge. The wing segments are divided into streamwise strips and the body segments into rings that are subsequently divided into panels. The lifting surface segments inside the fuselage consist of only one strip. These so-called lift-carry-over strips carry the same doublet strengths as the adjacent (wetted) strip. Thus, these strips account for the phenomenon that an isolated body, which normally does not generate much lift, does so when wings are added to provide for the necessary circulation.

The panel method in its present shape caused some problems at supersonic Mach numbers. First of all, severe convergence problems were encountered. Two reasons can be given for this. In supersonic flow the potential in a certain point is only affected by sources and doublets that are situated inside a so-called domain of influence, that is, the forward (projected) Mach cone with the point under consideration as its apex. Sources and doublets that are not in this domain of influence do not contribute and the corresponding positions in the matrix of aerodynamic influence coefficients are therefore zero. The Mach lines which describe the surface of the Mach cones, are "reflected" inside of the fuselage. In case the afterbody is tapered and closed, the concentration of reflecting Mach lines becomes very high toward the rear, causing interference problems with the source distribution on the fuselage panels. This effect causes large differences between the solutions, which leads to bad convergence of the iteration process. This problem was solved by extending the cylindrical part of the fuselage, thus leaving the afterbody open-ended. In this way the Mach lines reflect only a few times, reducing the spurious internal

---

interference. The afterbody drag is neglected in this way and is therefore added with the use of the methods of [ref.2].

The second problem is related to the build-up of the matrix of aerodynamic influence coefficients. The matrix is partitioned according to the segment partition of the panelling scheme. This implies that velocities on a certain segment, induced by elements of that same segment are placed in the same diagonal block. The system of equations is solved by using a block-Jacobi iterative procedure. To facilitate this, the diagonal blocks are inverted separately. A condition for (fast) convergence of iterative methods is, that the elements on the diagonals are dominant. However, in supersonic flow, the influence exerted by one panel segment on elements of other segments can be quite large. Due to the extensive subdivision in segments, this causes non-diagonal elements to be large as well, and consequently the diagonal elements are no longer dominant. In practice this leads to very bad convergence or even divergence of the iteration process. In principle, the problem can be solved in two ways: first, the panel distribution can be adapted such, that elements which exert a large influence on each other are grouped into the same segments or that larger segments are used, such that all large values are located in one diagonal block. This would, however, require a complete re-evaluation and reprogramming of the panelling process (for instance joining the fuselage segments above and below the lifting surfaces). Instead, a different approach was chosen. The iterative solution methods were completely discarded, and instead the whole system of equations was solved by using straightforward LU-decomposition with partial pivoting. To do this, the original subroutine SOLVE of the NLRAERO program was replaced by a new routine that reads the matrices from the NLRAERO output files, carries out the LU-decomposition, solves the system and redirects the solution to the correct files. The disadvantage of this method is, that more memory capacity is needed and that more calculations have to be performed, leading to an increase in CPU time that is not balanced by the time-saving effect that iterations no longer have to be performed. The major advantage however, is that a solution is always obtained.

Finally, it was found that in some cases the panel distribution can have a paramount influence on the calculated supersonic flow properties. In case of wing-body combinations, it was found that totally unrealistic results (like negative lift gradients) were obtained if too few panels were used. This is caused by large areas of vacuum pressure on the lower side of the fuselage (the pressures as calculated by the panel program are checked to see whether they are within the interval between the vacuum pressure and the stagnation pressure and, if necessary, cut off to either of these values). By using a sufficient number of panels, this problem could be largely solved. In case vacuum pressures were still present, they occurred symmetrically on the (cylindrical) afterbody, and therefore had no influence on the calculated forces and moments. However, adding a horizontal stabilizer again led to unrealistic results. Due to the vacuum pressures on the afterbody high pressures were induced on the lower side of the stabilizer (cut off to stagnation pressure) while low pressures were induced on the upper side (cut off to vacuum). This causes unrealistically high increases in total configuration lift and obviously totally unreliable pitching moment coefficients.

---

This problem could not be satisfactorily solved. Constraining the values of the wing thickness ratio, wing incidence or Mach number did not lead to better results. It was found, that a decreasing wing incidence as well as an increasing Mach number led to more pressure coefficients being cut off to vacuum, leading to irregular variations in the lift coefficient. Variation of the panel distribution on the different configuration segments did not solve the problem either.

Because of this, it was decided not to incorporate the horizontal stabilizer in the panel method. Instead, the influence of the stabilizer is taken into account by means of the methods provided by the USAF Stability and Control DATCOM ([ref.3]) and added to the NLRAERO results. For each configuration two runs are performed, one with zero fuselage angle of attack and one with a small positive angle of attack. In the subsonic case, the original solution routine with the block-Jacobi procedure is used. In the supersonic case, the direct method must be applied. By repeating this procedure for a large number of configurations, regression surfaces can be obtained for values like the lift gradient, lift coefficient at zero fuselage angle of attack, wave drag coefficient and induced drag as a function of the design parameters. The randomly generated configurations are subjected to geometric constraints on their design variables. This ensures for instance that all fuselages have the same volume and that no unrealistic designs are included that can disturb the accuracy of the approximations.

## 10.4 Accuracy of Aerodynamic Approximations

By analyzing a number of test cases, it was established that especially the zero degrees angle of attack lift coefficient and the wave drag coefficient as calculated from the regression surfaces were very inaccurate for supersonic flow. Although, as can be deduced from the following table, the accuracy of the regression for the mentioned parameters is not very high, approximate values that differ by a factor two or three from the exact values, such as sometimes occurred during these test runs, were totally unexpected. It is a well known fact that optimizers take advantage of approximation errors by directing the search towards regions of the design space where the applied analysis is no longer valid, resulting in large differences between the computed values and the actual ones ([ref.6]). However, the presently employed approximate analysis method is supposed to be a global approximation, valid throughout the entire design space and subject to the same geometrical constraints as the designs to be optimized.

---

	supersonic aerodynamic coefficients					
	$Kp_{wf}$	$Kp_w$	$C_{L_{\alpha_i=0}}$	$C_{D_0}$	$C_{M_{\alpha_i=0}}$	$C_{M_\alpha}$
number of samples	200	200	200	200	200	200
MCC <sup>2</sup> [%]	80.30	93.32	91.58	92.80	83.42	89.09
F test	3.72	12.77	9.94	11.78	4.59	7.46
standard error of regression [% of mean response]	7.26	4.02	15.87	27.91	22.67	9.40
% of residuals in Normal distribution	98.0	98.5	99.0	98.5	99.0	98.0

	subsonic aerodynamic coefficients				
	$Kp_{wf}$	$C_{L_{\alpha_i=0}}$	k	$C_{M_{\alpha_i=0}}$	$C_{M_\alpha}$
number of samples	200	200	200	200	200
MCC <sup>2</sup> [%]	98.27	99.19	99.52	99.00	99.50
F test	51.93	112.35	188.79	90.37	182.70
standard error of regression [% of mean response]	1.74	5.44	1.44	5.76	1.69
% of residuals in Normal distribution	99.0	99.5	99.0	99.0	98.5

Based on 13 variables (105 coefficients).

The above results refer to the regression equations of the square roots of the functions to be approximated. The square root trick was applied to the pitching moment coefficients as well, based on their absolute values (since the pitching moment coefficient about the nose of the airplane is always negative). To obtain the approximate function values from these polynomials, the answers are to be squared and, in case of the pitching moment coefficients, multiplied by -1.

**Table 21** Statistics pertaining to regression of aerodynamic coefficients

The difference in accuracy between subsonic and supersonic approximations is expected, since supersonic approximations will always be less accurate because of the previously mentioned numerical discontinuities. Despite the mentioned adaptations to

the calculation method, some problems remained however. In order to study the matter of approximating supersonic aerodynamic parameters, 500 random designs were generated, subject to geometrical constraints to exclude unrealistic or geometrical impossible configurations. Of these 500 designs, 13 could not be evaluated because of a singular matrix of aerodynamic influence coefficients. A further 91 designs were discarded because of negative lift or drag coefficients, positive pitching moment about the fuselage nose, negative zero-lift drag or negative lift gradients. It was not further investigated if these results were caused by numerical errors and if they could be avoided by adapting the panelling scheme. Since the sample generation, panelling and aerodynamic evaluation are completely automated, the same panelling scheme is applied to every configuration (that is, same number of strips per segment and same number of panels per strip). The remaining 396 samples were used to investigate the influence of the number of samples on the parameters that measure the accuracy of the regression surfaces, such as the Multiple Correlation Coefficient Squared, the F-test for statistical significance and the standard error of the regression.

The result of this investigation clearly showed the contradicting requirements imposed on the number of samples on which the regression is based as far as the  $MCC^2$  and the F-test are concerned. This was already discussed in chapter 5. Decreasing the number of samples to which the regression surfaces are fit will increase the value of  $MCC^2$ , but this is trivial since it is inherent to its definition. This is why it is very dangerous to test the usability of a regression just by checking its  $MCC^2$  since this value can always be made equal to one by choosing the number of samples equal to the number of terms in the approximation or vice versa. The statistical significance of the regression becomes less as the number of samples decreases, although some fluctuation occurs once the system of equations becomes less overdetermined. It was further observed that the standard error of the regression relative to the mean response attains a value which is almost independent of the number of samples on which the regression is based. Occasionally, for a number of samples slightly in excess of the number of coefficients needed, the extra samples lie close to the exact fit, resulting in a very small error. It was established, however, that such regression surfaces are of very little use. The order of the samples was reversed as well, but this did nothing to change the presented conclusions.

The main reason for the bad regression was found to be the fact that the results of the panel method at zero fuselage angle of attack were sometimes unreliable. It was established that the lift and pitching moment coefficient at two degrees fuselage angle of attack were much better predicted by response surfaces than those at zero degrees fuselage angle of attack. Two sets of data were used to test this observation more thoroughly. First, based on the presently used set of data obtained at zero and at two degrees fuselage angle of attack new regression surfaces were generated for the lift and pitching moment coefficient at two degrees fuselage angle of attack. These results are presented in the following table. The other regression surfaces naturally are the same as presented before.

---

	supersonic aerodynamic coefficients					
	$K_{p_{wf}}$	$K_{p_w}$	$C_{L_{\alpha_i=2^\circ}}$	$C_{D_0}$	$C_{M_{\alpha_i=2^\circ}}$	$C_{M_\alpha}$
number of samples	200	200	200	200	200	200
MCC <sup>2</sup> [%]	80.30	93.32	89.19	92.80	84.86	89.09
F test	3.72	12.77	7.54	11.78	5.12	7.46
standard error of regression [% of mean response]	7.26	4.02	8.10	27.91	11.21	9.40
% of residuals in Normal distribution	98.0	98.5	99.5	98.5	98.5	98.0

**Table 22** Statistics pertaining to regression of supersonic aerodynamic coefficients

It is clear that the standard errors of the regression as a percentage of the mean response are about halved and now have values around 10%.

A second set of data was generated for two and four degrees fuselage angle of attack. Of the 200 generated samples, 14 were discarded because of unrealistic results. This set of data did not result in better regression surfaces. The standard errors of the regression as a percentage of the mean response were slightly better for all gradients, but slightly worse for the lift and pitching moment coefficient at two degrees fuselage angle of attack. The error for the zero-lift drag coefficient was about 32%.

The following conclusions are drawn. Although the accuracy of the supersonic regression surfaces remains a problem, a sufficient accuracy can be realized once the lift and pitching moment coefficients at two degrees fuselage angle of attack are considered. With the lift and pitching moment gradients, the values at zero degrees are easily found. However, the relation between the zero-lift drag and the design variables is so noisy that it is not possible to generate second-order polynomials that approximate this relation with sufficient accuracy. Implementing the panel method directly into the analysis for supersonic Mach numbers is not an option because of the relatively large CPU time required for one aerodynamic evaluation and the large amount of optimization iterations involved in a multidisciplinary optimization. Despite the fact that three dimensional and interference effects are captured best by a panel method, it was decided to base the zero-lift drag calculations on semi-empirical relations. The main drag components in supersonic flow are the friction drag and induced drag which are calculated with the use of semi-empirical relations anyway.

Since the use of semi-empirical methods to calculate total configuration lift gradients and pitching moments is very cumbersome and restricted to rather simple planforms whereas in many cases data is presented graphically or should be obtained from test results nonetheless, it was decided to accept the achievable accuracy of the regression

surfaces of the other parameters.

The subsonic analysis will be based completely on the approximating polynomials, since the accuracy of the subsonic regression is very high, whereas the analysis of configurations designed for supersonic Mach numbers at subsonic conditions leaves hardly any alternative but the use of a panel method.

## 10.5 The Influence of Optimal Camber on the Cruise Glide Ratio

According to the supersonic area-rule, the wave drag due to volume of a configuration in supersonic flow is determined by its area distribution. By a wise choice of the area distribution this drag component may be minimized. The most well-known example of such area-ruling is the *coke-bottle* fuselage shape to reduce wave drag due to wing-fuselage interference. The supersonic drag due to lift can be minimized too, by applying camber and twist to optimize the pressure distribution of the wing. Ideally, the total configuration wave drag can be minimized by optimizing both the fuselage camber and area distribution and the wing camber and distribution of thickness and twist. For a conceptual design phase, however, this would introduce a huge amount of extra variables. Since the present configurations for simplicity are defined by fuselages with cylindrical center-sections and flat wings, it is possible that the level of aerodynamic efficiency required for this type of aircraft cannot be realized. Therefore, in this paragraph, a quantification of the influence of optimal camber and twist on the induced drag is carried out. The large difference between the drag of a fuselage with cylindrical center-section and a fuselage designed for minimum wave drag at the same volume and length (Sears-Haack body) is shown as well. A simplified optimization study is performed for illustrative purposes and a comparison is made between the theoretical optimal results and the results as calculated by the NLRAERO program.

In [ref.7], Jones presented a relation for the theoretical minimum drag due to lift of thin, slender wings in supersonic flow:

$$D_L = \frac{L^2}{\pi q b^2} + \frac{1}{2}(M^2 - 1) \frac{L^2}{\pi q l^2} \quad (225)$$

in which  $b$  is the overall span and  $l$  the overall length of the wing and  $q$  represents the dynamic pressure. The first term of this equation is the theoretical induced drag of high aspect ratio wings developing full leading-edge suction with an elliptical spanwise lift distribution. The second term is the theoretical supersonic drag due to lift for a wing with an elliptical streamwise pressure distribution. Although the theoretical minimum drag as predicted by equation (225) can only be approached by an optimal choice of the camber and twist distribution whereas the equation is only valid for wings of very narrow proportions lying close to the center of the Mach cone, in [ref.10] it is nonetheless used to predict the drag due to lift of arbitrary wings by correcting the value predicted by equation (225) with a factor  $K_w$  to account for

diversions from the elliptical loading.

Since it has already been suggested in industry that in order to realize the values for the cruise glide ratio that are required for supersonic cruising aeroplanes with a range of about 10,000 km, it is necessary to apply optimum camber and twist, the purpose of this paragraph is to investigate the influence of such optimization of the pressure distribution.

Equation (225) can be rewritten as follows, for wings:

$$C_{D_L} = \frac{C_L^2}{\pi A} \left( 1 + 2 \left( \frac{\beta s_w}{l_w} \right)^2 \right) \quad (226)$$

First, this equation is compared to the non-optimal induced drag for flat delta wings:

$$C_{D_L} = \frac{C_L^2}{\pi A} (2E - \sqrt{1 - m^2}) \quad (227)$$

This equation was derived previously in chapter 9. The elliptical integral of the second kind,  $E$ , can be approximated as follows:

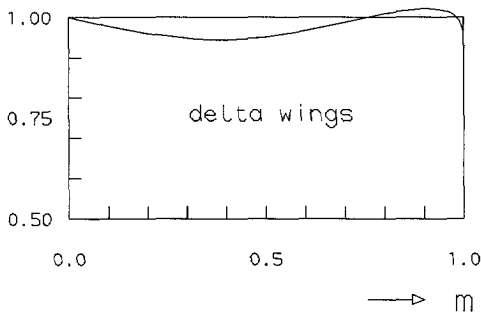
$$E = 1 + \left( \frac{\pi}{2} - 1 \right) m^\eta, \quad \eta = 1.226 + 0.15\pi(1 - \sqrt{m}) \quad (228)$$

The parameter  $m$  is the leading-edge parameter, defined as the ratio of the Mach angle and the complement of the leading-edge sweep angle. The parameter  $\beta s_w/l_w$ , or box parameter, in equation (226) is equal to the leading-edge parameter for a delta wing. Thus, for a pure delta wing, the ratio between the theoretical minimum induced drag and the induced drag of a flat uncambered wing with full leading-edge suction in supersonic flow is equal to:

$$\frac{(C_{D_L})_{\text{opt}}}{(C_{D_L})_{\text{f.s.}}} = \frac{1 + 2m^2}{2E - \sqrt{1 - m^2}} \quad (229)$$

This ratio is plotted in Figure 39 against the leading-edge parameter  $m$ . The best improvement achievable by applying optimal wing camber and twist to a pure delta wing appears to be only a reduction in induced drag to about 94% of the theoretical full suction induced drag of a flat wing. According to the figure, for some values of the leading-edge parameter, the ratio exceeds one, the maximum being about 102%.





**Figure 39** Ratio of optimal induced drag to induced drag with full leading-edge suction for delta wings

This is partly caused by the fact that equation (227) is derived from a relation by Brown ([ref.1]) for the leading-edge suction force on flat delta wings in supersonic flow:

$$C_T = \frac{\pi A}{4E^2} \sqrt{1 - m^2} \sin^2 \alpha = K_T \sin^2 \alpha \quad (230)$$

in which the elliptical integral of the second kind is approximated -as explained previously- by equation (228). Perhaps more importantly, the slender wing theory does not apply for wings with near-sonic leading-edges, rendering the accuracy of Jones' equation doubtful for values of  $m$  close to 1.

A general way of writing equation (227) is:

$$C_{D_L} = C_L^2 \left( \frac{1}{K_p} - \frac{K_T}{K_p^2} \right) \quad (231)$$

in which  $K_p$  represents the lift gradient in potential flow and  $K_T$  represents the suction force coefficient which follows from equation (230). It is emphasized that this relation applies to delta wings, but it is assumed to be valid for derived planforms too. The lift gradient  $K_p$  in supersonic flow according to [ref.9] is given by:

$$K_p = \frac{2\pi m}{E\sqrt{M^2 - 1}} \quad (232)$$

Incidentally, this relation applies to wings with subsonic and supersonic leading-edges, as long as the leading-edge parameter for a supersonic leading-edge is kept

equal to one. In that case equation (232) changes into the well known Ackeret equation for linearized supersonic flow.

For a pure delta wing, this relation becomes:

$$K_p = \frac{\pi A}{2E} \quad (233)$$

which resulted in equation (227). Now substituting equation (232), the following relation is obtained for the induced drag of a flat wing of arbitrary planform developing full leading-edge suction:

$$C_{DL} = \frac{C_L^2}{\pi A} \left( E - \frac{A\beta\sqrt{1-m^2}}{8m} \right) \frac{A\beta}{2m} \quad (234)$$

For the ratio between the optimal induced drag and the induced drag of a flat wing with full leading-edge suction this yields:

$$\frac{(C_{DL})_{opt}}{(C_{DL})_{fs}} = \frac{1 + 2 \left( \frac{\beta s_w}{l_w} \right)^2}{\frac{A\beta}{2m} \left( E - \frac{A\beta\sqrt{1-m^2}}{8m} \right)} \quad (235)$$

This equation is only valid for isolated wings, since in case of the presence of a fuselage, equation (230) is based on the net wing area and the suction parameter should therefore be corrected for the gross wing. Studying this relation, one can deduce that for a given Mach number, leading-edge parameter and aspect ratio, the ratio of optimum induced drag to full suction induced drag decreases with decreasing slenderness ratio  $s_w/l_w$ . This can be achieved by increasing the tip chord and decreasing the root chord accordingly on a simple swept wing. It will now be investigated of what order the improvement in induced drag due to optimum wing camber and twist will be for realistic planforms.

Naturally, the improvement will be largest for wings with a sonic leading-edge, since such wings no longer develop the suction force and by definition have the highest induced drag. First a pure delta wing is examined at a Mach number of two. The aspect ratio of such a delta wing with a sonic leading-edge is 2.309. It can be easily verified that the optimum induced drag is 95.5% of the induced drag of the flat wing. Now, increasing the tip chord and reducing the root chord in order to keep the aspect ratio constant for comparison, it can be verified that for a slenderness ratio of 0.5 ( $s_w/l_w$ ), the optimum induced drag amounts 79.6% of the induced drag for a flat wing

(Figure 40). For a slenderness ratio of 0.4 this percentage reduces to only 62.4%. In order to limit the trailing-edge sweep angle to values not higher than the leading-edge sweep angle, for a given aspect ratio,  $\frac{2}{3 \tan \Lambda_{le}}$  is the smallest permissible slenderness ratio, or 0.38 in this example. The optimum induced drag will then amount to 59.4% of the induced drag for the flat wing. This is the maximum improvement attainable for a wing of aspect ratio 2.309 with a sonic leading-edge at Mach equals two. Wings with a negative notch ratio (increased root chord) cannot have a smaller slenderness ratio at the given aspect ratio.

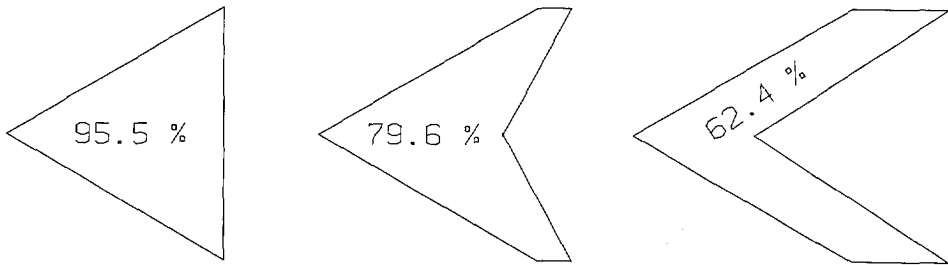


Figure 40 Three planforms of aspect ratio 2.309 with a sonic leading-edge at Mach = 2

The values for the optimal induced drag for the four slenderness ratios mentioned, that is, 0.577 (delta), 0.5, 0.4 and 0.38 are respectively:  $0.413 C_L^2$ ;  $0.345 C_L^2$ ;  $0.270 C_L^2$  and  $0.257 C_L^2$ . For the flat wing the induced drag would be  $0.433 C_L^2$  for all four cases.

The example shows that optimizing the camber and twist distributions of wings can indeed result in large improvements of the induced drag. It is obvious that the greatest advantage is achieved for wings with high slenderness ratios. However, the slenderness ratio of a realistic wing is usually constrained by geometrical, aerodynamical and structural requirements. In order to find out what induced drag reductions are at best achievable for a realistic wing design, equation (235) was optimized for a simple swept wing defined by four variables, being the overall length  $l_w$ , the semi span  $s_w$ , the tip chord  $c_t$  and the notch ratio  $a$ . Four constraints were imposed on the optimization ensuring that the box ratio  $\beta_{s_w}/l_w$  stays smaller than one (a prerequisite for the use of equation (226)), the leading-edge remains subsonic, the trailing-edge sweep angle cannot become larger than the leading-edge sweep angle and limiting the planform parameter  $r$  to a certain minimum. This planform

parameter, which is defined as:

$$r = \frac{S}{2s_w l_w}$$

(236)

and which can be regarded as the percentage of the box area that is covered by the wing, represents the aerodynamic and structural constraints that are imposed on a realistic wing design. The optimization was carried out by the genetic algorithm GENESIS which was treated earlier. The results of the optimization are presented in the following table, for a minimum planform parameter of 0.4 and 0.3 (the planform parameter of the final planform of the previous example is 0.33).

	$r_{min}=0.4$	$r_{min}=0.3$
$l_w$	56.26	56.84
$s_w$	24.52	25.87
$c_t$	13.78	12.02
$a$	0.44	0.61
$s_w/l_w$	0.436	0.455
$A$	2.166	3.027
$m$	1	1
$\frac{(C_{D_L})_{opt}}{(C_{D_L})_{f.s}}$	72.7%	54.7%

Table 23 Planforms for maximum induced drag reduction by camber and twist optimization

The actual planform shapes are illustrated in Figure 41. It is emphasized that these are not planforms for minimum induced supersonic drag, but planforms for which maximum induced drag reduction can be achieved by applying optimum wing camber and twist. Apparently, the reductions that are achievable for realistic wings are very large indeed, which leads to the conclusion that optimization of the pressure distribution is an important means to increase the supersonic glide ratio and should therefore not be left out of account in the analysis of a supersonic cruising aeroplane.

To investigate this matter further, an optimization study was performed to derive a wing-fuselage combination with maximum supersonic lift-to-drag ratio, while at the same time being able to cover a range of at least 10,000 km. The wing-fuselage combination is defined by a total of 9 design variables, being the overall length of the

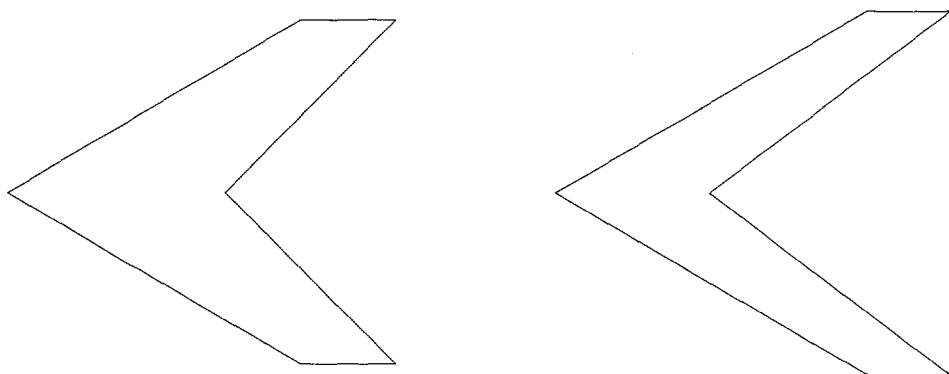


Figure 41 Planform shapes for maximum induced drag reduction by optimal cambering and twisting

net wing  $l_w$ , the semi span of the net wing  $s_w$ , the tip chord  $c_t$ , the notch ratio  $a_1$ , the location of a kink in the leading-edge  $a_2$  and  $a_3$ , the wing relative thickness, the fuselage length and the fuselage diameter. The cruise lift coefficient is a design variable too. The wing relative thickness is taken into account in order to be able to include wing zero-lift drag and make later comparison with the NLRAERO panel program possible. Also, by taking wing volume into account, it is possible to include the range constraint since in the present example it is assumed that all fuel is stored in the wing and that the total amount of fuel depends therefore on the wing volume (75% of this volume is used as fuel tank). The wing airfoil sections are defined by a symmetrical double-parabolic arc and therefore have zero leading-edge radius.

The fuselage is defined by a cylindrical center-section and two equal parabolic fore- and afterbodies. The length of the cylindrical section is derived from the fuselage length and diameter in such a way that the total fuselage volume is constant, that is, equal to  $1.65 \text{ m}^3$  per passenger (this includes volume for construction, pantries, galleys, toilets and the flight deck). No fuel is stored in the fuselage.

In this simplified analysis, the influence of the powerplant is completely neglected. In order to be able to calculate the range, it is assumed that an overall powerplant efficiency of 50% can be realized. No powerplant drag, tailplane drag or trim drag is added. The cruise Mach number is kept equal to 2.0 and the number of passengers is fixed at 250. The wing-loading in this example is fixed as well, at a value of  $5000 \text{ N/m}^2$ .

Nine geometric constraints are imposed on the design:

- 1) fore- and afterbody slenderness ratio greater than 3
- 2) total wing span less than 70 m
- 3) strake leading-edge sweep angle less than  $80^\circ$

- 4) outer wing leading-edge sweep angle greater than  $0^\circ$
- 5) trailing-edge sweep angle smaller than outer wing leading-edge sweep angle
- 6) net wing root chord less than fuselage center-section
- 7) root thickness less than half fuselage diameter
- 8) fuselage nose angle less than Mach angle
- 9) wing box ratio less than 1

Furthermore, the planform parameter  $r$  is constrained to 0.4 in order to obtain realistic wing planforms.

The total configuration drag is split up into the following components: friction drag, volume drag of the fuselage and the wing (zero-lift wave drag) and the wave drag due to lift. It is assumed that the wing incidence is zero and that the fuselage does not contribute to the lift.

The friction drag of the wing-fuselage combination is calculated as follows, assuming a fully turbulent boundary layer:

$$C_{F_i} = \frac{0.455}{\log(\text{Re})^{2.58}} \quad (237)$$

in which the subscript  $i$  refers to incompressible flow. To obtain the value for compressible flow, the following relation is used (see Appendix):

$$C_{F_c} = \frac{C_{F_i}}{(1 + 0.144M^2)^{0.65}} \quad (238)$$

The Reynolds number for the wing is based on the mean geometric chord, whereas for the fuselage it is based on the fuselage length. Since the equations yield the friction coefficient for a flat plate, the value for the fuselage, being a relatively strongly curved body, is corrected with a factor  $\frac{2}{\sqrt{3}}$ . The friction coefficients are multiplied by the wing respectively fuselage wetted area and divided by the gross wing area.

The wave drag due to volume of the fuselage is related to the wave drag of a Sears-Haack body by means of the following formula:

$$C_{D_{W_f}} = K_0 \frac{128 \text{Vol}_f^2}{\pi S l_f^4} \quad (239)$$

A Sears-Haack body has minimum wave drag for a given volume and length and in equation (239) the factor  $K_0$  is to account for deviations from this minimum. By a wise design of the total configuration cross-section distribution (area-ruling) the factor  $K_0$  can be reduced to values even below one. For the present optimization it is set at one.

The wave drag due to volume of the wing is calculated by means of the USAF Stability and Control DATCOM procedure ([ref.3]). For double-parabolic airfoil sections this drag contribution becomes:

$$C_{D_{W_w}} = \frac{16 \left( \frac{t}{c} \right)^2 S_{bw_e}}{3 \beta S} \quad (240)$$

for wings with a supersonic leading-edge and:

$$C_{D_{W_w}} = \frac{16 \left( \frac{t}{c} \right)^2 S_{bw_e}}{3 \tan \Lambda_{le_o} S} \quad (241)$$

for wings with a subsonic leading-edge. The area  $S_{bw_e}$  refers to the exposed part of the basic wing, which is obtained by extending the leading- and trailing-edge of the outer wing panel to the fuselage. For round-nosed airfoils, an extra leading-edge drag component must be added, according to:

$$C_{D_{le}} = 1.28 \frac{M^3 \cos^6 \Lambda_{le_{bw}}}{1 + M^3 \cos^3 \Lambda_{le_{bw}}} \left[ \frac{2r \left( \frac{2s_{bw} - d_f}{\cos \Lambda_{le_{bw}}} \right)}{S} \right] \quad (242)$$

in which  $r$  is the nose radius and  $s_{bw}$  is the semi span of the basic wing (from the fuselage centerline to the wing tip). In case the nose radius changes over the span the value in equation (242) should be that of the mean chord of the exposed planform.

The numerator of the term between brackets will be recognized as the exposed frontal area of the basic wing leading-edge. Since the airfoil sections used in this example are sharp-nosed, the leading-edge drag will be zero.

Finally, the drag due to lift is calculated by using the previously presented relation for the theoretical minimum induced supersonic drag according to Jones (equation (226)).

The results of the optimization, again using the GENESIS genetic search algorithm, are presented in the following table. Initially, the gross wing area was not constrained resulting in an extremely large wing, which in combination with the assumed wing loading of  $5000 \text{ N/m}^2$  leads to a totally unrealistic result. Therefore, the gross wing area was constrained to  $1000 \text{ m}^2$ .

$l_w$ (net wing)	52.7879 m
$s_w$ (net wing)	20.5279 m
$c_t$	6.6833 m
$a_1$	0.0865
$a_2$	0.8456
$a_3$	0.2884
$l_f$	78.1720 m
$d_f$	2.7297 m
$t/c$	0.0147
$C_L$	0.1036
$L/D$	14.8037
$W_{\text{fuel}}/W_0$	0.3091
$R$	10532 km
$C_{D_0}$	0.003645

$k$	0.3125
$C_{D_F}$	0.003274
$C_{D_w}$	0.0003703
$C_{D_L}$	0.003354
$S$	$999.9166 \text{ m}^2$
$S_{bwe}$	$277.4817 \text{ m}^2$
$s_w$	21.8928 m
$l_w$	57.0956 m
$W_0$	509816 kg
$h$	12727 m
$\Lambda_{le_i}$	$72.4^\circ$
$\Lambda_{le_o}$	$12.9^\circ$
$l_{cyl}$	61.7025 m

Table 24 Optimal configuration



The optimal configuration is sketched in Figure 42.

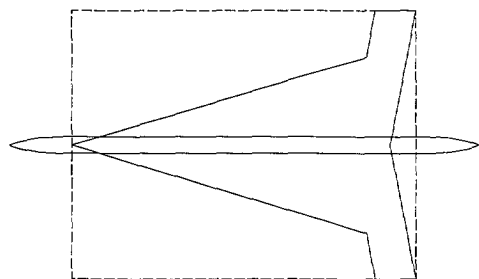


Figure 42 Optimal L/D configuration

The configuration was also evaluated by means of the NLRAERO panel method. As explained previously, the friction drag and afterbody drag should be added to the results. According to [ref.2], the drag of a parabolic afterbody can be written as follows:

$$C_{D_{ab}} = \frac{7\pi d_f^4}{24 l_{ab}^2 S} \quad (243)$$

This equation is based on the theory of slender bodies and is valid under the assumptions that the incoming flow is parallel to the cylindrical centerbody and that the fuselage is long enough to neglect interference with the forebody.

The results from the panel method, including the previously presented friction drag and the afterbody drag according to equation (243) then yield:

$$C_{D_w} = 0.001596$$

$$C_{D_F} = 0.003274$$

$$k = 0.5103$$

$$\text{or: } C_D = 0.004870 + 0.5103 \cdot C_L^2$$

For a  $C_L$  of 0.1036 this yields a  $C_D$  of 0.01035 and a glide ratio of 10.01.

This value should not be compared indiscriminately to the value obtained from the optimization, since the difference is not only due to the fact that a flat wing is considered here. Also, the fuselage lift and lift-dependent drag, which did not feature in the analysis of the optimization, have a large influence (the lift coefficient of 0.1036 is not completely generated by the wing in this case). Furthermore, there is a large

difference in the values of the zero-lift wave drag as calculated by the panel method and that calculated by equations (239), (240) and (241) (the optimal wave drag is more than four times lower). A Sears-Haack body of the given length and volume would have a diameter of 3.377 m leading to an afterbody slenderness (half of the total length divided by the diameter) of 11.6. The present afterbody has a slenderness of only 3. This shows that quite considerable drag reductions can be achieved by applying diversions from the classical cylindrical fuselage. Fortunately, the contribution of the wave drag to the total drag is relatively small (about 5% - 15%) whereas the friction drag amounts to almost half of the total drag.

In chapter 9 relations were presented for the lift and drag coefficient for wings developing leading-edge suction, leading-edge vortex flow or a combination of both. For the lift and drag coefficients of wings developing full leading-edge suction, including the influence of the fuselage, the following relations apply:

$$C_L = K_p \alpha + K_T \alpha^3 + K_{pf} (\alpha_i - \alpha_i^*) \quad (244)$$

$$C_D = (K_p - K_T) \alpha^2 + K_{pf} (\alpha_i - \alpha_i^*)^2 + K_{pf} \alpha_i^* (\alpha_i - \alpha_i^*) \quad (245)$$

in which  $\alpha_i^*$  is the fuselage angle of attack at which the fuselage generates zero-lift.

From the results of NLRAERO, the wing lift gradient was calculated as 1.717 per radian and the fuselage lift gradient as 0.1586 per radian. The value for  $\alpha_i^*$  was calculated as -0.018 rad (about -1°).

Since the outer wing panel leading-edge is supersonic, no suction force is generated here and therefore only the inner suction force coefficient has to be calculated:

$$K_{Ti} = \frac{\pi A}{4E_i^2} \sqrt{1 - m^2} \left( \frac{l_i}{s_w} \right)^2 \quad (246)$$

based on the net wing area, in which  $l_i$  is the semi span of the inboard panel of the net wing.

Based on the gross wing area, this value becomes 0.3339.

Substituting the values for the wing lift gradient, the fuselage lift gradient, the suction force coefficient and the fuselage angle of attack for zero fuselage lift in equations (244) and (245) the following values are obtained for an angle of attack of  $3.1^\circ$ :

$$\begin{aligned}C_L &= 0.1041 \\C_{D_L} &= 0.004658 \\C_{D_0} &= 0.004870\end{aligned}$$

which yields a  $C_D$  of 0.009528 and a glide ratio of 10.93, which is 75% of the optimal value according to Jones' equation.

For double-delta wings, equation (232) for the lift gradient changes into:

$$K_P = \frac{2\pi}{S\sqrt{M^2-1}} \left( \frac{m_i S_i}{E_i} + \frac{m_o S_o}{E_o} \right) \quad (247)$$

With  $S_o = 85.646 \text{ m}^2$ ,  $S_i = 914.353 \text{ m}^2$ ,  $m_o = 1$  and  $m_i = 0.549$  this yields:

$$K_P = 1.65, \text{ based on the gross wing.}$$

Substituting this value and the previously presented value for  $K_T$  into equation (231) shows that the minimal induced drag is 65% of the flat wing induced drag (substitution of the actual lift gradient of 1.717 would yield a value of 66%). The actual ratio of the theoretical minimum induced drag and the induced drag caused by the wing is 72%, which is caused by the fact that the wing generates a lift coefficient of only 0.09258, the rest being generated by the fuselage.

Now using the following relation (general shape of equation (235)):

$$\frac{(C_{D_L})_{\text{opt}}}{(C_{D_L})_{f.s.}} = \frac{1 + 2 \left( \frac{\beta s_w}{l_w} \right)^2}{\pi A \left( \frac{1}{K_P} - \frac{K_T}{K_P^2} \right)} \quad (248)$$

as an estimation for the possible reduction through optimal camber and twist of the induced drag component generated by the wing yields:

$$C_{D_L} = 0.003244$$

which yields a  $C_D$  of 0.008114 and a glide ratio of 12.83.

Finally, the influence of separation of the flow around the leading-edge is investigated. Since the presently considered airfoils are double-parabolic, the nose radius equals zero. According to the Carlson leading-edge thrust concept and the Polhamus analogy -as explained in chapter 9- this implies that no leading-edge suction force will develop but that a leading-edge vortex will occur instead, causing an extra normal force on the airfoil equal to the theoretical suction force, perpendicular to the leading-edge. The lift and drag coefficient are given by:

$$C_L = K_p \alpha + K_v \alpha^2 + K_{pf} (\alpha_f - \alpha_f^*) \quad (249)$$

$$C_D = K_p \alpha^2 + K_v \alpha^3 + K_{pf} (\alpha_f - \alpha_f^*)^2 + K_{pf} \alpha_f^* (\alpha_f - \alpha_f^*) \quad (250)$$

in which:

$$K_v = \left( \frac{\partial C_T}{\partial \alpha^2} \right) \frac{1}{\cos \Lambda_{le}} \quad (251)$$

which yields for the inner wing panel (strake):

$$K_v = 1.105$$

For an angle of attack of  $3.1^\circ$  the following results are obtained:

$$\begin{aligned} C_L &= 0.1072 \\ C_{D_L} &= 0.005808 \\ C_{D_0} &= 0.004870 \end{aligned}$$

which results in a  $C_D$  of 0.01068 and a glide ratio of 10.04.

The results derived in this paragraph are summarized in the following table:

	optimal	NLRAERO	present method		
			full suction	optimal	full vortex
$C_L$	0.1036	0.1036	0.1041	0.1041	0.1041
$C_{D_0}$	0.003645	0.004870	0.004870	0.004870	0.004870
$C_{D_L}$	0.003354	0.005477	0.004658	0.003244	0.005808
L/D	14.80	10.01	10.93	12.83	10.04

N.B. Results do not include powerplant drag, tailplane drag or trim drag.

Table 25    Summary of results with different methods of analysis

As explained in paragraph 10.4, in the actual supersonic aerodynamic analysis used in the optimization, the zero-lift drag coefficient will not be based on NLRAERO results, but on the USAF Stability and Control DATCOM ([ref.3]). For pointed afterbodies as used in the present application, no base drag occurs, and the body drag consists of the nose drag, afterbody drag and interference drag of the nose and center-section on the afterbody. Since, in the present application, the nose section and the afterbody section are equal, only the afterbody drag has to be calculated and multiplied by two. The relation for the supersonic drag for a pointed parabolic afterbody of circular cross section according to [ref.3] is exactly the same as the relation according to [ref.2], that is, equation (243).

The interference drag is presented in graphical form in [ref.3]. For a nose section length equal to the afterbody length, one single curve applies which yields the interference drag as a function of the afterbody length, the fuselage diameter and the length of the cylindrical center-section. This curve is cut in three parts which are approximated by the following parabolic relations:

$$\left(\frac{l_{cyl}}{l_{ab}}\right) < 0.5:$$
$$C_{D_{ab(nc)}} = \frac{\pi d_f^4}{16 l_{ab}^2 S} \left( 1.3217 - 2.9786 \left(\frac{l_{cyl}}{l_{ab}}\right) + 2.4762 \left(\frac{l_{cyl}}{l_{ab}}\right)^2 \right) \tag{252}$$

$$0.5 < \left( \frac{l_{cyl}}{l_{ab}} \right) < 1.5:$$

$$C_{D_{ab(nc)}} = \frac{\pi d_f^4}{16 l_{ab}^2 S} \left( 0.7369 - 0.7102 \left( \frac{l_{cyl}}{l_{ab}} \right) + 0.2175 \left( \frac{l_{cyl}}{l_{ab}} \right)^2 \right) \quad (253)$$

$$\left( \frac{l_{cyl}}{l_{ab}} \right) > 1.5:$$

$$C_{D_{ab(nc)}} = \frac{\pi d_f^4}{100 l_{ab}^2 S} \quad (254)$$

Using these relations in combination with equations (240), (241) and (242), the wave drag of the configuration analyzed in this example becomes:

$$C_{D_w} = 0.001711$$

which yields:

$$C_{D_0} = 0.004985$$

Using this value instead of the previously used value of 0.004870, the glide ratio with full suction becomes 10.79, with correction for optimal twist and camber it becomes 12.65 and with full vortex flow it becomes 9.93. It is noted here that the zero-lift drag coefficient as calculated by the methods of [ref.3] does not include the wing incidence as a variable. In reality there is a small influence, since at zero-lift, the fuselage will have a slightly negative angle of attack, whereas the wing will have a small positive angle of attack.

## 10.6 Conclusions

From the results discussed in this chapter, two main conclusions can be drawn. First, the use of regression surfaces to represent supersonic data generated by the NLRAERO program, results in intolerable inaccuracies. This is partly caused by numerical problems with the NLRAERO program and partly because of discontinuities resulting from the modelling of supersonic flow. For the lift and pitching moment coefficient at zero degrees angle of attack, this problem was solved by deriving the polynomials for two degrees angle of attack which led to better accuracies. For the zero-lift drag, the problem was not solved, however. The zero-lift

drag analysis will therefore be completely based on semi-empirical methods like those of [ref.2] and [ref.3].

Secondly, applying optimal camber and twist can lead to very large reductions in supersonic drag due to lift as compared to flat wings. For realistic planforms, induced drag reductions to about 55% of the induced drag of a flat wing can be achieved. This in turn results in important improvements of the supersonic glide ratio, which affects the range parameter. It can not yet be concluded from this chapter, however, whether the application of camber and twist is a prerequisite for supersonic aeroplanes with a transpacific capability.

## 10.7 References

1. Brown, C.E.; *Theoretical Lift and Drag on Thin Triangular Wings at Supersonic Speeds*; NACA Rept. No. 839, November, 1964
2. ESDU Engineering Sciences Data Unit, Vol.3b
3. Hoak, D.E.; *USAF Stability and Control DATCOM*; 2<sup>nd</sup> ed., 1968
4. Hoeijmakers, H.W.M.; *A Panel Method for the Determination of the Aerodynamic Characteristics of Complex Configurations in Linearized Subsonic and Supersonic Flow*; National Aerospace Laboratory NLR, NLR TR 80124 U, part 1,2,3, December 1980
5. Hoeijmakers, H.W.M.; National Aerospace Laboratory NLR, Memorandum AT-83-005 U, November 1983
6. Hutchison, M.J.; Huang, X.; Mason, W.H.; Haftka, R.T. and Grossman, B.; *Variable-Complexity Aerodynamic-Structural Design of a High-Speed Civil Transport Wing*; AIAA 92-4695, September 1992
7. Jones, R.T.; *The Minimum Drag of Thin Wings in Frictionless Flow*; Journal of the Aeronautical Sciences, Vol.18, No.2, February 1951
8. Middel, J.; *Development of a Computer Assisted Toolbox for Aerodynamic Design of Aircraft at Subcritical Conditions with Application to Three-Surface Aircraft*; Ph.D. dissertation, ISBN 90-6275-768-5 / CIP, Delft University Press, 1992
9. Schlichting, H. and Truckenbrodt, E.; *Aerodynamik des Flugzeuges*; Springer Verlag, Berlin 1960
10. Torenbeek, E.; *On the Conceptual Design of Supersonic Cruising Aircraft with Subsonic Leading Edges*; Delft Progress Report, 8 (1983), November 1982

## Appendix Compressibility Correction for Friction Drag Coefficient

### Notation

- C coefficient  
 $C_f$  friction drag coefficient  
M Mach number
-

N	power in friction drag equation
Re	Reynolds number
r	recovery factor
T	temperature
V	velocity
x	lengthwise coordinate

$\mu$	viscosity coefficient
$\rho$	density
$\tau$	shear stress
$\omega$	power

### subscripts

i	incompressible
c	compressible
ref	reference quantity in compressible flow
w	wall

### Compressibility Correction for Friction Drag Coefficient

The following is largely based on [ref.1] and [ref.2].

The incompressible local friction drag coefficient is usually given by:

$$C_{f_i} = \frac{C}{Re^N} \quad (255)$$

in which  $C = 0.0592$  and  $N = 0.2$  for a turbulent boundary layer (determined experimentally).

Since the pressure is constant in the boundary layer of a flat plate, the following applies for an ideal gas:

$$\frac{\rho_{ref}}{\rho_{\infty}} = \frac{T_{\infty}}{T_{ref}} \quad (256)$$

in which the reference quantities indicate compressible flow. The relation between the



viscosity coefficient and the temperature can be written as follows:

$$\frac{\mu_{\text{ref}}}{\mu_{\infty}} = \left( \frac{T_{\text{ref}}}{T_{\infty}} \right)^{\omega} \quad (257)$$

in which the power  $\omega$  is approximately 0.75.

Now, the shear stress on the wall in compressible flow can be related to that in incompressible flow as follows:

$$\tau_w = \frac{1}{2} \rho_{\text{ref}} V_{\infty}^2 \frac{C}{\left( \frac{\rho_{\text{ref}} V_{\infty} x}{\mu_{\text{ref}}} \right)^N} = \frac{1}{2} \rho_{\infty} V_{\infty}^2 \frac{C}{\left( \frac{\rho_{\infty} V_{\infty} x}{\mu_{\infty}} \right)^N} \left( \frac{T_{\infty}}{T_{\text{ref}}} \right)^{1-N(\omega+1)} \quad (258)$$

Thus, the relation between the friction coefficient in compressible flow and that in incompressible flow can be written as:

$$C_{f_c} = \frac{C_{f_i}}{\left( \frac{T_{\text{ref}}}{T_{\infty}} \right)^{0.65}} \quad (259)$$

The reference temperature in compressible flow relative to the ambient temperature depends on the wall temperature (and therefore on the Mach number). In [ref.1] this relation is given as:

$$\frac{T_{\text{ref}}}{T_{\infty}} = 1 + 0.72 \left( \frac{T_w}{T_{\infty}} - 1 \right) \quad (260)$$

Neglecting heat transfer, the wall temperature equals:

$$T_w = T_{\infty} \left( 1 + r \frac{\gamma - 1}{2} M_{\infty}^2 \right) \quad (261)$$

Taking the recovery factor  $r$  equal to 1, this yields:

$$\frac{T_{ref}}{T_{\infty}} = 1 + 0.144 M_{\infty}^2 \quad (262)$$

### References

1. Hayes, W.D. and Probstein, R.F.; *Hypersonic Flow Theory*; Academic Press, New York, 1959
  2. Truitt, R.W.; *Fundamentals of Aerodynamic Heating*; The Ronald Press Company, New York, 1960
-

# 11. The Propulsion Model

## 11.1 Summary

This chapter describes the development of an analysis procedure for the off-design performance of bypass engines destined to power supersonic cruising transport planes. Most published methods suitable for an evaluation of the off-design characteristics of engines in the conceptual stages of the design treat turbojet engines, or turbofan engines with separated exit flow. Since bypass engines (mixed exit flow) are especially advantageous at low bypass ratios, they are often applied on supersonic aircraft. Considering mixed exit flows introduces an extra condition that must be fulfilled in the off-design analysis, reducing the range over which the engine parameters can be varied. Most published off-design methods require a fixed working point for one of the turbines or a fixed nozzle throat area, which imposes constraints on the operational envelope of the engine. This causes a reduction in the flexibility of the design process, that cannot be afforded in the design of a supersonic cruising civil transport. Furthermore, flexibility is important because of the use of a genetic optimization algorithm, since this kind of algorithm tends to make large "jumps" through the design space. This should not result in termination of the analysis, but instead the optimization algorithm should as much as possible be provided with results to "learn" from. Finally, most methods do not take the intake design into account. Since engines designed for supersonic cruising airplanes require spilling of excess air during off-design operations, which results in a loss of installed thrust, and since the influence of the intake efficiency on the performance of the engine is quite large, it is necessary to incorporate the intake design as well. All this calls for the development of a new analysis procedure.

Some typical problems encountered in supersonic propulsion are discussed in this chapter. An analytical procedure, developed by the author, to determine the off-design working point of a bypass engine is presented. The procedure does not impose any constraints on the working points of the turbines. To account for the phenomenon of excess air, the program is equipped with a detailed intake routine in order to provide for the necessary amount of air massflow and to calculate the intake efficiency and drag. The result of this procedure is a single, closed-form expression, based on the assumption that the fan only compresses the bypass air, that yields an off-design working point of the engine at a given Mach number, altitude and low pressure turbine pressure ratio.

## 11.2 Notation

$A$	area
$a$	local speed of sound
$c_p$	specific heat at constant pressure
$D$	drag
$d$	diameter
	height

---

---

$f$	high pressure spool function
$g$	gravitational acceleration
$h$	altitude
$I$	momentum
$I_{sp}$	specific impulse
$M$	Mach number
$\dot{m}$	massflow
$a/f$	air-to-fuel ratio
$p$	pressure
$R$	universal gas constant
$T$	temperature
	thrust
$v$	velocity
$W$	work
$\gamma$	specific heat ratio
$\delta$	flow deflection angle
	intake ramp deflection angle
$\varepsilon$	total pressure ratio
$\eta$	efficiency
$\eta_r$	recovery factor (intake efficiency)
$\lambda$	bypass ratio
$\nu$	Prandtl-Meyer function
$\Pi$	pressure ratio
$\sigma$	shockwave angle
$\Phi$	coefficient in off-design working point equation
$\varphi$	engine total temperature ratio relative to design point value
	flow function
$X$	coefficient in off-design working point equation
$\Psi$	coefficient in off-design working point equation

**superscript**

*	critical
---	----------

**subscripts**

(numbers refer to station positions defined in Figure 44, unless indicated otherwise)

a	additive
ab	afterburner
bleed	bleed air
c	core
	compressor
capt	captured
cc	combustion chamber
corr	corrected

---

des	design point
e	exit
eng	engine
exp	expansion
f	fan
	fuel
g	combustion gases
i	in front of $i^{\text{th}}$ shockwave
ii	behind $i^{\text{th}}$ shockwave
inst	installed
int	intake
m	mechanical
max	maximum
min	minimum
mix	mixer
mom	momentum
net	net
pol	polytropic
pres	pressure
s	spillage (air)
	shockwave
t	total
	turbine
	nozzle throat

### abbreviations

HP	high pressure
HPT	high pressure turbine
LP	low pressure
LPT	low pressure turbine
SFC	specific fuel consumption
SST	Supersonic Transport
TET	turbine entry (total) temperature

## 11.3 Introduction

The propulsive device exerts a paramount influence on the entire airplane design. Not only does it determine the amount of thrust available at critical points in the aircraft's flight envelope, like cruise, climb and take-off, but also it influences range performance through the fuel consumption and powerplant drag under different operating circumstances, as well as the total airplane weight and center of gravity position.

Because of the large number of interdisciplinary links departing from the propulsion discipline, it is very important to incorporate the powerplant design in any multivari-

ate airplane optimization study. This is especially true for the optimization of a supersonic cruising transport airplane, as in this case the engine must operate throughout largely different regimes of flight altitudes and Mach numbers, with conflicting demands. Therefore it is necessary to take the engine design into account not just by optimizing a scale factor relative to a selected baseline engine, but in a more detailed manner, thus creating a powerplant design that will meet all its demands while simultaneously yielding an optimal aircraft design.

In the present study, the intake area, the bypass ratio and the compressor pressure ratio were chosen as design variables. As an engine will be sized for a certain design condition, the design Mach number and altitude are design variables as well. The turbine entry temperature was not chosen as a design variable, but is generally fixed at a maximum allowable value with respect to turbine material properties, unless a lower value must be chosen because of constraints imposed on the compressor. Since in the present model, the total pressures of the hot and the cold flow at the point of mixing are put equal, the engine is fully defined in this way (in reality, this assumption requires variable geometry in the exhaust and mixer).

Furthermore, the engine performance depends on the actual, off-design Mach number and altitude. Thus, in the present study, seven variables determine the delivered thrust and fuel consumption, five of which are design variables. The values of the design variables are to be determined during the design or optimization process of the airplane. The off-design Mach number and altitude are input from several performance algorithms that need the thrust and specific fuel consumption at certain points in the aircraft flight envelope. Therefore, during the analysis of the engine, these variables are assumed to be known. The intake area remains constant throughout the calculations. The design compressor pressure ratio and the design bypass ratio are fixed at the design point condition but during the off-design operations the compressor pressure ratio and the bypass ratio change. The way in which the design point calculations are carried out will be explained in paragraph 11.6.

Because of the relatively low Mach numbers involved ( $< 2.5$ ), no ramjets, but only turbo engines are considered.

Since the present chapter (and in fact the entire study) deals with conceptual design, with the accent on the evaluation of multidisciplinary optimization techniques, the compressor characteristics are not taken into account. The compressor and fan working points during off-design operation are calculated from the power equilibrium and massflow conservation requirements plus a condition to enable mixing of the hot and cold flow without losses. In this conceptual stage the occurrence of compressor surge or stall is not taken into account. One additional constraint is imposed, however. The engine frontal area is chosen in such a way that the maximum occurring entry Mach number does not exceed 0.5, which is assumed to be a good guess for the largest Mach number that can be handled by the fan and compressor without dramatic loss of efficiency. The use of supersonic fans is at present only studied in conceptual studies and will not be considered here. The same applies for variable cycle engines with duct burning. Throughout the calculation procedure the

---

efficiencies of the rotating devices are kept constant at a predetermined value, in this case 95%, whereas the mechanical efficiencies of the spools are assumed to be 98%. Lastly, the combustion chamber efficiency is put at 90% with a 2% total pressure drop.

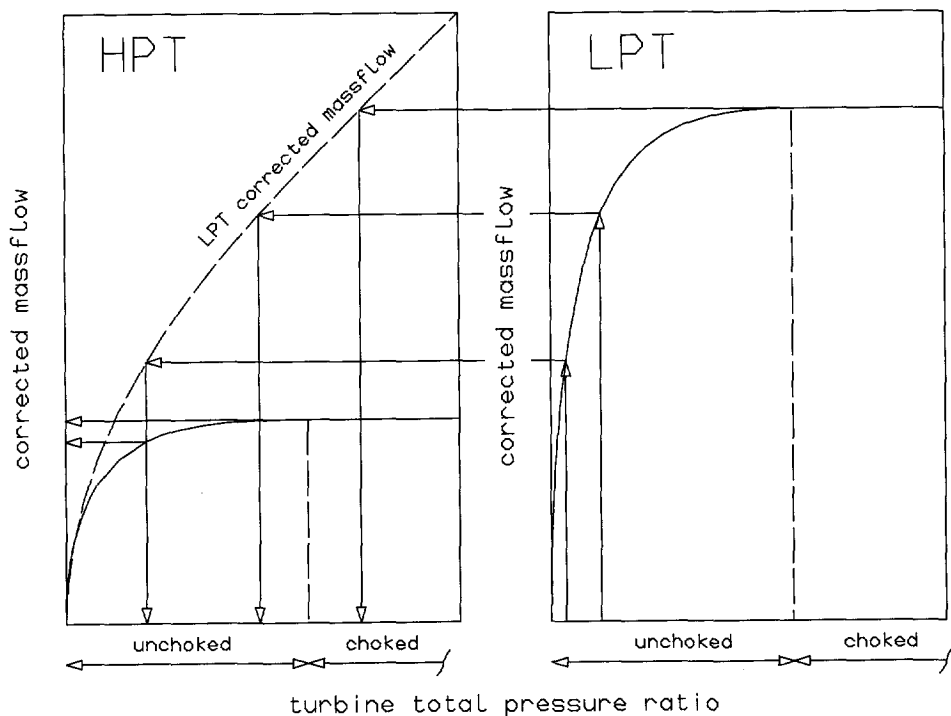
It is assumed that the low pressure compressor, or fan, only compresses the bypass air, whereas the core air is exclusively compressed by the high pressure compressor (model [ref.4], in Figure 44 this is indicated by hatching the outer part of the fan only). The aft-fan concept is an exact application of this model.

## 11.4     Some Typical Off-design Problems of Supersonic Propulsion

While design point calculations of turbo engines proceed in a rather straightforward manner, as certain parameters are chosen as design variables and therefore have a known value in the design point, off-design calculations are much more difficult. The all important process of matching the compressor and turbine works, as well as the massflow through potential throats like the intake, compressors, turbines and the nozzle can only be carried out if the off-design behavior of these components is known. As far as the intake is concerned, in this study its off-design characteristics will be predicted by a rather detailed intake analysis. The combustion chamber temperatures are considered low enough (less then 1500 K and less than 2000 K for the afterburner) to justify the negligence of real gas influences, and therefore no special off-design calculations are performed for the combustion parameters.

The off-design characteristics of the rotating devices, however, can only be determined with the use of compressor and turbine maps which yield the corrected mass flow through the component under consideration as a function of the total pressure ratio and the rotational speed. To each combination of the total pressure ratio and the presented corrected mass flow corresponds a certain rotational speed of the spool. Beyond a certain critical value of the total pressure ratio, however, the rotational speed of the spool and the total pressure ratio no longer influence the corrected massflow, which will thereafter stay constant at its maximum value. At that point, the turbine is said to be choked. Contrary to a throat in a jetpipe, the critical total pressure ratio for a choked turbine is somewhat higher than the value of 1.9 which applies to a polytropic nozzle efficiency of 95% and a specific heat ratio of 1.33 for the combustion gases. This is caused by the fact that the turbine rotates and a good estimation for the critical total pressure ratio of a turbine is 2.5. Since compressor and turbine maps are generally not available at the conceptual stage of the design process, some special measures must be taken to facilitate the off-design calculations.

In some publications ([ref.8], [ref.9], [ref.11]), the corrected mass flow  $\frac{\dot{m}\sqrt{T_t}}{P_t}$  through the exhaust nozzle is fixed, which in turn fixes the total pressure ratio of the turbine throughout the off-design regime. Under this condition, a simple relation can be



**Figure 43** Generalized turbine characteristics

deduced for the off-design compressor pressure ratio. For a double spool engine, it is in principle sufficient that the low pressure turbine be choked (if the corrected massflow through the nozzle is fixed as well, this leads to a fixed operating point for both turbines). This is illustrated by Figure 43, which shows the relation between the so-called generalized turbine characteristics of the low pressure turbine (LPT) and the high pressure turbine (HPT). These generalized characteristics provide the relation between the corrected massflow and the total pressure ratio of a turbine independent of its rotational speed. This relation can be approximated by a simple formula which will be discussed later on, since the method presently discussed ([ref.8], [ref.9], [ref.11]) is independent of the exact shape of the turbine characteristic as the analysis is fixed at one point.



Referring to the station definition of Figure 44, the following observation can be made:

$$\frac{\dot{m}_t \sqrt{T_{5_t}}}{P_{5_t}} = \frac{\dot{m}_t \sqrt{T_{4_t}}}{P_{4_t}} \frac{P_{4_t} \sqrt{T_{5_t}}}{P_{5_t} \sqrt{T_{4_t}}} \quad (263)$$

which yields:

$$\frac{\dot{m}_t \sqrt{T_{5_t}}}{P_{5_t}} = \frac{\dot{m}_t \sqrt{T_{4_t}}}{P_{4_t}} \left( \frac{P_{4_t}}{P_{5_t}} \right)^{1 - \frac{\eta_t(\gamma_s - 1)}{2\gamma_s}} \quad (264)$$

in which the polytropic turbine efficiency was used. The turbine massflow used in these expressions is the sum of the core air massflow and the fuel massflow.

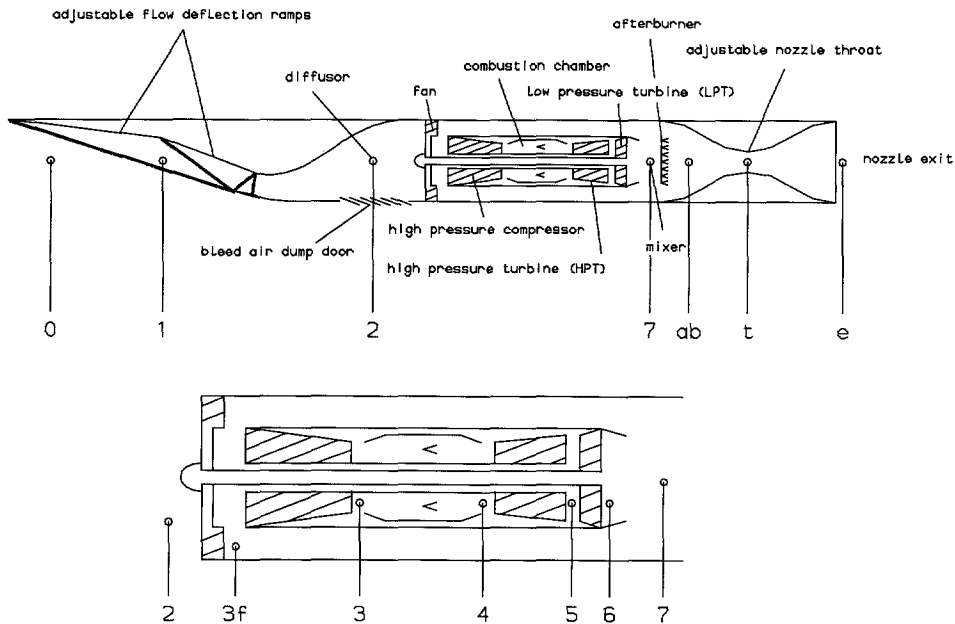


Figure 44 Definition of engine components and station numbers

This relation is presented in Figure 43 by the dashed line and represents the link between the two generalized characteristics (variable turbine vanes are not considered in this study). It can now easily be established, that in case the corrected massflow

through the exhaust nozzle, which is linked to the LPT corrected massflow by a similar relation, is constant, or in case the LPT operates in the choked condition throughout the operational regime of the engine, the working point of the high pressure turbine in terms of corrected massflow, total pressure ratio and -via the polytropic turbine efficiency- the total temperature ratio too, are fixed. Since the work delivered by the high pressure turbine must equal the work performed by the high pressure compressor, the following relation -usually referred to as the power equation- couples the performances of the high pressure turbine and the high pressure compressor, using a mechanical efficiency (no engine bleed and cooling considered):

$$\dot{m}_c c_p (T_{3_t} - T_{2_t}) = \eta_m \dot{m}_c \left(1 + \frac{1}{a/f}\right) c_{p_g} (T_{4_t} - T_{5_t}) \quad (265)$$

or:

$$\frac{T_{3_t}}{T_{2_t}} - 1 = \eta_m \frac{c_{p_g}}{c_p} \left(1 + \frac{1}{a/f}\right) \frac{T_{4_t}}{T_{2_t}} \left(1 - \frac{T_{5_t}}{T_{4_t}}\right) \quad (266)$$

The subscript c refers to the core air flow.

When the power equation is divided by its design point value, and when the HPT operates at its fixed working point throughout the off-design situation, the following relation remains:

$$\frac{\left(\frac{T_{3_t}}{T_{2_t}} - 1\right)}{\left(\frac{T_{3_t}}{T_{2_t}} - 1\right)_{des}} = \frac{\left(1 + \frac{1}{a/f}\right)}{\left(1 + \frac{1}{a/f}\right)_{des}} \frac{\left(T_{4_t}/T_{2_t}\right)}{\left(T_{4_t}/T_{2_t}\right)_{des}} = \frac{\left(1 + \frac{1}{a/f}\right)}{\left(1 + \frac{1}{a/f}\right)_{des}} \varphi \quad (267)$$

After substituting the polytropic relation between the high pressure compressor total pressure ratio and the total temperature ratio:

$$\frac{T_{3_t}}{T_{2_t}} = \epsilon_c^{\frac{\gamma-1}{\eta_c \gamma}} \quad (268)$$

in which the polytropic compressor efficiency is used, the following simple relation remains for the off-design high pressure compressor total pressure ratio (hereafter

called compressor pressure ratio):

$$\varepsilon_c = \left[ 1 + \varphi \frac{\left(1 + \frac{1}{a/f}\right)}{\left(1 + \frac{1}{a/f}\right)_{\text{des}}} \left( \varepsilon_{c_{\text{des}}}^{\frac{\gamma-1}{\eta_c \gamma}} - 1 \right) \right]^{\frac{\eta_c \gamma}{\gamma-1}} \quad (269)$$

When the conditions in front of the compressor are known for the Mach number and altitude under consideration -from the relations describing the ICAO standard atmosphere and the intake performance- all variables through the engine up to the conditions behind the low pressure turbine are known for a chosen value of the combustion chamber exit total temperature (at station 4, also referred to as TET, or turbine entry temperature). For a given value of this temperature, the compressor pressure ratio is known, together with the compressor temperature ratio, yielding the temperature in front of the combustion chamber. From the chosen TET follows the fuel mass flow,  $\dot{f}$ , (through the combustion chamber efficiency with constant  $c_p$ ) and with the core air mass flow the air-to-fuel ratio  $a/f$ . This value should again be substituted in the relation for the off-design compressor ratio since the  $a/f$  value in the off-design was up to this point unknown. This process should be repeated until convergence of  $a/f$ , which, since its influence is not very large, will happen at all times (although the number of iteration steps is in some cases unexpectedly large; that is, about 25).

From the fixed working point of the high pressure turbine follow the total pressures and temperatures at stations 4, 5 and 6. Since the conditions in front of both turbines are at this point fully known, the massflow captured by the intake can be adapted to the demands of both turbines. (This is essential; since the massflow captured by the intake is usually larger than the massflow required by the turbines, the amount of captured intake air must be adapted by spilling the excess air upstream of the turbines, preferably in front of the fan. This procedure will be treated extensively in paragraph 11.7.2).

In case both turbines operate in the choked condition, no further knowledge about the generalized turbine characteristics is necessary but the maximum corrected massflow, which in this case is equal to the corrected massflow through either component at the design point. If either one of the turbines operates in a subcritical condition however, the functional relation between the unchoked turbine massflow and the total pressure ratio must be known. According to [ref.3] this relation can be approximated as follows:

$$\dot{m}_{\text{corr}} = \dot{m}_{\text{corr,max}} \sqrt{1 - \frac{(\Pi_t^* - \Pi_t)^2}{(\Pi_t^* - 1)^2 (\Pi_t)^2}} \quad (270)$$

in which the superscript "\*" denotes the critical value (in this case 2.5) and  $\Pi_t$  denotes the turbine total pressure ratio. This distinctive notation is only used here, so that no further specification of the station numbers is necessary since equation (270) applies to both turbines.

As explained, usually the off-design corrected mass flow through the turbine will be too large according to equation (270) and the excess air must be dumped, leading to intake drag and hence a loss of installed thrust.

The method as described so far has a few disadvantages, however, which reduce its applicability to analyze supersonic engines. Firstly, a fixed corrected mass flow through the exhaust requires a fixed throat area. However, because of independently changing Mach numbers, pressures and temperatures throughout the operational regime, situations may occur at which the nozzle throat is no longer choked. (The opposite case in which the throat is choked but the presented corrected massflow is greater than the maximum corrected massflow will usually not occur in practice, since the turbine(s) will be the critical factor. Because of this, excess air must be spilled and the resulting massflow through the nozzle will lead to a corrected massflow that is too low instead of too high).

In order not to let the nozzle throat constrain the operational envelope of the engine, a variable throat in a convergent-divergent exhaust pipe with constant exit area was chosen for in the present study, which enables the throat to be adjusted to the varying conditions at different off-design points. This implies that the throat will always be at the critical condition, such that a supersonic exit flow is guaranteed, while simultaneously complying with the conservation of massflow requirement. Specifically, this enables free variation of the turbine pressure ratios.

Furthermore, a variable nozzle throat area is necessary in case an afterburner is used, because of the increase in total temperature and fuel mass flow. The feasibility of an afterburner in a civil transport plane is somewhat doubtful, however, with respect to airport noise. It seems impossible to comply with FAR 36 requirements once an afterburner is selected during take-off. On the other hand, however, the use of an afterburner during supersonic transition might still be an option, since environmental noise is no longer of concern during that phase of the flight. In the present method the afterburner option is still maintained, but not used.

Assuming then the use of a variable exhaust nozzle, the method previously described should be adapted, since the working point of the low pressure turbine will move along its characteristic and therefore there is no fixed working point for the high pressure turbine, unless the low pressure turbine is choked. If this is not the case, the throat area must be known in advance and the working point of the high pressure turbine (although now varying throughout the off-design regime) can be computed linking the characteristics of the nozzle, low pressure turbine and high pressure turbine together from the nozzle on forward through the engine. For a chosen value of the TET, the compressor pressure ratio can be determined from the high pressure spool power equation. Obviously, the previously deduced formula based on [ref.11]

---

((269)) relating the off-design compressor pressure ratio to the design point value is now of no use, since it was derived under the assumption of a constant operating point for the high speed turbine.

Unfortunately, for bypass engines under off-design conditions, it is very hard to comply with either of the two prerequisites of the low pressure turbine being choked at all times or the throat area being known in advance. The reason for this problem is the mixing condition. In order to mix the hot core flow and the cold bypass flow without loss of thrust, it is desirable that the velocities ([ref.4]) of both flows are equal. Furthermore, the static pressures on the streamline separating the hot and the cold flow prior to mixing are equal (Kutta condition). This mixing condition is sometimes modelled by just taking care that the total pressures of the two flows are equal ([ref.4] and [ref.10]). Since in reality the Kutta condition applies, it should be realized that small differences between the total pressures will occur (if the Mach numbers of the hot and the cold flow are kept small by a sensible choice of the engine area). In case the condition of equal total pressures will be enforced in combination with the Kutta condition, variable mixer geometry will be required (see Appendix B, chapter 15).

Applying the procedure outlined before, the conditions throughout the engine up to the point where mixing occurs are known. Now the total pressure behind the fan follows directly from the low pressure spool power equation:

$$\lambda c_p (T_{3f_t} - T_{2_t}) = \eta_m \left(1 + \frac{1}{a/f}\right) c_{p_g} (T_{5_t} - T_{6_t}) \quad (271)$$

As the system up to now is completely determined, the mixing condition will be violated, or, in case it is enforced by putting the total pressure behind the fan equal to the total pressure behind the low pressure turbine in the power equation, the low pressure turbine pressure ratio will not correspond to the presumed working point. In case a choked condition was presumed for the LPT, enforcing the mixing condition might result in a subcritical LPT pressure ratio.

Since it can be concluded from these observations that an iterative procedure will be necessary anyhow, it is resolved that the best procedure will be to start off with a selected low pressure turbine pressure ratio, without imposing any more demands on the operating points of either the turbines or the nozzle. After the conditions upstream of the LPT are determined, the mixing condition will be enforced, in such a way that it corresponds with the initially selected LPT operating point. This can be done in two different ways. First of all, in case the total pressures behind the fan and the low pressure turbine are to be equal (while simultaneously the low pressure turbine pressure ratio is fixed) it can be deduced from the low pressure spool power equation, that the bypass ratio  $\lambda$  must change. This implies that the fan pressure ratio will change. In the present modelling, it is assumed that the required fan pressure ratio variations are indeed possible. However, if those variations turn out to be too large during a more detailed design phase, the use of variable fan vanes might be

considered. Subsequently it might even be necessary to impose bounds on the fan pressure ratio and therefore on the bypass ratio (see Appendix C, chapter 15).

Depending on the off-design operating conditions of the engine, the massflow through the engine is redistributed over the core and the bypass duct. It should be realized, that in case the bypass ratio decreases, air from the bypass duct is redirected through the core. This can only be realized up to a limited value depending on the amount of air the turbines can handle under the given conditions. If more air is sent through the core than can be handled by the turbines, the excess air must be dumped (either through a dump door in front of the fan, or by spillage in front of the intake, which will be discussed later on). This leads to the previously mentioned loss of installed thrust.

The other way of implementing the mixing condition is to add a mixer to the model, in which the two flows are mixed at different total pressures, but in which the Mach numbers of both flows are adapted to ensure equal static pressures. Both off-design procedures will be discussed in detail later on.

There is another reason why the formula for the off-design compressor pressure ratio (equation (269), fixed operating point of the HPT) is of limited use. This will be explained in the following example, which also sheds some more light on the phenomenon of excess air.

The corrected massflow through the high pressure turbine can be expanded as follows:

$$\frac{\dot{m}_t \sqrt{T_{4t}}}{P_{4t}} = \frac{\dot{m}_2}{(1+\lambda)} \left( 1 + \frac{1}{a/f} \right) \frac{\sqrt{T_{4t}}}{\epsilon_c \epsilon_{cc} P_{2t}} =$$

$$\frac{\frac{M_0 A_0 \sqrt{\gamma}}{\sqrt{RT_0} (1+\lambda)} \left( 1 + \frac{1}{a/f} \right) \sqrt{T_{4t}}}{\epsilon_c \epsilon_{cc} \eta_r \left( 1 + \frac{\gamma-1}{2} M_0^2 \right)^{\frac{\gamma}{\gamma-1}}} \quad (272)$$

in which the intake recovery factor  $\eta_r$  (intake efficiency, total pressure ratio over the intake) and the total pressure loss over the combustion chamber  $\epsilon_{cc}$  are used. In the deduction of this equation the airflow through the engine is -at least for the time being- taken equal to that captured by the intake (area  $A_0$ ).

Dividing this expression by the design point value, while for the moment restricting the engine operations to the stratosphere and neglecting the influence of the change of

fuel-to-air ratio and intake efficiency, assuming that the bypass ratio is kept constant yields:

$$\frac{\left( \frac{\dot{m}_t \sqrt{T_{4_t}}}{P_{4_t}} \right)}{\left( \frac{\dot{m}_t \sqrt{T_{4_t}}}{P_{4_t}} \right)_{\text{des}}} = \frac{\epsilon_{c_{\text{des}}}}{\epsilon_c} \sqrt{\frac{T_{4_t}}{T_{4_{t_{\text{des}}}}}} \frac{M_0}{M_{0_{\text{des}}}} \frac{A_0}{A_{0_{\text{des}}}} \frac{\left( 1 + \frac{\gamma-1}{2} M_{0_{\text{des}}}^2 \right)^{\frac{\gamma}{\gamma-1}}}{\left( 1 + \frac{\gamma-1}{2} M_0^2 \right)^{\frac{\gamma}{\gamma-1}}} \quad (273)$$

If, for instance, an engine is considered which was designed for a cruise Mach number of 2.5, at which design point the high pressure turbine would be choked, then the following applies if the engine would be operated at an off-design free stream Mach number of 1.5:

$$\frac{\left( \frac{\dot{m}_t \sqrt{T_{4_t}}}{P_{4_t}} \right)}{\left( \frac{\dot{m}_t \sqrt{T_{4_t}}}{P_{4_t}} \right)_{\text{des}}} = 2.8 \frac{\epsilon_{c_{\text{des}}}}{\epsilon_c} \sqrt{\frac{T_{4_t}}{T_{4_{t_{\text{des}}}}}} \frac{A_0}{A_{0_{\text{des}}}} \quad (274)$$

Normally, it is endeavored to keep the turbine entry temperature as high as possible in order to be able to inject a large fuel mass flow. This temperature is limited with respect to the strength, fatigue and creep properties of the Titanium alloy of which the turbine blades are manufactured. A state of the art value for the maximum turbine entry temperature is about 1350 K. For a second-generation supersonic transport plane a value of 1500 K and perhaps even as high as 1600 K might be anticipated. In the present model the design value of the TET (cruise condition) is put at its maximum allowable value (which in this study is put at 1500 K).

Returning to the previous formula, in case -for the sake of discussion- it is assumed that the turbine is still choked in the off-design condition under consideration and if the turbine entry temperature would be kept at its maximum allowable value, then it follows that the compressor pressure ratio must become almost three times as high as its design value. In terms of formula (269),  $\phi$  would be greater than one. This is unacceptable, since the value for the off-design compressor pressure ratio is constrained as well. It cannot -or only with a very small margin- become greater than its design value. Furthermore, a second constraint is imposed on the compressor pressure

ratio in order to prevent the compressor exit temperature from exceeding about 850 K (this value is based on the Olympus engine which powers Concorde, and it might be expected that for a second-generation SST temperatures of up to 950 K are achievable). The reasons for this limit are again the material properties; the value being so much lower than that for the turbine is caused by the fact that compressor blades are very difficult to cool at present.

Since the compressor pressure ratio cannot be adapted sufficiently to comply with equation (274), either the bypass ratio must change or the excess air must be dumped (or both). As will be explained later on, this can be done in two different ways; either by adapting the shockwave system in the intake by means of the adjustable flow deflection ramps, causing a decrease of  $A_0$  with respect to the design condition, or by opening a dump door in front of the compressor (equation (274) was derived under the assumption of constant massflow through the compressor and turbine). Although it is possible to limit the amount of excess air by diminishing the turbine entry temperature somewhat, this cannot solve the problem completely since there must always be sufficient margin to add fuel. The maximum allowable compressor pressure ratio at the design Mach number of 2.5 in order to keep the compressor exit temperature at 850 K is 6.35 (based on a compressor polytropic efficiency of 95%). With this value at  $M=1.5$ , the compressor exit temperature is about 550 K. Diminishing the turbine entry temperature to 1200 K would still require about 60% spillage air. Equation (269) for the off-design compressor pressure ratio as a function of  $\phi$  obviously is of no use for this kind of off-design operation.

Taking the considerations discussed in this paragraph into account, the procedure applied for the off-design calculations of a bypass engine will consist of an iterative procedure which determines the compressor pressure ratio and the turbine entry temperature and in a second iterative loop the air-to-fuel ratio. The procedure will check these variables to ensure that their maximum allowable values are not exceeded and if necessary will enforce the constraints on the operation of the turbines with reference to the maximum allowable corrected massflow by dumping any excess air.

## 11.5 Intake Model

A very important part of supersonic powerplant design, is the design of the intake. The main task of the intake is to decelerate the captured supersonic air to subsonic speed at the least possible loss of total pressure. This total pressure loss is measured by means of the ratio of the total pressures immediately behind the intake and immediately in front of it. This ratio is referred to as the recovery factor or intake efficiency and is a direct measure of the increase in entropy in the intake. It is very important that this recovery factor be as high as possible, as it is of large influence on the available thrust (a rule of thumb states, that each percent loss of intake efficiency results in a one and a quarter percent loss of thrust). Compression in the inlet usually takes place by means of a number of oblique shockwaves which will result in a weaker final normal shockwave (which takes care of the final deceleration to subsonic speed) than would be the case with a pitot inlet. This results in a better intake

---



efficiency. The larger the number of oblique shockwaves for a given Mach number, the higher the intake efficiency. The limit is reached for isentropic compression, which can be achieved by a concave intake ramp.

Compression of the intake air can take place in different ways: external compression, internal compression or a mixture of both. Usually, the intake operates at the critical condition in the design point, in order to avoid spillage drag. This implies that the external waves must all hit the intake lip. If the final shockwave is to be a normal one, and if no detached internal shockwaves may occur, this implies that the external lip deflection angle will increase fast with an increasing number of oblique shockwaves. As a shockwave will emanate from the lip as well, a large lip deflection angle will cause a large pressure to act on the external lip area. The lip drag can amount to 80% of the total powerplant drag. This large lip drag is the reason that the fully isentropic or Oswatitsch intake is not applied in practice. Furthermore, the Oswatitsch intake is only suitable for one free stream Mach number, and therefore requires variable geometry. (The principle is at least to some extent realized in Concorde, however, where part of the intake flow is compressed isentropically by leading it past a curved ramp. Instead of terminating the shockwave system with a normal shock, in Concorde the final deceleration to subsonic speed takes place by means of the so-called strong solution of an oblique shockwave, which enables a smaller lip angle [ref.7]). The disadvantage of internal compression on the other hand is, that the intake will become very long, which implies a larger weight and a higher friction drag. It was decided not to take the number of intake shocks as a design variable (also because it is a discrete variable), but to fix it at a reasonable compromise, that is, two external and one internal oblique shock.

In this application, only two-dimensional intakes are considered. This decision is not based on physical consideration, but made for simplicity of calculation. In industry, the choice between two-dimensional or rotation-symmetrical intakes is still a source of constant discussion. The American supersonic cruise projects invariably are equipped with rotation-symmetrical inlets with adjustable (translating) cones or spikes to generate oblique shocks, with as a practical example the Lockheed SR-71 *Habu*, or *Blackbird* as it is generally known. The Anglo-French Concorde uses a two-dimensional intake with adjustable flow deflection ramps.

The present intake model was developed by the author (in cooperation) for a conceptual design study of a hypersonic cruising transport plane [ref.2]. It was extended to account for some conditions only encountered at low supersonic speeds (like the occurrence of strong solutions and detached oblique shockwaves). The model is based on the fact that the intake houses two adjustable ramps and has a fixed intake lip (Figure 44). The intake is sized for two criteria. Firstly, the external shocks should hit the intake lip, which implies that no spillage drag occurs. Secondly, the intake recovery factor is optimized for the design condition. According to [ref.6], for an intake with an arbitrary number of oblique shockwaves and one normal shock, the intake efficiency is maximal if the total pressure ratio over each oblique shock is the same. For a given total pressure ratio over the oblique shockwave system and number of oblique shocks, the total pressure drop over each shock is known. Then one value

---

of  $M_i \sin \sigma$  can be calculated which applies to each shock, using the following equation (a well known relation from elementary oblique shockwave theory):

$$\eta_s = \left( \frac{(\gamma+1)M_i^2 \sin^2 \sigma}{2 + (\gamma-1)M_i^2 \sin^2 \sigma} \right)^{\frac{\gamma}{\gamma-1}} \frac{1}{\left( 1 + \frac{2\gamma}{\gamma+1} (M_i^2 \sin^2 \sigma - 1) \right)^{\frac{1}{\gamma-1}}} \quad (275)$$

The value for  $M_i \sin \sigma$  is computed using the Newton-Raphson iterative procedure. Thus, the shockwave angles are only depending on the Mach number in front. The Mach number in front of the first shock is known since it is equal to the free stream Mach number. The Mach number in front of the other shocks are equal to the Mach number behind the preceding shock, which follows from another elementary shockwave formula:

$$M_{ii} = \sqrt{\frac{(\gamma+1)^2 M_i^4 \sin^2 \sigma - 4(M_i^2 \sin^2 \sigma - 1)(\gamma M_i^2 \sin^2 \sigma + 1)}{2\gamma(M_i^2 \sin^2 \sigma - 1) + \gamma + 1} \left[ (\gamma-1)M_i^2 \sin^2 \sigma + 2 \right]} \quad (276)$$

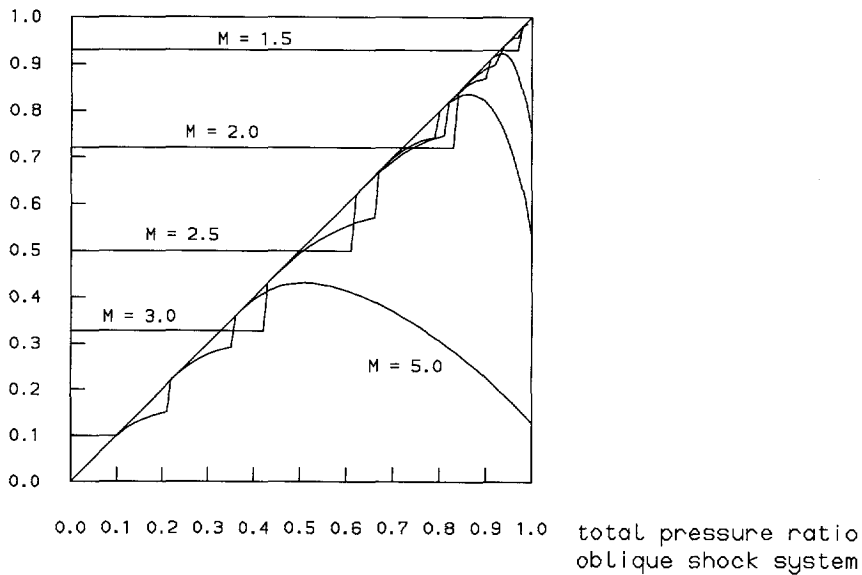
Once all Mach numbers are determined, the shockwave angles can be calculated from the obtained value of  $M_i \sin \sigma$ . The corresponding deflection angle can be obtained from:

$$\tan \delta = \frac{2(M_i^2 \sin^2 \sigma - 1)}{\tan \sigma \left[ (\gamma+1)M_i^2 - 2(M_i^2 \sin^2 \sigma - 1) \right]} \quad (277)$$

The deflection angle of the third shock is the intake lip internal angle. Two degrees are added to this value as construction thickness to yield the lip external angle.

Since the Mach number behind the last oblique shock is known, the total pressure drop over the terminating normal shock can be calculated. The intake recovery factor then follows easily by multiplying this value by the total pressure drop over the oblique shockwave system. This procedure then can be repeated for a different total pressure ratio over the oblique shockwave system, until the optimal intake efficiency is obtained for the cruise condition. In Figure 45 the recovery factor is drawn as a function of the design total pressure ratio of the oblique shockwave system for different Mach numbers. The kinks indicate the points where the combination of Mach number and required pressure ratio over the shock are such, that a strong solution is obtained, resulting in a subsonic Mach number behind the shock and in practice often a detached shockwave. At such points the number of oblique shocks is reduced by one and the procedure repeated. Obviously, the attainable number of shockwaves decreases with the free stream Mach number.

intake efficiency

**Figure 45** Intake recovery factor

With the imposed condition that the external shocks should hit the intake lip in the design point, all ramp dimensions can be determined from the known shockwave and ramp deflection angles and the intake height which can be obtained from the intake area.

The only parameters that influence the intake design are the intake height and the cruise Mach number. The intake height is related to the intake area by means of the required Mach number in the diffuser in front of the fan. This Mach number is usually limited to a value of about 0.5, since most fans at present are unable to cope with faster airflows. Putting the mass flow immediately behind the normal shockwave equal to that immediately in front of the fan, it follows:

$$\frac{M_2}{\left(1 + \frac{\gamma-1}{2} M_2^2\right)^{\frac{(\gamma+1)}{2(\gamma-1)}}} = \frac{M_0 A_0 / A_2}{\eta_r \left(1 + \frac{\gamma-1}{2} M_0^2\right)^{\frac{(\gamma+1)}{2(\gamma-1)}}} \quad (278)$$

If for a constant intake area and engine frontal area the free stream Mach number is decreased, the right-hand-side term increases as the term depending on the Mach number increases faster than the recovery factor. Realizing that the Mach number

behind the normal shock is subsonic, this implies that this Mach number increases too. Therefore, if the engine frontal area was chosen in such a way as to yield a fan entry Mach number of 0.5 for a certain free stream Mach number, then a decrease of this Mach number will lead to an unacceptably high fan entry Mach number. This implies that the turbofan diameter should be based on Mach 1. It is assumed, that the sides of the nacelle are parallel as in Figure 44. Since in this stage of the calculations the nozzle exhaust area is not yet known, but the engine must fit in the nacelle, the following must apply to the engine diameter:

$$A_2 = 1.34 A_0 \text{ (from equation (278), after substituting } \gamma=1.4, M_0=1 \text{ and } M_2=0.5)$$

which yields:

$$d_{\text{eng}} = 1.306 \sqrt{A_0} \quad (279)$$

since the engine frontal area  $A_2$  is circular. As the intake area is rectangular, the following applies for the intake height:

$$d_{\text{int}} = 0.766 \sqrt{A_0} \quad (280)$$

For a given intake area and cruise Mach number therefore, the intake height, the position and angle of the lip and the dimensions of the ramps can be determined. Also, with an estimated fixed value for the expansion slope of the diffuser, the total intake length and wetted area can be calculated, which is needed for the friction drag calculation.

It must be realized, that the intake optimization procedure constitutes an implicit constraint on the optimization of the aircraft. For instance, in case the aircraft is optimized for minimum take-off weight, it might be that an intake with a smaller external lip angle than would follow from the above procedure, would justify a somewhat smaller engine for a certain critical design condition. The decrease in engine weight might compensate for the possible increase in mission fuel weight as a result of the non-optimal intake recovery factor. In principle, the intake lip angle could be kept at zero. This would, however pose constraints on the maximum deflection of the intake flow, in order to prevent a detached internal shockwave. In turn, this would require a relatively sharp first shockwave to prevent spillage, resulting in a long intake. Especially at lower Mach numbers, this would cause spillage drag. This is a typical feature of optimization; an optimal design is valid only for the methods used to calculate it.

As the lengths of the ramps as well as the position and angle of the intake lip are known, the intake can be adjusted for off-design Mach numbers. Again it is endeavored to let the external shocks hit the intake lip in order to avoid spillage drag, which determines the desired shockwave angle. Especially when the intake is designed for a relatively high Mach number, this shockwave angle under off-design

conditions can be too small. At each Mach number there is a minimum attainable shockwave angle, corresponding to a deflection angle approaching zero. In this intake routine, however, it is assumed that even an undeployed intake ramp will cause a minimum flow deflection of five degrees. In case the minimum attainable shockwave angle exceeds the value needed to hit the intake lip, spillage occurs.

Since the Mach number behind the first shock and thus in front of the second one, is even smaller than the free stream Mach number, the same problem might occur for the second oblique shock (if applicable), even when the first shock does hit the intake lip. The latter case would cause the two shocks to intersect, thus generating a vortex layer. This vortex layer should not enter the intake for the risk of Ferri buzz. Should this situation occur, the deflection of the second ramp will be taken equal to that of the first ramp thus eliminating the second shock. Depending on the deflection angle of the first ramp and the lip angle, this might result in an expansion flow on the inner side of the lip. Although such an expansion will not lead to a loss of total pressure as the expansion flow is isentropical, it will increase the Mach number in front of the normal shock, with consequences for the intake efficiency. In case both external shocks cannot hit the intake lip, the second external shock is eliminated as well, since otherwise the amount of spillage drag would be too large.

If a strong solution is obtained for either one of the external shocks, no further shocks (normal nor oblique) are generated. This will lead to a lower intake efficiency than might have been obtained with more oblique shocks, but on the other hand it might reduce the lip drag since the flow in front of the lip is subsonic in this case and no shock will emanate from the lip. If a strong solution is obtained for the internal shock, it will be replaced by a normal one.

The internal shockwave angle follows from the conditions behind the last external shock and the fixed internal lip angle. In case the resulting deflection angle is positive, the corresponding shockwave angle can be determined by using the regula-falsi method to find a solution for the shockwave angle, between zero and the maximum attainable shockwave angle corresponding to the deflection angle under consideration, in such a way that equation (277) is satisfied. In case the resulting deflection angle is negative, the Mach number behind the expansion waves is calculated from the Prandtl-Meyer equations:

$$\delta_{\text{exp}} = v(M_{ii}) - v(M_i) \quad (281)$$

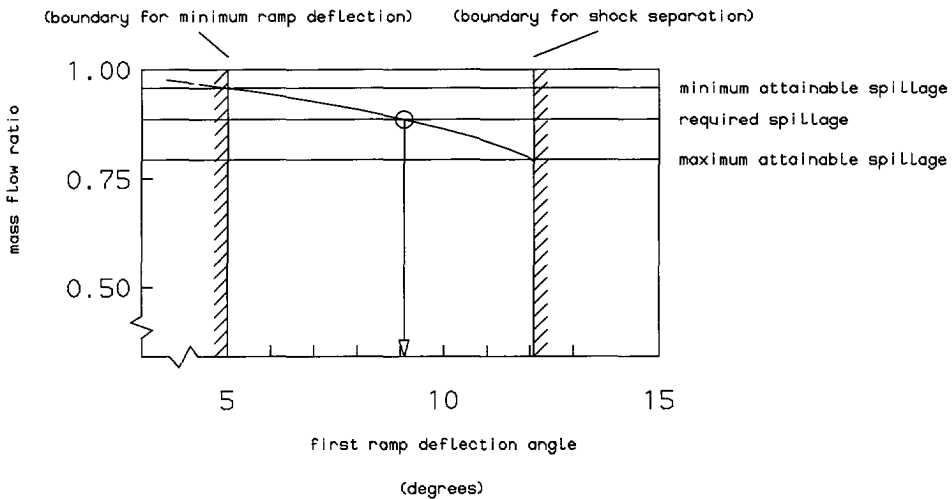
in which  $v$  is known as the Prandtl-Meyer function:

$$v(M) = \sqrt{\frac{\gamma+1}{\gamma-1}} \arctan \sqrt{\frac{\gamma-1}{\gamma+1} (M^2 - 1)} - \arctan \sqrt{M^2 - 1} \quad (282)$$

The expansion angle and the Mach number in front of the expansion are known. The

Mach number behind the expansion is calculated using the regula-falsi method since it must be greater than the Mach number in front of the expansion and smaller than an assumed large upper bound. This value was taken as three times the lower bound.

When spillage air is deliberately generated in order to provide the high pressure turbine with the required massflow, first of all the maximum possible deflection angle corresponding to the considered free stream Mach number is calculated. If this deflection is not sufficient for the required amount of spillage air, then the first ramp will be deflected to this maximum and the rest of the excess air is dumped. The Mach number behind the first shock will then be subsonic. If the required amount of spillage air is smaller than the maximum attainable amount, then the required shockwave angle can be found with the use of the regula-falsi method. The unknown shockwave angle must lie between the design point value (no spillage air, in which case the difference between the available spilled massflow and the required spilled massflow is negative) and the maximum attainable shockwave angle (in which case the difference between the available spilled massflow and the required spilled massflow is positive). This procedure is illustrated in Figure 46. This example applies to an intake designed for  $M=2$  at an off-design Mach number of 1.5. The design shockwave angle for this intake is 36.2 degrees. At  $M=1.5$ , the deflection angle corresponding to this shockwave angle is smaller than the minimum value of 5 degrees. Therefore, at  $M=1.5$  the no-spillage situation which by definition exists at the design point (mass flow ratio equals 1) cannot be realized and some minimum spillage occurs. The maximum attainable amount of spillage is achieved by a ramp deflection of about 12 degrees, corresponding to a shockwave angle of 66.6 degrees. This is a strong solution and hence the Mach number behind the shock is subsonic.



**Figure 46 Shockwave angle for required massflow**

In case the amount of excess air is larger than the maximum amount that can be spilled by the intake, the remaining excess air is bled off (dumped by means of a dump door in front of the fan, see paragraph 11.7.2). Furthermore, as will be explained in paragraph 11.7.2 as well, a large increase in installed thrust can be achieved by spilling less air by the intake and dumping more.

## 11.6 Design Point Procedure

The design point calculations proceed as follows. From a call to an atmospheric routine based on the ICAO standard atmosphere, the conditions in front of the intake follow. The intake routine designs the optimal intake for the cruise condition and thus yields the intake efficiency as well as the dimensions of the intake and the intake ramps. Since the compressor pressure ratio at the design point is a design variable (and therefore fixed during the engine performance analysis) and since the turbine entry total temperature (which is the combustion chamber exit total temperature) is fixed at its state of the art maximum, the total pressure and temperature ratio over the compressor as well as the air-to-fuel ratio are known. Also the work of the high pressure compressor is known and therefore the total pressure and temperature behind the high pressure turbine (station 5) can be calculated.

As said before, in this calculation model it is assumed that all core air is compressed by the high pressure compressor which is driven by the high pressure turbine, whereas all the bypass air is compressed by the low pressure compressor (fan) which is driven by the low pressure turbine. Regardless of the off-design mixing procedure chosen, the matching of the fan and the low pressure turbine at the design point takes place under the assumption that the fan exit total pressure must equal the low pressure turbine total pressure in order for mixture to take place without losses ([ref.4], [ref.10]). Since this unknown total pressure must be between the fan entry total pressure (at which value the available work would be positive and the required work zero) and the low pressure turbine entry total pressure (at which value the available work would be zero and the required work positive) the matching equation can be solved by means of the regula-falsi iterative procedure:

$$\lambda c_p T_{2t} \left( \left( \frac{P_{6t}}{P_{2t}} \right)^{\frac{\gamma-1}{\eta_c}} - 1 \right) = \eta_m \left( 1 + \frac{1}{a/f} \right) c_{p_g} T_{5t} \left( 1 - \left( \frac{P_{6t}}{P_{5t}} \right)^{\frac{(\gamma_s-1)\eta_t}{\gamma_s}} \right) \quad (283)$$

Throughout the program, it is endeavored to solve the equations by the regula-falsi method, rather than by the Newton-Raphson method. The major advantage of the regula-falsi method is -contrary to the Newton-Raphson method- that it will always find a solution, it cannot diverge. It is, however, necessary to specify two values between which the solution must lie and this is not possible in all cases. Furthermore, sometimes convergence can be quite slow, in which case special measures are taken to speed things up.

The nozzle exhaust area can be determined according to two different criteria. First, it can be designed for ideal expansion in the design point (which, in principle, would be the cruise condition; diversions from this policy will be treated later on) which results in maximum net thrust for the given engine. For moderate supersonic cruise speeds (Mach number  $< 2.5$ ) and consequently cruise altitudes within the stratosphere, this is acceptable as the exhaust area will stay within reasonable margins. Secondly, the exhaust area could be incorporated as a design variable, which implies fixing it at its design value resulting in nonideal expansion at the design condition. In case at relatively low corrected massflows -such as would occur at take-off, for instance- supersonic exit flow is generated, very low exit pressures would result from the increased exit area of engines designed for ideal expansion at the cruise condition. This results in negative pressure thrust that is not compensated by the increased momentum thrust. A smaller exhaust area would obviously reduce this problem, be it at the cost of a smaller cruise net thrust.

In reality, however, the amount of overexpansion is limited by the occurrence of flow separation in the divergent part of the jetpipe. A constraint can be put on the exit pressure by means of the Summerfield criterion, which states that the exit pressure may not be lower than approximately 0.35 times the ambient pressure. In the program an algorithm is implemented that adjusts the exhaust area until this criterion is met. The adjustment of the exhaust area in this context must be regarded not as a mechanical adjustment, but as a simulated reduction of the exhaust area due to flow separation. The algorithm will reduce the exhaust area with large steps until the criterion is met, and subsequently increase it again with small steps until the criterion is just violated by a small margin.

Applying the Summerfield criterion to overexpanded nozzle flows eliminates the influence of the exhaust area on the net thrust, since it limits the extent to which expansion can take place. Therefore, the nozzle exhaust area will be designed for ideal expansion at the design point and the option to include it as a design variable, though it is still contained in the program, will not be used.

After the design point calculations have been completed, both generalized turbine characteristics are fixed according to equation (270), regardless of the fact if the turbines are choked or not. Since for each turbine a combination of the total pressure ratio and the corrected massflow is known, which constitutes a point on the generalized characteristic, the whole characteristic is fixed. The ratio of the LPT maximum corrected massflow and the HPT maximum corrected massflow depends on the HPT total pressure ratio and on both total pressure ratios relative to their critical value. In

---



case both turbines are choked this ratio becomes:

$$\frac{\left( \frac{\dot{m}_t \sqrt{T_{5_t}}}{P_{5_t}} \right)_{\max}}{\left( \frac{\dot{m}_t \sqrt{T_{4_t}}}{P_{4_t}} \right)_{\max}} = \left( \frac{P_{4_t}}{P_{5_t}} \right)^{1 - \frac{\eta_t(\gamma_t - 1)}{2\gamma_t}} \quad (284)$$

The most important assumptions in the design-point procedure are summarized hereafter :

- intake in design condition (no spillage) (changes during off-design operation)
- turbine entry temperature at maximum (changes during off-design operation)
- bypass ratio at design value (changes during off-design operation)
- compressor pressure ratio at design value (changes during off-design operation)  
(this value is limited by the maximum achievable compressor exit temperature)
- intake area at design value (constant)
- exit area designed for ideal expansion at cruise condition
- total pressures of hot and cold flow equal in mixer throughout operational envelope
- all efficiencies and combustion chamber pressure ratio constant throughout operational envelope

## 11.7 Off-design Procedure

As explained previously, the procedure starts with the selection of the low pressure turbine total pressure ratio (at a small step from its minimum value of 1). From the equation of the generalized turbine characteristic (equation (270)) fixed by the design point calculations, follows the LPT corrected massflow. This amount is first divided by the maximum HPT corrected massflow and the corresponding HPT total pressure ratio is computed according to the dashed line in Figure 43 (equation (264)). If this pressure ratio is supercritical then the HPT total pressure ratio is found, if not, then equation (264) is solved with the regula-falsi method, since the wanted pressure ratio must then lie between 1 and the critical value (2.5). As soon as both the LPT total pressure ratio and the HPT total pressure ratio are known, the HPT corrected massflow is known as well as both total temperature ratios.

### 11.7.1 HP Compressor and Turbine Matching

Now the combination of the off-design compressor pressure ratio, the high pressure

turbine entry temperature and, if applicable, the amount of excess air are determined. Combining the previously presented equations for the HPT characteristic (272) and the high pressure spool power equation (266) and equation (268) yields:

$$(1+\lambda) \left( \frac{\dot{m}_t \sqrt{T_{4t}}}{P_{4t}} \right) \epsilon_{cc} \sqrt{\eta_m c_{pg} \left( 1 - \frac{1}{T_{4t}/T_{5t}} \right)} P_{2t} = \frac{\sqrt{\epsilon_c^{\frac{\gamma-1}{\gamma}} - 1}}{\epsilon_c} = f(\epsilon_c) \quad (285)$$

$$\dot{m}_2 \sqrt{1 + \frac{1}{a/f} \sqrt{c_p T_{2t}}}$$

This expression was derived by substituting the turbine entry temperature from the power equation into the equation for the generalized characteristic. After a call to the intake routine (which initially does not spill any excess air, apart possibly from a small amount of minimum spillage, as was mentioned in the discussion of the intake routine), the left-hand-side term is completely known for a chosen initial value of the LPT total pressure ratio. The bypass ratio can initially be put equal to the design point value. In case the engine parameters are chosen such that the total pressures of the hot and the cold flow at the point of mixing are equal (mixing condition, see paragraph 11.7.3), then the bypass ratio will vary and equation (285) should be re-evaluated after the analysis of the mixing process. However, as will be explained in paragraph 11.7.3, it was established that the mixing condition, both power equations and the massflow requirement can be transformed into one equation from which the off-design bypass ratio and the amount of excess air follow directly. Therefore, equation (285) does not feature as such in the analysis procedure.

The right-hand-side term, which is a function of the compressor pressure ratio only, is drawn in Figure 47. For a polytropic compressor efficiency of 95%, it has a maximum of about 0.2447 at 1.72. There are obviously two solutions to the equation, one below 1.72 and one in excess of 1.72. It appears that the small solutions are not very useful, since such low compressor pressure ratios result in very low values for the net thrust, leaving only a very marginal installed thrust. The large solutions are obtained as follows. Since the left-hand-side term is known it can be checked whether this value is below the maximum right-hand-side function value of 0.2447. If it is, a solution can be obtained. To this end, the compressor pressure ratio is increased with large steps, until the function  $f$  has become smaller than the left-hand-side term. At that point, the equation can be solved using the regula-falsi method, since the wanted compressor pressure ratio lies between the present value and 1.72. Once the compressor pressure ratio is determined, the turbine entry temperature can be readily obtained from the high pressure spool power equation (265). Then, the air-to-fuel ratio,  $a/f$ , can be computed and with this value the procedure is repeated until convergence of  $a/f$ . The necessity of this iteration can be argued, since the influence of the  $a/f$  ratio is not very large, whereas the computational effort can be considerable. It is therefore questionable if the cost in computational speed is balanced by the benefit of exact compliance with the equations.

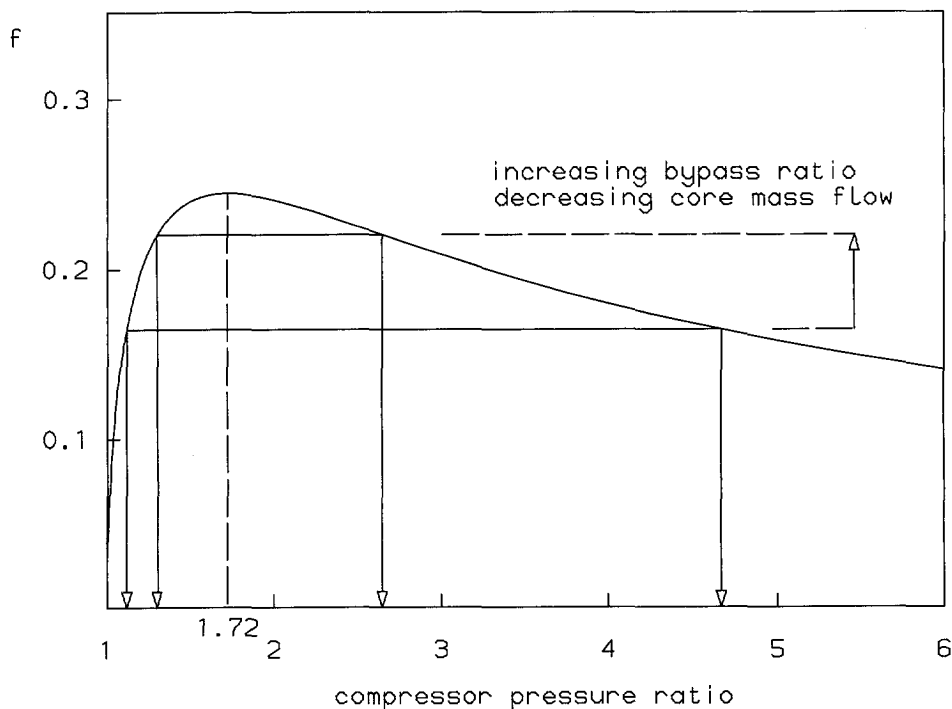
In case the left-hand-side term is above the maximum value of  $f(\epsilon_c)$ , no solution can be found. Since most of the variables in the left-hand-side term are fixed, there are in effect only two solutions to this problem. First of all, the bypass ratio could be decreased, until the left-hand-side term would be slightly below the maximum function value of the right-hand-side term. It was established that this alternative caused problems, either with the convergence of  $a/f$  or with the convergence of the bypass ratio iteration process, that imposes the mixing condition. Therefore, the bypass ratio is kept constant at this stage of the procedure and another, less obvious remedy is applied. The only variables that are not yet fixed in the left-hand-side term of equation (285) are the high pressure turbine total temperature ratio and the related corrected massflow. These variables increase or decrease with the low pressure turbine total pressure ratio, which is the initial variable for the procedure. This variable is -as explained previously- increased in a stepwise fashion, and the only way to change the left-hand-side term is to increase this value. This would also increase the left-hand-side term at first, and thus seems just the opposite thing to do. However, when the high pressure turbine operates at very low pressure ratios, the result of the procedure will be an extremely large value for the compressor pressure ratio (see Figure 47). Since this value -as indicated before- is limited by the smallest of the design value and the value which yields the maximum allowable compressor exit temperature, this causes a large amount of excess air, which has to be dumped. This decreases the core massflow which increases the left-hand-side term. (From Figure 47 it follows easily that this will decrease the compressor pressure ratio). Therefore, the cause of the large value of the left-hand-side term is the low value of the LPT total pressure ratio, and the right thing to do is to break off the calculations and proceed with a higher value of the LPT pressure ratio.

Once solutions are obtained for the compressor pressure ratio and the turbine entry temperature, these values are checked to establish whether their maximum values are violated. If so, the variable of concern is set to its maximum and the other variable is recalculated from the high pressure spool power equation. If this variable exceeds its maximum as well, it is also fixed at this value and the other variable is recalculated, leading to a value smaller than its maximum (an increase of the compressor pressure ratio leads to an increase of the turbine entry temperature, and vice versa). Since at this point, the values are no longer in accordance with the HPT generalized characteristic, the core mass flow is adapted: the difference between the presented mass flow and the demanded massflow is referred to as excess air which will be dumped. The process of dumping excess air will be explained in the next paragraph.

### 11.7.2     Spillage or Dumping of Excess Air

For a propulsive device that has to operate in such a broad regime of Mach numbers and altitudes as an engine destined to power a supersonic transport plane, it is virtually impossible to avoid dumping of excess intake air at all points in the operational envelope. Since the engine design is always a compromise between the amount of thrust generated and the amount of powerplant drag that has to be overcome at a cost of increased (specific) fuel consumption, it is very important to incorporate a detailed intake model in a multivariate optimization study. The details

---



**Figure 47** High pressure spool function

of this intake model were elaborated in paragraph 11.5.

In principle the excess air can be dumped in two ways: either it can be bled off (into the free air by means of a dump door) which will cause a loss of momentum and thus of net thrust, or it can be spilled by means of the shockwave system in the intake. This will cause additive or spillage drag, which implies a loss of installed thrust.

In case the excess air is spilled by the first external shockwave of the inlet (Figure 48), only a part of the full captured mass flow  $\dot{m}_0$  will flow through the engine:

$$\dot{m}_2 = \dot{m}_0 \left( 1 - \frac{A_s}{A_0} \right) \quad (286)$$

The generated net thrust in this case will be:

$$T_{\text{net}} = (\dot{m}_2 + \dot{m}_f) v_e - \dot{m}_2 v_0 + A_e (p_e - p_0) \quad (287)$$

with:

$$\dot{m}_f = \frac{\dot{m}_2}{(1 + \lambda) a / f} \quad (288)$$

The spilled air will exert a pressure force on the internal flow tube which results in an additive or spillage drag term equal to:

$$D_a = A_s (p_1 - p_0) \quad (289)$$

Using the definition of the net thrust according to equation (287) (resulting from the momentum analysis of the internal flow), the powerplant drag will be the sum of the additive drag and the external drag, the latter of which is the sum of the lip drag (wave drag) and the friction drag. These three drag components plus the net thrust together constitute the total momentum balance of the engine. One should be warned however, that just the net thrust and not the total or installed thrust features in the definition of the specific fuel consumption, since the inserted fuel mass flow was only used to generate net thrust and not to overcome the external drag forces. It would therefore be more appropriate to add these terms to the total airplane drag than to subtract them from the net thrust.

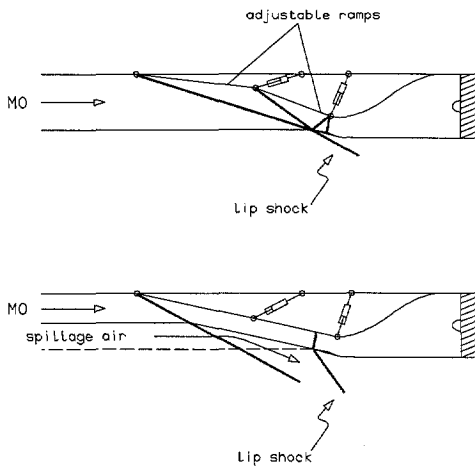


Figure 48 Supersonic intake spillage

An alternative would be, to bleed off the excess air upstream of the high pressure turbine by means of a dump door. In practice this will not involve a large door being deflected which would result in a large amount of drag, but the opening of a series of louvers that let out the excess air at a small angle with the free stream direction. (The least loss of momentum would of course result in case the flow is dumped at zero angle with the free stream direction. This is considered impossible, but the angle can be kept very small. Since the loss of momentum in axial direction is proportional to the cosine of the dump angle and this angle is considered small, the calculations will be based on a zero dump angle). The right spot to do this would be in front of the fan, because in that case no fuel would be added to heat air that is to be dumped anyway, while on the other hand the energy balance between the compressor and the turbine stays intact, which otherwise would have led to a loss of total pressure in the nozzle and thus (with the same expansion ratio) a loss of pressure thrust. The bleed air that is dumped through the dump door is assumed to exit at the local speed of sound. This assumption is justified, since usually the total pressure recovery is such, even after the flow has passed through several shockwaves, that if the bleed air is exposed to ambient pressure the flow will isentropically accelerate to at least Mach 1. Using this procedure the net thrust will be:

$$T_{\text{net}} = (\dot{m}_2 + \dot{m}_f)v_e + (\dot{m}_0 - \dot{m}_2)a_e - \dot{m}_2 v_0 - (\dot{m}_0 - \dot{m}_2)v_0 + A_e(p_e - p_0) \quad (290)$$

The momentum loss in this case is an implicit form of drag which is accounted for in the net thrust. In principle there is no spillage drag in this case, only the lip and friction drag are to be added.

Both alternatives of disposing of excess air lead to different values for the installed drag. Not only will the lip drag be different for both cases, since in the first case compression on the lip takes place by means of two shockwaves as opposed to only one in the second case (Figure 48) and not only will the intake efficiency differ slightly leading to a slightly different exhaust pressure, also, the additive drag in the first case does not equal the momentum loss in the second case. More specific even, it will be shown that the momentum loss in the second alternative is usually larger than the spillage drag of the first alternative. The momentum loss of case two as compared to case one is equal to:

$$\Delta I = (\dot{m}_0 - \dot{m}_2)(v_0 - a_e) = \dot{m}_0 v_0 \left( \frac{A_s}{A_0} \right) \left( 1 - \frac{a_e}{v_0} \right) \quad (291)$$

With:

$$\frac{a_e}{v_0} = \frac{\sqrt{\frac{2}{\gamma+1} + \frac{\gamma-1}{\gamma+1} M_0^2}}{M_0} \quad (292)$$

the following relation can be obtained:

$$\Delta I = p_0 A_s M_0 \gamma \left( M_0 - \sqrt{\frac{2}{\gamma+1} + \frac{\gamma-1}{\gamma+1} M_0^2} \right) \quad (293)$$

With the use of the well known relations for oblique shockwaves, the spillage drag of case one can be rewritten as follows:

$$\begin{aligned} D_a &= A_s (p_1 - p_0) = \frac{2 A_s p_0 \gamma}{\gamma+1} (M_0^2 \sin^2 \sigma - 1) = \\ &= \frac{A_s \gamma p_0 M_0^2 \tan \delta \tan \sigma}{1 + \tan \delta \tan \sigma} \end{aligned} \quad (294)$$

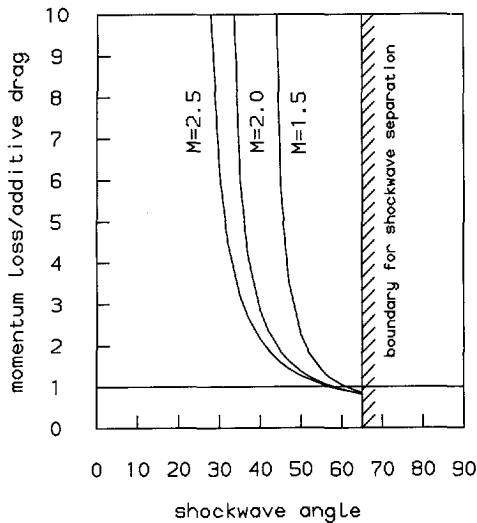
The ratio between the momentum loss and the spillage drag can then be written as:

$$\frac{\Delta I}{D_a} = \left( 1 - \frac{\sqrt{\frac{2}{\gamma+1} + \frac{\gamma-1}{\gamma+1} M_0^2}}{M_0} \right) \left( 1 + \frac{1}{\tan \delta \tan \sigma} \right) \quad (295)$$

Observing this relation, it can first be concluded that for a constant free stream Mach number the ratio between momentum loss and spillage drag will decrease with increasing flow deflection angle (and thus shockwave angle). The shockwave angle is limited to a certain maximum value with regard to shockwave separation. The corresponding maximum deflection angle decreases sharply with decreasing Mach number (for a Mach number of 2.2 the maximum attainable deflection angle is about 26 degrees whereas for a Mach number of 1.5 this value becomes about 12 degrees). This imposes a maximum on the amount of excess air that can be spilled by means of the intake shock system (hereafter this will be called maximum spillage).

In Figure 49 equation (295) is drawn as a function of the shockwave angle. It appears that only for very large shockwave angles dumping air can be more favorable than

spilling. Naturally, if more excess air needs to be spilled than is possible without allowing the first external shock to separate, the rest of the air must be dumped. As shown previously, most spillage air will occur at relatively low Mach numbers, where the maximum attainable deflection angle is low.



**Figure 49** Comparison of intake drag caused by dumping versus spillage of excess air

In the analysis above, the influence of the change in shockwave angle on the lip drag and the pressure thrust was not taken into account. The change in lip drag will not be very significant as can be easily deduced from a shockwave table, but the influence of the intake efficiency on the pressure thrust can be quite large. In situations with a large amount of excess air, for instance in the case of an engine designed for a high supersonic Mach number operated at a low supersonic Mach number, an expansion flow can occur at the intake lip when a large amount of the excess air is spilled. This leads to a drag reduction.

The above analysis applies to equal amounts of excess air, that is, a comparison is made between the spillage drag resulting from a certain amount of spilled air and the momentum loss resulting from the same amount of air when it is entirely dumped. Even if spilling the excess air will in most cases be more advantageous than dumping it, it can be worthwhile to dump part of the air. This is a direct consequence of the intake drag depending non-linearly on the amount of air spilled. In Figure 50 the amount of spillage drag and the amount of momentum loss (which is a form of intake drag) are sketched as a function of the percentage of the excess air which is spilled respectively bled off. It is clear that the amount of excess air can be divided in spillage air and bleed air in such a way that the resulting additive drag or intake drag is minimized.



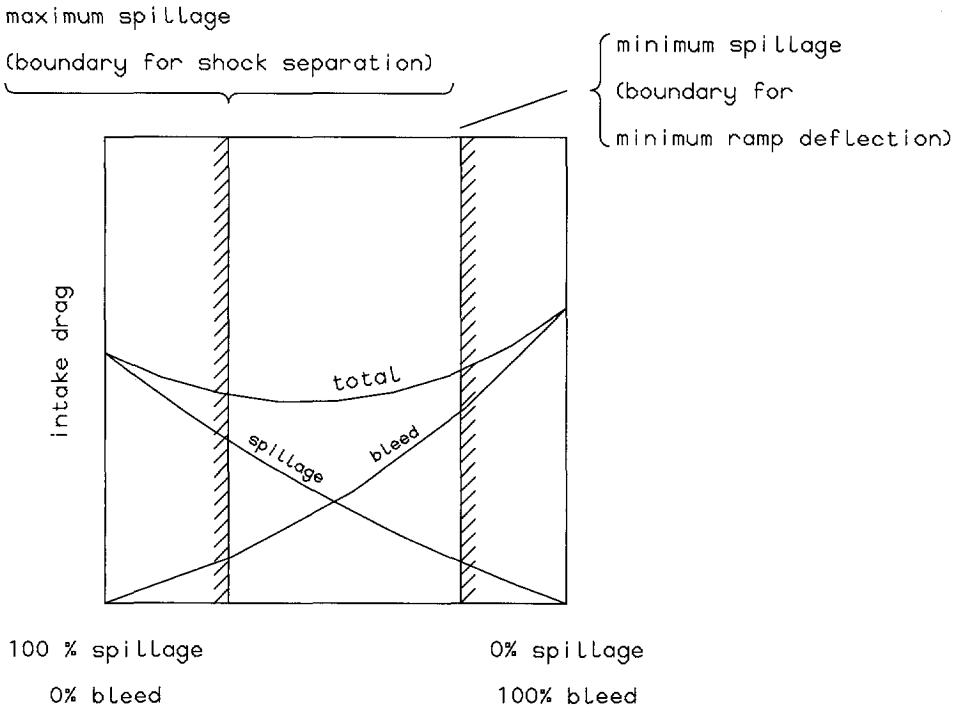


Figure 50 Optimization of amounts of spillage and bleed air

There are two bounds on the amount of spillage air that should be incorporated: firstly, the previously mentioned maximum attainable spillage, which is related to the angle at which shockwave separation occurs. Secondly, it is acknowledged that the flow will always be deflected to some extent, even if the first intake ramp would not be deployed. The air that encounters the upper edge of the intake will at all times experience some disturbance, and be deflected over an oblique shock. The angle of this shock is in the present model assumed to correspond with a deflection angle of 5 degrees. In case the corresponding shockwave does not hit the intake lip, spillage will occur (hereafter this kind of spillage will be called minimum spillage). It should be realized that this kind of spillage depends on the intake design. In some cases these bounds will be active and the true minimum in intake drag cannot be attained.

In the program, the condition for minimum intake drag is approximated as follows. After a complete off-design calculation has been performed, the total amount of spilled air will be divided into equal parts. The maximum allowable amount of spillage air will thus be constrained to a certain percentage of the original amount of spillage air and the off-design calculations are repeated. Remaining excess air is dumped. The best combination of spillage air and bleed air is chosen. It should be

realized that during this procedure, the sum of the spillage air and the bleed air does not stay constant. Varying the amount of spillage air will change the intake efficiency and therefore the total pressure at station 4. This total pressure features in the corrected massflow through the high pressure turbine and therefore affects the amount of excess air.

It was found that the picture sketched in Figure 50 only has theoretical validity. The situation in reality differs dramatically from this idealized model due to a number of influences, one of which is the earlier mentioned effect of the change in intake efficiency caused by adapting the shockwave angles. A second effect is related to the intake model, as explained in paragraph 11.5. It may occur, that the second shock cannot hit the intake lip because the Mach number behind the first shock has become too low. This might especially happen with intakes designed for high Mach numbers and therefore small shockwave angles. This would imply that spillage occurs over two shockwaves which results in a large amount of spillage while secondly the risk exists that the first and second shock will intersect in front of the intake causing Ferri buzz. In case the second shock cannot hit the intake lip, therefore, the intake model will eliminate the second shock. This will result in an abrupt change of intake efficiency.

[kN]

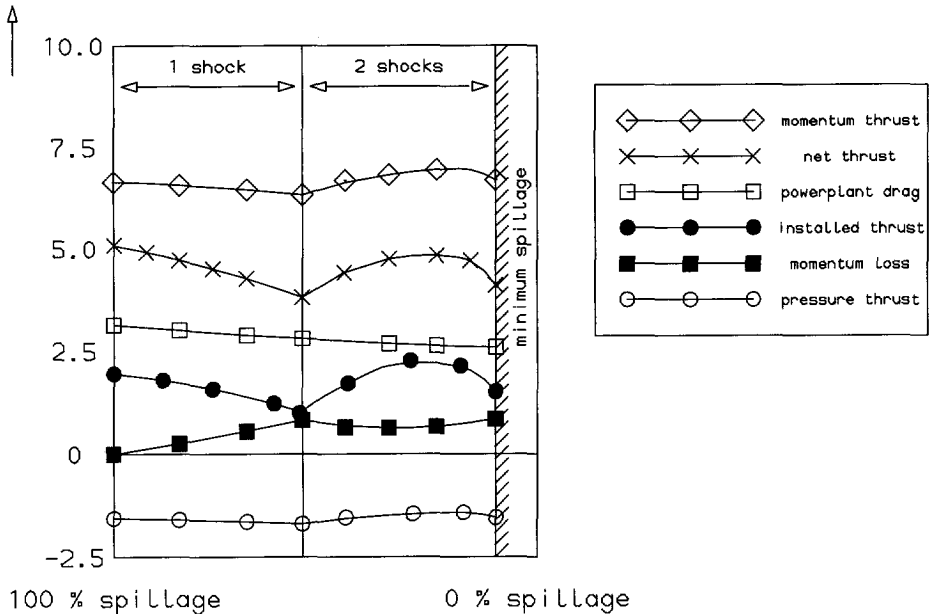


Figure 51 Thrust and powerplant drag for different amounts of spillage air

In Figure 51 the net thrust, the pressure and momentum contribution to the net thrust, the powerplant drag (which includes the spillage drag), the momentum loss due to the dumping of excess air and the installed thrust are plotted as a function of the amount of spillage air. With less spillage, the shockwave angle of the first shock decreases while the Mach number behind the first shock increases, until abruptly, a second shock is generated. This causes the break in the curves. As can be seen from the figure, the spillage drag continually decreases with decreasing amount of spillage air. The momentum loss due to dumping excess air will first increase from zero. It is obvious, that even before the second shock appears, the picture does not correspond to Figure 50. The decrease in spillage drag is far less than the increase in momentum loss. This is caused by the decrease of the intake efficiency, which in turn causes a decrease of the total pressure at station 4. Since the corrected massflow through the HPT is fixed, this requires a decrease of the intake air as well. This leads to an additional amount of dumped air (in other words, the sum of the spilled air and the dumped air does not remain constant). In the example of Figure 51 which refers to an intake designed for  $M=2.5$  at an off-design Mach number of 2.0, the maximum amount of spillage air was 5.2 kg/s. Decreasing this in the first step to 4.8 kg/s requires 0.54 kg/s to be dumped. The decrease in intake drag amounts 67.9 N whereas the increase in momentum loss amounts 122.6 N, resulting in a 54.7 N loss of installed thrust.

Furthermore, the change of intake efficiency and the related decrease of massflow through the engine costs 13.4 N of pressure thrust (which has a negative contribution in this example, due to overexpansion) and 33.2 N of momentum thrust. The net thrust includes the momentum loss of the dumped air. The total balance of decreasing the amount of spillage air by 8% is a reduction of the installed thrust with 101.3 N or 5% (the installed thrust decreases from 1984 N to 1847 N).

The optimal distribution of the excess air will therefore be full spillage and no dumping. However, in the right-hand part of the figure, as the second intake shock appears, a second optimum occurs. This optimum is a result of an optimum in the intake efficiency. The installed thrust at the optimum amounts 2201 N, or an increase with respect to the local optimum at full spillage of 13%.

To be able to obtain the absolute maximum in installed thrust, the following procedure is applied. Since the abrupt increase in the number of intake shocks only occurs with decreasing amount of spilled air, at first all excess air is dumped. In subsequent steps, the permitted percentage of spillage air is increased. This implies that the program runs through Figure 51 from the right to the left. Still, situations can occur in which the absolute optimum is obtained for 100% spillage or 100% bleed. In these cases the number of shockwaves does not change but the reduction in drag due to spillage is offset by the increased momentum loss or vice versa.

---

At subsonic free stream Mach numbers, the amount of spillage air cannot be controlled. The air that cannot be handled by the high pressure turbine will simply flow around the intake. Referring to Figure 52 the following expression can be derived for the additive drag in the subsonic case, by applying momentum conservation on the inner streamtube. To avoid confusion, the reader's attention is called to the fact that the pressure at infinity is denoted  $p_0$  whereas the pressure at the intake area  $A_0$  is denoted  $p_1$ .

$$\begin{aligned}
 D_a &= p_1 A_0 (1 + \gamma M_1^2) - p_0 A_1 \left( \frac{A_0}{A_1} + \gamma M_0^2 \right) \\
 &= p_0 A_1 \left( \frac{p_1 A_0}{p_0 A_1} (1 + \gamma M_1^2) - \left( \frac{A_0}{A_1} + \gamma M_0^2 \right) \right) \\
 &= p_0 A_1 \left( \frac{M_0}{M_1} \sqrt{\frac{T_1}{T_0}} (1 + \gamma M_1^2) - \left( \frac{A_0}{A_1} + \gamma M_0^2 \right) \right) \\
 &= p_0 A_1 \left( \frac{M_0}{M_1} \sqrt{\frac{1 + \frac{\gamma-1}{2} M_0^2}{1 + \frac{\gamma-1}{2} M_1^2}} (1 + \gamma M_1^2) - \left( \frac{A_0}{A_1} + \gamma M_0^2 \right) \right) \quad (296)
 \end{aligned}$$

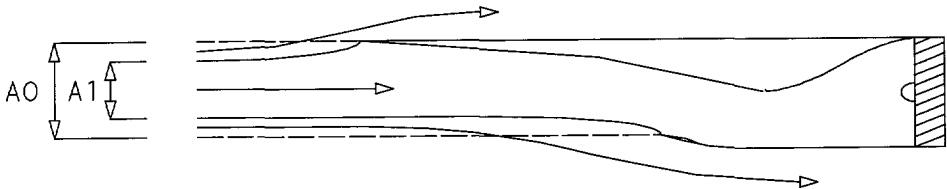


Figure 52 Subsonic spillage

The only unknown in equation (296) is the Mach number at the intake,  $M_1$  which can be determined from the amount of spillage air with the use of the mass conservation requirement:

$$\frac{A_0}{A_1} = \frac{M_0}{M_1} \left( \frac{1 + \frac{\gamma-1}{2} M_1^2}{1 + \frac{\gamma-1}{2} M_0^2} \right)^{\frac{\gamma+1}{2(\gamma-1)}} \quad (297)$$

It can be easily verified that for  $A_0/A_1$  equals 1,  $M_1$  equals  $M_0$  and the additive drag equals zero. For  $A_1$  equals one half of the intake area and a free stream Mach number of 0.5, it follows that  $M_1$  equals 0.2225 and the additive drag amounts 5% of  $p_0 A_0$ . For an area ratio between zero and one,  $M_1$  lies between zero and the free stream Mach number. Thus, for a given amount of excess air,  $M_1$  can be determined by a regula-falsi procedure, after which the amount of additive drag follows easily.

### 11.7.3      Mixing the Hot and the Cold Flow

Once the compressor pressure ratio and the turbine entry temperature have been determined, all variables upstream of the low pressure turbine are known. The total pressure behind the LPT follows immediately from the initially chosen total pressure ratio. The temperature and pressure behind the fan follow from the low pressure spool power equation. As indicated previously, the hot and the cold flow cannot be mixed just like that, without a resulting loss of installed thrust. Either the engine is forced to operate in such a way that the total pressures of both flows are equal at the point of mixing, or a mixer is added to the model in which the flows are mixed at different total pressures, but equal static pressures (Kutta condition). In the second case, a mixer must be incorporated in the model. In the first case, the power equation is no longer satisfied and the work required by the fan must be adapted to the work delivered by the low pressure turbine. The only way this can be done is to redirect air from the bypass stream to the core or vice versa (the engine will adapt itself in this way to the work delivered by the LPT), by an appropriate change of the fan pressure ratio. Since in reality the Kutta condition applies on the streamline separating the hot and the cold flow prior to mixing, the simultaneous enforcement of equal total pressures will require variable mixer geometry; alternatively, small differences between the total pressures will occur (see Appendix B, chapter 15).

Two different versions of the off-design analysis program were developed, that only differ in the manner in which the mixing condition is applied. The two alternatives will be discussed separately hereafter. The constant area mixer model eventually was abandoned, since it was not flexible enough in combination with the present model.

#### Constant Area Mixer

In the constant area mixer, the two flows are mixed at different total pressures, but

under the condition that the static pressures at the end of the splitter plate that shields the hot flow from the cold flow must be the same (Kutta condition). The model carries out a momentum analysis in order to determine the exit total pressure of the mixed flow. It is assumed that the total area of the incoming flow equals the total area of the outgoing flow (constant area mixer). In this way, any pressure forces on the walls of the mixer are neglected. The following analysis is largely based on [ref.5].

Using the familiar expression for the total pressure:

$$P_t = P \left( 1 + \frac{\gamma-1}{2} M^2 \right)^{\frac{\gamma}{\gamma-1}} \quad (298)$$

and applying this on the total pressures at stations 3f and 6 while enforcing the Kutta condition, the following expression can be obtained:

$$\frac{P_{3f,t}}{P_{6,t}} = \frac{\left( 1 + \frac{\gamma-1}{2} M_{3f}^2 \right)^{\frac{\gamma}{\gamma-1}}}{\left( 1 + \frac{\gamma_g-1}{2} M_6^2 \right)^{\frac{\gamma_g}{\gamma_g-1}}} \quad (299)$$

Once both total pressures are known, the Mach number of the cold flow at the entry of the mixer can be written as a function of the hot flow mixer entry Mach number:

$$M_{3f} = \sqrt{\frac{2}{\gamma-1} \left[ \frac{P_{3f,t}}{P_{6,t}} \left( 1 + \frac{\gamma_g-1}{2} M_6^2 \right)^{\frac{\gamma_g}{\gamma_g-1}} \right]^{\frac{\gamma-1}{\gamma}} - 1}} \quad (300)$$

It is clear, that in order to avoid the Mach numbers to become negative (reverse flow) or greater than one (choked mixer intake) there are bounds imposed on the ratio of the total pressures behind the fan and the LPT. The lower limit below which value no

usable solution can be obtained is:

$$\left( \frac{P_{3f_t}}{P_{6_t}} \right)_{\min} = \frac{1}{\left( \frac{\gamma_g + 1}{2} \right)^{\frac{\gamma_g}{\gamma_g - 1}}} \approx 0.54 \quad (301)$$

An absolute upper limit is:

$$\left( \frac{P_{3f_t}}{P_{6_t}} \right)_{\max} = \left( \frac{\gamma + 1}{2} \right)^{\frac{\gamma}{\gamma - 1}} \approx 1.89 \quad (302)$$

Initially, as the starting value of the low pressure turbine pressure ratio is still low, the ratio of the total pressure behind the fan and behind the LPT will be smaller than the minimum value. In that case, the calculations are stopped and started all over again with a new LPT pressure ratio (increased with a small step). As long as the LPT is not choked, an increase of the LPT total pressure ratio will cause an increase in the HPT total pressure and total temperature ratio. This in turn will cause a decrease of the compressor pressure ratio (according to equation (285)), or in case this value exceeds its maximum, a decrease of the turbine entry temperature. Therefore, either the total temperature or the total pressure at station 4 is decreased. Since the HPT total temperature ratio and total pressure ratio were increased, this implies a decrease of the total temperature at station 5. It is recalled now, that the LPT work is proportional to:

$$T_{5_t} \left( 1 - \left( \frac{P_{6_t}}{P_{5_t}} \right)^{\frac{(\gamma_g - 1)\eta_t}{\gamma_g}} \right)$$

Although the total temperature at station 5 decreases, the effect is smaller than the increase of the term between brackets. Therefore, the LPT work will increase, leading in turn to an increase of the total pressure behind the fan. Thus, the ratio of the total pressures behind the fan and the LPT is increased.

Once the value of the LPT pressure ratio has become such, that the maximum limit of the total pressure ratio becomes violated, the calculation is finished.

To be able to calculate both Mach numbers, a second equation is needed. This equation is provided in the shape of the mass conservation requirement:

$$\frac{\dot{m}\sqrt{T_t}}{pA} = M \sqrt{1 + \frac{\gamma-1}{2} M^2} \sqrt{\gamma/R} \quad (303)$$

Applying this relation to both the hot and the cold flow entering the mixer, the following expression can be obtained:

$$\frac{A_{3f}}{A_6} = \frac{\lambda}{(1 + \frac{1}{a/f})} \frac{M_6}{M_{3f}} \sqrt{\frac{\gamma_g T_{3ft} (1 + \frac{\gamma_g - 1}{2} M_6^2)}{\gamma T_{6t} (1 + \frac{\gamma - 1}{2} M_{3f}^2)}} \quad (304)$$

In the design point analysis, the total pressures behind the fan and the LPT are put equal, while the mixer entry Mach number of the hot flow is put at 0.5 (this yields for the mixer entry Mach number of the cold flow: 0.54). The ratio of the mixer intake areas can then be calculated. This value will be fixed throughout the off-design calculations.

In the off-design analysis, with the known ratios of the areas and the total pressures, the combination of the hot and cold mixer entry Mach numbers is determined such that both equations are satisfied. If no such solution is possible within the interval  $<0,1>$ , the procedure is cut off and a new off-design analysis is started with a slightly increased LPT total pressure ratio.

Applying the momentum conservation law to the hot and the cold flow entering the constant area mixer at their respective Mach numbers, and the mixed flow leaving it at an as yet unknown Mach number, the following relations can be obtained:

$$I_{3f} + I_6 = I_7 \quad (305)$$

with:

$$I = pA(1 + \gamma M^2) \quad (306)$$



This relation can be rewritten as:

$$I = \dot{m} \sqrt{\frac{RT_t}{\gamma \varphi(M, \gamma)}} \quad (307)$$

in which:

$$\varphi(M, \gamma) = \frac{M^2}{(1 + \gamma M^2)^2} \left( 1 + \frac{\gamma - 1}{2} M^2 \right) \quad (308)$$

Substituting this expression in the momentum equation for both the hot flow and the cold flow, the following expression for the mixed flow function can be obtained:

$$\varphi_7 = \frac{\left( \lambda + 1 + \frac{1}{a/f} \right)^2 \frac{T_{7t}}{\gamma_7}}{\left( \lambda \sqrt{\frac{T_{3ft}}{\gamma \varphi_3}} + \left( 1 + \frac{1}{a/f} \right) \sqrt{\frac{T_{6t}}{\gamma_g \varphi_6}} \right)^2} \quad (309)$$

with:

$$\varphi_7 < \frac{1}{2(\gamma_7 + 1)} \quad (310)$$

The specific heat ratio and the total temperature of the mixed flow follow from the mixing relations:

$$c_{p_7} = \frac{c_{p_g} + \frac{\lambda}{\left( 1 + \frac{1}{a/f} \right)} c_p}{1 + \frac{\lambda}{\left( 1 + \frac{1}{a/f} \right)}} \quad (311)$$

in which it is assumed that the gas constant,  $R$ , does not change.

$$T_{7t} = \frac{T_{3ft} + T_{6t} \frac{c_{pg}}{c_p} \frac{(1 + \frac{1}{a/f})}{\lambda}}{\frac{c_{pg}}{c_p} \left( 1 + \frac{(1 + \frac{1}{a/f})}{\lambda} \right)} \quad (312)$$

Once  $\phi_7$  has been calculated, the Mach number of the mixed flow immediately behind the mixer can be determined:

$$M_7 = \frac{\sqrt{2\phi_7}}{\sqrt{(1 - 2\gamma_7\phi_7) - \sqrt{1 - 2(\gamma_7 + 1)\phi_7}}} \quad (313)$$

Finally, the total pressure behind the mixer can be obtained:

$$P_{7t} = \epsilon_{\text{mix}} P_{6t} \left( 1 + \frac{\lambda}{(1 + \frac{1}{a/f})} \right) \sqrt{\frac{\gamma_g T_{7t}}{\gamma_7 T_{6t}}} \left( \frac{1}{1 + \frac{A_{3f}}{A_6}} \right) \frac{M_6}{M_7} \frac{\left( 1 + \frac{\gamma_7 - 1}{2} M_7^2 \right)^{\frac{\gamma_7 + 1}{2(\gamma_7 - 1)}}}{\left( 1 + \frac{\gamma_g - 1}{2} M_6^2 \right)^{\frac{\gamma_g + 1}{2(\gamma_g - 1)}}} \quad (314)$$

In this expression,  $\epsilon_{\text{mix}}$  is the total pressure ratio over the mixer, which will be put equal to that of the combustion chamber, that is, 0.98.

It was found, that the usable margin for the above procedure was extremely small. The ratio of the total pressures behind the fan and the low pressure turbine lies between the minimum and maximum allowable value for just a very small range of LPT pressure ratios. In some cases, this even leads to values for  $\phi_7$  that exceed the upper limit, thus corresponding to supersonic mixer exit Mach numbers.

Because of this, it can happen that a certain engine will provide no thrust at certain points of its operational envelope while in other points a small margin does exist. To avoid this problem it was decided not to use the model with the constant area mixer any further in this application but instead to resort to the other alternative to solve the mixing problem.

### Adjusting the bypass ratio

The other alternative to comply with the mixing condition is to realize that the engine

will adapt itself to a given situation of pressure and temperature ratios over the different components by redistributing the presented airflow over the bypass duct and the core in such a way as to satisfy the power equations and massflow requirements of the turbines. This, then, implies that if the engine parameters are set such that the total pressures of the cold and hot flow become equal at the point of mixing, the bypass ratio will change. In principle, this will require a re-evaluation of equation (285), leading in turn to a change in the compressor pressure ratio and consequently an iterative procedure is needed.

In the process of developing the program, however, it was established that a single formula can be derived to determine the off-design working point of the engine, since -as indicated before- the total pressure and temperature ratios of all components are known once an initial value for the LPT total pressure ratio has been chosen. For the works of the fan and the low pressure turbine per unit of core massflow, the following relations hold:

$$\frac{W_{fan}}{\dot{m}_c} = \lambda c_p T_{2t} \left( \left( \frac{P_{3f_t}}{P_{2t}} \right)^{\frac{\gamma-1}{\gamma \eta_c}} - 1 \right) \quad (315)$$

$$\frac{W_{LPT}}{\dot{m}_c} = \eta_m \left( 1 + \frac{1}{a/f} \right) c_{pg} T_{5t} \left( 1 - \left( \frac{P_{6t}}{P_{5t}} \right)^{\frac{(\gamma_s-1)\eta_s}{\gamma_s}} \right) \quad (316)$$

Putting the total pressure behind the fan equal to that behind the LPT, to comply with the mixing condition, the following relation can be derived:

$$P_{3f_t} = P_{6t} = P_{2t} \epsilon_c \left( \frac{P_{4t}}{P_{3t}} \right) \left( \frac{P_{5t}}{P_{4t}} \right) \left( \frac{P_{6t}}{P_{5t}} \right) \quad (317)$$

Substituting this in the relation for the fan work yields:

$$\frac{W_{fan}}{\dot{m}_c} = \lambda c_p T_{2t} \left( \epsilon_c^{\frac{\gamma-1}{\gamma \eta_c}} \left[ \left( \frac{P_{4t}}{P_{3t}} \right) \left( \frac{P_{5t}}{P_{4t}} \right) \left( \frac{P_{6t}}{P_{5t}} \right) \right]^{\frac{\gamma-1}{\gamma \eta_c}} - 1 \right) \quad (318)$$

Rewriting the high pressure spool power equation, the following relation for the total temperature at station 5 can be obtained:

$$T_{5_t} = \frac{c_p T_{2_t} \left( \epsilon_c^{\frac{\gamma-1}{\gamma\eta_c}} - 1 \right)}{\eta_m c_{pg} \left( 1 + \frac{1}{a/f} \right) \left( \frac{T_{4_t}}{T_{5_t}} - 1 \right)} \quad (319)$$

Substituting this expression in the relation for the LPT work yields:

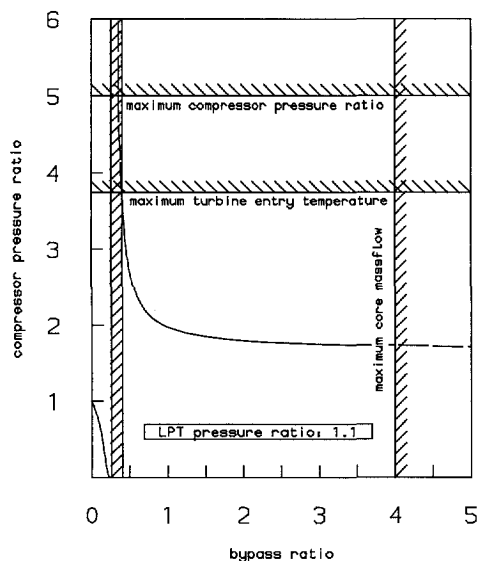
$$\frac{W_{LPT}}{\dot{m}_c} = \frac{c_p T_{2_t} \left( \epsilon_c^{\frac{\gamma-1}{\gamma\eta_c}} - 1 \right) \left( 1 - \left( \frac{P_{6_t}}{P_{5_t}} \right)^{\frac{(\gamma_s-1)\eta_t}{\gamma_s}} \right)}{\left( \frac{T_{4_t}}{T_{5_t}} - 1 \right)} \quad (320)$$

Now putting the fan work (318) and the LPT work (320) equal to each other, the following expression is derived:

$$\epsilon_c = \left[ \frac{\lambda - \frac{1 - \left( \frac{P_{6_t}}{P_{5_t}} \right)^{\frac{(\gamma_s-1)\eta_t}{\gamma_s}}}{\frac{T_{4_t}}{T_{5_t}} - 1}}{\left[ \left( \frac{P_{4_t}}{P_{3_t}} \right) \left( \frac{P_{5_t}}{P_{4_t}} \right) \left( \frac{P_{6_t}}{P_{5_t}} \right) \right]^{\frac{\gamma-1}{\gamma\eta_c}} \cdot \lambda - \frac{1 - \left( \frac{P_{6_t}}{P_{5_t}} \right)^{\frac{(\gamma_s-1)\eta_t}{\gamma_s}}}{\frac{T_{4_t}}{T_{5_t}} - 1}} \right]^{\frac{\gamma\eta_c}{\gamma-1}} \quad (321)$$

This equation gives the relation between the compressor pressure ratio and the bypass ratio such that both the high pressure spool power equation and the mixing condition are satisfied. In the equation, all temperatures and pressures are written as ratios that

are known as soon as an initial value for the low pressure turbine total pressure ratio is chosen (the ratio of the total pressures at station 4 and 3 is the total pressure drop over the combustion chamber, which has been assumed constant). In Figure 53 and Figure 54 an example of the equation is shown for two different values of the LPT total pressure ratio. These figures should not be interpreted as the off-design variations of the compressor pressure ratio and the bypass ratio, but as a curve on which the solution must be located for the off-design point under consideration. Hereafter it will be shown that only one single point on this curve will be practical.



**Figure 53** Mixing condition:  
relation between compressor pressure ratio and bypass ratio

As can be easily seen from the formula, the function has a singularity which has been indicated in the figures by a thin dashed line. To the left of this singularity, the values for the compressor pressure ratio are lower than one and therefore of no practical use. To the right, the value of the compressor pressure ratio decreases very rapidly from

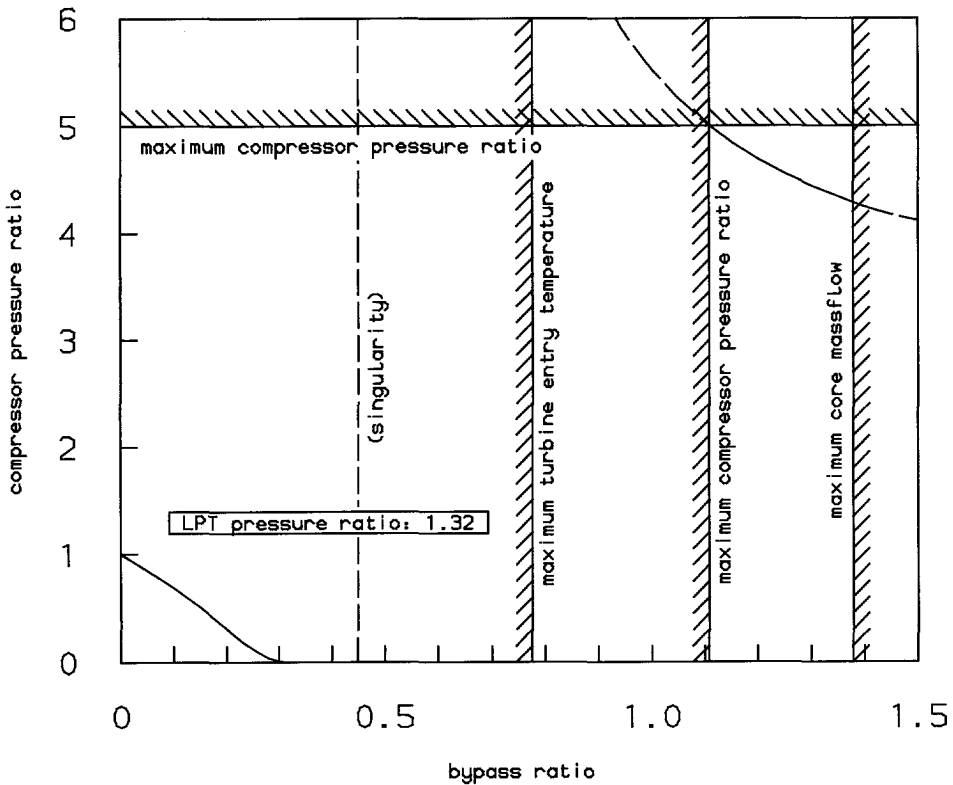
infinity and approaches the asymptotical value  $\epsilon_c \rightarrow \frac{P_{3t}}{P_{6t}}$ , which implies that for a

bypass ratio of infinity, the fan does not work. Three constraints are imposed on the compressor pressure ratio. First of all, it cannot exceed its design value. Secondly, the temperature at station 3 may not exceed its limit (in this case 850 K). A third constraint on the compressor pressure ratio is imposed in order to limit the turbine entry temperature (maximum combustion chamber temperature). Using the high

pressure spool power equation, the following relation for the maximum compressor pressure ratio with respect to the TET can be deduced:

$$\varepsilon_{c_{\max}} = \left( \frac{T_{4_{t_{\max}}} c_{p_s} \left(1 + \frac{1}{a/f}\right) \left(1 - \frac{T_{5_t}}{T_{4_t}}\right) \eta_m}{c_p T_{2_t}} + 1 \right)^{\frac{\eta_c}{\gamma-1}} \quad (322)$$

This constraint depends on the fuel massflow, and will therefore change during the iterative procedure to solve  $a/f$ .



**Figure 54** Mixing condition:  
relation between compressor pressure ratio and bypass ratio

Applying the outlined procedure throughout the aircraft's flight envelope -as explained without any knowledge about the compressor and fan characteristics- could pose severe demands on their flexibility. Once an engine design has been fixed and the off-design range of the pressure ratio and corrected massflow combinations is known, it must be established whether a compressor or fan design can be realized with an acceptable performance map. If the required flexibility is too high, the use of variable compressor vanes might be necessary, especially in case the most extreme values of the pressure ratio require a different number of compressor stages. Even with the use of a variable fan or compressor, the achievable level of flexibility is limited and bounds may have to be imposed on the range of the pressure ratio, especially of the fan. For the present method of analysis, these can be translated into bounds on the bypass ratio, which in turn might influence take-off performance. Such considerations, however, are outside the scope of conceptual design (see also [ref.1] for a treatise on compressor/fan flexibility demands by conflicting off-design performance requirements of a conceptual supersonic cruise aircraft design). However, the influence of fan pressure ratio limits on take-off performance is investigated for the final engine design in Appendix C of chapter 15.

Previously, a relation was deduced between the bypass ratio, the compressor pressure ratio and the presented massflow that satisfied the high pressure spool power equation and the high pressure turbine massflow requirement (equation (285)). Actually, there are now two equations available with two unknowns for a specified amount of presented massflow. The two equations are combined as follows to yield the massflow through the engine:

$$\dot{m}_2 = \Phi(1 + \lambda) \sqrt{\left( \frac{\lambda - X}{\Psi\lambda - X} \right)^{\frac{2\gamma\eta_c}{\gamma-1}} - 1} \quad (323)$$

with:

$$\Phi = \frac{\left( \frac{\dot{m}_t \sqrt{T_{4t}}}{P_{4t}} \right) \epsilon_{cc} \sqrt{\eta_m c_{pg} \left( 1 - \frac{1}{T_{4t}/T_{5t}} \right) P_{0t}}}{\sqrt{1 + \frac{1}{a/f} \sqrt{c_p T_{2t}}}} \eta_t \quad (324)$$

$$X = \frac{\left( 1 - \frac{P_{6t} \frac{(\gamma_s - 1)\eta_t}{\gamma_s}}{P_{5t}} \right)}{\left( \frac{T_{4t}}{T_{5t}} - 1 \right)} \quad (325)$$

$$\Psi = \left[ \left( \frac{P_{4t}}{P_{3t}} \right) \left( \frac{P_{5t}}{P_{4t}} \right) \left( \frac{P_{6t}}{P_{5t}} \right) \right]^{\frac{\gamma-1}{\gamma\eta_c}} \quad (326)$$

Equation (323) which for a given LPT total pressure ratio depends on the bypass ratio only, combines the high pressure spool power equation, the low pressure spool power equation, the high pressure turbine massflow requirement and the mixing condition.

In Figure 55 and Figure 56 the massflow is shown as a function of the bypass ratio for two values of the LPT pressure ratio. The constraints on the compressor pressure ratio can easily be translated to constraints on the bypass ratio by:

$$\lambda = \frac{\left( \epsilon_c^{\frac{\gamma-1}{\gamma\eta_c}} - 1 \right) \left( 1 - \frac{P_{6t} \frac{(\gamma_s - 1)\eta_t}{\gamma_s}}{P_{5t}} \right)}{\left( \frac{T_{4t}}{T_{5t}} - 1 \right) \left[ \left( \frac{P_{4t}}{P_{3t}} \right) \left( \frac{P_{5t}}{P_{4t}} \right) \left( \frac{P_{6t}}{P_{5t}} \right) \right]^{\frac{\gamma-1}{\gamma\eta_c}} \epsilon_c^{\frac{\gamma-1}{\gamma\eta_c}} - 1} \quad (327)$$

Each value for the bypass ratio corresponds via this formula to a compressor pressure ratio according to Figure 53 and Figure 54, and each value of the compressor pressure ratio corresponds to a certain function value in Figure 47. Once the bypass ratio is determined, the massflow follows. It is obvious from this figure, that an upper bound on the bypass ratio exists as well. This follows from the fact that the massflow through the engine can never be greater than the presented massflow, that is, the



captured massflow minus a possible amount of minimum spillage, related to the minimal occurring flow deflection in the intake at small ramp angles. The absolute upper limit of the bypass ratio follows from equation (285) for the maximum value of  $f(\epsilon_0)$  and  $\dot{m}_0$ . However, since the bypass ratio and the compressor pressure ratio are related by formula (321) as well, the maximum available massflow may well be exceeded for a lower bypass ratio.

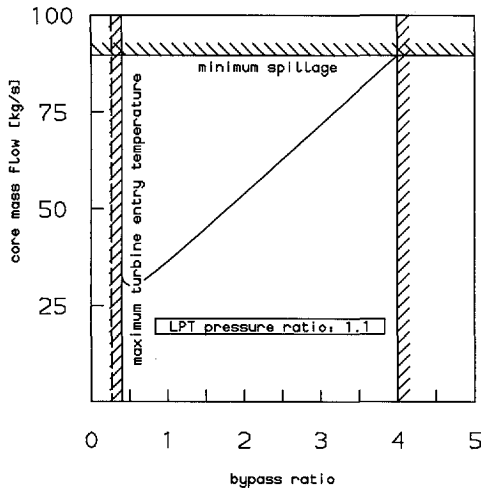


Figure 55 Determination of off-design working point

Between the most critical lower and upper limit, the bypass ratio can be varied to yield the required massflow and indirectly the compressor pressure ratio and the turbine entry temperature. From the figures it can be deduced, that with increasing low pressure total pressure ratios, the margin between the upper and lower limits on the bypass ratio decreases. Once the lower limit has exceeded the upper limit, the procedure is terminated. In theory it would be possible to find a solution where only minimum spillage or no spillage at all occurs. From the figures it follows however, that the corresponding bypass ratios are extremely large and the compressor pressure ratios extremely low. It was established, that the influence of the compressor pressure ratio on the engine performance by far exceeds the influence of the amount of excess air. In other words, reducing the amount of intake drag by reducing the compressor pressure ratio does not pay off. For a given value of the LPT pressure ratio, the maximum installed thrust is obtained for the lower limit of the bypass ratio.

After the bypass ratio, the compressor pressure ratio and the turbine entry temperature have been determined, an iterative procedure is started to find the air-to-fuel ratio,  $a/f$ . This procedure is based on the regula-falsi method, and during each step the method explained above to determine the bypass ratio, the compressor pressure ratio and the TET is repeated.

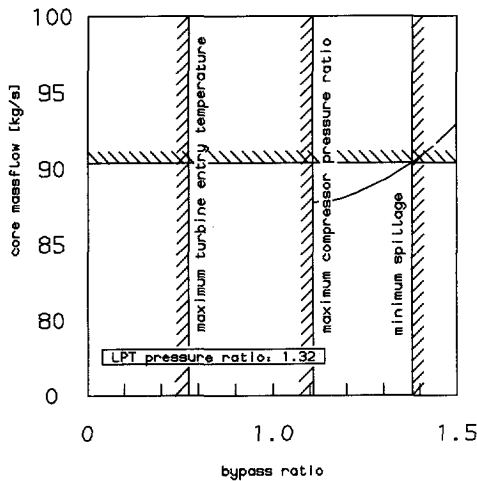


Figure 56 Determination of off-design working point

One small problem remains to be taken care of. According to equation (324), the coefficient  $\Phi$  depends on the intake efficiency. After the excess air as determined by equation (323) is spilled, the intake efficiency has changed, leading in turn to a change of  $\Phi$ . The only function of this coefficient is to satisfy the high pressure turbine massflow requirement. The change of the intake efficiency will not influence the calculated values for the bypass ratio and compressor pressure ratio but only the amount of excess air which in turn will affect the intake efficiency. This requires a small iterative procedure as well, also based on the regula-falsi method. During the iteration only pressure levels through the engine are affected, but all ratios remain constant. After convergence, all total pressures and temperatures upstream of the mixer are known.

In some cases, no solution can be found because the intake efficiency changes abruptly during the iteration. This is a result of the elimination of a shock in case its predecessor has a strong solution. It should be realized that this is no numerical problem; no solution exists because of a physical discontinuity. The occurrence of this problem is recognized in the program as soon as the iteration "jumps" over the solution (during the regula-falsi procedure the solution should be approached from one side). Should this problem occur, the solution will be chosen for which the intake spills too little and the dump door will take care of the rest of the excess air (dumping air will not influence the intake efficiency).

Of the mixing procedure of page 241, only the formulas that determine the total temperature and the specific heat ratio after mixing are retained. After this has been done, the total corrected massflow that enters the jetpipe (including fuel massflow and, optionally, the afterburner fuel massflow) is known. The throat area required to

ensure supersonic exit speeds can be found by applying the mass conservation requirement which, after some rewriting yields:

$$A_t = \frac{\frac{\dot{m}_2}{1+\lambda} \left( \lambda + 1 + \frac{1}{a/f} \right) \sqrt{\frac{T_{7t} R}{\gamma_7}} \left( 1 + \frac{\gamma_7 - 1}{2} M_t^2 \right)^{\frac{(2-\eta_{pol})\gamma_7 + \eta_{pol}}{2(\gamma_7-1)\eta_{pol}}}}{P_{7t} M_t} \quad (328)$$

in which the polytropic nozzle efficiency is used. The Mach number in the throat,  $M_t$ , depends on this nozzle efficiency. For a value of 95%, it is equal to 0.9667.

The same formula can be used for the conditions at the exit, only in this case the area is known (from the design point calculations) and the Mach number must be determined. This equation cannot be solved analytically. Instead, the Newton-Raphson iterative procedure is used.

Once the exit flow Mach number has been determined, the static pressure at the exhaust can be calculated from:

$$P_e = \frac{P_{7t}}{\left( 1 + \frac{\gamma_7 - 1}{2} M_e^2 \right)^{\frac{\gamma_7}{\eta_{pol}(\gamma_7-1)}}} \quad (329)$$

Now that all conditions throughout the engine are known, the net thrust can be calculated. The loss of momentum caused by dumped excess air is incorporated in the net thrust.

The net thrust thus becomes:

$$T_{net} = \dot{m}_e v_e - \dot{m}_2 v_0 - \frac{\dot{m}_{bleed} v_0}{M_0} \left( M_0 - \sqrt{\frac{2}{\gamma+1} + \frac{\gamma-1}{\gamma+1} M_0^2} \right) + A_e (P_e - P_0) \quad (330)$$

It is emphasized that  $\dot{m}_2$  is the massflow that really enters the engine and is subsequently distributed over the core and the bypass duct. The air which has entered the intake, but is subsequently dumped through the dump door is denoted by  $\dot{m}_{bleed}$ . As explained previously, this air will exit at the local sonic speed. The spillage air, which has not entered the intake at all, but has bypassed the intake altogether does not feature in the definition of the thrust. The associated spillage drag will be added separately, together with the lip drag (wave drag) and the friction drag.

The massflow at the exit equals the sum of the bypass air and the core air with added fuel:

$$\dot{m}_e = \frac{\dot{m}_2}{1+\lambda} \left( \lambda + 1 + \frac{1}{a/f} \right) \left( 1 + \frac{1}{a/f_{ab}} \right) \quad (331)$$

In case no afterburner is selected, the last term, of course, will be one.

From the air-to-fuel ratio, the fuel massflow that has been added in the combustion chamber and possibly in the afterburner, can be computed. The specific impulse  $I_{sp}$ , which is the reciprocal value of the specific fuel consumption (SFC) can then be calculated:

$$I_{sp} = \frac{T_{net}}{\dot{m}_f g} \quad (332)$$

As will be explained in chapter 12, it was deemed impossible to comply with the take-off noise requirements while simultaneously providing sufficient take-off thrust with an engine equipped with a fixed nozzle exit area, designed for ideal expansion at supersonic flight in the stratosphere. Therefore, it was decided that a variable exit area is needed in order to realize ideal expansion throughout the operational envelope.

## 11.8 References

1. Albers, M.; Hertel, J. and Schaber, R.; *Antriebskonzepte für Zukünftige Überschallverkehrsflugzeuge*; Jahrbuch der Deutschen Gesellschaft für Luft- und Raumfahrt e.V. (DGLR), 93-03-044, Bonn, 1993
2. Bos, A.H.W. and Oving, B.A.; *Conceptual Design of a Hypersonic Cruising Transport Aircraft* (in Dutch); Masters thesis, Faculty of Aerospace Engineering, Delft University of Technology, December 1991
3. Cordes, G.; *Strömungstechnik der Gasbeaufschlagten Axial Turbine unter besonderer Berücksichtigung der Strahltriebwerksturbine*; Springer Verlag, Berlin, 1963
4. Houtman, C.J.; *Vliegtuiggasturbines*; collegedictaat i54, Technische Universiteit Delft, faculteit der Werktuigbouwkunde en Maritieme Techniek, 1987
5. Mattingly, J.D.; Heiser, W.H. and Daley, D.H.; *Aircraft Engine Design*; AIAA Education series, New York 1987
6. Oswatitsch, K.; *Pressure Recovery for Missiles with Reaction Propulsion at High Supersonic Speeds (the Efficiency of Shock Diffusers)*; NACA TM-1140, 1947
7. Rettie, I.H. and Lewis, W.G.E.; *The Design and Development of an Air Intake for a Supersonic Transport Aircraft*; AIAA paper 67-752

8. Saravanamuttoo, H.I.H.; *A Rapid Method for the Matching of Two-Spool Turbojets*; Canadian Aeronautics and Space Journal, October 1970
  9. Saravanamuttoo, H.I.H.; *A Rapid Matching Procedure for Twin-Spool Turbofans*; Canadian Aeronautics and Space Journal, October 1972
  10. van der Velden, A.J.M.; *Aerodynamic Design and Synthesis of the Oblique Flying Wing Supersonic Transport*; Ph.D. dissertation, Stanford University, SUDAAR 621, June 1992
  11. Wittenberg, H.; *Prediction of Off-Design Performance of Turbojet and Turbofan Engines*; AGARD CP 242, May 1978
-



# 12. The Environmental Impact of a Supersonic Transport Aircraft

## 12.1 Summary

The increased environmental awareness over the last decades has led to a series of strict regulations concerning aircraft noise and air pollution. For supersonic transports (SST's), it is especially difficult to comply with these regulations because of the sonic boom problem and the relatively high combustion chamber temperatures and cruise altitudes. Since it is already very difficult to comply with the safety and performance requirements, because of the conflicting demands during subsonic and supersonic operation, taking the environmental constraints into account as well will leave only a very small constrained design space, if any at all. Even worse, since many of the environmental constraints are dictated by politics, they tend to be very susceptible to changes. Until a world-wide basis for regulation will be agreed upon, this makes the design of a second-generation supersonic transport a very uncertain enterprise.

It is clear that a design study, and especially an optimization study will produce unrealistic results if the environmental constraints are not taken into account. Unfortunately, since the mechanisms behind such subjects as pollution and noise are not yet fully understood there rests a large amount of uncertainty in the prediction of the environmental impact of a certain design. Furthermore, often very detailed and specialized analyses are required, which are beyond the scope of a conceptual design study. Because of this, simple relations will be used to incorporate the very important impact of the environmental constraints on the design. It is expected that these models will yield accurate trends in terms of dependencies on conceptual design variables rather than accurate absolute values.

The three topics of maximum concern for a future supersonic transport are:

- 1) pollutant emissions with special reference to ozone-layer depletion
- 2) take-off noise
- 3) sonic boom

## 12.2 Notation

$A_0$	intake capture area
$A_t$	throat area
$F$	fuel mass flow
$h$	altitude
	Planck's constant
$K_L$	lift factor
$K_p$	pressure amplification factor
$K_s$	shape factor
$M$	Mach number

---

	catalyst (in chemical reaction equations)
$p_{3_t}$	combustion chamber entry pressure (see Figure 44, chapter 11)
$\Delta p_{\max}$	maximum bow shock overpressure
$T$	thrust
$T_{3_t}$	combustion chamber entry temperature (see Figure 44, chapter 11)
$t$	time
$u_e$	exit jet velocity
$W$	maximum take-off weight
$\delta$	ratio of ambient pressure to sea level static pressure
$\epsilon_c$	compressor pressure ratio
$\lambda$	bypass ratio
$\nu$	frequency

### subscripts

0	sea level
cr	cruise
des	design
f	fuselage
inst	installed

### abbreviations

CPR	Compressor Pressure Ratio
EINO <sub>x</sub>	NO <sub>x</sub> Emissions Index
EPNLdB	Effective Perceived Noise Level [dB]
SST	Supersonic Transport

In accordance with tradition, weights are given in kilograms, although -according to the SI system- the unit of weight is Newton and the unit of mass is kilogram.

## 12.3 Pollutant Emissions

Although the aircraft contribution to the total air pollutant emissions is at present very small, it is expected that this contribution will rapidly increase due to the increase in air traffic. The following table (from [ref.14]) shows the result of a predictive study for Los Angeles International Airport, including the control of other polluting sources than aircraft and the expected growth of air traffic. The fast rising tendency is clear.



	1970	1975	1980
CO	1.0	5.8	13.6
HC	1.0	2.7	2.5
NO <sub>x</sub>	0.7	2.8	5.7

**Table 26**    Aircraft contributions in percentage of total air pollutant emissions

The emissions of main concern for aircraft are carbon monoxide, unburned hydrocarbons and nitrogen oxides (NO<sub>x</sub>). SO<sub>x</sub> and metal compound emissions are generally very low for aircraft fuels. This is caused by the fact that aircraft fuels are very clean and almost free of contaminants.

CO and unburned hydrocarbons affect the global climatological balance by enhancing the natural greenhouse effect. These emissions occur especially at low power settings and are a direct result of ineffective combustion. Therefore, reductions in this field are to be expected in improvements of the combustion process which is not affected by general conceptual design variables.

The main attention in this study will be given to the prediction of NO<sub>x</sub> emissions since these emissions are of special importance for supersonic transports because of their destructive influence on the ozone layer. Because of the relatively high temperatures in the combustion chambers of supersonic cruising aircraft the NO<sub>x</sub> emissions are expected to be high in comparison to subsonic aircraft, while furthermore these emissions are discharged at a critical altitude (above 13 km). The amount of NO discharged as a fraction of fuel burnt is directly affected by some conceptual design variables.

As is commonly known, the ozone layer performs an important function in absorbing ultraviolet radiation. Damaging the ozone layer will cause a dangerously increased level of UV-radiation which will in turn cause an increase in the cases of skin cancer. Presently, much attention is being paid to the so-called hole in the ozone layer over the poles and northern Europe and North-America. Although it is not known at the moment to what extent the present global ozone distribution is caused by natural mechanisms and to what extent air pollution is the cause, it is commonly agreed that NO plays an extremely important role in the ozone balance.

The formation of stratospheric ozone is described by the Chapman mechanism (1930) which describes the breaking up of oxygen molecules into free atoms by photolysis and the catalytic reaction of these atoms with oxygen molecules creating ozone [ref.9]:



Inverse reactions are the destruction of ozone by photolysis and by free oxygen atoms:



Furthermore, much of the lower stratosphere ozone is brought there by vertical air motions. The photochemical reactions depend on the solar radiation and absorption balance and depend therefore on the latitude.

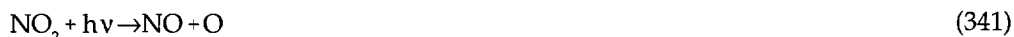
It was not until 1961 that, due to newly obtained values for the rate constants of the Chapman reactions, a large discrepancy was discovered between the global amount of ozone and the amount predicted by the Chapman model. Actually, based on the new rate constants, the Chapman reaction would double the world inventory of ozone after approximately 12 days. It was only in 1968 that the strong destructive influence on stratospheric ozone was ascribed to reactions with natural nitric oxides.

The most important reactions are:



Determination of the rate constants of these reactions led to the acknowledgement of two separate reactions:

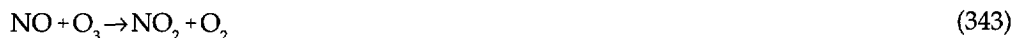
1) a neutral reaction:



---

no net reaction

2) the catalytic cycle for ozone destruction:



The rate of this reaction is about 10,000 times as fast as that of the similar Chapman reaction (336).

The difference between the Chapman model and the model including the  $\text{NO}_x$  reaction is shown in Figure 57, in which the ratio between the rate of ozone destruction and ozone formation is given. In the model including the  $\text{NO}_x$  reaction, the amount of natural stratospheric  $\text{NO}_x$  was adapted such as to yield the global values of ozone for March 22 1973. It was found that the amount required correlated well with the few known measurements at the time.

Although large uncertainties still exist with respect to the amounts of natural and pollution related  $\text{NO}_x$ , estimates in [ref.9] indicate that a fleet of 500 SST's with the same emission index as Concorde would approximately double the amount of natural  $\text{NO}_x$  in the stratosphere.

According to [ref.9] the emissions of future supersonic transports should at least be less than 20 g  $\text{NO}_x$  per kg fuel (Emission Index  $\text{EINO}_x < 20$ ). According to the two predictive models discussed hereafter, this is almost exactly the amount of the Olympus engine which powers Concorde.

In 1972, a publication by Lipfert ([ref.10]), showed for the first time the excellent correlation of the sea-level  $\text{NO}_x$  emissions for a large number of engines with the combustion chamber entry temperature (Figure 58). This correlation is still the most

---

important reference in many prediction models.

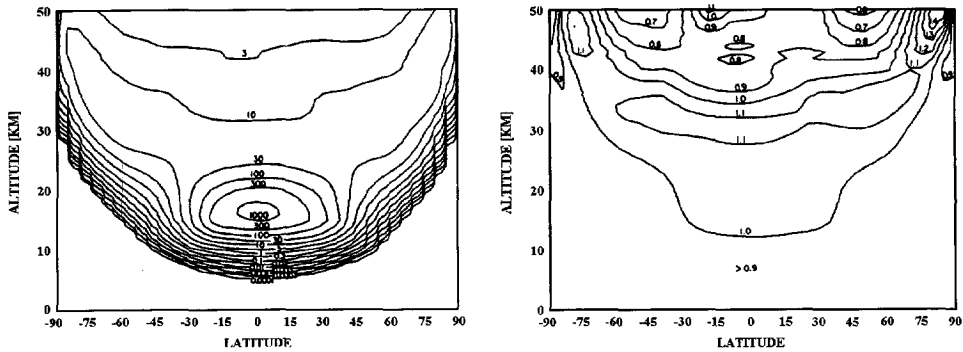


Figure 57 Ratio of ozone formation rate to rate of destruction for Chapman model (left) and model including  $\text{NO}_x$  (right) ([ref.9])

Since the ozone destruction by  $\text{NO}_2$  goes via the creation of  $\text{NO}$ , it is assumed that the  $\text{NO}$  emission index is a sufficient indication of the impact of a certain design on the ozone layer [ref.2]. According to [ref.2], there are two mechanisms that account for the formation of  $\text{NO}$ ; firstly the so-called prompt  $\text{NO}$  formation which is caused by the formation of free radicals and secondly the modified Zeldovitch mechanism:



In [ref.2] the  $\text{O}$ ,  $\text{N}_2$  and  $\text{N}$  concentrations are assumed to be equal to their equilibrium concentrations at the flame temperature. With the use of the rate constants of the modified Zeldovitch reaction, which also depend on flame temperature, the formation rate of  $\text{NO}$  can be deduced. In [ref.2] the expression for the  $\text{NO}$  formation rate is integrated from  $t=0$  sec. where the  $\text{NO}$  concentration equals that resulting from the prompt  $\text{NO}$  formation, to an average residence time of the reactants. It is found that the flame temperature, which is strongly affected by the combustor inlet temperature and pressure, has a major influence on  $\text{NO}$  formation. This was already established by Lipfert [ref.10] who showed the correlation between the sea-level  $\text{NO}_x$  emission index and the combustor entry temperature for a large number of different engines. The so-called Lipfert plot is reproduced in Figure 58.

Because of this strong correlation with flame temperature, the analysis of [ref.2] is based on a near stoichiometric fuel-to-air ratio, that is, an equivalence ratio of 0.9 (the equivalence ratio is defined as the fuel-to-air ratio as a fraction of the stoichiometric fuel-to-air ratio). Most publications agree that this yields the maximum flame temperature. The method of [ref.2] calculates the flame temperature from the combustor inlet temperature and pressure and estimates the average residence time by correlating the model to a representative result from the Lipfert plot. The resulting

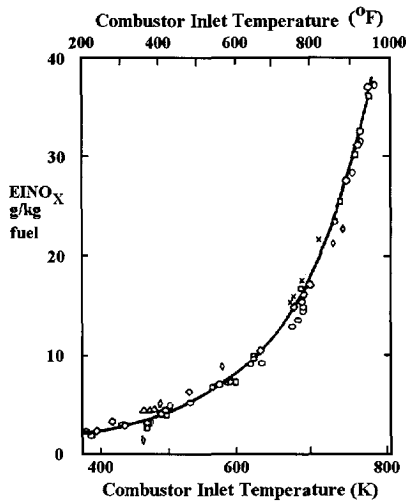


Figure 58 Lipfert  $\text{NO}_x$  correlation with combustor entry temperature ([ref.2])

value of 0.6 msec. is assumed to be valid for all cases.

Since the model of [ref.2] shows excellent agreement with the Lipfert plot and the model does not take additional variables into account, it seems acceptable to use the Lipfert correlation as the predictive model for the  $\text{NO}_x$  emission level. Figure 59 shows this correlation of the [ref.2] model with the Lipfert plot which exhibits an almost linear behavior once the  $\text{NO}_x$  emission level is plotted on a logarithmical axis. According to [ref.2] and other models the  $\text{NO}_x$  emission level correlates with the square root of the ambient pressure, which yields the following model:

$$\text{EINO}_x = 10^{\left(1 + 0.0032(T_{3t} - 581.25)\right)} \sqrt{\delta} \quad (347)$$

In [ref.2] the correlation with the square root of the pressure is investigated for two different equivalence ratios; 0.6 and 0.9. It was found that the results for an equivalence ratio of 0.9 tend to underestimate the  $\text{NO}_x$  emissions but the results for an equivalence ratio of 0.6 correlated very well with the square root of the pressure. Because future engines are expected to operate with equivalence ratios close to 0.6, the model is assumed to be usable.

A different model referred to in [ref.15] yields:

$$\text{EINO}_x = 0.000366 \sqrt{p_{3t}} e^{T_{3t}/200} \quad (348)$$

According to these models, the  $\text{NO}_x$  emission level is determined by the flight Mach

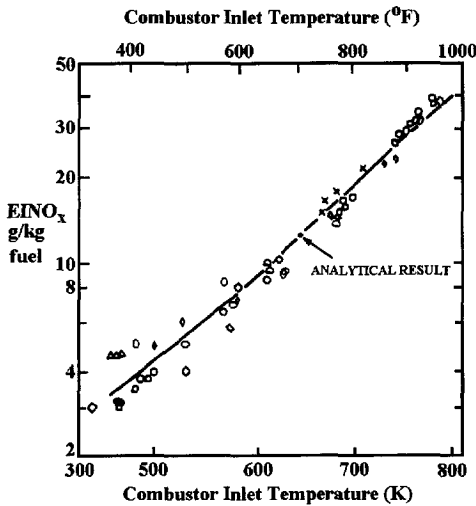


Figure 59 Correlation between Lipfert plot and model of [ref.2]

number, altitude, intake recovery factor, compressor pressure ratio and compressor efficiency.

However, since the major influence is exerted by the flame temperature, important improvements can be realized by reducing flame temperature, for a given combustion chamber entry temperature. This can be achieved without imposing any constraints on the conceptual design variables like those that describe the cruise condition.

Among the most effective measures to reduce flame temperature at constant combustion chamber entry temperature are:

- 1) water injection
- 2) pre-mixing
- 3) non-stoichiometric fuel-to-air ratios

According to the results of [ref.2], very large reductions in  $\text{NO}_x$  emissions can be achieved by injecting water in the combustion chamber. On the other hand, several other publications doubt the practical possibility to inject water during cruising (the most critical phase for  $\text{NO}_x$  emissions) because of the large amounts of water that have to be carried.

Pre-mixing of fuel and air and control of the flame shape are currently investigated.

The use of lean mixtures, especially mixtures with equivalence ratios of about 0.6 are thought to be very effective in reducing flame temperature.

It seems that major improvements in  $\text{NO}_x$  emissions can be achieved by focusing on

the combustion chamber, rather than by imposing constraints on the cruise condition.

In Figure 60, the  $\text{NO}_x$  emission level is plotted as a function of altitude for an engine that compares to the Olympus 593 (compressor pressure ratio = 11,  $M = 2.0$ , compressor efficiency = 0.85, intake recovery factor = 0.937). Three different models are shown: the results of the model of [ref.2], results of the model from [ref.15] and results from the present model based on the Lipfert plot. For the cruise altitude of Concorde, 18 km according to [ref.2], the model of [ref.2] yields an emission level of about 18 g/kg fuel, the model of [ref.15] yields about 22 g/kg fuel and the present model yields about 23 g/kg fuel.

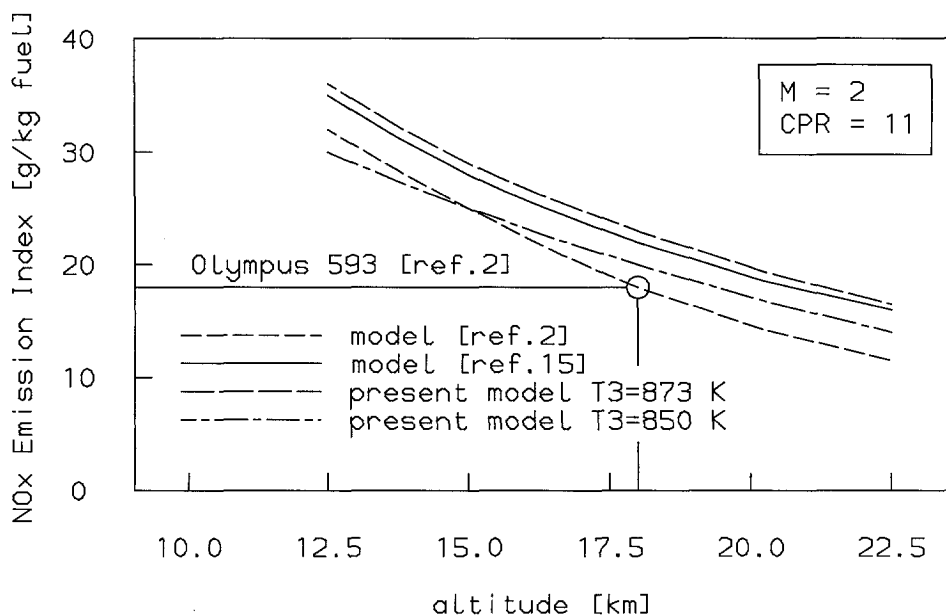


Figure 60 Comparison of three prediction methods

Actually, the compressor pressure ratio of the Olympus engine was chosen in order to limit the maximum compressor temperature to about 850 K (the above presented values yield 873 K) because of material properties of the compressor. The results of Figure 60 were based on 873 K. For the present model a line for 850 K is plotted as well, showing once again the large influence of the combustion chamber entry temperature on  $\text{NO}_x$  emissions.

It is interesting to notice that, according to the present model, the maximum  $\text{NO}_x$  emission level that can occur at about 13 km is 35 g/kg fuel if the combustor entry temperature is limited to 873 K and 30 g/kg fuel if it is limited to 850 K. According to

[ref.15], the altitude of 13 km is critical since below this altitude the methane-smog reaction creates ozone from  $\text{CH}_4 + \text{NO}_x$  faster than it is destroyed by the previously discussed catalytic cycle.

From Figure 60 it follows also that  $\text{NO}_x$  emission will decrease with increasing altitude. Since it is likely that the compressor temperature constraint will be active during cruising (in fact limiting the cruise compressor pressure ratio), the present model will effectively put a lower constraint on the cruise altitude.

Since important reductions in the  $\text{NO}_x$  emission levels can be achieved by reducing the flame temperature -as discussed previously- it seems acceptable to impose an  $\text{NO}_x$  emission constraint during the conceptual design phase equal to that of Concorde, as was also suggested in [ref.9], that is, 20 g/kg fuel. It may be anticipated that further reductions beyond this level can be achieved during a more detailed design phase.

## 12.4 Take-off Noise

For a supersonic transport, the take-off noise requirements are very difficult to comply with. This is partly caused by the fact that the induced drag of a highly-swept or delta wing is relatively high and the lift gradient relatively low, leading to high lift-off angles of attack. Together with the sharp leading-edges which cause leading-edge vortices, the take-off drag can be quite high. This calls for a large amount of thrust during take-off (Concorde uses reheat !).

The amount of noise is measured in terms of dB Effective Perceived Noise Level (EPNLdB) which includes weight factors on the higher frequencies and those frequencies that are especially annoying to the human ear. Although an accurate prediction of the various noise contributions to the level of EPNL is rather difficult and far beyond the scope of this study, it is clear that the noise of the exhaust jet is by far the largest contribution. Currently, almost every study into noise abatement regards the use of noise suppressors as an unrealistic option, because the effectivity is still very doubtful, the construction would be very heavy and the amount of noise reduction as compared to the amount of installed thrust loss is unproportionally small. In any case, the required amount of noise reduction cannot be achieved by the current studies of noise suppressors. It is therefore deemed necessary to reduce the exit jet velocity during take-off. There are many different solutions to achieve this.

---



The following table gives the maximum allowable values for the EPNLdB during take-off, according to the Federal Aviation Regulations, part 36 [ref.6]:

W (lbs.)	EPNLdB
600,000	108
300,000	103
150,000	98
75,000	93

Table 27    Maximum allowable EPNLdB during take-off (FAR 36)

According to Figure 61 (from [ref.15]) this implies that the jet velocity should be reduced at least below 600 m/s. Below 500 m/s very significant reductions of the jet noise can be achieved. In fact, [ref.1], [ref.12] and [ref.13] state that a jet velocity of 400 m/s will probably be needed in order to comply with the FAR 36 stage 3 requirements. In [ref.8], the required exit jet velocity is estimated at 450-500 m/s. Calder and Gupta [ref.4] present a jet velocity of 2000 ft/s (600 m/s) as a target for a second-generation supersonic transport. According to Figure 61 this would result in a EPNLdB of 105 dB. Other contributions to the noise level are still to be added. It is therefore assumed in this application that constraining the exit jet velocity to 500 m/s at take-off will be sufficient guarantee that the noise level during take-off will be in compliance with regulations. This will lead to larger engine diameters than would be necessary to provide the required take-off thrust if take-off noise would not be an active constraint. It should be realized, that reducing the exit jet velocity will cause another important contribution, that is, the fan noise, to be more dominantly present in the total noise level. This effect will not be taken into account.

A very discomforting consideration, which is nevertheless very important, is that a future supersonic transport plane will have to compete with contemporary subsonic airplanes and that regulations will probably follow trends in subsonic engine technology. The development of subsonic engines with even higher bypass ratios than those of today might result in exit jet velocities even below 300 m/s. Such reflections will not be taken into account in this study.

Many conceptual studies into engine technology for a future second-generation supersonic transport are concerned with the problem of reducing the exit jet velocity during take-off by some sort of variable geometry or cycle variability. Such engines are referred to in literature as Variable Cycle Engines. The term applies as well to designs which attempt to reduce subsonic noise and specific fuel consumption by using a relatively high bypass ratio and applying duct-burning during supersonic flight.

Two major lines of thought can be discerned at present, both of which have in common that an additional airflow is supplied during the take-off phase in order to

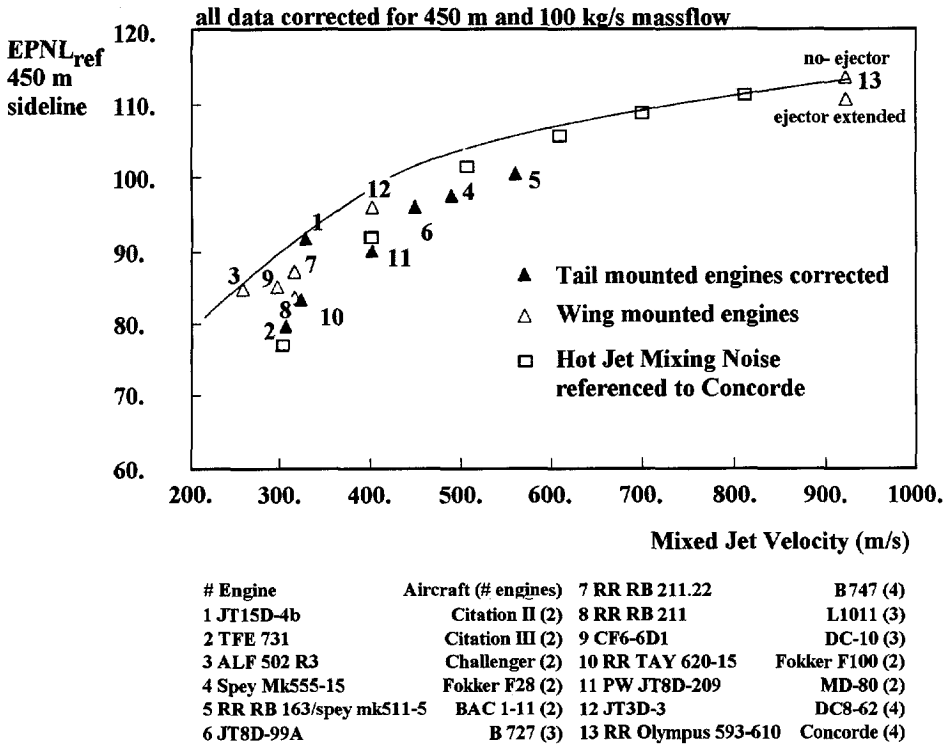


Figure 61 Effective Perceived Noise Level as a function of mixed exhaust jet velocity ([ref.15])

reduce the exit jet velocity. The solution forwarded by Rolls-Royce and SNECMA consists of a tandem-fan of which the rear one is situated partly in the "shade" of the frontal fan in order to keep the engine diameter within limits ([ref.11] and [ref.12]). During the take-off phase and the lower subsonic flight phase a secondary airflow is submitted to the second fan through a secondary inlet door situated between the two fans. This concept is usually referred to as the *flow multiplier* concept.

The second solution is known as the *mixer-ejector* turbofan and this option is mainly favored by the American companies like Pratt & Whitney and General Electric, although Rolls-Royce have published an engine study falling in this category as well. The mixer-ejector principle injects the secondary airflow only in the jetpipe, that is, behind the low pressure turbine. The main difference with the flow multiplier concept consists of the fact that the cold airflow is only compressed by a single device. Other versions of this concept are discussed in [ref.8], amongst others the double-bypass solution, in which air compressed by the high pressure compressor is submitted to the bypass flow, and the turbine-bypass option for pure jets. According to [ref.8], the mixer-ejection and duct burning options still rely too much on additional noise suppression systems and the required intake diameter in most applications tends to

be relatively high.

Because of the complexity and weight of these concepts, MTU suggest to provide the required flexibility just by applying an adjustable nozzle throat to facilitate a varying bypass ratio. In a subsequent design phase it will be established whether the required compressor and fan operating envelope in terms of pressure ratio-corrected mass flow combinations can be achieved by the detailed compressor and fan design. In case the required fan or compressor flexibility is too large, the use of variable compressor vanes might be considered ([ref.1]). Since this concept has many similarities with the engine analysis procedure used in the present application (chapter 11) and because of the complexity and required detail of the variable cycle solutions, the applicability of the variable throat area/ variable bypass ratio model to reduce the exit jet velocity during take-off will be studied here.

The engine performance will be evaluated using the method outlined in chapter 11 as a function of seven input variables, being the design Mach number and altitude, the design compressor pressure ratio, the design bypass ratio, the intake area and the off-design Mach number and altitude. As explained in chapter 11, the off design working point of the engine subject to the low pressure and high pressure spool power equations, the turbine massflow requirements and the condition for loss-free mixing can be expressed in one single relation between the off-design bypass ratio and the core airflow (or vice-versa). Once an operating point is chosen according to this relation, all engine parameters are known. It was established in chapter 11, that a large value of the core massflow (that is, close to the captured massflow) results in a relatively large value of the bypass ratio and a small value for the compressor pressure ratio. This in turn leads to very small values for the installed thrust (despite the fact that a core massflow close to the value of the captured massflow diminishes the amount of spillage (drag)). Therefore, normally, in order to maximize the amount of installed thrust at each off-design point, the bypass ratio is chosen at its minimum allowable value (which is prescribed by the maximum allowable compressor pressure ratio, or maximum allowable turbine entry temperature). This results in the maximum amount of excess air which in supersonic flow will be spilled either in front of the intake by the shockwave system or in front of the fan by a dump door. In subsonic flow the excess air will simply flow around the intake.

In order to constrain the exit jet velocity during take-off, however, it might be interesting to choose a larger value for the bypass ratio and accept the resulting loss of thrust. Furthermore, since it might become necessary during a more detailed design phase to impose a limit on the fan pressure ratio -as explained in chapter 11- this too might require the choice of a higher bypass ratio during take-off, causing a deterioration of take-off performance on the one hand but a noise reduction on the other hand.

The consequences of such a procedure were investigated and the results summarized in the following table (a value of 0.2 was assumed for the take-off Mach number). The first column in each example represents the maximum installed thrust solution (minimum allowable bypass ratio). Since the following tables are purely meant as an

---

illustration of the influence of the bypass ratio on the exit jet velocity, fan pressure ratio constraints are not imposed. However, an analysis of the influence of fan pressure ratio limitations on take-off performance is carried out for the final engine design in Appendix C of chapter 15.

$M_{des}=2.5$ $h_{des}=20$ km
$A_0=1$ m <sup>2</sup> $\epsilon_{c,des}=5$ $\lambda_{des}=0.7$

F [kg/s]	1.34	0.49	0.37	0.21
$A_t$ [m <sup>2</sup> ]	0.166	0.303	0.338	0.345
$\lambda$	0.26	0.69	0.90	1.25
$\epsilon_c$	5.0	3.6	3.2	2.4
$u_e$ [m/s]	1019	590	532	470
$T_{inst}$ [N]	31600	15043	11001	4295

Table 28 Influence of cycle variability on exhaust jet velocity and thrust (fixed exhaust area)

$M_{des}=0.8$ $h_{des}=11$ km
$A_0=1$ m <sup>2</sup> $\epsilon_{c,des}=25$ $\lambda_{des}=0.7$

F [kg/s]	1.70	0.50	0.30	0.037
$A_t$ [m <sup>2</sup> ]	0.168	0.286	0.330	0.344
$\lambda$	0.36	0.89	1.29	7.5
$\epsilon_c$	9.4	5.2	4.0	1.2
$u_e$ [m/s]	1032	606	514	395
$T_{inst}$ [N]	51498	18616	10755	646

(Pratt & Whitney PW-100) (The design turbine entry temperature in this example was adapted to the actual value of the PW-100, that is, 1550 K instead of 1500 K).

Table 29 Influence of cycle variability on exhaust jet velocity and thrust (fixed exhaust area)

$M_{des}=2.0$ $h_{des}=15$ km
$A_0=1$ m <sup>2</sup> $\epsilon_{c,des}=5$ $\lambda_{des}=5$

F [kg/s]	0.54	0.41	0.31	0.21
$A_t$ [m <sup>2</sup> ]	0.279	0.296	0.323	0.342
$\lambda$	2.41	2.74	3.24	3.98
$\epsilon_c$	5.0	4.2	3.6	2.9
$u_e$ [m/s]	620	567	520	472
$T_{inst}$ [N]	19979	14874	11280	6833

Table 30    Influence of cycle variability on exhaust jet velocity and thrust (fixed exhaust area)

The results of these tables are presented together with a relation from [ref.11] (a correlation for different engines including the Olympus) in Figure 62.

According to these results, the loss in thrust is quite large. In the first two examples, reducing the exit jet velocity to about 500 m/s reduces the installed thrust with a factor of 2 respectively 2.7. At that point the first two engines still deliver thrust in the same order of magnitude as the third design. Apparently, it is possible by choosing a relatively high design point bypass ratio to reduce the take-off exit jet velocity, be it at the cost of both take-off and cruise thrust (the installed design thrust of the three examples is respectively 20087 N, 55755 N and 13420 N).

It seems further, that reducing the exit jet velocity below 500 m/s is especially costly in terms of installed thrust loss. The installed thrust loss caused by reducing the exit jet velocity to about 500 m/s could still be counteracted by increasing the intake area by a factor two or three. However, such measures are totally unacceptable if the exit jet velocity were to be brought down to about 400 m/s.

It was investigated whether an engine design could be obtained using the procedure described above, with a take-off thrust in excess of 150 kN (based on a 33% take-off thrust-to-weight ratio and a 200 metric ton class airplane; for instance Concorde) and an exhaust jet velocity below 500 m/s. It should be realized that these constraints are not very severe for a second-generation supersonic airplane; it is expected that the take-off weight might well be in excess of 200 metric ton whereas an exhaust jet velocity of 500 m/s might still require additional noise suppressors.

Using the GENESIS genetic optimization program [ref.7], an evaluation of the design space was performed, using an objective function that adds a penalty to designs that

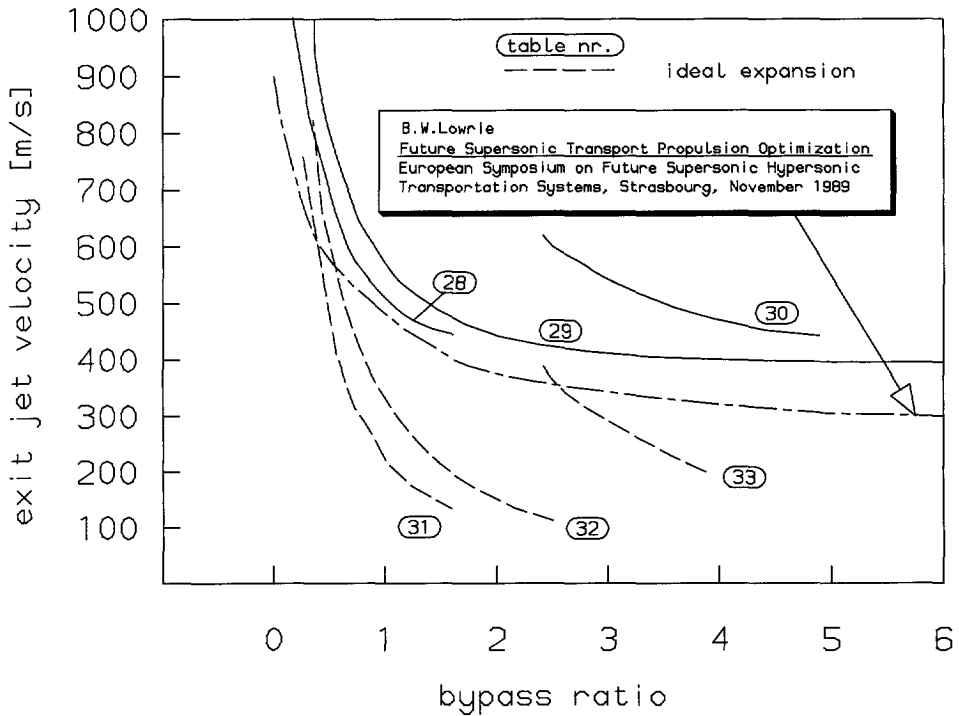


Figure 62 Relation between take-off exit jet velocity and bypass ratio

perform below the requirements. An upper bound of  $4 \text{ m}^2$  was imposed on the intake area and it was found that the best designs obtained all had this bound active but that none of the best designs complied with the requirements. After 5000 function evaluations, the best designs had a design Mach number of 1.8, an intake area of  $4 \text{ m}^2$ , a design compressor pressure ratio of 16.5, a design bypass ratio of 1.35 and a take-off bypass ratio parameter of almost zero. (If the value of this parameter is zero, the bypass ratio is chosen at its minimum, which implies maximum spillage, if its value is one, the bypass ratio is put at its maximum, which implies minimum spillage).

Because of the active upper bound on the intake area, a second run was performed with an upper bound of  $9 \text{ m}^2$ . Again, in this case, all best designs had an active upper bound on the intake area. The requirements were in this case just about satisfied. The best designs are described by the following variables: design Mach number of 1.9, intake area of  $9 \text{ m}^2$ , design compressor pressure ratio of 14.5, design bypass ratio of 3.9 and take-off bypass ratio parameter of 11%. The optimum values of the design variables have therefore clearly shifted. The best design obtained had a take-off thrust of 150 kN at an exit jet velocity of 558 m/s (different results can be obtained if an other penalty function is applied).

Obviously, a design with an intake area of more than 9 m<sup>2</sup> is totally unrealistic. It must therefore be concluded that complying with the performance and environmental requirements at the same time is not possible for this kind of engine.

Therefore a different approach was decided upon. It was established that the main cause for the problems is the fact that -especially during take-off- a large amount of overexpansion occurs in the exhaust nozzle. This is due to the fact that the exhaust area is designed for ideal expansion under cruise conditions and subsequently kept constant. In case, however, a variable exhaust nozzle would be applied, it would be possible to realize ideal expansion (that is, maximum net thrust) at each operating point. This does not only yield larger thrust, but lower exhaust jet velocities as well.

This is illustrated in the following tables which are the same as the previously presented ones, except that in this case ideal expansion is enforced by adapting the exhaust area.

$M_{des}=2.5$ $h_{des}=20$ km
$A_0=1$ m <sup>2</sup> $\epsilon_{c,des}=5$ $\lambda_{des}=0.7$

F [kg/s]	1.34	0.49	0.37	0.21
A <sub>t</sub> [m <sup>2</sup> ]	0.166	0.303	0.338	0.345
A <sub>e</sub> [m <sup>2</sup> ]	0.181	0.314	0.379	0.488
λ	0.26	0.69	0.90	1.25
ε <sub>c</sub>	5.0	3.6	3.2	2.4
u <sub>e</sub> [m/s]	765	333	262	173
T <sub>inst</sub> [N]	37624	20557	15814	8041

Table 31    Influence of cycle variability on exhaust jet velocity and thrust (ideal expansion)

$M_{des}=0.8$ $h_{des}=11$ km
$A_0=1$ m <sup>2</sup> $\epsilon_{c,des}=25$ $\lambda_{des}=0.7$

F [kg/s]	1.70	0.50	0.30	0.174
$A_t$ [m <sup>2</sup> ]	0.168	0.286	0.330	0.353
$A_e$ [m <sup>2</sup> ]	0.207	0.288	0.371	0.503
$\lambda$	0.36	0.89	1.29	1.83
$\epsilon_c$	9.4	5.2	4.0	2.9
$u_e$ [m/s]	823	369	254	168
$T_{inst}$ [N]	59164	24619	15374	8107

Table 32 Influence of cycle variability on exhaust jet velocity and thrust (ideal expansion)

$M_{des}=2.0$ $h_{des}=15$ km
$A_0=1$ m <sup>2</sup> $\epsilon_{c,des}=5$ $\lambda_{des}=5$

F [kg/s]	0.54	0.41	0.31	0.21
$A_t$ [m <sup>2</sup> ]	0.279	0.296	0.323	0.342
$A_e$ [m <sup>2</sup> ]	0.279	0.306	0.358	0.442
$\lambda$	2.41	2.74	3.24	3.98
$\epsilon_c$	5.0	4.2	3.6	2.9
$u_e$ [m/s]	388	322	263	195
$T_{inst}$ [N]	25996	20210	15902	10247

Table 33 Influence of cycle variability on exhaust jet velocity and thrust (ideal expansion)

The results of these tables are plotted in Figure 62 as well.

With this procedure another optimization was performed using the GENESIS algorithm. With the same upper bound of 9 m<sup>2</sup> on the intake area it was found that designs could be obtained which complied with the imposed requirements. An extra



penalty was imposed for the intake area in order to find the smallest suitable engine. The best designs are defined by the following values: design Mach number of 1.8 or 1.9, intake area of  $4.3 \text{ m}^2$ , design compressor pressure ratio of 7.4, design bypass ratio of 0.7 or 0.8, take-off bypass ratio parameter of 15.5%.

It can thus be concluded that an engine for a supersonic cruising airplane has to be equipped with a variable exhaust nozzle in order to be able to comply with both performance and environmental constraints. Since the optimal choice of the engine parameters is determined by their influence on the aircraft parameters and variables as well (and vice versa) the bypass ratio during take-off is regarded as a design variable too. The exit jet velocity will be constrained to 500 m/s and the five engine design variables plus the take-off bypass ratio parameter will be determined by an optimization procedure such as to comply with all performance and environmental constraints while simultaneously yielding the best airplane design.

## 12.5      Sonic boom

Almost immediately after the introduction of Concorde, a public rage developed in the United States about the sonic boom problem. Although few people have ever heard a sonic boom and blind tests have established that sonic booms were rated as less annoying than for instance "normal" traffic noise, public opposition (strongly encouraged by U.S. politicians) led to a ban on overland supersonic flights in U.S. airspace. Soon other countries followed (including Saudi-Arabia which led to the cancellation of Concorde's Middle East line). Today a world-wide ban on overland supersonic flights exists.

Because of the inefficiency of supersonic transports at subsonic speeds such a ban on supersonic overland flight has very serious consequences for the range performance. The number of routes available for the operation of an SST is dramatically reduced by this measure, which will make the economical viability even more uncertain than is already the case. This implies that serious attention needs to be paid to improving the subsonic efficiency of a second-generation SST in comparison to the first generation. There are indications that the subsonic range parameter of the new generation of designs can be equal or higher than the supersonic range parameter, implying that the subsonic range will at least be equal to the supersonic range. This of course depends on the specifications imposed concerning subsonic range performance; it is said that Aerospatiale require their *Alliance* project to fly the same distance at subsonic speeds as at supersonic speeds.

In the past, much attention has been given to the possibility to reduce the bow shock overpressure of the shockwave generated by an overflying supersonic airplane. The trade-off between a minimum wave drag configuration and a configuration with minimum shock wave overpressure was a well-known problem. Today, however, it is realized that every pressure signature that differs from the basic N-shape is unstable, which implies that the sonic boom strength of a certain design depends mainly on its weight, cruise altitude, cruise Mach number and fuselage length.

---

The need to implement a sonic boom constraint can be argued. If the world-wide ban on supersonic overland flight will be maintained in future (which no doubt will be the case), then the fact should be accepted that a certain part of the flight will be flown at subsonic speed. It is therefore more realistic to impose a subsonic range constraint (for instance 50% of the supersonic range) than to constrain the cruise conditions for a certain aircraft weight in order to reduce the sonic boom. New regulations about the effect of overflying SST's on ships or ocean wildlife (especially whales), however, might impose a constraint on the maximum overpressure of the bow shock nonetheless.

Carlson, in [ref.5] introduced a simplified method to predict the maximum bow shock overpressure for Concorde-like configurations, based on an N-shaped pressure signature formed by a bow-shock and a linear expansion followed by the same shock strength to restore ambient pressure.

The maximum overpressure is given by the following formula:

$$\Delta P_{\max} = 2K_p K_s \sqrt{P_{cr} P_0} (M^2 - 1)^{1/8} h_{cr}^{-3/4} l_f^{3/4} \quad (349)$$

in which  $K_p$  and  $K_s$  are a pressure amplification factor and a shape factor respectively.

In Carlson's model, the pressure amplification factor can be read from a figure, in the present application it is approximated for two different Mach numbers by the following expressions:

$$\begin{aligned} M=1.2: \quad K_p &= 0.9707 + 0.01336h + 0.0002957h^2 \\ M=2.0: \quad K_p &= 0.9943 + 0.004684h + 0.0001852h^2 \end{aligned}$$

with  $h$  in km.

The shape factor is based on the Concorde configuration and is approximated as a function of the lift parameter  $K_L$  as follows:

$$K_s = 0.04943 + 3.1949K_L - 16.2148K_L^2$$

The lift parameter according to Carlson is defined as:

$$K_L = \frac{\sqrt{M^2 - 1} W}{1.4 p M^2 l_f^2} \quad (350)$$

in which  $p$  is the ambient pressure in Pa and  $W$  the airplane weight in N.

Applying this very simple analysis to Concorde during cruise flight yields the following results:

Concorde  $l_f$ : 62.1 m  
W: 185065 kg  
M: 1.97  
h: 15.8 km  
p: 10682.8 Pa

These values yield for the coefficients:

$K_L$ : 0.01376  
 $K_s$ : 0.09335  
 $K_p$ : 1.1145

which in turn yield:  $\Delta p_{\max}=123$  Pa.

For a cruise Mach number of 2.0 at 18 km altitude ( $\delta=0.07407$ ) the coefficients attain the following values:

$K_L$ : 0.01939  
 $K_s$ : 0.1053  
 $K_p$ : 1.1386

yielding:  $\Delta p_{\max}=108$  Pa.

The actual bow-shock overpressure of Concorde is about 75-80 Pa which shows that the error of the simplified analysis is quite large. However the method is expected to yield accurate trends in the change of the overpressure as a function of the design parameters like cruise Mach number, altitude and airplane weight.

According to [ref.3], based on a human response study, unlimited overland supersonic flight might be accepted for bangs not louder than 65 dBA. Bangs with a loudness below 72 dBA might still be acceptable to two thirds of the population. Therefore, in [ref.3] this noise level was proposed as an upper level for supersonic flight through restricted corridors. The peak pressure of an N-shaped wave corresponding to these noise levels depends on the rise-time (see Figure 63, [ref.3]; the rise-time determines the frequency and therefore the dBA level). For a rise-time of 15 ms, the limit according to the corridor criterion would be about 100 Pa. For a more representative rise-time of 10 ms the maximum overpressure must be below about 1.5 psf or 72 Pa. In [ref.17], this same limit is used but in [ref.15] and [ref.16], the sonic boom constraint for a supersonic transport plane design study is put at twice this value (144 Pa), which is quite high in comparison with Concorde.

Since, in the light of the previous discussion, overland supersonic flight will not be allowed anyway, in the present study the limit of 72 Pa will be used. In [ref.13] it is suggested that with an optimal choice of the volume distribution maximum

---

overpressures could be reduced to values as low as 1 psf, or about 50 Pa. In comparison with other publications, this seems hardly likely. No constraint is imposed on the maximum overpressure during transition. A simple analysis shows that the maximum altitude where transition can take place is about 15 km (based on a wing loading of 4000 N/m<sup>2</sup>, a subsonic Mach number of 0.9 and a maximum attainable lift coefficient of 0.6, which is quite high). Flying at Mach 1.4, 15 km with an aircraft in the 300 metric ton class it can be verified that the maximum bow shock overpressure is still in the order of 150 Pa. Furthermore, during acceleration a so-called "superboom" can occur because of focussing of the shockwaves reaching the earth. It seems therefore almost unavoidable to let acceleration to supersonic speeds take place in special areas where loud bangs are allowed.

The same kind of analysis can be carried out for the cruise flight. In case the minimum cruise Mach number is fixed at 1.6, the maximum length of the airplane at 100 m and the maximum cruise altitude at 20 km an airplane in the 300 metric ton class produces a shock wave with a peak pressure of at least 147 Pa, with all these constraints active. Even if this value is compensated for the apparent overestimation of the method by about 50%, this would imply a peak pressure of 98 Pa. Only if the cruise altitude is increased to about 24 km the peak pressure complies with the corridor criterion. Cruise altitudes of more than 20 km however, are not realistic because of considerations regarding safety after explosive decompression (terrorist attacks) and cruise lift coefficient (a wing area of 1000 m<sup>2</sup> would still require a cruise lift coefficient of 0.3 in the present example). It is therefore very unlikely that the sonic boom constraint will be met.

Because the definition of dBA sound levels does not take the low frequencies into account (below 20 Hz, the so-called infrasound region) a separate criterion must be defined that measures the annoyance caused by such frequencies, since it has been reported that low frequency waves (which are also less well damped) can cause structures to vibrate and rattle. According to [ref.3] this criterion is met for all waves that comply with the corridor requirement (maximum overpressure must be less than 2.2 psf which is 107 Pa).

## 12.6 Conclusions

Recent experiences in industry have indicated that it is of the utmost importance to take environmental issues into account if a realistic and feasible second-generation supersonic transport design is to be obtained. It seems that the introduction of environmental constraints in the design and optimization process will further diminish the already extremely small feasible design space. The fact that most constraints do change constantly due to political decisions, makes it very difficult indeed to come up with a feasible design, let alone an optimized design, and the duration of such a design is generally not very long. It is therefore of the utmost importance to introduce as many design variables as possible in order to increase the degrees of freedom. Unfortunately the number of independent variables in the conceptual design phase is limited and the solution should be found in optimizing

---

dBA

## N-wave Sensitivities Required to Meet Loudness Goal

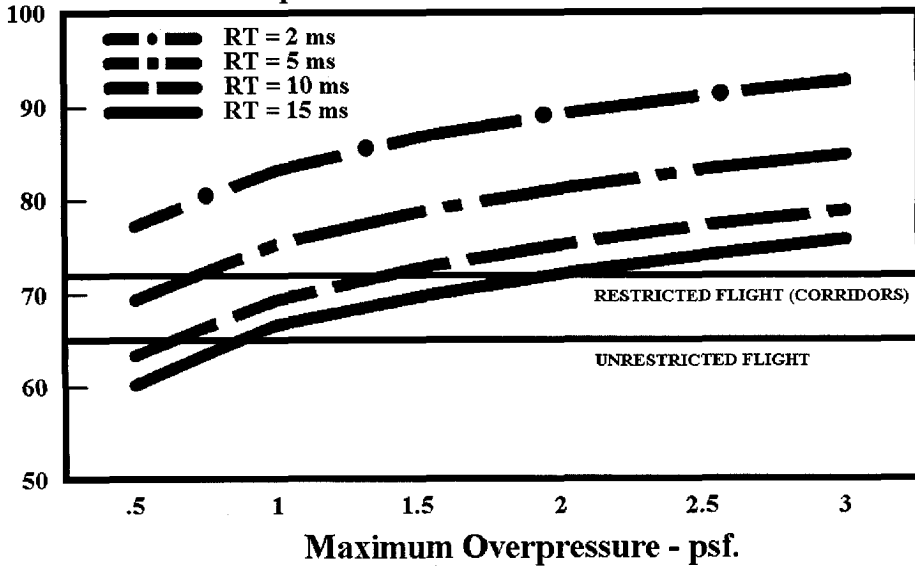


Figure 63 N-wave loudness versus maximum overpressure ([ref.3])

wing airfoil section shapes, camber and twist as well as the fuselage cross section distribution. In other words, it seems that the preliminary design level should be taken into account in the optimization as well, which is very difficult because of the sharp increase in the number of independent variables from about 30 in the conceptual level to several hundreds in the preliminary design level. This will produce great difficulties in the optimization procedure.

A remarkable trend has already been reported in the design of supersonic transports subject to environmental constraints ([ref.16]). Compliance with the three mentioned environmental constraints of greatest concern as well as a number of new safety requirements seems to require the reduction of the cruise Mach number to values as low as 1.6. This is far less than the target for most currently undertaken design studies.

## 12.7 References

1. Albers, M.; Hertel, J. and Schaber, R.; *Antriebskonzepte für zukünftige Überschallverkehrsflugzeuge*; Jahrbuch der Deutschen Gesellschaft für Luft- und Raumfahrt e.V. (DGLR), 93-03-044, Bonn, 1993

2. Blazowski, W.S.; Walsh, D.E. and Mach, K.D.; *Operating and Ambient Condition Influences on Aircraft Gas Turbine NO<sub>x</sub> Emissions*; Journal of Aircraft Vol.12, No.2, February 1975
  3. Brown, J.G. and Haglund, G.T.; *Sonic Boom Loudness Study and Airplane Configuration Development*; AIAA 88-4467 September, 1988
  4. Calder, P.H. and Gupta, P.C.; *Future SST Engines with Particular Reference to Olympus 593 Evolution and Concorde Experience*; Aeronautical Journal, June 1976, pp. 235-252
  5. Carlson, H.W.; *Simplified Sonic-Boom Prediction*; NASA TP 1122, March 1978
  6. Federal Aviation Regulations, Vol.III, part 36; Department of Transportation, Federal Aviation Administration, Washington D.C., 1969
  7. Grefenstette, J.J.; *A User's Guide to GENESIS*; Version 5.0, October 1990
  8. Habrard, A.; *The variable-Cycle Engine*  
*A Solution to the Economical and Environmental Challenge of the Future Supersonic Transport*; Proceedings of the European Symposium on Future Supersonic Hypersonic Transportation Systems, Strasbourg, 6-8 November 1989
  9. Johnston, H.S. and Whitten, J.; *Reaction of Ozone with Nitrogen Oxides at High Altitudes*; Atmospheric Pollution by Aircraft Engines, AGARD Conference Proceedings CP-125, April 1973
  10. Lipfert, F.W.; *Correlation of Gas Turbine Emissions Data*; ASME paper 72-GT-60, March 1972
  11. Lowrie, B.W.; *Future Supersonic Transport Propulsion Optimization*; Proceedings of the European Symposium on Future Supersonic Hypersonic Transportation Systems, Strasbourg, 6-8 November 1989
  12. Lowrie, B.W. and Portjoie, E.; *Two Variable Engine Cycle Concepts for Second Generation Supersonic Transport*; SAE Technical Paper 901892, Aerospace Technology Conference and Exposition, Long Beach, California, October 1990
  13. Mizuno, H.; Hagiwara, S.; Hanai, T. and Takami, H.; *Feasibility Study on the Second Generation SST*; AIAA 91-3104
  14. Sawyer, R.F.; Cernansky, N.P. and Oppenheim, A.K.; *Factors Controlling Pollutant Emissions from Gas Turbine Engines*; Atmospheric Pollution by Aircraft Engines, AGARD Conference Proceedings CP-125, April 1973
  15. van der Velden, A.J.M.; *Aerodynamic Design and Synthesis of the Oblique Flying Wing Supersonic Transport*; Ph.D. dissertation, Stanford University, SUDAAR 621, June 1992
  16. van der Velden, A.J.M.; *Multi-disciplinary SCT Design Optimization*; AIAA 93-3931, August 1993
  17. van der Velden, A.J.M.; *Sonic Boom of the Oblique Flying Wing*; Journal of Aircraft Vol.31, No.1, Jan.-Feb. 1994
-

# 13. The Mission Analysis

## 13.1 Summary

Due to the large flight envelope of aircraft designed to cruise at supersonic speeds, many conflicting demands are imposed on the design. The aircraft must not only perform well at its design point, it must be able to attain the design point in an acceptable manner as well. Amongst others, the occurrence of spillage air, the shift of the aerodynamic center position during transition to supersonic speeds and back, trim drag and the transonic drag rise are effects that can cause the performance of engines and aircraft that were designed for supersonic speeds to deteriorate dramatically under off-design conditions. Therefore, a mission analysis is performed in order to calculate the amount of mission fuel burnt from take-off until the cruise condition is attained. For supersonic aircraft -because of the reasons mentioned above- this can be quite a large quantity. Referring to [ref.1], it was found that no less than one third of the total mission fuel was burnt during the accelerated climb to cruise speed and altitude. For Concorde, this amount is 38%.

## 13.2 Notation

A	factor in square root equation
a	speed of sound
B	factor in square root equation
C	rate of climb
	factor in square root equation
D	drag
F	fuel flow
$\mathcal{J}$	merit function
f	function
g	gravitational acceleration
h	altitude
$h_e$	energy height
I	integral
$I_{sp}$	specific impulse
$\overline{I_{sp}}$	average specific impulse
L	lift
M	Mach number
n	horizontal acceleration
SEP	Specific Excess Power (= static rate of climb)
T	thrust
t	time
V	velocity
W	weight
x	set of independent variables
y	function

---

$\gamma$  climb angle

### subscripts

0 at take-off  
 1 at beginning of cruise flight  
 cr cruise  
 est estimated  
 f fuel  
 final  
 max maximum  
 S static

## 13.3 Conditions for Climb with Minimal Fuel Consumption

It is obvious that the amount of fuel burnt, depends largely on the way in which the ascent path is flown. This implies that, if the ascent is performed in a more economical way, a better optimal design is achievable. Therefore, in order to optimize the complete aircraft design, the ascent must be optimized as well. The optimization of the performance of a given airplane design is a dynamic optimization problem, which is an altogether different kind of optimization than that dealt with in the present work. Apart from the time parameter, also control variables like the flight path angle, the angle of attack and the throttle setting feature in it. To avoid the necessity to optimize all those variables, a simple energy state model was used, which together with the assumption that the engine operates at its maximum rating, reduces the problem to finding the optimal relation between speed and altitude during the accelerated ascent.

The choice of the merit function for the ascent path optimization is a difficult one, as it has a large influence on the final optimal design, as explained earlier. To be absolutely certain that the chosen path is the optimal one in the sense of yielding the optimal aircraft design, the parameters that describe it should be taken into account as design variables in the overall optimization problem. Instead, an educated choice was made for the ascent merit function, which is, to optimize the ascent path for minimal fuel consumption (another option would have been minimal flight time). Thus, the ascent merit function becomes:

$$\text{minimize } W_f = \int_{h_1}^{h_2} \frac{F dh}{C} \quad (351)$$

From the calculus of variation follows the Euler-Lagrange equation, which is the



optimality condition for a merit function of the shape  $I = \int_{x_0}^{x_1} \mathcal{F}(x, y, y') dx$  ([ref.7]):

$$\frac{d}{dx} \left( \frac{\partial \mathcal{F}}{\partial y'} \right) - \frac{\partial \mathcal{F}}{\partial y} = 0 \quad (352)$$

Writing:

$$C \left( 1 + \frac{V}{g} \frac{dV}{dh} \right) = \frac{(T-D)V}{W} = C_s \quad (353)$$

the merit function becomes:

$$\mathcal{F} = \frac{\left( 1 + \frac{V}{g} \frac{dV}{dh} \right) F}{C_s} = f \left( V, \frac{dV}{dh}, h \right) \quad (354)$$

For a given throttle setting  $C_s = f(V, h)$ , whereas  $V = f(h)$  is the function to be determined.

The condition for the flight path with minimum fuel consumption now becomes:

$$\left( \frac{\partial (C_s/F)}{\partial V} \right)_h - \frac{V}{g} \left( \frac{\partial (C_s/F)}{\partial h} \right)_V = 0 \quad (355)$$

Introducing the energy height:

$$h_e = h + \frac{V^2}{2g} \quad (356)$$

this simplifies to:

$$\left( \frac{\partial(C_s/F)}{\partial V} \right)_{A_t} = 0 \quad (357)$$

Therefore, at each energy height level, the value for the airspeed must be determined, at which the static rate of climb (or specific excess power) divided by the fuel consumption is maximal. The altitude corresponding to this speed and energy height follows easily.

### 13.4 Mission Analysis Procedure

The mission analysis routine divides the regime between the speed and altitude immediately after take-off and the cruise condition into equal steps of energy height. Since, naturally, the weight at the beginning of the cruise flight is not yet known at this point of the calculations, and consequently the cruise altitude is unknown as well, it is necessary to make an estimation of the amount of fuel needed to attain the cruise condition, using the following relation ([ref.6]):

$$\ln \left( \frac{W_0}{W_1} \right) = \frac{M_{cr} a_{cr}}{I_{sp} g} \left( 1 + \frac{1}{\bar{n}_0 L/D} \right) \quad (358)$$

This relation was derived under the assumption of constant  $T/D$  and vertical equilibrium during the accelerated climb. Under these assumptions the last term is constant and the average acceleration can be based on the initial (take-off) value according to:

$$\bar{n}_0 = \left( \frac{T}{W} \right)_0 - \frac{1}{(L/D)_0} \quad (359)$$

The average specific impulse is defined as:

$$I_{sp} = \int_0^{V_1} I_{sp} dV \quad (360)$$

According to [ref.6], the specific impulse is constant with  $V$  as long as the speed is smaller than 10% of the circular velocity ( $M=2.5$  in the stratosphere yields  $V/V_c = 2.5 \times 295/7900 = 0.09$ ). Therefore, the value at take-off is used for the specific impulse

as well. The values that are given in [ref.6] for the specific fuel consumption (reciprocal of specific impulse) are based on data from the 1960's with special reference to launcher systems (turbojets using reheat and no turbofan data). A typical value for the specific impulse of a turbofan is 4500 s. With a L/D ratio of 4 at take-off, a cruise Mach number of 2 and an average acceleration of 0.25, this yields a ratio of  $W_1/W_0 = 0.973$ .

Substitution of equation (359) in equation (358) yields after some rewriting:

$$W_{1_{\text{est}}} = W_0 e^{-\frac{M_\infty a_\infty}{I_p g} \left( 1 + \frac{1}{\left( \frac{T}{W} \right)_0 \left( \frac{L}{D} \right)_0^{-1}} \right)} \quad (361)$$

For Concorde the following values apply ([ref.2]):

$$\left( \frac{T}{W} \right)_0 = 0.33$$

$$\left( \frac{L}{D} \right)_0 = 4.0$$

yielding  $n_0 = 0.08$

The average specific impulse for Concorde equals about 1800 s. (Concorde uses reheat during take-off and transition).

Substituting these values in equation (361) yields:  $W_{1_{\text{est}}} = 0.86 W_0$ .

The fuel reserve amounts to 16% of the total fuel weight which implies 19% of the mission fuel. With a maximum zero fuel weight of 50% of the maximum take-off weight (*Jane's All the World's Aircraft Directory*) it can be established that 33,3% of the mission fuel is burned during the accelerated climb to cruise speed and altitude. The actual value for Concorde is 37.7%.

With the estimated initial cruise weight and the cruise Mach number, together with the fuselage cruise angle of attack, the estimated cruise altitude follows from the aerodynamics routine.

At each level of energy height, the speed is varied from the optimal value of the previous step to the speed corresponding to the optimal altitude of the previous step, thus ensuring an absolute ascending and accelerating path. A speed interval around the transonic regime is left out of consideration, but will be treated at a later stage of

the procedure (it is very improbable that the optimal point would lie inside this region). The reason for this is that it is not possible to obtain an accurate drag prediction for Mach numbers roughly between 0.9 and 1.4. Only the peak value of the zero-lift drag and a very rough prediction of the lift-dependent drag can be obtained. This implies, that once the Mach number has reached the maximum allowed subsonic value (of 0.9), it will be kept constant, resulting in the troposphere to decreasing speed with increasing altitude. This has been done in order to extend the subsonic branch to larger values of the energy height. Of course, no optimization along these values of energy height is possible since the speed can no longer be increased.

The maximum attainable speed is the cruise speed. It was decided not to take the cruise Mach number as a constraint, since this could result in a speed reduction once this limit would be reached in the troposphere. Particularly at supersonic speeds, this could result in large increases in altitude from one energy height step to the next one. Since it is assumed throughout the mission analysis that the values for lift and drag as well as thrust remain constant between the different steps in energy height, too large altitude changes cannot be allowed. Prohibiting deceleration during the supersonic part of the mission limits the maximum possible increase in altitude to the difference between the cruise energy height and the energy height after take-off divided by the number of steps of the analysis. The number of steps was chosen such that for the presently studied types of airplanes this value does not exceed 2 km.

Finding in this fashion optimal points for each energy height, a series of speed-altitude combinations is obtained which form an approximation to the optimal path, see Figure 64.

---

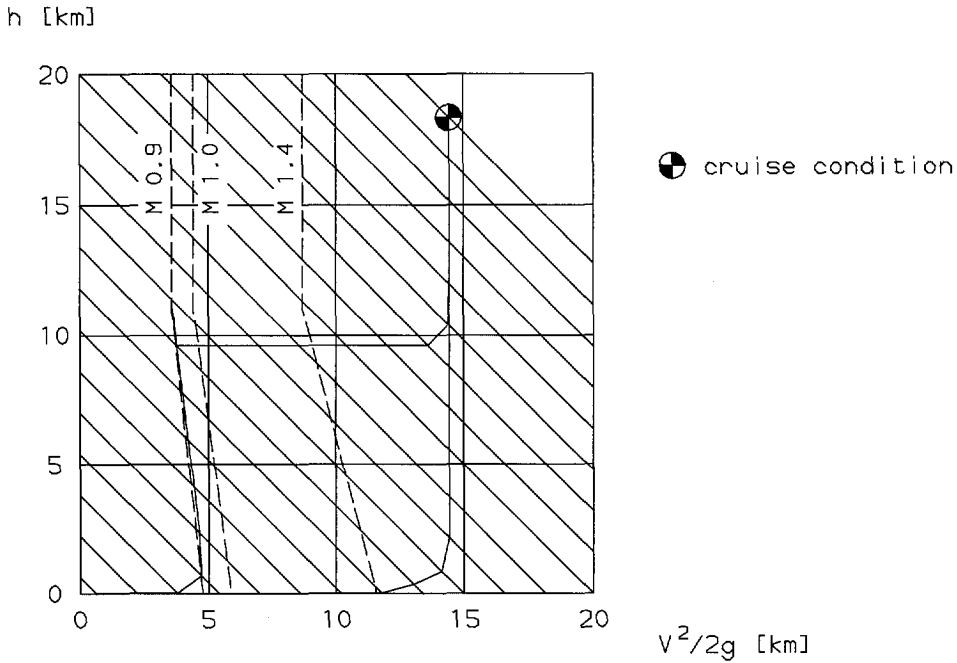


Figure 64 Mission analysis

A safety margin must be included to constrain the climb angle during the accelerated climb to cruise conditions. The climb angle at a certain point of the analysis is related to the velocity and specific excess power in that point and the increase in velocity and/or altitude to attain the next point of the analysis:

$$\sin \gamma = \frac{dh}{V dt} = \frac{SEP dh}{V dh_e} \quad (362)$$

which follows after substitution of

$$\frac{dh}{dt} = \frac{SEP}{\left(1 + \frac{V}{g} \frac{dV}{dh}\right)} = \frac{SEP}{\left(\frac{dh_e}{dh}\right)} \quad (363)$$

Since, in the outlined procedure, the mission analysis is carried out for a discrete

number of energy height values, the derivative  $\left(\frac{dh_e}{dh}\right)$  is not known (only  $\left(\frac{\Delta h_e}{\Delta h}\right)$ ). Furthermore, since it is not known during the determination of the optimal ascent path what the increase in velocity and/or altitude will be, variables like  $dv$  and  $dh$  must vanish from the expression.

It follows easily that

$$\frac{\Delta h_e}{\Delta h} = 1 + \frac{1}{2g} \frac{\Delta(V^2)}{\Delta h} \quad (364)$$

With  $\Delta(V^2) = \Delta V(V_1 + V_2)$  and  $\frac{dh_e}{dh} = 1 + \frac{V}{g} \frac{\Delta V}{\Delta h}$  this yields after some rewriting:

$$\frac{dh_e}{dh} = \frac{\Delta h_e}{\Delta h} - \frac{(\Delta V)^2}{2g\Delta h} \quad (365)$$

From the condition:

$$\frac{SEP}{V \left(\frac{dh_e}{dh}\right)} < \sin \gamma_{\max} \quad (366)$$

it follows:

$$\Delta h_{\max} = \frac{\sin \gamma_{\max} V_1}{SEP} \left( \Delta h_e - \frac{(\Delta V)^2}{2g} \right) \quad (367)$$

From the definition of the energy height it follows:

$$\Delta V = \sqrt{2g(h_{e_2} - h_2)} - V_1 \quad (368)$$

yielding:

$$h_{2\max} = h_1 + \frac{\sin \gamma_{\max} V_1}{SEP} \left( \Delta h_e - \frac{\left( \sqrt{2g(h_{e_2} - h_2)} - V_1 \right)^2}{2g} \right) \quad (369)$$

Inspection of the first and second derivatives of  $h_{2\max}$  with respect to  $h_2$  shows that with positive  $\Delta V$ ,  $h_{2\max}$  continuously increases with increasing  $h_2$  whereas the second derivative is always negative. To determine the maximum allowable altitude for the next point in the analysis,  $h_{2\max}$  is put equal to  $h_2$ . After rewriting the equation the following result is obtained:

$$h_{\max} = \frac{\sqrt{B^2 - 4AC} - B}{2A} \quad (370)$$

with:

$$A = g^2 (SEP - \sin \gamma_{\max} V_1)^2 \quad (371)$$

$$B = 2g \sin^2 \gamma_{\max} V_1^4 - 2fg (SEP - \sin \gamma_{\max} V_1) \quad (372)$$

$$C = f^2 - 2 \sin^2 \gamma_{\max} V_1^4 g h_{e_2} \quad (373)$$

and:

$$f = g SEP h_1 - g V_1 \left( h_1 + \frac{V_1^2}{g} \right) \sin \gamma_{\max} \quad (374)$$

In these expressions only the velocity, altitude and specific excess power at the current point of the analysis and the energy height at which the next point will be analyzed appear. Based on [ref.5], the maximum permitted angle of climb will be put at  $15^\circ$ .

At subsonic speeds, once the maximum allowed subsonic Mach number is attained, and at supersonic speeds, once the cruise speed is attained, a problem may occur, since the change in speed from the previous to the current value of energy height and consequently the change in altitude is prescribed. The only way in which the climb

angle can be constrained in this case is to adapt the throttle setting of the previous step (it is assumed that this will not influence the optimal point of the previous energy height):

$$\sin \gamma = \frac{\text{SEP} \left( \Delta h_e - \frac{\Delta(V^2)}{2g} \right)}{V \left( \Delta h_e - \frac{(\Delta V)^2}{2g} \right)} \quad (375)$$

It can be established easily, that once  $\Delta V$  equals zero (steady flight),

$$\sin \gamma = \frac{\text{SEP}}{V} = \frac{T-D}{W} \quad (376)$$

If the climb angle exceeds its constraint, the thrust should be reduced.

For supersonic aircraft, at those energy height levels which include supersonic speeds, two optimal points are found, resulting in a subsonic branch and a supersonic branch. At a certain energy height the supersonic optimum will be the best of the two. This will be the moment for transition from the subsonic branch to the supersonic one. Transition at constant energy height would in theory be optimal. It can be shown, that this corresponds to the minimum flight time solution and since it can be assumed that the exchange between potential and kinetic energy takes place in zero time, this would also imply a zero fuel consumption transition. In practice however, flying at constant energy height, requires the adjusting of the throttle setting to the varying amount of drag:

$$\frac{dh_e}{dt} = \frac{(T-D)V}{W} = C_s \quad (377)$$

Furthermore, it can be shown that the path corresponding to this kind of transition is a cycloid, which consists of a sharp dive and pull-up manoeuvre, which is unsuitable for a passenger aircraft. A solution would be, to keep the flight path angle constant during transition. This, however, yields the vertical dive as the optimal solution, which is even more impractical. Finally, the flight path angle could be constrained to a reasonable value. Considering the maximum permissible cabin floor slope and the maximum permissible rate of descent with respect to passenger comfort, only very small angles of descent would be realizable because of the high speeds involved. Therefore, it was decided to let transition take place at constant altitude, at the value corresponding to the point where the subsonic branch is left.

As these calculations are performed within the framework of an overall aircraft



optimization procedure, it should be checked whether transition can take place at the determined altitude at all. Based on a transonic peak drag estimation, carried out according to the USAF Stability and Control DATCOM procedure ([ref.3]), the maximum attainable transition altitude based on 95% maximum thrust is determined for the design under consideration. The induced drag coefficient  $k$ , cannot be determined by the DATCOM procedure. Instead a representative value of 0.25 is chosen, based on figure E5 of [ref.4]. Should the desired transition altitude be in excess of this maximum attainable transition altitude, then transition will take place at this maximum.

### Fuel trim

As explained before, the rearward shift of the aerodynamic center during transition to supersonic speeds poses a special design problem, as on the one hand the aircraft has to possess acceptable stability characteristics throughout the flight envelope, while on the other hand too large an increase in supersonic trim drag has to be avoided. An excellent compromise between these two contradicting demands, is to allow the center of gravity position to shift as well. This could be achieved in two ways: first of all a special trim tank could be used, usually positioned in the rear fuselage (BAe-Sud Aviation Concorde, Rockwell B1B), to and from which fuel can be transferred. It should still be decided however, what volume this tank should contain. An alternative arises from the fact that the wings of a supersonic transport airplane are usually strongly swept, which suggests that the use of a fuel trim technique, which simply determines the order in which the wing tanks are emptied, could just as well provide for a center of gravity shift.

In the present study, the fuel tanks are divided into one inboard and one outboard tank, and three different fuel trim techniques are considered. According to the first one, the fuel center of gravity stays at its place throughout the entire flight; according to the second one, the tanks are emptied from the inner tank outward, which implies a rearward shift of the fuel center of gravity, while the third technique is exactly the opposite. Applying either the second or the third technique, results in the fuel center of gravity staying at its place during the ascent, while it moves toward the center of gravity of the second tank to be emptied during the first stage of the cruise flight, where it stays during the rest of the flight, which continues until 10% of the total fuel weight is left. The optimal choice for a given engine design depends on the location of the tanks with respect to the aircraft center of gravity and on which constraints are active or critical (for instance, whether subsonic longitudinal stability is critical or supersonic directional stability). This also implies that different choices of the fuel trim technique result in different optimal designs. As the choice of the fuel trim technique is a discrete parameter, separate optimization studies have to be performed.

The most rearward center of gravity location on the ground is used to determine the main undercarriage position. In most cases this center of gravity position will occur for the totally fuelled aircraft without passengers (like after an emergency evacuation prior to take-off). With the determined undercarriage position and an estimated value for the lift-off rotation angle, the undercarriage length necessary to prevent tailsrape

---

at rotation is estimated. From this value, the minimum spanwise distance of the engine nacelles from the fuselage is determined, a value that is important in the calculation of the ceiling after engine failure.

### 13.5 References

1. Bos, A.H.W. and Oving, B.A.; *Conceptual Design of a Hypersonic Cruising Transport Plane* (in Dutch); Master's thesis, Delft University of Technology, Faculty of Aerospace Engineering, December 1991
  2. Calder, P.H. and Gupta, P.C.; *Future SST Engines with Particular Reference to Olympus 593 Evolution and Concorde Experience*; Aeronautical Journal, June 1976, pp.235-252
  3. Hoak, D.E.; *USAF Stability and Control DATCOM*; 2<sup>nd</sup> ed., 1968
  4. Nicolai, L.M.; *Fundamentals of Aircraft Design*; University of Dayton, Ohio, 1975
  5. Torenbeek, E.; *Synthesis of Subsonic Airplane Design*; Delft University Press, 1988
  6. Wittenberg, H.; *Range Performance of Hypervelocity Vehicles*; De Ingenieur, Vol.81, Nr.43, October 1969
  7. Wittenberg, H.; *Advanced Topics in Aircraft Performance Estimation and Optimization*; Report LR-9A, Delft University of Technology, Faculty of Aerospace Engineering, June 1986
-

# 14. The Analysis Program

## 14.1 Summary

The purpose of the analysis program is to calculate the values of the objective function and the constraints as a function of the design variables. In this context, it can be regarded as a black-box with the design variables as input and the objective function and constraints as output, and this is exactly as it will be treated by the optimization program, regardless of the optimization method used. The programming language is FORTRAN 77.

The program consists of a number of subroutines, which perform calculations in the field of the major disciplines involved, such as mass prediction, aerodynamics, propulsion, mission and performance, although no actual decomposition in the sense of chapter 3, is performed. The routines will be called sequentially and no efforts to facilitate parallel computation were undertaken.

## 14.2 Notation

$A$	airfoil section area
$a$	speed of sound
$a_1$	notch ratio
$a_2$	leading-edge kink location parameter
$a_3$	leading-edge kink location parameter
$C_L$	lift coefficient
$C_M$	pitching moment coefficient
$C_N$	yawing moment coefficient
$C_Y$	sideforce coefficient
$C_{M_0}$	zero-lift pitching moment coefficient
$c$	chord
$\bar{c}$	mean aerodynamic chord
$D$	drag
$d$	diameter
$f$	airfoil section area fill ratio
$I_{sp}$	specific impulse
$i$	incidence
$L$	lift
$l$	length
$l_h$	tail arm
$M$	Mach number
$N$	number
$R$	range
$r$	nose radius
$S$	gross area
$s$	semi span

---

	height of wing section
$t$	thickness
$V$	volume
$W$	weight
$x$	lengthwise coordinate (from fuselage nose unless indicated otherwise)
$x_{wf}$	aerodynamic center position of wing-fuselage combination
$y$	coordinate
$\bar{y}_{fuel}$	fuel spanwise center of gravity position
$\alpha$	angle of attack
$\beta$	sideslip angle
$\delta$	stabilizer deflection
$\Lambda$	sweep angle
$\lambda$	taper ratio

### subscripts

1	wing segment inboard of kink at beginning of cruise flight
2	wing segment outboard of kink at end of cruise flight
c	cylindrical section
cg	center of gravity
cr	cruise
eff	effective
f	fuselage
fuel	fuel
g	mean geometric
h	horizontal tailplane
le	leading-edge
n	nose section
np	neutral point
pax	passengers
r	root
t	tip
tank	fuel tank
te	trailing-edge
v	vertical tailplane
w	wing
wf	wing-fuselage combination

### abbreviations

HSAS	Hard Stability Augmentation System
SUMT	Sequential Unconstrained Minimization Techniques

### 14.3     Design Variables and Constraints

The airplane configuration to be optimized will be defined as a wing-body-tailplane combination and a separate engine design. This implies that no powerplant-airframe interference effects will be considered. The powerplant drag -as explained in chapter 11- will be calculated independently.

The fuselage is described by only two variables: the fuselage length,  $l_f$ , and diameter  $d_f$ . The fuselage is defined as a cylindrical section with equal parabolic nose and tail sections. The fuselage volume is kept constant for a given number of passengers according to:  $V=1.65 N_{\text{pax}} \text{ m}^3$ . This volume requirement includes space for toilets, pantries and flight deck, as well as construction space. It is assumed that all fuel is carried in the wing. Since the total volume of the defined fuselage is equal to:

$$V = \frac{\pi d_f^2}{60} (16l_n + 15l_c) \quad (378)$$

it can easily be deduced that the length of the nose (and tail) section must equal:

$$l_n = \frac{1}{14} \left( 15l_f - \frac{99N_{\text{pax}}}{\pi d_f^2} \right) \quad (379)$$

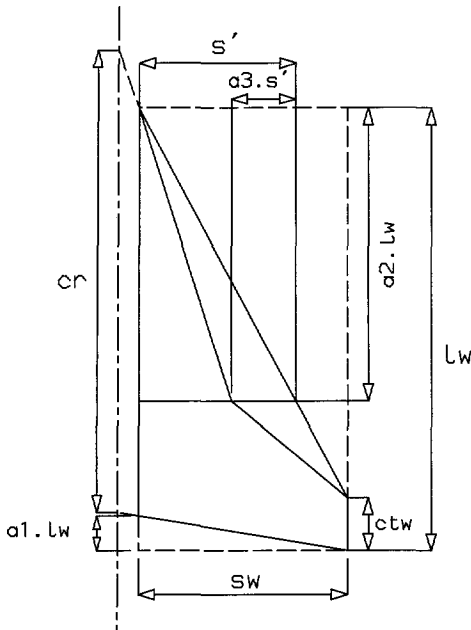
$$l_c = l_f - 2l_n \quad (380)$$

The wing geometry -as illustrated in Figure 65- is defined by the following variables:

- $l_w$     length of net wing planform box
- $s_w$     semi span of net wing planform box
- $c_{t_w}$     tip chord
- $x_{le_w}$     longitudinal location of net wing leading-edge
- $a_1$     notch ratio
- $a_2$     kink ratio
- $a_3$     kink ratio
- $i_w$     root incidence
- $t/c$     thickness-to-chord ratio
- $r/c$     leading-edge radius

The wing is uncambered, with a constant spanwise thickness-to-chord ratio and nose radius. The airfoil sections are symmetrical and defined as:

$$y = k_1\sqrt{x} + k_2x + k_3x\sqrt{x} + k_4x^2 \quad (381)$$



**Figure 65** Wing geometry definition

The coefficients  $k_1$ ,  $k_2$ ,  $k_3$  and  $k_4$  are given by the following equations as a function of the  $t/c$  and  $r/c$  of the wing:

$$k_1 = \sqrt{2r}$$

$$k_2 = 2 \left( \frac{t}{c} \right) - (1 + 2\sqrt{2}) k_1 \frac{\sqrt{c}}{c}$$

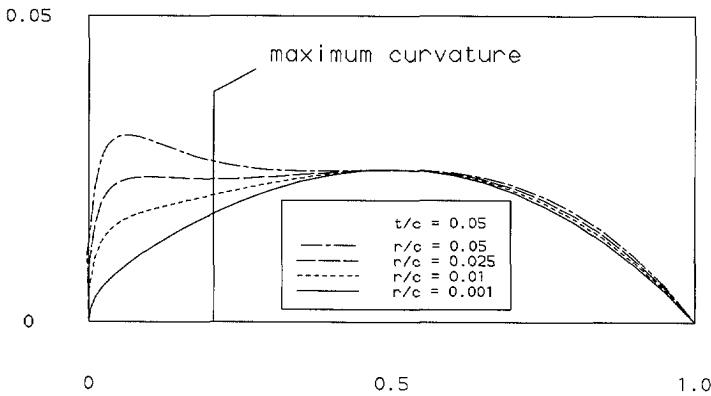
$$k_3 = 2(1 + \sqrt{2}) \frac{k_1}{c}$$

$$k_4 = \frac{-2}{c^2} (k_1\sqrt{c} + t)$$

The maximum curvature for these airfoil sections occurs at  $x = \frac{c}{2(1+\sqrt{2})} \approx 0.207c$ .

Substituting this value in the expression for the second derivative of  $y$  yields the following condition to prevent the maximum value of the second derivative to exceed zero, that is, to prevent the occurrence of inflection points:

$$\frac{r}{c} \leq 4.698 \left( \frac{t}{c} \right)^2, \text{ see also Figure 66.}$$



**Figure 66** Influence of nose radius on profile shapes

This condition is implemented as a geometrical constraint in the optimization.

The fuel volume of a tank segment (both wing halves) is calculated by means of the following equation:

$$V_{\text{fuel}} = \frac{2}{3} f l_{\text{tank}} (4c_g^2 - c_r c_t) \quad (382)$$

in which  $c_r$ ,  $c_t$  and  $c_g$  are the root, tip and mean geometric chord of the tank segment respectively. The length of the tank segment,  $l_{\text{tank}}$ , is measured perpendicular between the tank root and tip chord. The factor  $f$  is the fill ratio of the airfoil section area, defined according to the following relation:  $A_{\text{eff}} = 0.75 f c^2$  (it is assumed that three

quarters of this area can be used effectively as tank area), which yields for  $f$ :

$$f = 0.75 \left( \frac{2}{3} \left( \frac{t}{c} \right) + \frac{1}{5} (3\sqrt{2} - 4) \sqrt{\left( \frac{r}{c} \right)} \right) \quad (383)$$

The fuel spanwise center of gravity position (as needed by the mass prediction method in order to calculate the bending moment relief due to the fuel weight) is given by:

$$\bar{y}_{\text{fuel}} = \left( \frac{c_g^2 + \frac{1}{2} c_t^2}{4c_g^2 - c_r c_t} \right) l_{\text{tank}} \quad (384)$$

The horizontal tailplane as well as the vertical tailplane are described by four variables:

$l_h, l_v$	length of planform box
$s_h, s_v$	semi span of planform box
$c_{t_h}, c_{t_v}$	tip chord
$x_{l_{e_h}}, x_{l_{e_v}}$	longitudinal position of tailplane

The taper ratio of both the horizontal and vertical tailplane are kept constant at 0.5.

Further design variables are the cruise Mach number and angle of attack, as well as the design Mach number and altitude of the powerplant installation and the design compressor pressure ratio, the design bypass ratio, the intake capture area and the take-off bypass ratio parameter. This last mentioned variable is a measure of the amount of cycle variability necessary to comply with the take-off noise requirements. Thus, the total number of independent design variables amounts 28.

The constraints imposed on the optimization problem can be divided into two categories: the geometrical constraints and the performance constraints. The geometrical constraints are to ensure that no unrealistic values for the design variables are obtained during the optimization process and that no values are obtained for which the analysis methods used are not valid. Furthermore, they take care that no geometrically unrealistic designs are generated which are of no practical use in the first place, but which can also cause serious errors during the analysis calculations.

The way in which these constraints are treated depends on the optimization method employed. In case a gradient method is used, the geometrical constraints should be treated in the same way as the performance constraints. It should be realized that some gradient methods, notably the gradient projection methods -like for instance



sequential quadratic programming- or methods using augmented objective functions (penalty functions) like Sequential Unconstrained Minimization Techniques (SUMT) if they work from the infeasible region and direct search methods like Monte Carlo or genetic algorithms, have no protection against violations of constraints. With gradient methods these violations only occur on nonlinear constraints which form a convex set (in this case the tangent to the constraint points into the infeasible region) and they are normally small (or can be kept small by imposing move limits, see also chapter 2). It might be necessary to reduce the constraint tolerance or create a sufficient safety margin on those constraints that will cause errors during the further analysis of the design, should they be violated.

In the case of methods using penalty functions, a different approach should be followed. SUMT methods working from the infeasible region are useless and methods for which it is uncertain whether they perform from the interior or exterior such as the Augmented Lagrangian Method (chapter 2), should be avoided. Optimization methods that positively act as interior methods will not pose any problems. Direct search methods obtain their power from the fact that they scan the complete design space, implying that many infeasible designs will be considered, whereas these infeasibilities can be quite large. In these cases it is necessary to break off the analysis once a single infeasible geometrical constraint is encountered and to set the augmented objective function at a certain large value. This value should be chosen wisely in relation to the penalty imposed on a violated performance constraint, since too large a value will cause premature convergence in case of a genetic algorithm, since the fitness of an occasional feasible design will be much larger than the fitness of the majority of the infeasible designs, which will subsequently be deemed unfit for reproduction. On the other hand, the value should not be too small either, or the optimizer will favor a design with violated geometrical constraints to a design with slightly violated performance constraints.

If the initial population contains many designs that do not comply with the geometrical constraints, this procedure should not be employed. In case of a minimization problem, a genetic optimization method will compare the fitness value of the designs to the largest value encountered in a specified number of generations. With the previously outlined procedure, the largest value encountered will be the value for a geometrically infeasible design. If a large number of the designs is geometrically infeasible, their fitness value will be taken as zero and the probability that any of these designs will be retained in subsequent generations will be zero as well. This will result in very rapid convergence towards the few designs in the population that are geometrically feasible, but by no means optimal. This kind of situation might especially occur in case the feasible design space is extremely small. Random generation of the initial population then might result in a very large number of geometrically infeasible designs. Two remedies can be employed. First of all, instead of putting geometrically infeasible designs at a large value, they can be treated in the same way as the other constraints, that is, assigning them a value proportional to the violation, after which the further analysis is broken off. In this case only one design will have a feasibility of zero and premature convergence becomes less likely. The second way to cope with the problem is not to generate the initial population at

---

random, but to generate samples subject to the geometrical constraints separate of the genetic algorithm and to take these samples as the initial population. Both remedies are employed in this study.

In case the optimization is performed on an approximate analysis like response surfaces, the samples on which the approximations are based should comply with the geometrical constraints, which means that they have to feature in the sample generation as well. The geometrical bounds and constraints imposed on the design are summarized in the following tables.

variable			comments
$d_f$	lower bound	3.0 m	
	upper bound	$\left(2.0625 \frac{N_{\text{pax}}}{\pi}\right)^{1/3}$	upper and lower bound for $l_f$ equal
$l_f$	lower bound	$\frac{1}{15} \left( \frac{99 N_{\text{pax}}}{\pi d_f^2} + 42 d_f \right)$	nose and tail section slenderness > 3
	upper bound	$\frac{99 N_{\text{pax}}}{8 \pi d_f^2}$	cylindrical section > 0 m
		100 m	
$s_w$	upper bound	$35 - \frac{1}{2} d_f$	span < 70 m
$c_{t_w}$	lower bound	1 m	
	upper bound	$\frac{1}{2} l_w$	$\Lambda_{le_2} > \Lambda_{te}$
$a_3$	lower bound	0	
	upper bound	$1 - \frac{l_w - c_{t_w}}{5.7 s_w}$	$\Lambda_{le_1} < 80^\circ$
$a_1$	lower bound	$1 - \frac{l_f - \frac{1}{7} \left( 15 l_f - \frac{99 N_{\text{pax}}}{\pi d_f^2} \right)}{l_w}$	root chord at wing-fuselage intersection smaller than cylindrical section
		$-\frac{1}{2}$	
	upper bound	$\frac{1}{2}$	
$a_2$	lower bound	0	
	upper bound	$1 - \frac{c_{t_w}}{l_w}$	$\Lambda_{le_2} > 0^\circ$
		$\frac{(c_{t_w} - l_w(1 - a_1))(l_w - c_{t_w})}{(c_{t_w} + l_w(a_1(1 - a_3) - 1))l_w}$	$\Lambda_{le_2} > \Lambda_{te}$

Table 34 Geometric constraints

variable			comments
$x_{le_w}$	lower bound	$\frac{1}{14} \left( 15l_f - \frac{99N_{pax}}{\pi d_f^2} \right)$	wing leading-edge intersects behind fuselage nose section
	upper bound	$l_f - \frac{1}{14} \left( 15l_f - \frac{99N_{pax}}{\pi d_f^2} \right) - l_w(1 - a_1)$	wing root at intersection within cylindrical section
		$l_f - l_w$	tips may not protrude behind fuselage
$i_w$	lower bound	0	
	upper bound	$\arcsin \left( \frac{d_f}{2c_{r_w}} \right) (*)$	wing must fit in lower half of fuselage
$t/c$	lower bound	0.005	
	upper bound	$\left( \frac{d_f}{2c_{r_w}} \right) (*)$	wing must fit in lower half of fuselage
$r/c$	lower bound	0	
	upper bound	$4.698 (t/c)^2$	no inflection points in airfoil contour
$x_{le_h}$	lower bound	$x_{le_w} + l_w(1 - a_1)$	stabilizer leading-edge cannot intersect with wing trailing-edge
		$x_{le_w} - l_h + c_{t_h} + l_w \left( 1 + a_1 \left( \frac{s_h}{s_w} - 1 \right) \right)$	

(\*)  $c_{r_w}$  is the wing root at the fuselage centerline:

$$c_{r_w} = l_w(1 - a_1) + \frac{d_f}{2s_w} \left( \frac{l_w - c_{t_w}}{(1 - a_3)} - a_1 l_w \right) \quad (385)$$

**Table 35** Geometric constraints

The vertical fin geometry is described in the same way as the horizontal stabilizer. With the exception of the lower bounds on the position of the leading-edge, the same bounds that apply to the horizontal stabilizer are valid for the vertical fin as well. An

extra lower bound is imposed to position the vertical fin behind the cylindrical fuselage section:

$$x_{lev} > \frac{99 N_{\text{pax}}}{14 \pi d_f^2} - \frac{l_f}{14} \quad (386)$$

Two extra constraints are imposed on the wing-fuselage geometry:

- 1) the height of the rotated wing root section for the upper and lower part of the airfoil section may not exceed one quarter of the fuselage diameter to prevent its cutting through the fuselage or cabin.

$s < 1/4 d_f$  with:

$$s = \sqrt{y^2 + \left(\frac{1}{2} c_{r_w} - x\right)^2} \sin \left( i_w + \arctan \left( \frac{y}{\frac{1}{2} c_{r_w} - x} \right) \right) \quad (387)$$

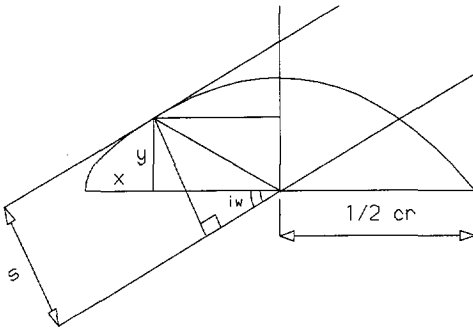
in which  $y$  and  $x$  are the coordinates of the point where the tangent to the contour equals the wing incidence (see Figure 67).

- 2) the cruise Mach number may not exceed the value for which the Mach angle is smaller than the fuselage nose angle:

$$M_{\alpha} < \sqrt{1 + \frac{\left(15 l_f - \frac{99 N_{\text{pax}}}{\pi d_f^2}\right)^2}{196 d_f^2}} \quad (388)$$

Some of the constraints summarized in the preceding tables are more important than others since some constraints actually exclude geometrically improbable designs, while other constraints are most important in case of sample generation to ensure that the designs are realistic.

The second category of constraints is that of the performance constraints, which are to ensure that the final optimized design complies with certain safety regulations and performs according to the specifications. In the present study, the following constraints were considered important for a second-generation supersonic transport aircraft:



**Figure 67** Determination of height of rotated root profile

- 1) The take-off field length to screen altitude should be comparable to current subsonic aircraft, that is, less than or equal to 3500 m.
- 2) The landing speed (with the use of vortex generator flaps) should not exceed the value of Concorde, being 83 m/s.
- 3) The aircraft should be able to serve on the trans-pacific routes, implying that the range with a supersonic cruise segment should at least be 5000 nautical miles, or over 9000 km.
- 4) Since it seems, at present, to be unavoidable that certain segments of the flight will have to be flown at subsonic speeds, the range with a fully subsonic cruise flight at Mach 0.9 should be at least 5000 km.
- 5) The available amount of installed thrust at cruise conditions should be sufficient to allow small manoeuvres, either in case of emergency or as directed by air traffic control, for which five percent excess thrust will be required.
- 6) The rate of climb at sea-level with all engines operating should be at least 25 m/s.
- 7) After engine failure, the aircraft should at least be able to reach a ceiling of 4600 m (defined by a rate of climb of 1 m/s).
- 8) The aircraft should have sufficient static longitudinal stability with stick fixed during the most critical subsonic operating condition. (The use of an artificial stability system may permit negative stability margins of more than 5%)
- 9) The aircraft should have sufficient static longitudinal stability with stick fixed during the most critical supersonic operating condition.
- 10) The aircraft should have sufficient static directional stability during the most critical supersonic operating condition.
- 11) The aircraft should still be able to accelerate after engine failure at screen altitude. This constraint is based on Concorde experience, since this type of aircraft tends to fly at the backside of the power curve during take-off, where speed instability occurs. Engine failure after take-off may lead to unstable deceleration unless sufficient power is available to accelerate.
- 12) Although it is expected that overland supersonic flights will be prohibited (constraint 4) there is indication that limits will be imposed on overseas

supersonic flight as well. Therefore, the maximum bow shock overpressure reaching the earth's surface should be less than 72 Pa (cruise flight; "corridor criterion").

- 13) With respect to ozone-layer depletion, the emissions of nitric oxides ( $\text{NO}_x$ ) during cruise flight are not allowed to exceed 20 g/kg fuel.
- 14) In order to comply with take-off noise regulations, the take-off jet velocity should not exceed 500 m/s.

Especially the environmental constraints are very tricky, since they tend to have a major impact on the design but on the other hand are very difficult to evaluate. Furthermore, since they are dictated by politics they are sensitive to changes in public opinion (sonic boom problem !). Also, the last constraint, regarding airport noise, which already is very difficult to comply with for supersonic aircraft might become more severe, since it is likely that regulations will follow trends in subsonic airplane technology. At present, take-off jet velocities of as little as 350 m/s are possible for subsonic engines.

The subroutines called by the analysis program are concerned with the following disciplines:

- geometry
- mass prediction
- aerodynamics
- propulsion
- take-off analysis
- mission analysis
- cruise analysis
- climb performance analysis
- stability analysis
- sonic boom prediction

The mass prediction model, the aerodynamics model, the propulsion model and the mission analysis were treated in chapters 8, 9, 11 and 13 respectively. The environmental constraints concerning emissions and take-off noise according to chapter 12 are implemented in the propulsion analysis. The sonic boom analysis according to the Carlson model was explained in chapter 12 as well. The other disciplines mentioned will be treated in the following paragraphs.

## 14.4     Geometry

In the geometry routine, a number of geometrical parameters needed for subsequent analysis will be calculated. Among these parameters are sweep angles, areas, tank volume, chords etc. Since these parameters depend only on the design variables and consequently do not change during the analysis of a specific design, and since these parameters will be needed in many different analysis subroutines, they are stored in common blocks.

---

## 14.5 Determination of Take-off Field Length

The take-off field length is calculated in three phases: the ground run phase, the rotation phase and the airborne phase. Each phase is divided into a number of time steps, and for each step the lift, drag and thrust will be calculated to obtain horizontal and vertical accelerations. From these values follows the runway length travelled to the next step.

The minimum speed at which rotation can take place is determined at 1.05 times the stall speed and depends therefore on the airplane weight and wing area. It is assumed that at this speed rotation of the airplane is indeed achievable (minimum handling speed). As highly swept wings and especially delta wings only stall at very high angles of attack, if at all, the maximum lift coefficient is based on a maximum permissible angle of attack (in relation to runway visibility). After completion of the take-off calculations the program optionally proceeds by calculating the take-off field length for a number of higher rotation speeds in order to find the optimum. However, since the take-off field length calculations require input from the aerodynamics and propulsion routines at every time step, the calculations are rather tedious and this option might as well be left out in order to save computational time (which is very important in multidisciplinary optimization).

During the rotation phase, the aircraft is rotated at a constant pitch rate until the normal force on the runway vanishes. After take-off, the aircraft is rotated further, at a halved pitch rate, until the required second segment climb gradient is attained. From thereon the rotation of the airplane stops and it climbs and accelerates further until the screen height is reached. At the end of the take-off calculations, the take-off field length, as well as the airplane weight and velocity at screen height are known, data which is fed into the mission analysis routine.

The landing speed is calculated for the same maximum permissible angle of attack as for take-off, only in this case the wing leading-edge radius is artificially set to zero, thus simulating the deployment of vortex generating devices. This landing speed is considered to be sufficient indication of the airplane's landing performance.

## 14.6 Cruise Analysis

The cruise calculations are based on Bréguet's range equation. In case fuel trim is applied, the cruise flight is divided into two segments, for each of which an average center of gravity location is computed. The aircraft will be trimmed for this center of gravity position and the corresponding weight and altitude (constant cruise Mach number) and the trimmed lift-to-drag ratio will be computed. The powerplant drag as calculated by the propulsion routine will be incorporated in this ratio whereas the engine throttle setting will be adjusted to the total amount of drag (Bréguet's range equation was derived under the assumption of horizontal equilibrium). It is assumed

---



that this will not affect the specific impulse of the engine. Instead of making a detailed analysis of the fuel required for holding and diversion purposes, it is assumed that 10% of the total fuel volume is kept in reserve. Then, the range can be calculated from the Bréguet equation:

$$R = M_{cr} a_{cr} I_{sp} \frac{L}{D} \ln \left( \frac{W_1}{W_2} \right) \quad (389)$$

in which the subscript 1 refers to the beginning and 2 to the end of the cruise flight.

The cruise routine further checks whether the available amount of cruise thrust at a rating of 95% Maximum Continuous Thrust is sufficient to accelerate the airplane.

## 14.7 Climb Performance Analysis

Climb performance is evaluated at two different altitudes, that is, at sea-level to verify the rate of climb with all engines operating and at 4600 m to check whether this altitude can still be attained after engine failure (one-engine-out ceiling). At each altitude, the angle-of-attack is increased from zero to the maximum allowable angle-of-attack (which is set at 12 degrees) and the corresponding Mach number is calculated from the lift coefficient, airplane weight and pressure. For the given Mach number and altitude, the thrust is calculated from the propulsion routine and the specific excess power is calculated. It is assumed that this parameter yields an acceptable indication for the rate of climb. The procedure is repeated until the maximum value for the specific excess power is found.

## 14.8 Stability Analysis

### 14.8.1 Longitudinal Stability

For supersonic transport planes, stability and control poses some extra problems not encountered in subsonic transport design. Due to the rearward shift of the center of pressure during supersonic transition, trim drag can become quite large, which obviously has an undesirable effect on the range parameter. In case this effect is counteracted by designing the tailplanes for the cruise condition, creating just a satisfactory stability margin while cruising, a stability problem will occur during deceleration to subsonic speeds (although it is somewhat alleviated by the fact that at increasing Mach numbers, the effectiveness of tailplanes diminishes, which implies that the tailplanes designed for the cruise condition are relatively large). There are some possible solutions to this problem, the least radical one with respect to aerodynamic design and performance being the use of active control systems. This may permit inherently unstable flight conditions with possible negative stability

---

margins of up to 10 percent (McDonnell-Douglas High Speed Civil Transport M=3.2 Study [ref.2]). The use of the so-called Hard Stability Augmentation System (HSAS) on the Boeing 2707-300 Supersonic Transport design enabled a rearward shift of the center of gravity by 5% by positioning the engines further to the rear. This new engine location was more favorable with respect to wing-powerplant interference and resulted in a significant improvement of the cruise lift-to-drag ratio. Furthermore, it was possible to reduce the length of the fuselage section in front of the wing by 150 inches ([ref.5]).

Another well-known solution, but with far more influence on airplane performance, is the transfer of fuel from and to a special trim tank, usually placed in the tail (BAe-Sud Aviation Concorde and Rockwell B1B).

Furthermore, a lot can be done by moving the wing and tailplanes or by applying a certain amount of camber to the wing. One way or another, it is clear that in order to optimize a supersonic transport design, it is of vital importance that stability and control aspects are taken into account, and therefore an estimation of the pitching moment coefficient and stability derivatives as a function of the design variables must be available.

The NLRAERO panel program used in this application calculates the pitching moment coefficient by integration of the pressure coefficient over the aircraft surface. As explained earlier, the suction force is corrected by the Carlson attainable thrust concept. It will be assumed that the influence of the suction force correction on the pitching moment coefficient is negligible. Furthermore it is assumed that the derivative of the vortex lift with respect to the angle of attack is negligible as well.

For the linearized pitching moment coefficient the following relation holds:

$$C_M = C_{M_0} + C_{L_{wf}} \frac{x_{cg} - x_{wf}}{\bar{c}} - C_{L_h} \frac{S_h}{S} \frac{l_h}{\bar{c}} \quad (390)$$

which can be written as:

$$C_M = C_{M_{\alpha_i=0}} + C_{L_{wf_a}} \alpha_i \frac{x_{cg} - x_{wf}}{\bar{c}} - C_{L_{h\alpha}} (\alpha_i + \delta) \frac{S_h}{S} \frac{l_h}{\bar{c}} \quad (391)$$

in which  $x_{wf}$  is the longitudinal position of the wing-fuselage aerodynamic center, while  $l_h$  is the longitudinal distance from the center of gravity to the horizontal stabilizer aerodynamic center.

For the longitudinal stability derivative with stick fixed it follows:

$$C_{M_\alpha} = C_{L_{wf_\alpha}} \frac{x_{cg} - x_{wf}}{\bar{c}} - C_{L_{h_\alpha}} \frac{S_h l_h}{S \bar{c}} = C_{L_\alpha} \frac{x_{cg} - x_{np}}{\bar{c}} \quad (392)$$

in which  $x_{np}$  is the longitudinal position of the stick fixed neutral point. From the panel method follows the pitching moment gradient without stabilizer:

$$C_{M_{wf_\alpha}} = C_{L_{wf_\alpha}} \frac{x_{cg} - x_{wf}}{\bar{c}} \quad (393)$$

as well as the pitching moment coefficient at zero fuselage angle of attack. The total configuration stability margin can now be calculated:

$$\frac{x_{cg} - x_{np}}{\bar{c}} = \frac{C_{M_{wf_\alpha}} - C_{L_{h_\alpha}} \frac{S_h l_h}{S \bar{c}}}{C_{L_{wf_\alpha}} + C_{L_{h_\alpha}} \frac{S_h}{S}} \quad (394)$$

The pitching moment coefficient can then be written as:

$$C_M = C_{M_{\alpha=0}} + \left( K_{P_{wf}} + C_{L_{h_\alpha}} \frac{S_h}{S} \right) \frac{(x_{cg} - x_{np})}{\bar{c}} \alpha_f - C_{L_{h_\alpha}} \frac{S_h l_h}{S \bar{c}} \delta \quad (395)$$

in which  $K_{P_{wf}}$  is the lift gradient as obtained from the panel method. The value for  $C_{M_{\alpha=0}}$  in equation (395) is the pitching moment at zero fuselage angle of attack as obtained from the panel method plus the pitching moment caused by the vortex lift. It is assumed that the vortex lift acts on the leading-edge of the mean aerodynamic chord.

For a given angle of attack and Mach number, the stability margin can be computed for a fixed stabilizer area. Furthermore, the deflection of the stabilizer necessary for pitching moment equilibrium can be determined. Then the total trimmed lift coefficient and the altitude follows (for a given airplane weight). At this point the lift coefficient has to be recalculated as there is a small influence of the Reynolds number. After the altitude corresponding to the given angle of attack, Mach number and airplane weight has converged, the trim drag can be calculated, as well as the friction drag coefficient. In this way, the total configuration lift and drag, including leading-edge suction, vortex lift and drag, wave drag, friction drag and trim drag can be

determined.

#### 14.8.2 Directional Stability

The design of the vertical tailplane often becomes critical for supersonic aircraft. Since the wing does not contribute to the directional stability, the destabilizing effect of the fuselage will have to be counteracted solely by the vertical tailplane. As the effectivity

of the tail decreases fast with increasing Mach number (Ackeret:  $C_{L_\alpha} \propto \frac{1}{\sqrt{M^2 - 1}}$ ) this

usually results in relatively large tailplanes. This, in turn will result in increased zero-lift drag throughout the flight envelope. Referring to [ref.1] which treats the conceptual design of a transport plane cruising at high supersonic Mach numbers, it was found that the decrease of the destabilizing effect of the fuselage while accelerating from  $M=4$  to  $M=5$  only amounted approximately 8%, while simultaneously the decrease of the stabilizing effect of the vertical tailplane amounted to more than 20%. This same trend, though not quite as dramatically as in this example, will occur at lower supersonic Mach numbers too. As the tailplane design must also be such that it is able to counteract the yawing moment which occurs after engine failure -without excessive additional trim drag- this also poses a condition on the engine design. The advantage of multidisciplinary design is, that the critical design condition of each component shows immediately in the shape of a violated constraint, causing corrective action to be taken by the optimization routine in the next iteration.

The calculation of the directional stability margin is carried out in much the same way as the longitudinal stability margin, with the exception that the wing does not contribute and that no rudder deflections are considered (in supersonic cruise at zero side-slip angle there are no trim deflections necessary). The aerodynamic coefficients of the fuselage were in this case calculated with the use of the ESDU Data sheets ([ref.3]) instead of generating extra results with the panel method.

The directional stability margin can be expressed as follows:

$$\frac{x_{cg} - x_{np}}{\bar{c}} = \frac{C_{N_{f\beta}} - C_{Y_{v\beta}} \frac{S_v l_v}{S \bar{c}}}{C_{Y_{f\beta}} + C_{Y_{v\beta}} \frac{S_v}{S}} \quad (396)$$

in which the vertical tailplane lift gradient is determined in the same way as for the horizontal stabilizer, that is, with the use of the USAF Stability and Control DATCOM ([ref.4]).

## 14.9     References

1.     Bos, A.H.W. and Oving, B.A.; *Conceptual Design of a Hypersonic Cruising Transport Plane* (in Dutch); Master's thesis, Delft University of Technology, Faculty of Aerospace Engineering, December 1991
  2.     Douglas Aircraft Company; *Study of High-Speed Civil Transports*; NASA CR 4235, December 1989
  3.     ESDU Engineering Sciences DATA Unit; Vol.3c, No.89008, April 1989
  4.     Hoak, D.E.; *USAF Stability and Control DATCOM*; 2<sup>nd</sup> ed., 1986
  5.     Kehrer, W.T.; *Design Evolution of the Boeing 2707-300 Supersonic Transport*; AGARD CP 147 Vol.1, October 1973
-



# 15. The Optimization Procedure

## 15.1 Summary

This chapter outlines the method of optimization used, based on experience and results described in the preceding chapters. Several problems were encountered and their nature as well as the way in which they were solved are treated. Some of these problems are typically related to the method of optimization, especially with respect to the genetic optimization procedure, others are related to the approximate analysis or the analysis method itself. The problems related to the optimization procedure could all be solved or bypassed and it is shown that the method works and will produce (semi-)optimal designs, without the need to provide a baseline design first. Therefore, it is shown that the optimization procedure can actually be used to generate a conceptual (airplane) design from scratch, in a multivariate and multi-disciplinary way, instead of only adapting a baseline, designed according to traditional design methods.

## 15.2 Notation

$a$	speed of sound
$C_L$	lift coefficient
$C_M$	pitching moment coefficient
$C_N$	yawing moment coefficient
$D$	drag
$g$	gravitational acceleration
$I_{sp}$	specific impulse
$K_p$	lift gradient in potential flow
$L$	lift
$M$	Mach number
$R$	range
$T$	thrust
$W$	weight
$W_0$	maximum take-off weight
$\alpha$	angle of attack
$\beta$	sideslip angle
$\eta$	efficiency

### subscripts

0	at take-off, immediately after take-off
1	at beginning of cruise flight
cr	cruise
est	estimated
f	fuselage

---

tot	total
w	wing
wf	wing-fuselage combination

### abbreviations

MCC<sup>2</sup> Multiple Correlation Coefficient Squared

In accordance with tradition, weights are given in kilograms, although -according to the SI system- the unit of weight is Newton and the unit of mass is kilogram.

## 15.3 Description of Optimization Procedure

Based on experiences and conclusions presented in earlier chapters, it was decided to employ a hybrid optimization method consisting of a genetic algorithm in combination with a gradient method applied to approximate analysis. The main reason for this procedure is the combination of the favorable characteristics of both methods, that is, the robustness of the genetic algorithm (high probability to locate the global optimum) and the accuracy and efficiency of the gradient method.

Another reason to employ the genetic search algorithm is, that it can be helpful to provide initial designs that comply with the performance constraints. Since the feasible design space is extremely small for the present specification (even if few constraints are enforced) and the number of interdisciplinary couplings is rather large, creating a feasible design by means of classical design procedures is a very hard if not impossible task. By defining a penalty function that sums the squares of violated constraints, the genetic algorithm may be used to find one or more designs that comply with the performance constraints. These designs may then be used as baseline designs for the gradient-based optimization.

Some initial experience with genetic optimization has shown that the method is not as simple and successful as is sometimes suggested. The best values found by the algorithm strongly depend on the objective function (penalty function) chosen, especially in relation to the penalty imposed on unrealistic or geometrically infeasible designs. Premature convergence continues to be a problem, even if the ranking procedure for reproduction is applied instead of proportional selection. Furthermore, convergence tends to be extremely slow and is generally achieved by a very small change to the best design encountered, which in most cases has been produced in an early generation and kept in the population during a large number of generations since. In these cases, the population consists of a large number of very bad designs and a single acceptable one, which is occasionally improved after several thousand function evaluations.

Another disadvantage of genetic algorithms turned out to be the fact that the choice of the optimization parameters (like the crossover probability and the mutation probability), as well as the selection procedure employed, is purely intuitive. No rules are available that can guide the user in these choices. Therefore the process of a

---



successful genetic search involves a great deal of trial-and-error. The only exception is, that in case after many function evaluations and restarts with new initial populations, the search has converged and no further improvement can be obtained, it sometimes helps to put the mutation probability at a larger value. If this value is taken too high, however, the search will degenerate into a basic random search which will not result in better designs in case of a strongly constrained problem.

In case a design can not be evaluated completely, for instance because negative accelerations are calculated at some point of the analysis, the objective function is set at a very large value and the calculations are terminated. It was established that this procedure should be applied with great care since it can make progress very slow or even stop it altogether. It was found that in many occasions, the final population consists almost entirely of incomputable designs and an occasional feasible design with a relatively low, but still far from satisfactory value for the objective function. To reduce this problem, the value of the penalty function should as much as possible be proportional to the amount of constraint violation and if possible the analysis should be continued. In some cases it is necessary to define additional constraints to enable the evaluation of unusable designs. As an example, in case during the rotation of the aircraft in the take-off run the available thrust is not sufficient to counteract the drag increase, the take-off field length cannot be evaluated. Instead, it is put at a large number plus the distance covered until the acceleration becomes negative. The amount of fuel burnt during take-off is put at zero and the speed after take-off is put at a representative value. An extra penalty function proportional to the negative acceleration is then added to the objective function. This method of proportional penalty will enable the optimization algorithm to "learn" and take corrective action, whereas information about any favorable characteristics of the design under consideration is retained. It was established that this procedure drastically improved progress of the optimization.

Any optimization algorithm, but especially the genetic search algorithm, is an excellent debugging device. The algorithm will take advantage of any imperfection in the analysis method, by guiding the design towards regions of the design space where the procedures used are no longer valid or where certain parameters attain unrealistic values. This renders the development of an analysis program a very tedious task, since preferably for each dependent variable an interval should be defined within which its value is considered realistic, a virtually impossible task, especially since this would implicitly constrain the design space, which is against the purpose of multidisciplinary optimization. Furthermore, such intervals should be based on past experience which is hardly present for a supersonic transport in the first place. In practice, each time the optimization program has converged toward an unrealistic design, a remedy must be devised and the procedure restarted.

## 15.4     Approximate Analysis

As indicated previously, the main improvement in multivariate optimization is to be obtained by limiting the calculation time required for the design analysis, especially in

---

case direct search methods are used. In the present study, the three most time-consuming analyses are aerodynamics, supersonic propulsion and mass prediction. The aerodynamic parameters (obtained from a panel method operated independently of the analysis program) are approximated by regression surfaces as explained in chapter 10. It was established, that the propulsion analysis requires quite a lot of computational time for supersonic off-design Mach numbers. This is mainly caused by the more complex supersonic intake analysis and the extra iteration loop to optimize the amount of spillage and bleed air. To reduce computational time, regression surfaces were generated for the net thrust, specific impulse and powerplant drag as a function of the engine design variables. The supersonic propulsion analysis is only used in the evaluation of the cruise performance and the supersonic branch of the mission analysis and using the approximate instead of the exact propulsion analysis reduced the required CPU time for the cruise analysis by a factor three to four. Using a simple equation instead of the detailed mission analysis -as explained further on- this reduced the total CPU time by about 50%. For subsonic off-design Mach numbers, the propulsion analysis is so fast in comparison that the exact analysis is retained.

It was further established, that while using approximate supersonic propulsion analysis, the mass prediction took about one third of the total CPU time (again using a simple equation instead of the complete mission analysis) and about one sixth in case the exact supersonic propulsion analysis was used. Of course, this is caused by the double iteration toward maximum take-off weight and wing weight ratio. As explained in chapter 8, these iterations did not improve the accuracy of the prediction very much and were therefore removed from the procedure.

The main task of the mission analysis is to calculate the amount of fuel burnt during the accelerated climb to cruise speed and altitude. Since it takes relatively much computational time, the mission analysis was replaced by a simple relation from [ref.6], in order to obtain some initial designs. This same relation is also used in the mass prediction to estimate the weight at the beginning of the cruise flight in order to obtain an indication for the cruise altitude (see chapter 13):

$$W_{1_{est}} = W_0 e^{-\frac{M_{cr} a_{cr}}{I_{sp} g} \left( 1 + \frac{1}{\left( \frac{T}{W} \right)_0 \left( \frac{L}{D} \right)_0^{-1}} \right)} \quad (397)$$

A complete design evaluation, using -as explained- a mass prediction without iteration, approximate supersonic propulsion analysis and a simplified mission analysis, takes between about 60 and 100 CPU seconds on a *SUN SPARCstation IPC* (33 MHz). On the *Convex 3840* near-supercomputer this is about 2 to 4 times less, but in this case the cost of a genetic search (in which several thousand function evaluations have to be carried out) will become prohibitively large, rendering the gain in computational speed irrelevant (especially since calculations are carried out during the night).

## 15.5     Optimization Results

### 15.5.1     Genetic Optimization

A set of 50 designs, that comply with the geometric constraints, was generated randomly and used as an initial population in the genetic optimization. The objective function to be optimized was chosen as the sum of the squares of the violated constraints. From the very start, severe convergence problems occurred because of the large number of incomputable designs generated and the large constraint violations of the occasionally generated computable designs. Most causes of these problems were identified -as indicated earlier in this chapter- and treated accordingly. Large progress was made by imposing only one constraint at first and subsequently adding more, using the best designs obtained as the initial population of the next optimization run. Since it was established that a conflict existed between the supersonic range constraint and the take-off field length constraint it was decided first to impose only the supersonic range constraint (which could be easily satisfied) and then to add the take-off field length constraint. In a later stage the other constraints were added. Throughout the optimization it became clear that the sea-level climb constraint and the one-engine-out constraint were not at all critical. During the gradient-guided search these constraints therefore were not taken into account, although they were still monitored.

As the optimization procedure progressed and the optimal designs converged to a certain area of the design space, it became apparent that the critical constraints now had become the approach speed, the supersonic range and the subsonic stability margin, with the emphasis on the latter two. It was possible to satisfy either of these in combination with the other (less critical) constraints but not both at the same time.

It turned out to be virtually impossible to comply even with all normal performance constraints, let alone with the environmental constraints as well. Therefore it was decided not to take the environmental constraints into account for the time being. Only the subsonic range constraint (which is an implicit sonic boom constraint) was retained. Throughout the further optimization process (including the gradient-based optimization) the  $\text{NO}_x$  emissions constraint did not present any problems. The take-off exit jet velocity constraint and especially the sonic boom (peak pressure) constraint could not be satisfied, however.

Since the largest gain in objective function reduction is obtained in the first generations, whereas progress usually comes close to a halt as the optimization proceeds, there is no point in carrying out an excessive amount of function evaluations. A very effective procedure to speed things up proved to be using the best designs, obtained at the point where progress stagnated, as the initial population for a new optimization run. The rest of the initial population was generated at random, thus providing for "new blood" in the population. This method proved more effective than increasing the mutation probability.

---

After several thousand function evaluations, no dramatic changes to the design occurred anymore. At that point it was decided to resort to some fine-tuning by using the exact supersonic propulsion analysis. Since the mission analysis was still approximated by equation (397), the supersonic propulsion analysis was used only for the cruise calculations.

### **Result with Exact Propulsion and Simple Mission Analysis**

The best design obtained in this way had violated approach speed, supersonic range and subsonic stability margin constraints. The violations were of the following nature; an approach speed of 85.5 m/s instead of 83.0 m/s, a supersonic range of 8462 km instead of 9000 km and a subsonic stability margin of -9.9% instead of the permitted -5%. The maximum take-off weight was 349563 kg.

With the semi-optimal engine design variables thus obtained, a file was created in which the net thrust, specific impulse and powerplant drag were gathered for a large number of Mach number-altitude combinations. In this way the propulsion analysis could be replaced by a simple interpolation routine, thus speeding up the mission analysis considerably.

### **Evaluation of Optimum with Elaborate Mission Analysis and Exact Propulsion Analysis by Interpolation**

The elaborate mission analysis showed a transition altitude of 8 km and a weight at the beginning of the cruise flight of 95% of the maximum take-off weight. The supersonic range reduced to 8329 km, the other two critical constraints were not affected.

Finally, a new design space was defined, based on a  $\pm 10\%$  change of the semi-optimal design variables. Within this design space, new regression surfaces were derived for the supersonic aerodynamic parameters. The polynomials for the total configuration lift gradient, the wing lift gradient and the pitching moment gradient showed increased accuracy, but those of the lift coefficient and pitching moment coefficient at two degrees angle of attack were less accurate than the old ones. Therefore only the first three new polynomials were used.

### **Evaluation of Optimum with New Aerodynamic Approximations**

With this new approximate aerodynamic analysis the critical performance parameters changed to an approach speed of 83.5 m/s, a supersonic range of 7920 km and a subsonic stability margin of -10.4%. The maximum take-off weight changed to 349333 kg. This design will hereafter be referred to as **baseline** design.

There are three peculiarities concerned with the present design. First of all, the fuselage diameter has attained its lower limit of 3 m which is quite small, but sufficient for 4- or 5-abreast seating. It is very likely that the algorithm tries to minimize fuselage wave drag by choosing a large fuselage slenderness. It is quite

---

probable that the ensuing weight increase is not modelled adequately, since the fuselage slenderness ratio does not feature as such in the weight equation (chapter 8). However, increasing the fuselage diameter will further deteriorate performance, which already does not comply with the specifications.

The second problem is closely related to the first problem. Despite what was mentioned earlier in this work, that the vertical tailplane design for a supersonic transport is often critical because of the deteriorating tailplane efficiency with increasing Mach number, the vertical tailplane of the present semi-optimal design is extremely small (about 3% of the gross wing area). The reason for this problem is that the directional stability analysis carried out by means of the ESDU method [ref.1], appears to be very sensitive to the fuselage diameter. During a later phase of the optimization, a more severe directional stability constraint was imposed, the results of which will be treated later on.

The third peculiarity is that the wing has a supersonic leading-edge. In principle this might not be a problem, but merely a result of the design methodology employed here, that is, by not designing the airplane sequentially but by minimizing the performance constraint violations. Although the result might seem somewhat odd, it should be remembered that with the present design methodology the wing is not designed for the cruise condition, but to yield -in synergy with the other design variables- a design that complies best with all specifications imposed. Since the design space of a supersonic transport -as was already mentioned throughout this work- is very small and irregular, it is possible that the present wing is the only way to realize an acceptable design. A similar result was obtained in chapter 3 where a subsonic transport plane was optimized. Here the cruise angle of attack at the optimum exceeded zero, which implies that the fuselage was not positioned at its minimum drag condition, resulting in a sub-optimal cruise L/D ratio. Again, however, the cruise condition was not the optimality criterion.

Another possibility is, that the supersonic leading-edge is caused by the *hitchhiking* phenomenon inherent to genetic optimization (chapter 6), although this is not very likely. It was tried to increase the mutation probability, but this stopped progress altogether. In case hitchhiking should cause problems, the second stage of the hybrid optimization method -the gradient-based search- should take care of them.

Also there is a possibility that the optimization program tries to prevent the occurrence of leading-edge vortices at all cost. These leading-edge vortices, which have mainly been modelled to accurately predict the low-speed characteristics of high-speed configurations, in particular the landing performance constraint, have quite a large negative influence on the lift-to-drag ratio. Since no leading-edge flaps were modelled (except a simulation of a vortex generating device in the landing approach), the only way to prevent such vortices, or to keep them small, is to either increase the leading-edge radius or to choose a supersonic leading-edge with a small wing thickness-to-chord ratio to alleviate -at least to some extent- the resulting drag penalty (*Starfighter*-wing). Apparently, the nose radius could not be increased sufficiently, or the resulting constraint violations were larger than in the case of the other solution.

---

### 15.5.2 Gradient-based Optimization

Of the design obtained from the genetic optimization, referred to as the baseline design, seven variables were fixed, being the fuselage diameter and the six variables that define the engine design. This implies that the engine performance only depends on the off-design Mach number and altitude and can therefore be analyzed by interpolation from a prepared engine data file. This speeds up the analysis considerably and enables the use of the full mission analysis within reasonable computational time.

Within the new design space -as mentioned, defined by a  $\pm 10\%$  perturbation of the semi-optimal design variables- a large amount of random designs were generated, in order to derive regression surfaces for the constraints. Since seven design variables are now fixed, only 21 remain, implying that the regression surfaces are defined by 253 coefficients and at least that many designs have to be generated. From the original 14 performance constraints (see previous chapter), the three environmental constraints, the sea level climb constraint and the engine failure constraint were withdrawn, leaving only 9. The analysis used to generate the samples was the same as that used to check the baseline design, that is, it included the full mission analysis, with supersonic propulsion analysis by interpolation and the new aerodynamic polynomials.

About half of the generated designs were incomputable and therefore removed from the data set (the main causes being not enough thrust during the accelerated climb or too large a negative tailplane lift necessary for trimming, leading to a negative total lift).

The obtained regression surfaces had, without exception, excellent values for the F-test and the percentage of the residuals that are within a Normal distribution. Also the  $MCC^2$  values were all in excess of 98% and in some cases 99%. The standard errors of the regression for the different constraints are summarized in the following table.

constraint	1	2	3	4	5	6	7	8	9
standard error of regression [%]	1.2	0.9	3.9	1.3	12.1	14.2	4.4	2.6	8.2

- |   |                       |   |  |
|---|-----------------------|---|--|
| 1 | take-off field length | 6 | subsonic longitudinal stability margin   |
| 2 | approach speed        | 7 | supersonic longitudinal stability margin |
| 3 | supersonic range      | 8 | supersonic directional stability margin  |
| 4 | subsonic range        | 9 | engine failure after take-off constraint |
| 5 | cruise excess thrust  |   |  |

Table 36    Standard errors of regression of performance constraints

The gradient-based optimization was carried out with the use of the E04VCF routine provided by the Numerical Algorithms Group ([ref.4]), that is, a Sequential Quadratic Programming routine.

Since the purpose of the optimization is to minimize the performance constraint violations, the Kresselmeir-Steinhauser function (equation (59)) was chosen as the objective function. The geometrical constraints are therefore the only constraints imposed on the problem (together with the bounds on the variables that keep them within the 10% design space). Since the optimization algorithm only has to deal with second-order polynomial approximations to the real analysis, optimal solutions are generated within seconds.

Initially the semi-optimal design obtained from the genetic optimization was used as a baseline, later the optimization was sometimes restarted from the last obtained design. Also, some fine-tuning was applied by gradually increasing the controlling factor in the Kresselmeir-Steinhauser function, which draws the function closer to the constraint contours. Choosing the controlling factor too large, however, quickly caused overflow errors.

Each time an optimum has been obtained, it is checked with the exact analysis. Initially quite large errors occurred, despite the fact that the polynomials were very accurate, as indicated by their statistical parameters. The cause of this problem was identified as follows. Apparently, as indicated by the large amount of incomputable designs (about 50% of all randomly generated samples within the 10% design space), there are many "holes" in the design space. Since the incomputable designs were removed from the sample set, they are not reflected in the regression. However, the regression surfaces do cover the holes and consequently large errors are possible. Since optimization algorithms -as repeated several times throughout this dissertation-take advantage of imperfections in the analysis, they are likely to come up with optimal designs that prove to be incomputable once they are evaluated by the exact analysis.

The following remedy was successfully applied. Since most of the problems occur during the detailed mission analysis, it is possible to compute values for the constraints by using the simplified mission analysis (equation (397)) instead. This is consistent with remarks at the beginning of this chapter, that it is always preferable to express an incomputable design in values that -although not entirely accurate- are at least representative for the design under consideration. In this way an optimization routine can handle them and take corrective action. By adding the re-evaluated optimum to the sample set, new regression surfaces with increased accuracy could be generated and the optimization restarted. By repeating this process several times, the errors in the prediction were reduced gradually. Since the optimal designs were all situated close to each other, the regression became biased -but also very accurate- in this region. The standard error of the regression stayed about the same for all constraints, except that of the subsonic longitudinal stability margin which increased with increasing number of samples toward about 30%.

### **SQP Optimum Evaluated with Elaborate Mission Analysis and Exact Propulsion Analysis by Interpolation**

When finally no further improvement could be realized, the final optimal design was characterized by the following features (**design #1**). The critical constraints are still the supersonic range and the subsonic longitudinal stability margin, which could not be satisfied. Although either of them could be improved at the cost of the other, the best result had a supersonic range of 8681 km (as compared to 7920 km for the baseline design that resulted from the genetic search) and a subsonic longitudinal stability margin of -13.6% (as compared to -10.4% for the baseline design). The approach speed constraint is no longer critical. The maximum take-off weight of this design had increased to 386271 kg. Incidentally, the best of the randomly generated samples within the  $\pm 10\%$  design space had a supersonic range of 8549 km and a subsonic stability margin of -12.4%, with all other constraints satisfied.

#### **15.5.3 Evaluation and re-Optimization of Gradient-guided Search Result**

### **Genetic Optimization with Simulated Leading-Edge Flaps, Simplified Mission and Exact Propulsion Analysis**

The most conspicuous feature of the final design is the fact that the wing leading-edge is still supersonic, which rules out the possibility that this was caused by the hitchhiking phenomenon. In order to investigate whether a subsonic leading-edge results in a larger drag penalty -because of the occurrence of leading-edge vortices- than the extra wave drag resulting from a supersonic leading-edge, the attainable thrust coefficients (see chapter 9) were set to 1, thus simulating the presence of leading-edge flaps that prevent vortices. Only in the landing approach they are set to zero in order to simulate the deployment of a vortex flap (in case full leading-edge flaps are present, a vortex flap should be regarded as a small tooth that is deployed from the leading-edge of the flap). A genetic optimization was carried out using the simplified mission analysis, the exact supersonic propulsion analysis and the original aerodynamic regression equations. It was established that this did not result in

---



designs with a subsonic-leading edge, although a slightly better design could be obtained (**design #2**). An interesting feature of this design is that the maximum take-off weight is roughly a metric ton higher than that of the baseline design (according to the same analysis method, except for the attainable thrust factor), namely 448195 kg.

This design was compared to a second design with the same values for all design variables, except that the leading-edge and trailing-edge were adjusted such as to yield a subsonic leading-edge at the same span and gross wing area. Although this design had slightly smaller values for the lift-induced drag and the zero lift drag, the trim drag had increased dramatically whereas the negative tail load reduced the total lift coefficient considerably, leading to very poor lift-to-drag ratios. Apparently, the disadvantage of a highly-swept wing as perceived by the optimization algorithm, is the fact that the neutral point is located relatively far behind the center of gravity, leading to excessive trim drag and a large negative tail load.

### **Genetic Optimization with Simulated Fuel Trim, Simplified Mission and Exact Propulsion Analysis**

This problem is well known in supersonic airplane design, and several ways are known to alleviate it (see chapter 7, page 293 of chapter 13 and paragraph 14.8.1). The present method of analysis only comes halfway, by allowing negative stability margins of up to -5% thus enabling more rearward locations of the center of gravity. Although no trim tank is modelled in the present analysis, the fuel tanks may be emptied in such a way as to provide a beneficial center of gravity shift, as explained at the end of chapter 13. To investigate whether a procedure, according to which the fuel tanks are emptied from the inward out, will result in smaller constraint violations and perhaps a subsonic wing leading-edge, another genetic optimization procedure was carried out. A weight penalty is imposed to account for the presence of the required extra fuel system (proportional to the tank volume). This finally resulted in a design which complied sufficiently with all performance and stability constraints (35 km short on supersonic range, subsonic longitudinal stability margin of -5.8%). However, the small vertical tailplane and the supersonic leading-edge remain. Another noteworthy feature of the design is the fact that the horizontal stabilizer has a relatively large span. Apparently it has been designed for subsonic efficiency (large aspect ratio) in order to achieve sufficient stability. It was established that by increasing the leading-edge sweep of the horizontal stabilizer at constant aspect ratio while leaving all other variables constant, the supersonic drag could be reduced such that the airplane complied with all constraints (except for the sonic boom peak pressure and take-off noise), although only after relaxation of the subsonic stability constraint to -10% (**design #3**).

### **Reducing the Conflict between Subsonic Stability and Supersonic Range**

Reducing the conflict between the subsonic stability constraint and the supersonic range constraint should enable compliance with the environmental constraints as well (or at least result in better values for these constraints). Further alleviating the

---

subsonic stability constraint (to a minimum dictated by the possibilities of artificial stability augmentation) is an option, but will only bring limited results (especially since values below -10% will probably not be realistic). Optimizing the procedure by which the fuel tanks are emptied (by applying more tank segments) or defining a trim tank might have an important effect, since on the one hand it causes a reduction of the trim drag during cruising, whereas on the other hand it will enable larger wing sweep, causing a wave drag reduction. Such procedures, however, would introduce a large number of new variables and require an adaption to the analysis procedure.

Another possibility is to anticipate on a future optimization of the wing camber and twist distribution during a more detailed design phase. Without actually carrying out this optimization, the method of chapter 10 may be used to estimate the attainable induced drag reduction. This again is a kind of approximate analysis. During the detailed design phase more accurate approximations may be obtained which can be fed upwards to the present level of conceptual design and optimization for a second run. This is beyond the scope of the present work, however. Unfortunately, a second favorable effect that can be achieved by applying optimal camber cannot be modeled at present. By optimizing the wing camber and twist distribution it is possible to generate a pitching moment at a relatively low drag penalty, which therefore will reduce trim drag. Furthermore, this would also enable a larger wing sweep angle which will further reduce wave drag.

In order to reduce the conflict between the supersonic range performance and the subsonic longitudinal stability, the following of the above mentioned measures were taken. The subsonic longitudinal stability constraint was relaxed to -10%, the fuel trim technique was applied and the method described in chapter 10 was used to simulate the effect of an optimized lift distribution.

### **Imposing a More Severe Directional Stability Requirement**

As explained before, a more severe constraint was imposed on the vertical tailplane design, since the initial directional stability constraint (which only required the airplane to be marginally stable) resulted in a very small tailplane indeed. Although it is difficult to define an acceptable value for such a constraint, in this application a  $C_{N_\beta}$  value of 0.25 is required to ensure satisfying Dutch-roll behavior ([ref.5]). Without changing any other design variables, this would roughly result in an increase of the vertical tailplane area to about 10% of the gross wing area. It was established, however, that this led to a drastic deterioration of performance (again specifically of the critical subsonic stability and supersonic range constraints), mainly caused by the increase in tailplane weight and the resulting maximum take-off weight increase and unfavorable shift of the critical center of gravity position. Therefore, an additional optimization was carried out with this new constraint.

Using the method described in chapter 10 to simulate the effect of an optimal lift distribution did not prove very fruitful, since this method (based on [ref.3]) does not apply to wings with supersonic leading-edges. Apparently, the benefit of minimal lift-

---

induced drag was not sufficient to counteract the disadvantage of a higher trim drag as associated with higher leading-edge sweep angles. An extension of the method such that it is applicable to wings with supersonic leading-edges is very much needed, but beyond the scope of the present work. However, bearing in mind the large drag reductions that were shown in chapter 10 to be achievable through optimization of the pressure distribution of wings with subsonic leading-edges, it is very likely that by applying optimal camber and twist it is possible to increase the feasible design space and obtain better values for the take-off noise constraint. As explained before -and also shown in chapter 12- the sonic boom peak pressure constraint will be difficult to comply with anyway.

### **Evaluating the Final Design with the Exact Aerodynamic Analysis**

Eventually, the genetic optimization resulted in a design that complied with all constraints (except for the sonic boom and take-off noise constraint) (**design #4**). Since the aerodynamic analysis was based on a regression of results of the NLRAERO panel code, the final configuration was re-evaluated using the exact values as obtained from the NLRAERO program. This led to a disappointing result. The cruise lift-to-drag ratio as calculated by means of the approximate analysis had a value of 7.65 whereas the exact analysis showed a cruise  $L/D$  of only about 5, leading to a severe range reduction. The main reason for this problem is the fact that the regression overestimates the total configuration lift gradient and the pitching moment coefficient by about 25%, leading to an overestimation of the untrimmed cruise lift coefficient, whereas the horizontal stabilizer trim angle changes from  $+1.4^\circ$  to  $-2.9^\circ$ , leading to a drastic decrease of the trimmed  $L/D$  ratio. Therefore, it can be concluded that the optimization algorithm once again has led the search to an area of the design space where the method of analysis is less accurate.

To remedy this problem, it was decided to generate new supersonic aerodynamic polynomials based on results obtained in a decreased design space of a  $\pm 10\%$  perturbation of the design variables of the "final" design, just like it was done for the gradient-based optimization (it is recalled here that then the change in performance was not so large; the supersonic range only changed from 8329 km to 7900 km, in combination with a slight increase in maximum take-off weight). The statistic parameters of the new supersonic aerodynamic polynomials are presented in the following table:

---

	supersonic aerodynamic coefficients				
	$Kp_{wf}$	$Kp_w$	$C_{L_{\alpha_i=2^\circ}}$	$C_{M_{\alpha_i=2^\circ}}$	$C_{M_\alpha}$
number of samples	200	200	200	200	200
MCC <sup>2</sup> [%]	81.49	96.80	53.17	47.77	76.79
F test	9.07	62.41	2.34	1.89	6.82
standard error of regression [% of mean response]	2.42	0.87	5.95	8.24	2.96
% of residuals in Normal distribution	97.0	98.5	97.5	97.5	97.0

**Table 37** Statistics pertaining to regression of supersonic aerodynamic coefficients

Again it is noted that the regression for the lift coefficient and pitching moment coefficient at two degrees angle of attack are below standard. Nonetheless, only the new polynomial for the pitching moment coefficient at two degrees angle of attack was not adopted. The lastly obtained configuration was also evaluated with these new supersonic aerodynamic polynomials and compared to the NLRAERO results:

	approximate	NLRAERO
$Kp_{wf}$	2.3005	2.3835
$C_{L_{\alpha_i=2^\circ}}$	0.08532	0.09303
$C_{L_{\alpha_i=0}}$	0.005017	0.009830
$C_{M_\alpha}$	-6.1529	-6.0236
$C_{M_{\alpha_i=0}}$	-0.03411	-0.03247

**Table 38** Comparison of approximate and exact aerodynamic coefficients

Although these new polynomials are quite accurate, the slight underestimation of the lift gradient apparently results in a very large underestimation of the lift coefficient at zero fuselage angle of attack. The trimmed lift-to-drag ratio, which based on NLRAERO results amounts to a value of about 5, is further reduced to about 3 using these new polynomials. The influence of the lift gradient therefore is very large. Experience teaches, however, that although in supersonic flow the lift gradient and the lift coefficient at two degrees angle of attack as obtained from the panel method

are quite accurate, the lift coefficient at zero degrees angle of attack is not.

### Genetic Optimization with New Aerodynamic Approximations

With the new polynomials a second genetic optimization was performed within the mentioned reduced design space. This again resulted in a design that sufficiently complied with the performance constraints (**design #5**). Subsequent evaluation with the NLRAERO panel method again showed a difference which had a large impact on the cruise lift-to-drag ratio, although the difference was smaller this time. It was further established however, like reported in chapter 10, that the supersonic results of the NLRAERO program are very sensitive, especially to the number of panels. It is a well known fact that there exists an optimal panel distribution for which the numerical errors are minimal, but this distribution is difficult to derive, especially since in the present work the panel method is applied to a large number of different configurations whereas the panelling process is carried out automatically. Although the lift and pitching moment gradients are less susceptible to noise (and better predicted by the regression equations), the values for the lift and pitching moment coefficient at two degrees angle of attack are unreliable. In chapter 5 it was already mentioned that the sensitivity to noise of the supersonic NLRAERO results is a problem which can be overcome by using regression equations as a smooth approximation, be it at the cost of accuracy (see also [ref.2]). This implies however, that the benefit of verification of the optimal result with the use of the exact analysis becomes questionable since the accuracy of the "exact" analysis itself is doubtful. As the optimum obtained by approximate analysis should always be checked by the exact analysis (since violated constraints or excessive approximation errors may render the optimum useless), it must be concluded, that the NLRAERO panel method is not the right tool for the analysis of configurations in supersonic flow.

The following table shows the values as approximated by the regression equations, together with the values as obtained by the NLRAERO method, first for a panel distribution of 243 body panels and 171 wing panels at zero and two degrees angle of attack (1) and secondly for a panel distribution of 147 body panels and 105 wing panels at two and four degrees angle of attack (2).

	approximate	NLRAERO (1)	NLRAERO (2)
$K_{p_{wf}}$	2.6291	2.4634	2.6267
$C_{L_{\alpha_i=2^\circ}}$	0.1226	0.08043	0.06303
$C_{L_{\alpha_i=0}}$	0.03085	-0.00556	-0.02866
$C_{M_\alpha}$	-7.4604	-6.3767	-6.9348
$C_{M_{\alpha_i=2^\circ}}$	-0.2368	-0.2043	-0.1342
$(L/D)_{cr}$	7.6	5.8	6.5

Table 39 Comparison of approximate and exact aerodynamic coefficients

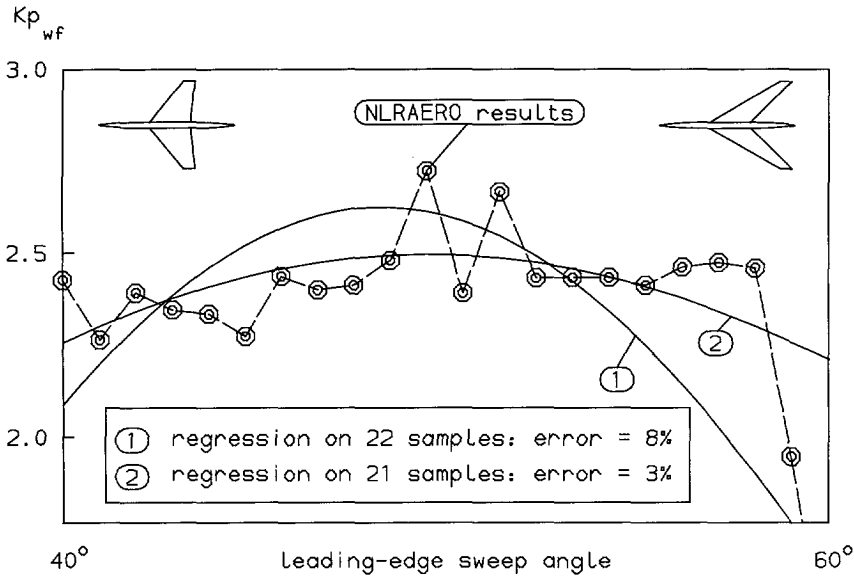


Figure 68 Total configuration lift gradient; NLRAERO results and regression ( $M=2.07$ )

For further illustration, the configuration designated **design #5** was analyzed with the NLRAERO program for different values of the length of the net wing planform box,  $l_w$ , such that the leading-edge sweep angle varied between  $40^\circ$  and  $60^\circ$  at constant aspect ratio and wing area (supersonic leading-edge). In Figure 68, Figure 69 and Figure 70 the total configuration lift gradient, the wing lift gradient and the lift coefficient at two degrees angle of attack are plotted, together with the least-squares regression in  $l_w$ .

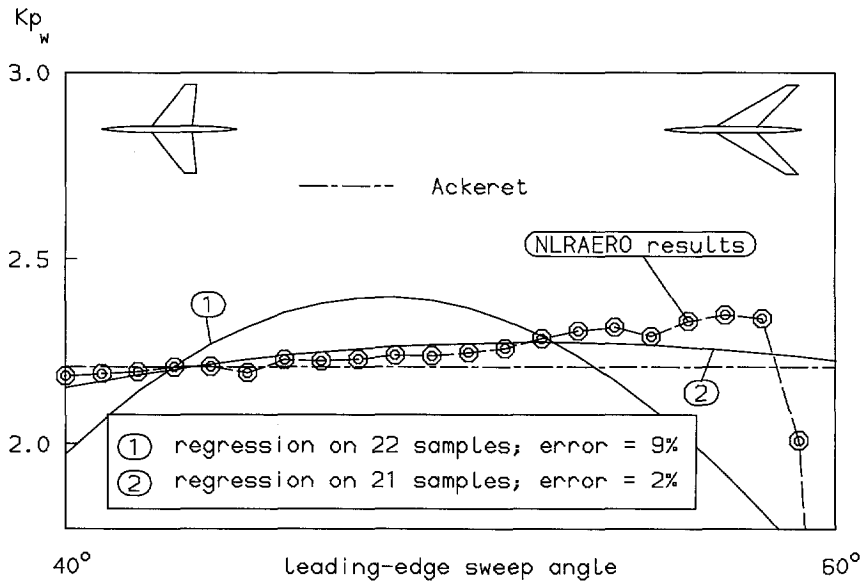


Figure 69 Wing lift gradient; NLRAERO results and regression (M=2.07)

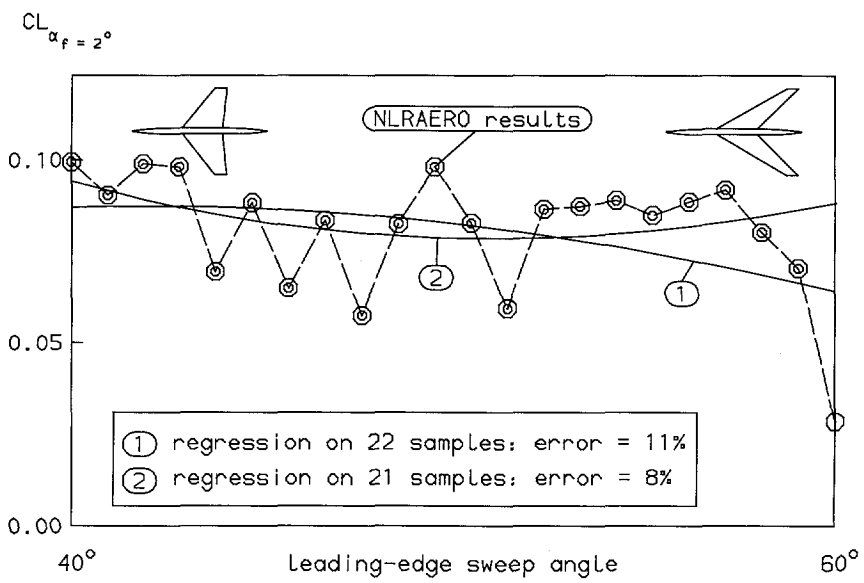


Figure 70 Lift coefficient at two degrees angle of attack; NLRAERO results and regression (M=2.07)

Since the last sample was particularly inconsistent with the other data, two regressions are shown, the first one based on all 22 samples, the second one based on all except the last sample. Since the leading-edges of all considered samples are supersonic, the wing lift gradient can be directly compared to the value according to Ackeret's theory (except for wing-fuselage interference), that is,  $C_{L_\alpha} = \frac{4}{\sqrt{M^2 - 1}}$ . This

line is drawn in Figure 69 too. The noisy character of the NLRAERO results is clear, although the fact that the wing alone results are quite accurate lends credibility to the remarks in chapter 10 that the occurrence of vacuum pressures on the body is the main cause of the problems.

### Fine-tuning the Final Design through SQP Optimization

Upon evaluation of the design designated **design #5** with the exact mission analysis (since the genetic search only uses an approximate mission analysis), it was established that somewhere along the accelerated climb to cruise conditions, the aircraft runs out of thrust. To create a more accurate optimum based on the exact design analysis, a final SQP optimization was carried out on a regression based on designs analyzed within the mentioned reduced design space by means of the exact mission and propulsion analysis (propulsion analysis by means of interpolation). In the same way as described previously, by constantly re-evaluating the "optimal" design and updating the regression, increasingly better designs were obtained. Since the regression is never exact, this process could be carried out indefinitely. Instead, at some point it was stopped, since it was felt that further optimization no longer contributed substantially to the conclusions and results of this work. The final design (**design #6**), as presented hereafter, did comply sufficiently with the imposed constraints, except that the subsonic longitudinal stability and the supersonic directional stability were not entirely satisfying. Further optimization might remedy this.

The design evolution of the designs mentioned in this chapter, is depicted in Figure 71. It should be mentioned that the genetic algorithm by itself did not come up with configurations having swept tailplanes, but after human interference it would sometimes keep this property in the population.

The final design is somewhat more elaborately presented in Appendix A, together with three-view drawings and trimmed polars for the subsonic and supersonic cruise.

It is surprising that an airplane that flies at approximately the same cruise Mach number as Concorde with a hardly superior lift-to-drag ratio is able to achieve a range increase of almost 50%, especially in view of the fact that no drastic improvements in the state of the art have been modeled. This applies in particular to the mass prediction model which is based on "old" technology, notably Aluminum structures and subsonic cruising aircraft, whereas in this field major improvements are to be expected for a second-generation supersonic transport (carbon reinforced materials). Therefore, the factors in the Bréguet range equation for both Concorde and the designs resulting from the present study are compared in the following table (10% of



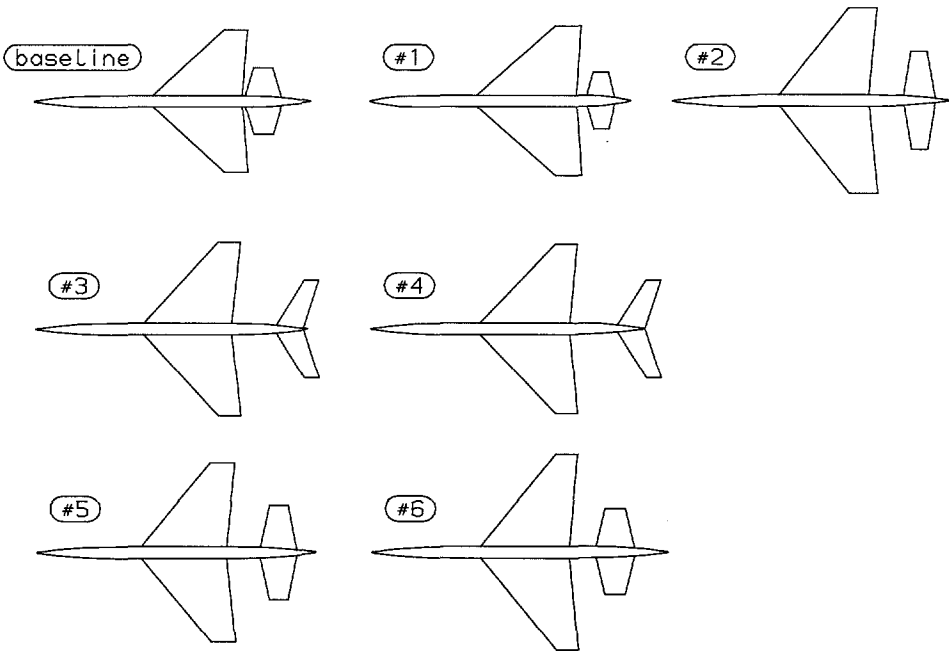


Figure 71   Design evolution

the total fuel weight is assumed as reserve for both calculations):

	Concorde	present study	factor
M	2.00	2.03	1.02
L/D	7.3	7.9	1.08
$I_{sp}$ [s]	3033	3290	1.08
$(\eta_{tot})$	41%	45%	
$\ln W_1/W_2$	0.45	0.54	1.20
R [km]	5839	8425 (climb excluded)	1.44

Table 40   Comparison between Concorde and present study range performance

Apparently, the main improvement is caused by the increase in specific impulse during cruising as well as during take-off (together this accounts for a range increase of almost one third). At virtually the same fuel fraction, the ratio between the take-off

weight and the weight at the beginning of the cruise flight is 0.86 for Concorde as compared to about 0.95 for the present design. This large difference is caused by the fact that Concorde uses reheat during the accelerated climb to cruise conditions.

## 15.6 References

1. ESDU Engineering Sciences DATA Unit; Vol.3c, No.89008, April 1989
2. Giunta, A.A.; Dudley, J.M.; Narducci, R.; Grossman, B.; Haftka, R.T.; Mason, W.H. and Watson, L.T.; *Noisy Aerodynamic Response and Smooth Approximations in HSCT Design*; AIAA 94-4367-CP.
3. Jones, R.T.; *The Minimum Drag of Thin Wings in Frictionless Flow*; Journal of the Aeronautical Sciences, Vol.18, No.2, February 1951
4. Numerical Algorithms Group; *Fortran Libraries Mk.12, Vol.4*
5. Torenbeek, E.; *Synthesis of Subsonic Airplane Design*; Delft University Press, 1988
6. Wittenberg, H.; *Range Performance of Hypervelocity Vehicles*; De Ingenieur, Vol.81, Nr.43, October 1969

## Appendix A Final Design Characteristics

In accordance with tradition, weights are given in kilograms, although -according to the SI system- the unit of weight is Newton and the unit of mass is kilogram.

### dimensions:

span:	51.84 m
length:	77.84 m
fuselage diameter:	3.13 m
gross wing area:	783.49 m <sup>2</sup>
horizontal stabilizer area:	183.61 m <sup>2</sup>
vertical stabilizer area:	64.00 m <sup>2</sup>
leading-edge sweep angle:	39.59°
wing thickness-to-chord ratio:	0.039
wing nose radius:	0
wing incidence:	0

**weights:**

wing:	58420 kg
fuselage:	15432 kg
powerplant:	44414 kg
fuel:	213107 kg
payload:	24450 kg
horizontal tail:	7262 kg
vertical tail:	2492 kg
equipment:	43100 kg
surface controls:	3365 kg
nose gear:	2229 kg
main gear:	16720 kg

maximum take-off weight:	430992 kg
weight at beginning of cruise:	408916 kg

**supersonic mission:**

transition altitude:	8332 m
cruise altitude:	
initial:	15317 m
average 1 <sup>st</sup> segment:	16082 m
average 2 <sup>nd</sup> segment:	17855 m

cruise Mach number:	2.03
cruise angle of attack:	3.2°

cruise L/D:	
1 <sup>st</sup> cruise segment:	7.91
2 <sup>nd</sup> cruise segment:	7.85

specific impulse:	3290 s
-------------------	--------

range:	
accelerated climb:	238 km
1 <sup>st</sup> cruise segment:	2075 km
2 <sup>nd</sup> cruise segment:	6350 km
total:	8663 km

**performance:**

take-off field length:	2910 m
take-off thrust (per engine):	322 kN
take-off specific impulse:	4023 s
approach speed:	74 m/s
subsonic range:	7755 km

---

**environmental aspects:**

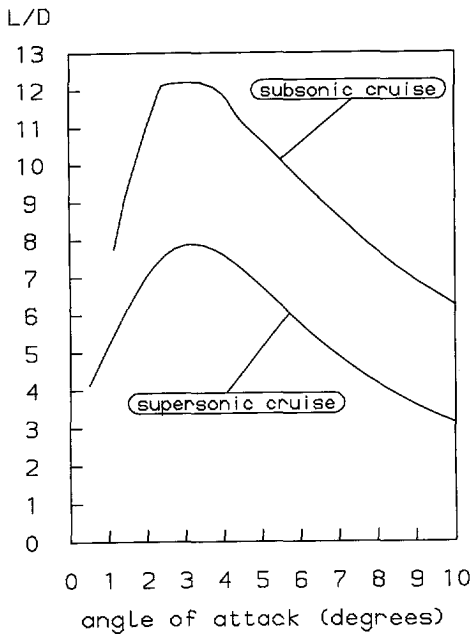
maximum shockwave overpressure:	120 Pa
exit jet velocity:	780 m/s
NO <sub>x</sub> emission:	19 g/kg fuel

**engine design variables:**

design Mach number:	1.76
intake (capture) area:	3.58 m <sup>2</sup>
compressor pressure ratio:	16.03
bypass ratio:	0.50

The main dimensions and a 3-view drawing of the final design are given in Figure 73 and Figure 74. The directional stability margin could be improved to a more satisfying value by increasing the vertical tailplane, at the cost of a small range deterioration. It is this configuration that is sketched in the figures.

The trimmed drag polars for the subsonic and supersonic cruise are shown in Figure 72.



**Figure 72** Trimmed drag polars for subsonic and supersonic cruise

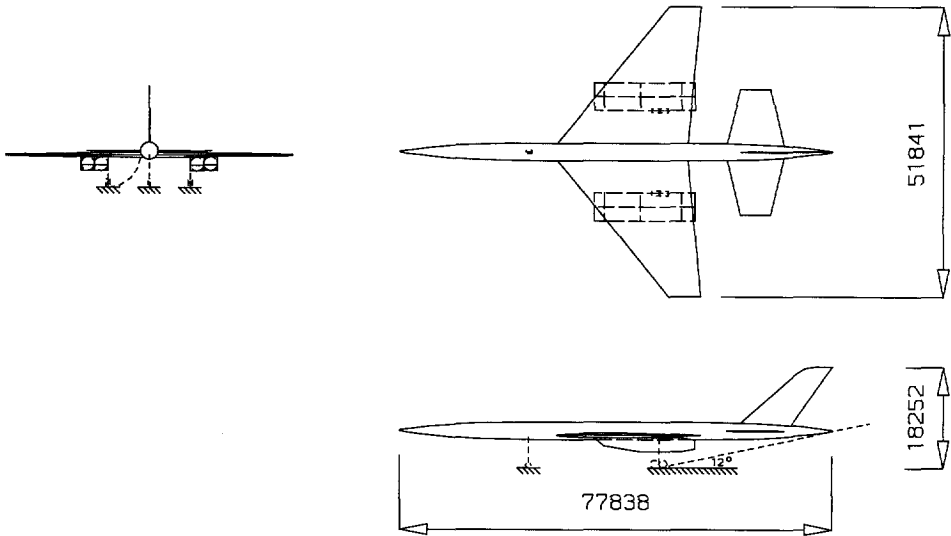


Figure 73 Final design: main dimensions

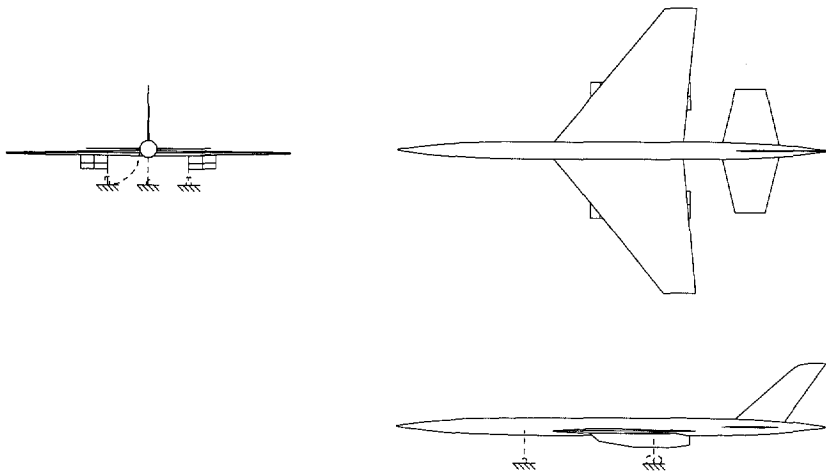


Figure 74 Final design: 3-view drawing

## Appendix B Evaluation of the Mixing Process

### Notation

A	area
M	Mach number
$\dot{m}$	mass flow
p	pressure
R	gas constant
T	temperature
$\epsilon$	pressure ratio
$\eta$	efficiency
$\gamma$	specific heat ratio
$\lambda$	bypass ratio

### subscripts

2	station number (diffusor)
c	cold flow
	compressor
cr	cruise
f	fan
g	combustion gases
h	hot flow
int	intake
t	total
to	take-off

### abbreviations

HPT	high pressure turbine
LPT	low pressure turbine
TET	turbine entry temperature

### Evaluation of the Mixing Process

The engine off-design performance was calculated on basis of the assumption that the mixing process is abstracted by putting equal the total pressures behind the fan and the low pressure turbine. In reality, the Kutta condition (equal static pressures) will apply on the streamline separating the hot and the cold flow prior to mixing. If the assumption of equal total pressures in the mixer is enforced together with the Kutta condition, it will be necessary to vary the mixer areas of the hot and the cold flow (under the constraint that their sum remains constant). Hereafter it is examined what kind of geometry adaptations this would require between the cruise condition and take-

---

off, for the final engine design.

From the relation for the total pressure, putting both the total pressures and the static pressures equal in the mixer, the Mach numbers of the hot and the cold flow are related as follows:

$$M_c = \sqrt{\frac{2}{\gamma-1} \left[ \left( 1 + \frac{\gamma_g - 1}{2} M_h^2 \right)^{\frac{\gamma_g(\gamma-1)}{\gamma(\gamma_g-1)}} - 1 \right]} \quad (398)$$

With this relation, the required areas for the hot and the cold flow in the mixer can be expressed as a function of the hot flow Mach number as follows:

$$A_c = \frac{\dot{m}_c \frac{\sqrt{T_{t_c}}}{p_{t_c}} \sqrt{\frac{R}{\gamma}} \left( 1 + \frac{\gamma_g - 1}{2} M_h^2 \right)^{\frac{\gamma_g(\gamma+1)}{2\gamma(\gamma_g-1)}}}{\sqrt{\frac{2}{\gamma-1} \left[ \left( 1 + \frac{\gamma_g - 1}{2} M_h^2 \right)^{\frac{\gamma_g(\gamma-1)}{\gamma(\gamma_g-1)}} - 1 \right]}} \quad (399)$$

$$A_h = \frac{\dot{m}_h \frac{\sqrt{T_{t_h}}}{p_{t_h}} \sqrt{\frac{R}{\gamma_g}} \left( 1 + \frac{\gamma_g - 1}{2} M_h^2 \right)^{\frac{\gamma_g+1}{2(\gamma_g-1)}}}{M_h} \quad (400)$$

The ratio of both areas can then be written as follows, when the total pressures of the cold and the hot flow are equal (neglecting for simplicity the fuel mass flow):

$$\frac{A_c}{A_h} = \lambda \sqrt{\frac{T_{t_c}}{T_{t_h}}} \sqrt{\frac{\gamma_g}{\gamma}} M_h \frac{\left( 1 + \frac{\gamma_g - 1}{2} M_h^2 \right)^{\frac{\gamma_g - \gamma}{2\gamma(\gamma_g-1)}}}{\sqrt{\frac{2}{\gamma-1} \left[ \left( 1 + \frac{\gamma_g - 1}{2} M_h^2 \right)^{\frac{\gamma_g(\gamma-1)}{\gamma(\gamma_g-1)}} - 1 \right]}} \quad (401)$$

The function of the hot Mach number on the right-hand-side of the equation is almost equal to one. This implies that the mixer area ratio is all but independent of the Mach numbers of the hot and the cold flow and the mixer area ratio in the cruise condition

relative to that during take-off can be written as:

$$\frac{\left(\frac{A_c}{A_h}\right)_{cr}}{\left(\frac{A_c}{A_h}\right)_{to}} = \frac{\lambda_{cr} \left( \sqrt{\frac{T_{tc}}{T_{th}}} \right)_{cr}}{\lambda_{to} \left( \sqrt{\frac{T_{tc}}{T_{th}}} \right)_{to}} \tag{402}$$

The total temperatures of the hot flow and the cold flow in the cruise flight are respectively 1036 K and 535 K whereas the bypass ratio is 0.51. During take-off these values are 1023 K and 447 K and the bypass ratio is 0.27. Since the engine frontal area was sized for a maximum permissible Mach number of 0.5, the sum of the mixer areas should equal 1.34 times the intake area  $A_0$  (3.58 m<sup>2</sup>), or 4.8 m<sup>2</sup>. During cruising the mixer area ratio is about 0.36 yielding a hot mixer area of 3.53 m<sup>2</sup> and a cold mixer area of 1.27 m<sup>2</sup>. During take-off the mixer area ratio is half of the cruise value (0.18) yielding a hot mixer area of 4.07 m<sup>2</sup> and a cold mixer area of 0.73 m<sup>2</sup>. Thus, the diameter of the hot duct changes about 15 cm in case the mixer is cylindrical. If the mixer is square, the height of the hot duct changes 22 cm since the engine diameter (width) equals 2.47 m (according to the relations on page 224).

The total massflow through the engine while cruising is 353 kg/s whereas the total pressure in front of the mixer is 215602 Pa. The Mach numbers of the flow, just before mixing occurs, can thus be calculated as about 0.15 (both hot and cold). The take-off massflow through the engine equals 478 kg/s. With a pre-mixing total pressure of 439309 Pa this yields a value of about 0.1 for the Mach numbers of the hot and the cold flow. The difference between the total pressure and the static pressure is therefore very small.

The following table summarizes the most important engine parameters during the cruise and the take-off condition (at screen height,  $M=0.34$ ). Since the engine was designed for a Mach number of 1.76, the compressor pressure ratio during cruising ( $M=2.03$ ) is smaller than the design value of 16.03 because of the constraint on the compressor exit temperature.

	$\lambda$	$\varepsilon_c$	$\varepsilon_f$	$\varepsilon_{HPT}$	$\varepsilon_{LPT}$	$\eta_{int}$	TET [K]	$\dot{m}_2$ [kg/s]
cruise	0.51	11.76	2.53	3.62	1.26	0.94	1481	353
take-off	0.27	14.17	4.00	2.99	1.16	1.00	1372	478

Table 41    Engine parameters during cruise and take-off condition



## Appendix C Evaluation of the Fan Design

### Notation

S	distance
T	thrust
u	velocity
$\varepsilon$	pressure ratio
$\lambda$	bypass ratio

### subscripts

c	compressor
e	exit
f	fan
inst	installed
to	take-off

### Evaluation of the Fan Design

As explained in chapters 11 and 12, the presently used variable-cycle model may impose too high demands on the flexibility of the fan design, which may require variable vanes. It was also acknowledged that the increase in flexibility that can thus be obtained is limited too. Therefore, it may be necessary to impose bounds on the range of the fan pressure ratio. Such bounds, however, are dependent on the number of fan stages of the final design and on the amount of flexibility that eventually can be realized by applying variable vanes. Extra fan stages will enable a higher pressure ratio, but will increase weight. This implies that a more detailed weight prediction is required, which is out of the scope of the present study.

Since the range of the fan pressure ratio can therefore only be determined during a more detailed design phase, no attempts were made to constrain the fan pressure ratio (and therefore the bypass ratio) in the present model. However, in this appendix, the influence of variations of the bypass ratio (and therefore of the fan pressure ratio) on take-off performance is shown for the final engine design. It should be remembered that in case the fan pressure ratio would be constrained to a value other than the one chosen in the final design, the engine weight and therefore the maximum take-off weight of the airplane would change too. This implies that the increases of the take-off field length as indicated in the following table should be regarded as global sensitivities.

The following table gives the results for increasing values of the bypass ratio parameter as defined in chapter 12 (if this parameter equals zero, the bypass ratio attains its minimum permissible value in order to maximize installed thrust, if it

---

equals one, the bypass ratio attains its maximum permissible value in order to minimize take-off noise at the cost of thrust). In the present study it was concluded that the take-off noise constraint could not be met and as a result this constraint was removed from the optimization; therefore, in the final design the bypass ratio attains its minimum permissible value, which corresponds to the first column.

$\lambda$	0.27	0.31	0.36	0.42	0.48	0.50
$\epsilon_f$	4.00	3.70	3.26	2.80	2.54	2.36
$\epsilon_c$	14.17	13.99	13.07	11.87	11.32	10.48
$u_e$ [m/s]	780	736	674	606	561	529
$T_{inst}$ [kN]	322 (100%)	312 (97%)	281 (87%)	243 (75%)	224 (70%)	200 (62%)
$S_{to}$ [m]	2910	2939	3191	$\infty$	$\infty$	$\infty$

**Table 42** Influence of fan pressure ratio limitations on take-off performance (ideal expansion)

According to this table, limiting the fan pressure ratio to a value of 3.26, which is close to that during cruising ( $\epsilon_f=2.53$ ), will cause a thrust reduction to 87% of the original amount, which in turn will lead to an increase of the take-off field length by about 300 m (which, however, will stay within specifications). Further reduction of the fan pressure ratio will lead to a shortage of thrust during rotation and therefore will require a re-optimization of the design.

# 16. Conclusions and Recommendations

## 16.1 Summary

Most of the conclusions drawn in this dissertation were discussed at the end of the chapters in which they were arrived at, and they will not be repeated here. A few conclusions, however, directly relate to either the problem of Multidisciplinary Design Optimization, or to the design of a second-generation supersonic transport.

## 16.2 Multidisciplinary Design Optimization

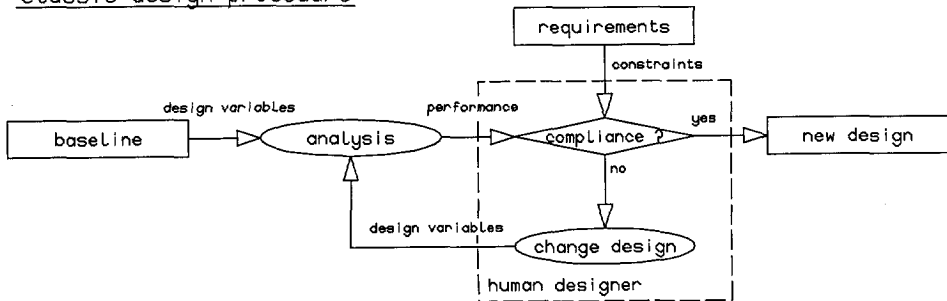
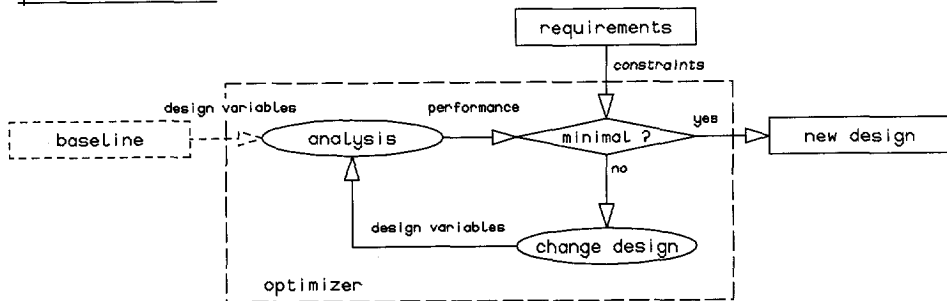
It is important to realize that peculiar results are not to blame on the method of optimization. If a design complies with the imposed constraints and provides a better value for the merit function than an earlier obtained design, the method can be deemed successful. Especially when, after many more function evaluations or gradient searches with different starting points, no further improvement can be realized, the final design -at least from an engineer's point of view- can be accepted as the optimum (although a mathematician will want some proof of this). In this sense, the presently adopted method was proven to be successful, since it provided a conceptual design from scratch, within the limitations of the method of analysis (regardless whether this analysis is correct or not). Therefore, the optimization method has become a practical design tool instead of just a method to adapt a baseline designed according to traditional methods. This is illustrated in Figure 75, in which the classical design procedure and the present method are depicted schematically.

Processes taking place within the human designer are very flexible and may be based on experience, knowledge and intuition. The weight assigned to certain constraints, as well as convergence criteria, requirements and so forth may be changed as the design process develops. Human qualities in the fields of interpretation, pattern recognition and the use of knowledge and experience are superior to those of computers. However, human capabilities to optimize/or to handle more than about three independent variables are very limited.

classic design procedure	present method
* sequential	* concurrent minimization of constraint violations
* baseline required	* no baseline required
* human designer in the loop (slow)	* human designer out of the loop (fast)
* design decisions taken by human designer	* design decisions taken by optimizer
* few design variables	* many design variables

Table 43    Comparison between classic design procedure and present method

---

classic design procedurepresent method

**Figure 75** Comparison between classic design procedure and present method

Both the classic design procedure and the present method treat the direct problem, that is, the performance of the design (fitness, merit function value) is analyzed for a fixed vector of design variables, which is subsequently adapted in an iterative fashion by either the human designer or the optimization algorithm. A different approach would be to treat the inverse problem, that is, to derive the values of the design variables from the required performance. The human designer is capable of carrying out the inverse process to a certain extent, based on his experience and knowledge, by creating a baseline design (configuration). However, when the design becomes increasingly complex and the requirements more conflicting (as is the case with the presently treated problem of designing a supersonic transport plane), the capabilities of the human designer fall short. This requires the development of knowledge-based design programs. In theory, neural networks should be capable of handling the inverse problem too.

An interesting view on the position of the human designer in the design process can be found in [ref.1] and [ref.2]. Here, the so-called *natural design cycle* is treated, in which the *analysis* in the left-hand-side of Figure 75 is carried out in real-time and thus virtually becomes part of the human designer. The issue of optimization is treated as well, but this too is based on the assumption that the analysis is carried out in real-time.

Should the final result not please the human designer, there are two possibilities. Either extra constraints should be imposed or the analysis method should be adapted (because it is incomplete, inaccurate, because it does not cover certain effects or influences accurately or because it contains one or more bugs). In both cases the optimization algorithm has taught the human designer a lesson, by taking advantage of the fact that the optimization problem is not well posed, and by pointing out the problem area.

An example of the first problem is the subsonic testcase of chapter 3. The final design has an almost straight wing of rather high aspect ratio and cruises at a Mach number of 0.55 in the stratosphere, whereas the baseline design had a swept wing and cruised at Mach 0.6 at 6200 m. The final design therefore has a cruise speed which is almost 100 km/h (or about 15%) lower than the baseline whereas the aspect ratio might be unrealistic from a structural point of view. If this is not what the designer had intended, then he should have imposed a minimum cruise speed or maximum flight time constraint and probably some structural constraints as well. Of course then the resulting optimum would have been worse, in terms of maximum take-off weight.

The second kind of problem was elaborately treated in the previous chapter.

The results and experiences with the present method clearly indicate that the most severe constraints imposed on a design are the limitations (mostly resulting from simplifications) of the analysis procedure itself. This suggests that a design methodology should be as multivariate, multidisciplinary and flexible as possible, without losing too much accuracy or requiring too much computational time. In other words, the computational time per independent design variable should be as low as possible. As always, the limit is computer speed, regardless of the method of optimization used. Losing too much accuracy will compromise the most important goal of MDO and concurrent engineering, that is, reducing the length and number of feedbacks in the design process. Since the calculation effort is the biggest problem in MDO whereas accuracy is the most important during the final fine-tuning of the optimum, the presently adopted method constitutes an acceptable compromise by saving computational time at the cost of accuracy during the global search while gradually increasing the accuracy of the approximate analysis during the local search.

## **16.3      The Design of a Second-Generation Supersonic Transport Aircraft**

A large amount of the problems encountered during the optimization, were a direct result of the nature of the testcase, that is, the design of a second-generation supersonic transport. Especially the fact that the aerodynamic parameters are very susceptible to computational noise, in combination with the often tedious analyses and the conflicting requirements, resulted in an optimization problem where accuracy, noise, computational time, and a very small design space with many "holes" where designs proved incomputable, were the major barriers blocking success. These

---

problems would not have been so severe in the case of a subsonic test case. It can therefore be concluded that the choice of a supersonic transport aircraft as a test case has been very educative, since it highlighted the major areas of concern in multivariate optimization, as well as in the design of a supersonic transport aircraft. The lessons learned concerning the latter, will be discussed in this paragraph.

The insufficient accuracy and sensitivity to noise of the supersonic aerodynamic parameters as obtained by the NLRAERO panel code have been responsible for many problems, which -as discussed- were not solved to full satisfaction. Since it was shown that especially the total configuration lift gradient and the trim lift and drag have a very large influence on the performance of the airplane, it might be necessary to resort to more advanced computational fluid dynamics codes that yield more accurate and less noisy results on which a better regression will be possible.

As indicated in chapter 8, the weight prediction method, for a number of reasons, is rather conservative. A significant part is based on statistical information that represents technology of the 1970's but is no longer representative for the 1990's. Especially in the field of new materials, important weight reductions can be achieved today. In chapter 8 it was explained that the method could only be adapted by very rough estimations (since for instance the percentage of the structure that would be constructed with new materials is not yet known) and that the author does not feel this would much enhance the accuracy of the prediction since the method by nature already has an error of perhaps 20% and furthermore is derived for subsonic airplanes, whereas many uncertainties concerning important weight contributions still remain.

The fact that the prediction most likely overestimates the maximum take-off weight has a significant influence on the final result. It was established that the available design space was extremely small, even without taking the environmental constraints into account. Since the weight of the airplane affects all performance parameters, a more accurate weight prediction might have led to a different result. As has been repeated many times throughout this dissertation, the accuracy of the prediction methods must be as high as possible or else constraints that appear to be just met, may turn out to be violated upon closer examination. This calls for the development of a mass prediction method for supersonic transports that relies less on statistical information.

The same considerations apply to the propulsion analysis, since to some state of the art parameters, values were assigned that may be on the conservative side too. For instance, relaxation of the turbine entry temperature constraint, inherently enforced during the propulsion analysis, would have created additional design space.

In the light of this discussion it is also of great importance to incorporate area-ruling in the design, preferably by means of approximate analysis, since otherwise a large number of extra variables will be introduced. This will also create additional design space and might enable compliance with the take-off noise constraint.

---

Another possibility to comply better with the take-off noise constraint is the use of full-span leading-edge flaps to avoid leading-edge vortices at all points of the flight envelope except during the landing approach. In that way the required thrust at low speeds can be reduced, allowing lower exit jet velocities. This possibility was investigated by carrying out yet another genetic optimization in which the attainable thrust coefficient was set at 1 except during the landing approach. Although the overall merit function as well as the environmental constraint values showed improvement, the performance constraints deteriorated drastically. By assigning weight factors to the different constraint violations this problem might be somewhat controlled, but the main cause of the problem, that is, too many opposing requirements leaving hardly any design space, will remain. The same problems that occurred in the initial phase of the optimization, when the problem of main concern was to find a single feasible design amongst a large amount of incomputable ones, show here as well. If too large a penalty is imposed on performance constraint violations, progress will come to a halt, whereas too large penalties on environmental constraint violations will reduce the merit function at the cost of performance deterioration.

It was shown that during subsonic cruise, a range parameter  $MI_{sp} \frac{L}{D}$  can be obtained that is at least equal to the value during supersonic cruise, implying that subsonic overland stretches will not affect range (of course total flight time and therefore economic viability will be affected). This quite remarkable result is said to apply to the *Alliance* project of Aerospatiale too, which is allegedly capable of covering the same range whether cruising at subsonic or at supersonic speeds. This would leave the sonic boom peak pressure over sea and the take-off noise as the main environmental problems threatening the supersonic transport, in combination with a number of mentioned technological problems like fuel combustion value deterioration, cabin air conditioning and safety after explosive decompression at cruise altitude. Of course, if more strict regulations will be enforced in the field of pollutant emissions, the  $NO_x$  constraint, which in this application was kept equal to the value of Concorde and which proved not critical, might become a problem too.

It has not been the purpose of this dissertation to arrive at a conclusion as to the feasibility of a second-generation supersonic transport. On the contrary, it has been shown that many uncertainties still exist and that a lot of research in a wide field of disciplines still lies ahead. Furthermore, it is clear that any conclusions pertaining to the design of a supersonic transport are biased by the limitations of the analysis from which they result. It has also been shown, however, that multidisciplinary design and optimization is an indispensable tool for a project as complex as the design of a supersonic transport plane. The author is firmly convinced that the technological problems can be solved with the use of MDO techniques. The main problems are caused by environmental issues, some of which may be impossible to solve because they are inherent to supersonic flight (sonic boom), while others may impose such a financial penalty on the supersonic transport that it will never be able to compete with a contemporary subsonic airplane.

---

In the early days, the question of main concern related to supersonic transport design was: "can it be done technologically ?". This question was convincingly affirmed by the designers of Concorde. Then it became: "can it be done economically ?" and the answer to that question still has to be found. Now the question has become whether it can be done within the environmental regulations, both technologically and economically.

The issue of the environmental impact of a second-generation supersonic transport has been predominantly present in this work, as it is in industry today. It has been shown that it is very difficult for a supersonic transport to comply with the imposed environmental constraints and that major adaptations to the design are therefore required. This will lead to a drastic deterioration of the airplane's efficiency. In this sense, the impact of the environment on the supersonic transport might even be bigger than the impact of the SST on the environment. Multidisciplinary Design Optimization techniques will be a necessary tool to solve the technological part of the problem. The feasibility issue will in the end be a financial matter and therefore a matter of regulation. If the tendency to incorporate the environmental issue in the design process is to continue, it is necessary (be it just for economical reasons) to come to a world-wide and stable set of regulations. Only then the economical viability and the risk of development can be assessed and compared to other (subsonic) concepts.

Whatever the outcome, because of the large influence of the environmental constraints on any design, environmental regulation will gradually start to determine the appearance of products of technology and therefore the appearance of the world. If there is a place in such a world for a supersonic transport aircraft of the second generation, remains as yet very uncertain.

Delft, May 13, 1996

## 16.4 References

1. Shahroudi, K.E.; *Development and Validation of a Computer Assisted Design Methodology for Gas-Turbine-Based Aircraft Engines*; Ph.D. dissertation, ISBN 90-407-1070-8, Delft University Press, 1994
  2. Shahroudi, K.E.; *Non-Random Adaptive Grid Method for High-Speed Optimization of Highly Dimensional, Badly Behaving Real-Time Functions*; ASME paper, 94-GT-487, June 1994
-



# Glossary

<b>analysis</b>	The process of determining the values of the objective function and constraints for a given vector of <b>design variables</b> .
<b>approximate analysis</b>	A fast and simple substitute for the slow and complex <b>analysis</b> usually associated with multidisciplinary design.
<b>bounds</b>	Upper and lower limits between which the <b>design variables</b> may be varied.
<b>decomposition</b>	The breaking up of either the design <b>analysis</b> or the <b>optimization</b> process itself into smaller subtasks that can be carried out concurrently and independently.
<b>dependent variables</b>	In the present work also referred to as <b>parameters</b> . These are variables that depend either directly or indirectly on the <b>independent (design) variables</b> and which therefore will change as the design develops. They are not directly controlled by the optimization algorithm however. During the <b>analysis</b> of a design, these variables are the unknowns, whereas the independent (design) variables are assumed to be known.
<b>design evaluation</b>	See <b>analysis</b> .
<b>design variables</b>	Also referred to as <b>independent variables</b> . These variables are directly controlled by the optimization algorithm whose task it is to find such a combination of values for the design variables as to yield the best value for the objective function. The design variables are input in the design <b>analysis</b> and as such considered as constants in the analysis methodology.
<b>direct methods</b>	Optimization methods that use the nonlinear constraints to calculate a search direction that is both usable and feasible by means of gradient information, as opposed to <b>penalty methods</b> that add a penalty term to the objective function in case a constraint is violated or threatens to be violated. Not to be confused with <b>direct search methods</b> .

---

<b>direct search methods</b>	Optimization methods that do not employ gradient information in the search for the optimum. Also known as non-gradient-guided optimization methods. Examples of this class of methods are the genetic algorithms and the Monte Carlo method.
<b>function evaluation</b>	See <b>analysis</b> .
<b>independent variables</b>	See <b>design variables</b> .
<b>move-limits</b>	Temporary <b>bounds</b> on the <b>design variables</b> imposed to safeguard the accuracy of local approximations to the <b>analysis</b> .
<b>optimization</b>	The process of finding the combination of <b>design variable</b> values that yields the best value for a predefined objective function, subject to constraints.
<b>penalty methods</b>	Methods of optimization that add a penalty term to the objective function proportional to the amount of constraint violation (exterior methods) or proportional to the reciprocal or logarithmic value of the constraint (interior method).
<b>parameters</b>	See <b>dependent variables</b> .
<b>parameter study</b>	A survey of the design space with the intention to establish the influence of changes to the independent (design) variables on the dependent variables.
<b>regression surfaces</b>	Simplified relations fitted through a number of <b>samples</b> by means of a least-squares procedure. Also known as <b>response surfaces</b> .
<b>response surfaces</b>	See <b>regression surfaces</b> .
<b>sample</b>	Combination of design point (vector of <b>design variables</b> ) and its corresponding value for the <b>dependent variable</b> of interest, obtained by carrying out the exact <b>analysis</b> .
<b>sensitivity analysis</b>	The determination of the derivatives of certain <b>dependent variables</b> of interest with respect to other dependent variables or <b>independent variables</b> in order to determine the sensitivity of certain characteristics with respect to changes to the design.

---

# Samenvatting

Het doel van dit proefschrift is het onderzoeken van methoden voor multivariabele en multidisciplinaire optimalisatie van conceptuele vliegtuigontwerpen, toegepast op het ontwerp van een supersoon transportvliegtuig (*Supersonic Transport, SST*, ook bekend als *High Speed Civil Transport, HSCT*). In die zin is het werk beschreven in dit proefschrift tweeledig: beschikbare methodieken, geschikt voor het optimaliseren van grote, complexe systemen worden gepresenteerd en geëvalueerd en tevens wordt een analyse procedure voor het doorrekenen van supersone transportvliegtuigen ontwikkeld.

Het grootste probleem met betrekking tot multivariabele optimalisatie is, hoe de informatie te verkrijgen, benodigd voor het naar het optimum leiden van de optimalisatie procedure, binnen een redelijke tijdsduur, gegeven het feit dat een multidisciplinaire ontwerp evaluatie met een voldoende nauwkeurigheid en voorspellende betrouwbaarheid bijzonder tijdsintensief is. De tot dusver bekende methodieken verkrijgen deze informatie ofwel door de gradiënten van de doelfunctie en de begrenzingen met betrekking tot de onafhankelijke ontwerpvariabelen te analyseren, ofwel door een zeer groot aantal ontwerp evaluaties uit te voeren.

Gezien het grote aantal ontwerpvariabelen en begrenzingen kan het aantal benodigde gradiënten tamelijk groot worden, terwijl deze gradiënten voor elk verbeterde ontwerp opnieuw moeten worden berekend totdat het optimum is bereikt.

Omdat de berekeningen voor een vliegtuig ontwerpanalyse nogal complex zijn en vaak iteratieve oplosprocedures bevatten, zijn er geen gesloten analytische relaties beschikbaar voor de benodigde gradiënten. In plaats daarvan moeten ze worden verkregen door eindige differentie (één variabele per keer verstoren en het hele ontwerp opnieuw berekenen) hetgeen zeer veel ontwerpevaluaties vergt (aantal ontwerpvariabelen + 1 voor elke optimalisatiestap in het geval van eenvoudige "forward differencing"). Er is vastgesteld dat de nauwkeurigheid van aldus verkregen gradiënten niet erg hoog is.

Enkele methodieken om, hetzij het proces van het berekenen van de gradiënten, hetzij het optimalisatieprobleem zelf, op te splitsen werden gepresenteerd in de jaren '80 en in dit onderzoek getest. In plaats van het berekenen van globale afgeleiden door het verstoren van het complete ontwerp, worden partiële afgeleiden gegenereerd door eindige differentie van subproblemen. Deze decompositie methoden zijn in het bijzonder goed geschikt voor parallele berekening en het is dan ook hier dat ze het grootste voordeel bieden. Het gebruik van decompositie methoden in combinatie met tweede-orde optimalisatie algoritmen was echter geen volledig succes. Ook blijft het probleem van de onnauwkeurige gradiënten.

Het gebruik van niet-gradiënt geleide algoritmen, met name de genetische algoritmen die tegenwoordig een snel stijgende populariteit ondervinden, is eveneens onderzocht in dit proefschrift. Het grote aantal vereiste ontwerpevaluaties en het probleem van voortijdige convergentie zijn onderkend als de grootste nadelen van dit type optimalisatie. Bovendien, in tegenstelling tot gradiënt methodes, is er geen garantie

---

dat het optimum is bereikt of zelfs zal worden bereikt. Aan de andere kant zal de methode waarschijnlijk een globaal optimum vinden, in plaats van te convergeren naar een lokaal optimum. Daarom is de methode robuuster dan een gradiënt geleide methode, tenzij voortijdige convergentie optreedt.

Omdat de rekeninspanning benodigd voor een multivariabel/multidisciplinair optimalisatieprobleem buitengewoon groot is, onafhankelijk van de gebruikte optimalisatie methode, moet de belangrijkste vooruitgang worden geboekt in het reduceren van de rekentijd, ofwel door gebruik te maken van hoge-snelheids computers (duur) ofwel door de tijd, benodigd voor een enkele ontwerpevaluatie, te reduceren. Het laatste kan worden bereikt door de analyse op slimme wijze te rangschikken, of door een snelle en simpele benaderende analyse te gebruiken, in plaats van de langzame, gedetailleerde. De eerste optie, om de analyse op een slimme manier te rangschikken omvat het tot een minimum beperken van terugkoppelingen die een iteratieve oplossing noodzakelijk maken, het gebruiken van slimme schattingen voor waarden die aan het begin van de analyse nodig zijn maar pas exact bekend zijn aan het eind, zonder opnieuw te itereren, kortom, het inruilen van nauwkeurigheid voor snelheid, vanzelfsprekend binnen redelijke grenzen.

De tweede optie, om een eenvoudige benaderende analyse te gebruiken in plaats van de uitgebreide, is op grote schaal gebruikt in dit onderzoek. Er zijn twee soorten benaderende analyse. De eerste soort is een lokale benadering, in het huidige ontwerppunt, die na iedere optimalisatiestap moet worden gecorrigeerd met behulp van de exacte analyse. Het eenvoudigste voorbeeld van deze vorm van benaderende analyse zijn de eerste-orde afgeleiden van het huidige ontwerp met betrekking tot de onafhankelijke ontwerpvariabelen zoals die door ieder gradiënt afhankelijke optimalisatie algoritme worden gebruikt.

De tweede vorm is globale benadering door middel van regressie vlakken, in het huidige onderzoek in de vorm van tweede-orde polynomen. Dit concept heeft vele voordelen. Niet alleen verkort het de benodigde rekentijd (hetgeen, zoals reeds eerder uitgelegd, voordelig is voor elke optimalisatie methodiek), de benadering kan bovendien van te voren worden gegenereerd en hoeft niet na elke optimalisatiestap weer te worden herzien. De benaderende analyse kan worden gebruikt om resultaten mee te nemen van disciplines die door aparte programma's worden geanalyseerd die niet binnen het optimalisatie algoritme kunnen worden geïmplementeerd of die eenvoudig te tijdsintensief zijn om te overwegen ze in een optimalisatiecyclus op te nemen. Globale benadering is ook een middel om analyses mee te nemen die op een geografisch afgelegen plaats (een andere afdeling of een ander bedrijf) worden uitgevoerd. Niet in de laatste plaats levert het gebruik van tweede-orde polynomen de relaties tussen de doelfunctie, de begrenzingen en de onafhankelijke ontwerpvariabelen als eenvoudige, gesloten functies. Gradiënt informatie kan daarom snel en gemakkelijk worden berekend.

In dit onderzoek wordt een hybride optimalisatie methodiek gebruikt, bestaande uit een globale evaluatie met behulp van een genetisch algoritme om de ontwerpruimte te verkleinen, gevolgd door een gradiënt afhankelijke optimalisatie om het exacte

optimum te lokaliseren. De globale optimalisatie maakt gebruik van de exacte analyse, waaruit alle overbodige iteraties zijn weggelaten en benaderende analyse om sommige disciplines mee te kunnen nemen. De gradiënt geleide optimalisatie maakt slechts gebruik van benaderende analyse, gegenereerd voor de gereduceerde ontwerpruimte, waarin het optimum zich bevindt, zoals dat door de globale optimalisatie werd gevonden. Op deze manier worden de goede eigenschappen van gradiënt afhankelijke methodes (efficiënt, exact) en globale zoekmethodes (robuust) gecombineerd.

De optimalisatie wordt uitgevoerd op een tweede-generatie supersoon transportvliegtuig (SST). De hoofdkenmerken van het analyse programma zijn:

- een voor veranderingen in het ontwerp gevoelige vleugelgewichts predictie methode
- het gebruik van een panelen methode voor het verkrijgen van subsone en supersonische aërodynamische eigenschappen van hoge-snelheids configuraties
- toepassing van de *"leading-edge suction analogy"* van Polhamus en het *"attainable thrust concept"* van Carlson om een gedeeltelijk losgelaten voorrandstroming en de daaruit resulterende voorrandwervels, belangrijk voor sterk gepijlde vleugels onder invalshoek, in rekening te brengen
- een procedure om de prestaties van dubbelstroomstraalmotoren, ontworpen voor supersoon vliegen, buiten het ontwerppunt te voorspellen, inclusief een gedetailleerde inlaatroutine om in het ontwerppunt een optimaal inlaattrendement te realiseren en om spilweerstand en gondelweerstand in de analyse mee te nemen
- een schatting van het effect van een tweede-generatie supersoon verkeersvliegtuig op het milieu



# Curriculum Vitae

The author was born January 11, 1968 in Winschoten, Holland. After completing primary school and high school in Emmen, he enrolled in 1986 at the Delft University of Technology, Faculty of Aerospace Engineering, specializing in aerodynamic aircraft design in the Design and Flight Mechanics group (*vakgroep a2L, Ontwerpen en Vliegmechanica*) during his final studies. In December 1991 he graduated as a Master of Science on a conceptual design study of a hypersonic cruising transport plane under direction of prof.ir. E. Torenbeek. In February 1992 he became a post-graduate student in the group of prof. Torenbeek.

---