AGARD-R-789

AGARD-R-789

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT
7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD REPORT 789

Stability in Aerospace Systems

(La Stabilité dans les Systèmes Aérospatiaux)

Papers presented at the Guidance and Control Panel Workshop on Stability in Aerospace Systems, held at ONERA in Mauzac (Toulouse), France from 23rd—25th June 1992.

NORTH ATLANTIC TREATY ORGANIZATION

Published February 1993

Distribution and Availability on Back Cover

AGARD-R-789

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT
7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD REPORT 789

Stability in Aerospace Systems

(La Stabilité dans les Systèmes Aérospatiaux)

Papers presented at the Guidance and Control Panel Workshop on Stability in Aerospace Systems, held at ONERA in Mauzac (Toulouse), France from 23rd—25th June 1992.

Note: Exceptionally, some authors have not submitted texts of their papers, only copies of the viewgraphs used during their presentations. AGARD apologises for the lack of continuity in these papers, and the poor quality of some, and hopes that readers will gain some understanding of the subject matter from them.

North Atlantic Treaty Organization
Organisation du Traité de l'Atlantique Nord

The Mission of AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

— Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;

— Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);

— Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;

— Improving the co-operation among member nations in aerospace research and development;

— Exchange of scientific and technical information;

— Providing assistance to member nations for the purpose of increasing their scientific and technical potential;

— Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced
directly from material supplied by AGARD or the authors.

Published February 1993

ISBN 92-835-0702-9

SPS

Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

Theme

Stability in guidance and control can be illustrated by the following examples:

— Air vehicles subjected to atmospheric instabilities (particularly downbursts and windshears), including flight control in downbursts, detection of downbursts from the ground or by airborne systems, and aircraft guidance to avoid them.

— Control of fighters at high angles of attack which may become unstable due to a bifurcation point beyond which their behaviour becomes unpredictable, though belonging to the deterministic domain.

— Control of laminar-turbulent transition for high performance aircraft (subsonic and supersonic), including improvement of stability by control of transition, and the impact of fully laminar flow on aircraft control.

The concept of stability also needs to be looked at from a theoretical point of view — for example, the impact of uncertainties in initial conditions on system behaviour. This has links with the concept of model robustness.

The programme was as follows:

1st Session. Fundamental aspects of stability with examples:
The dilemma of stability vs. manoeuvrability and its consequences on military and civilian aircraft; stability and controllability of non-linear systems; various types of stability of an equilibrium state.

2nd Session. Basic theoretical aspects and chaos:
Newtonian mechanics and thermodynamics; transition from stability to chaos; boundary layer control on a wing or fuselage, stabilisation of an unstable aircraft (decoupling of the controls of aircraft, helicopters, or convertibles).

3rd Session. Applications of aerospace techniques:
a) External parameters: the atmosphere, turbulence, windshears and downbursts; their detection and modelling (methods of detection — whether ground or air based — the information processing of the data collected); control laws when flying through atmospheric disturbances; criteria.

b) Control of the flow around wings and fuselage. C_L and C_D variations with boundary layer effects. Stability of a spin. Prediction of the instability and consequences for vehicle guidance.

Thème

Les problèmes de stabilité dans l'optique du guidage et du pilotage peuvent être illustrés par les exemples suivants:

— Véhicules aériens soumis aux instabilités atmosphériques (particulièrement les cisaillements de vents) incluant le contrôle du vol dans un "downburst" ainsi que la détection de ceux-ci, soit du sol, soit à bord de l'avion, et le guidage de l'avion pour éviter ces phénomènes.

— Contrôle des chasseurs à des très grandes incidences qui peuvent devenir instables à cause de "bifurcations" derrière lesquelles le comportement des véhicules devient imprévisible, bien qu'appartenant au domaine "déterminé".

— Contrôle de la transition laminaire-turbulent pour les avions à hautes performances (subsoniques ou supersoniques), incluant l'amélioration de la stabilité grâce au contrôle de la transition et l'impact d'un écoulement totalement laminaire sur le comportement de l'avion.

Le concept de stabilité doit être également considéré d'un point de vue théorique, par exemple, en examinant les conséquences des incertitudes sur les conditions initiales sur le comportement du système. Ce problème est lié, d'ailleurs, au concept plus général de la "robustesse" d'un modèle.

Le programme de l'atelier a été le suivant:

1ère Session. "Aspects fondamentaux de la stabilité et exemples"
Le dilemme stabilité/manoeuvrabilité et ses conséquences sur la stabilité des aéronefs civils et militaires; la stabilité et la pilotabilité des systèmes non-linéaires; différents types de stabilité d'un "point" d'équilibre (état).

2ème Session. "Aspects théoriques fondamentaux et chaos"
Mécanique newtonienne et thermodynamique; transition stabilité/chaos; contrôle de la couche limite sur une aile ou un fuselage, stabilisation d'un avion instable (découplage des commandes dans un avion, un hélicoptère ou un convertible).

3ème Session. "Applications aux techniques aérospatiales"

a) Les paramètres externes: l'atmosphère, la turbulence, les cisaillements de vents et les "downbursts"; leurs détecteurs et leur modélisation (méthode de détection — sol ou aéroportée — et traitement de l'information collectée); lois de contrôle en vol en atmosphères très turbulentes; critères.

b) Contrôle de l'écoulement autour des ailes et du fuselage; C_Z et C_X lorsqu'on prend en compte la couche limite; stabilité d'un avion en vrille; prédiction de l'instabilité et conséquences pour le guidage du véhicule.

Guidance and Control Panel

Chairman: Mr S. Leek
British Aerospace
Defence Dynamics Ltd
PO Box 19
Six Hills Way, Stevenage
Herts SG1 2DA
United Kingdom

Deputy Chairman: Mr J.K. Ramage
Chief, Flight Control
Advanced Development Branch
Wright Laboratory (WL/FIGX)
Wright-Patterson AFB,
OH 45433
United States

TECHNICAL PROGRAMME COMMITTEE

Chairman:	Dr M.J. Pelegrin	(FR)
Members:	Dr A. Benoît	(BE)
	Prof. R.C. Onken	(GE)
	Mr J. Bardal	(NO)
	Mr J.K. Ramage	(US)

HOST NATION COORDINATOR

Dr M.J. Pelegrin
Haut Conseiller à l'ONERA
ONERA/C E R T
BP 4025
2 avenue Edouard Belin
F-31055 Toulouse Cedex
France

PANEL EXECUTIVE

Commandant M. Mouhamad, FAF

Mail from Europe:
AGARD—OTAN
Attn: GCP Executive
7, rue Ancelle
F-92200 Neuilly-sur-Seine
France

Mail from US and Canada:
AGARD—NATO
Attn: GCP Executive
Unit 21551
APO AE 09777

Tel: 33(1)47 38 57 80
Telex: 610176 (France)
Telefax: 33 (1) 47 38 57 99

ACKNOWLEDGEMENTS/REMERCIEMENTS

The Panel wishes to express its thanks to the French National Delegates to AGARD for the invitation to hold this meeting in Mauzac and for the facilities and personnel which made the meeting possible.

Le Panel tient à remercier les Délégués Nationaux de la France près l'AGARD de leur invitation à tenir cette réunion à Mauzac et de la mise à disposition de personnel et des installations nécessaires.

Contents

SESSION I — FUNDAMENTAL ASPECTS OF STABILITY WITH EXAMPLES

Chairman: Dr A. Benoît (BE)

SESSION II — BASIC THEORETICAL ASPECTS AND CHAOS

Chairman: Mr S. Leek (UK)

---

* FDP contribution.
** FMP contribution.
† Not available at time of printing.

SESSION III – APPLICATIONS TO AEROSPACE TECHNIQUES
Chairman: Dr C.H. Houpis (US)

---

* FDP contribution.
** FMP contribution.

INTRODUCTION

by Dr M.J. PELEGRIN, Programme Committee Chairman

Is stability a measurable quantity - like mass or an identifiable quantity - like temperature? There are many definitions of stability, sometimes contradictory: in fact, it is a subjective quantity which should be defined in the context of the theme considered. The reference system in which the system evolves should be defined: stability may exist in a given reference system, but no longer exists in other reference systems. Stability seems to be a dominant factor for aircraft or missile control - or any type of vehicle. However, stability and manoeuvrability are two opposing factors which intervene in aircraft control: for civilian aircraft stability is the dominant factor; for military aircraft or missiles manoeuvrability is the dominant factor. The above are some of the reasons which led to the organization of a Workshop on "Stability" for the AGARD community.

Basically, stability is related to irreversibility which means energy dissipation for linear systems, but linear systems are very rare though they also often represent a suitable approximation of non linear systems. Stability is also a matter of accuracy. Let's take the earth's rotation: is it stable or unstable? This question has no meaning until the range of accuracy we are looking for, and in fact, the whole context is specified. Due to the accuracy of existing atomic clocks, it is demonstrated that daily variations are of the order of 1ms yearly or pluri-annual variations of the order of tens of ms, occur in a pseudo periodic manner. However, the angular velocity is necessarily decreasing on a long-range basis: this is mainly due to the water/earth friction of tides. In the pre-Cambrian period (400 M years) the day was 15 hours! What has been said about the angular velocity of the earth could also be said about the direction of the earth's momentum. At the pole the trace of the rotational vector moves continuously in a circle of about 2m in diameter. However, for all human activities the earth's rotation is considered (except by some astronomers) as stable.

Poincaré in the 1870s studied stability for non-autonomous and autonomous problems. Ljapounov in the 1900s introduced a way of proving whether or not stability was sufficient, but not the necessary conditions. Thereafter the behaviour of a system in the vicinity of an "equilibrium point" was studied in detail (Poincaré) and equilibrium points or "singularities" were classified as nodes-summits-focus-saddle.

A variety of possible behaviour in the vicinity of a point, the limit cycles, which can be stable or unstable, were introduced; they generalized the stability point by letting a periodic motion, normally of a small amplitude, around the stability point (in the phase plane or space).

Recently, new vocabulary has been introduced: strange attractor instead of equilibrium point or limit cycle.

All chapters of physics are affected by the concepts of stability. Mechanical systems were the first to be affected by Poincaré's approach. Theories and studies concern non linear systems. All that can be said about linear systems has been said at the present time. No global solution is expected. The robustness concept of a control system is an extension of linear systems studies. This concept is important for applications in industry or vehicle control. Robustness can be defined as the capacity (capability) of coping with the specified performances in spite of some unknown concerning the parameters which define the system to be controlled - or sometimes, the controller parameters themselves. Obviously, a linear differential equation with such uncertainties on coefficients is no longer a linear equation. In fact, even in the beginning of linear system studies "phase margin" and "gain margin" were used to compensate for some errors in the system description. Nowadays more elaborate techniques such as H_∞ optimization enables us to deal with multi-input, multi-output systems.

The stability concept is widely used in thermodynamics, at least before statistical thermodynamics came into being. Irreversibility - at least from a practical point of view - leads to the entropy concept, which simply says that "in an autonomous system entropy can only increase". Irreversibility is no longer accepted as a universal law in statistical thermodynamics, namely those who study "chaos". Maxwell's devil could operate ... if we wait an appropriate very long time.

We can probably say that the dilemma "stability-instability" has mainly progressed in fluid mechanics, aero and hydro dynamics and spectacular (in both meanings of the word) results arose from the Bernard's curls; they directly derive from a non-organized structure; order came up from disorder! If the heating of the fluid is adjusted, it can be stable as long as we do not modify the heat transfer (it is a non-autonomous system).

Nonetheless, in this domain, the aerodynamic flow around a wing can be stable though instability may locally appear in the boundary layer: during a "normal flight" the boundary layer becomes turbulent (i.e., locally unstable) somewhere between one-half or two-thirds of the wing chord. Buffet phenomena is due to the escape of curls from the boundary layer, a phenomena which should be avoided for aircraft performance and passenger comfort. Flutter phenomena which is very dangerous, is due to aero-elastic coupling between air-flow and wing elasticity: it appears when the frequencies of 2 modes converge to a unique frequency (normally the 1st bending and 1st torsion modes). This is typically a case of instability which can gradually arise from stability when some flight parameters vary (velocity and/or load factor).

Stability - or instability - has also a meaning in static structures (bridges */, buildings, dams, earth embankments, ship or aircraft structures, etc.) when loads reach a given value: this is the buckling. Even in the earth's crust, instability appears (earthquakes). Roughly speaking, it can be said that due to the tectonic plate motions, when, in a given location, the friction constraint is surpassed, the sliding effect or the elastic deformation of a part of ground is suddenly transformed into a jump.

Many other subjects could be mentioned. To conclude, I would like to briefly mention atmospheric conditions. Has stability a meaning? Probably not, though in many countries, like those in the temperate zone, weather has a certain degree of stability: if you say "Tomorrow the weather will be be like today's" you are not making a bad forecast! (The probability of success is well above 50% since in these countries weather does not change every day - a certain degree of stability exists.) Weather is a consequence of air movement over the world: it should be predictable as for any system for which the equations are known. Unfortunately, air movement is governed by partial derivative equations - they are known with a reasonable certainty -but a set of homogeneous initial conditions (3-D) is___ to be acquired even by meteorological satellites. Meteorologists proceed by region (they use some grids, ranging in size from a few kilometers to hundreds or thousands of kilometers) and try to start with coherent sets of initial conditions on the boundaries of these grids. The computer then solves the equations and they arrive at a correct(?) prevision for 24 or now 48 hours. Their goal is to achieve predictions with the same degree of accuracy for a 72-hour period before the end of the century.

This Workshop will consider these questions in depth during 3 days. The programme of the Workshop is given in the Table of Contents.

---

*/ The Tacoma suspension bridge which collapsed on 7 Nov 1940 was subjected to a relatively low wind (18.7 m/s). It was a typical phenomena similar to flutter (conjecture of 2 vibrating modes).

EXPOSÉ D'OUVERTURE

par

Marcel Bénichou
Président de l'Office National
d'Etudes et de Recherches Aérospatiales
BP 72 - 92322 Châtillon Cedex
France

Mesdames, Messieurs,

Dès la fin du 19ème siècle, le problème de la stabilité des aéroplanes a été au centre des préoccupations des concepteurs de machines et des scientifiques qui se penchaient sur leurs projets.

Cet intérêt ne s'est pas démenti au cours du siècle écoulé depuis, passant de la recherche de la stabilité "naturelle" des aéronefs, à la recherche du contrôle de machines instables, en d'autres termes à la recherche de la stabilité artificielle ; passant de la nécessité de réunir les conditions d'un pilotage praticable, puis sûr, à des ambitions plus diverses.

En fonction des missions des aéronefs, des critères empiriques de stabilité ont dû être définis, différents pour les avions de transport et pour les aéronefs de combat, du fait d'exigences contradictoires en termes de maniabilité et de stabilité. On se contentera de citer la recherche relativement récente du meilleur compromis maniabilité-stabilité aux très hautes incidences – dépassant 50° – pour accroître les possibilités du combat aérien.

D'autre part, les véhicules réels étant flexibles et non rigides, la maîtrise de l'aéroélasticité est vite devenue un défi majeur, avec des aspects statiques, et dynamiques tels que le flottement, défi d'autant plus grand que le domaine de vol s'élargissait sans cesse.

Enfin, les phénomènes de buzz d'entrée d'air ont limité les excursions à Mach élevé ou obligé à mettre en place des dispositifs permettant de les éviter.

On voit qu'à beaucoup d'égards, et bien plus que d'autres secteurs d'activité, l'aéronautique a dû et doit prendre en compte la notion d'instabilité.

J'ai dit "l'aéronautique", je devrais dire bien sûr "l'aéronautique et l'espace" puisque les instabilités de combustion des propergols dans les propulseurs des lanceurs spatiaux ont causé bien des soucis, et puisque dans la conquête de l'espace, en raison des délais de transmission qui peuvent atteindre plusieurs minutes, les systèmes spatiaux doivent être dotés d'une large autonomie et donc avoir recours à des architectures de contrôle boucle fermée dont la stabilité doit être assurée.

Pour surmonter les difficultés, et parallèlement à la recherche de palliatifs empiriques, les tentatives de modélisation des phénomènes, pour mieux les maîtriser, se sont développées, en prenant appui sur la théorie et les simulations, ce qui est une démarche classique. Mais au fur et à mesure que les difficultés étaient surmontées, l'accroissement recherché du domaine de vol, en vitesse, altitude et incidence, nécessitait de nouveaux efforts.

Les prouesses du calcul numérique, la miniaturisation des organes de puissance et l'augmentation de leur bande passante sont venues à point nommé pour permettre effectivement de contrôler l'instabilité dans ces domaines élargis.

Un vaste champ d'analyse reste néanmoins à défricher en grande partie ; celui du comportement humain et de l'interface homme-machine. La tâche de pilotage la plus fréquente est une tâche de poursuite de cible : au sens propre lorsqu'il s'agit de combat aérien, au sens figuré lorsqu'il s'agit par exemple de rejoindre une plate-forme d'atterrissage mobile telle que le pont d'un porte-avions par mer forte. Les limitations naturelles de l'homme et la variabilité de ses caractéristiques (fatigue, sommeil, distraction, etc.) font qu'on n'a pas fini de parler de pompage pilote.

De grands progrès ont donc été accomplis selon un processus faisant de plus en plus appel à la recherche appliquée.

Néanmoins, et en parallèle, on a assisté et on continue à assister à un effort soutenu dans le domaine de la recherche fondamentale et en particulier dans la recherche des règles du "chaos déterministe", auxquelles peuvent obéir les systèmes non linéaires, avec en particulier l'ambition de mieux comprendre au niveau le plus fin les phénomènes de turbulence et les écoulements instationnaires, sans se limiter à les modéliser au niveau macroscopique, ou à repousser l'apparition de la turbulence de couche limite par recours à des moyens empiriques.

Votre ordre du jour reflète les deux tendances : recherche fondamentale et recherche appliquée, de façon assez équilibrée d'ailleurs ; le chaos tient une place très importante dans les deux premières sessions. La troisième session, consacrée aux applications, se préoccupe essentiellement des problèmes de contrôle du vol et de turbulence avec un fort accent sur le vol à très haute incidence, domaine où les ex-soviétiques peuvent apporter demain le fruit de leur expérience, et qui répond à un besoin explicite des utilisateurs militaires.

***

Cependant, deux questions surgissent, à l'examen de cet ordre du jour.

La première concerne le passage au stade applicatif dans l'aéronautique et l'espace, des recherches concernant les phénomènes chaotiques : cette échéance est-elle proche ou lointaine ?

Si l'on considère par exemple la transition laminaire-turbulent, l'amélioration des connaissances, par la voie expérimentale, ne s'est pas accompagnée d'une amélioration des méthodes pratiques de prévision, qu'il s'agisse des transitions par by-pass, ou pour les transitions naturelles, de méthodes innovatrices par rapport à celles basées sur la loi dite "en e^n".

Pour ce problème et, plus généralement, pour une meilleure compréhension des phénomènes de turbulence, faudra-t-il attendre les progrès futurs de l'informatique pour en donner une représentation de plus en plus fine afin d'appréhender des tourbillons de plus en plus fins ? Ou une autre approche est-elle envisageable dès aujourd'hui ?

***

La seconde interrogation se situe sur un autre plan. Elle est liée à ce que j'appellerai l'intrusion de l'informatique dans le traitement en temps réel, sur les aéronefs, des problèmes de stabilité et, plus généralement, de commandes de vol.

Le progrès apparemment sans limite de la puissance des calculateurs conduit à faire prendre en charge par les dispositifs automatiques, amortisseurs ou amplificateurs, un nombre croissant de variables (non seulement position des gouvernes classiques, mais nombre de gouvernes, régime moteur, configurations, etc.) dans des domaines de vol de plus en plus grands, pour optimiser la manœuvrabilité ou les consommations.

Les risques augmentent alors : risque d'erreurs dans l'analyse des cas susceptibles d'être rencontrés et des phénomènes susceptibles d'intervenir ; risque d'erreurs dans la réalisation des logiciels ; à ces risques d'erreurs ou d'omissions s'ajoute le risque d'introduction de fausses solutions : la simulation numérique de phénomènes physiques régis par des équations aux dérivées partielles comportant des termes non linéaires, peut entraîner, du fait de la discrétisation, des comportement parasites, c'est-à-dire d'origine purement numérique. Certains centres de recherches tentent de progresser dans ce domaine. De façon plus banale il y a le risque de converger vers une solution qui n'est pas la bonne du fait de la non-unicité de solution même si les calculs sont parfaits. Ces risques ne sont pas seulement théoriques : des accidents et incidents récents, qui sont de nature à faire mettre en cause les logiciels, ont sensibilisé les bureaux d'étude des constructeurs.

Un grand hebdomadaire de la presse spécialisée, amplifiant le mouvement d'inquiétude, titrait récemment, sans ambages : "La complexité des logiciels embarqués menace la sécurité des avions".

On peut surtout se demander s'il suffira de faire des progrès dans la validation des programmes informatiques des calculateurs de bord, suivant les procédés actuellement éprouvés, ou si des réflexions plus profondes ne s'imposent pas.

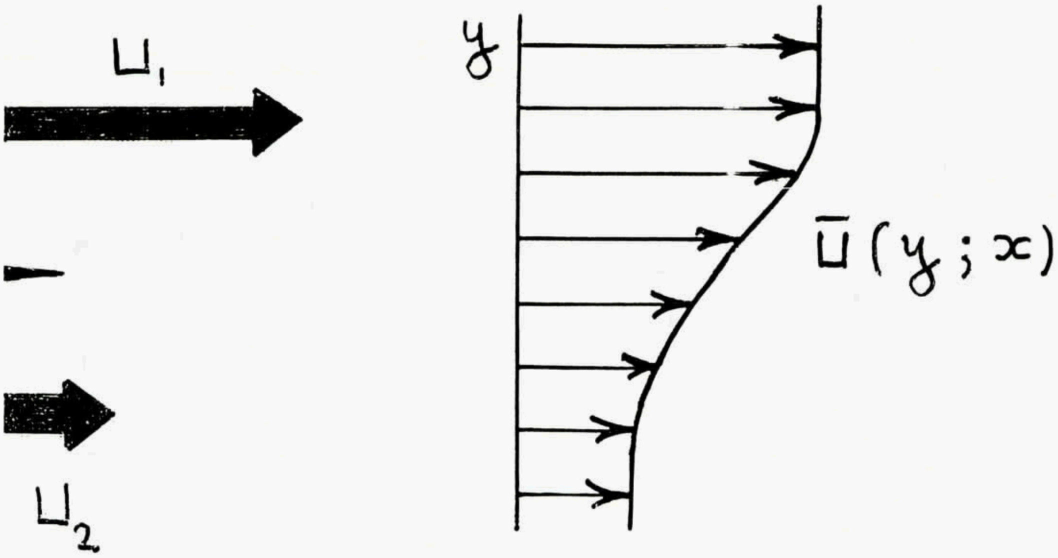
Je vous remercie de votre attention.

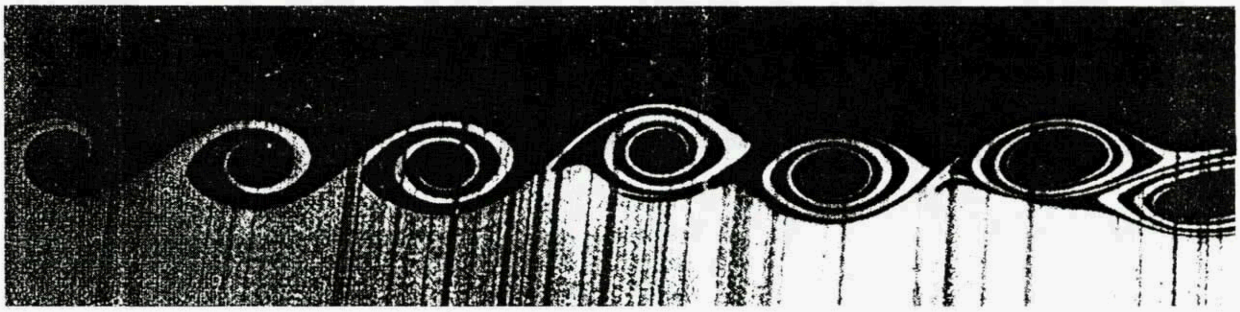
STABILITE HYDRODYNAMIQUE DES ECOULEMENTS CISAILLES
(HYDRODYNAMIC STABILITY OF FREE SHEAR FLOWS)

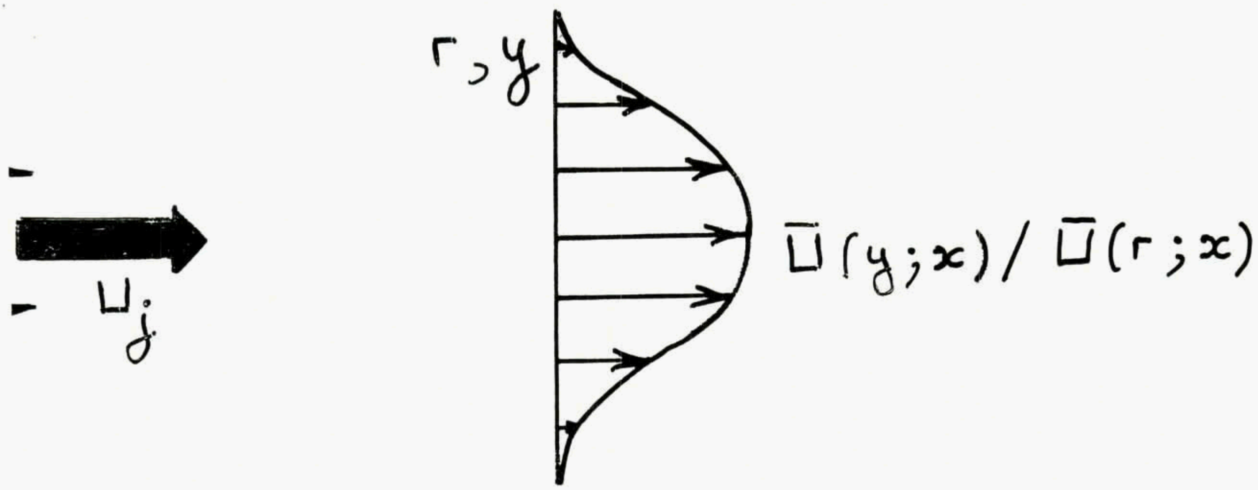
par

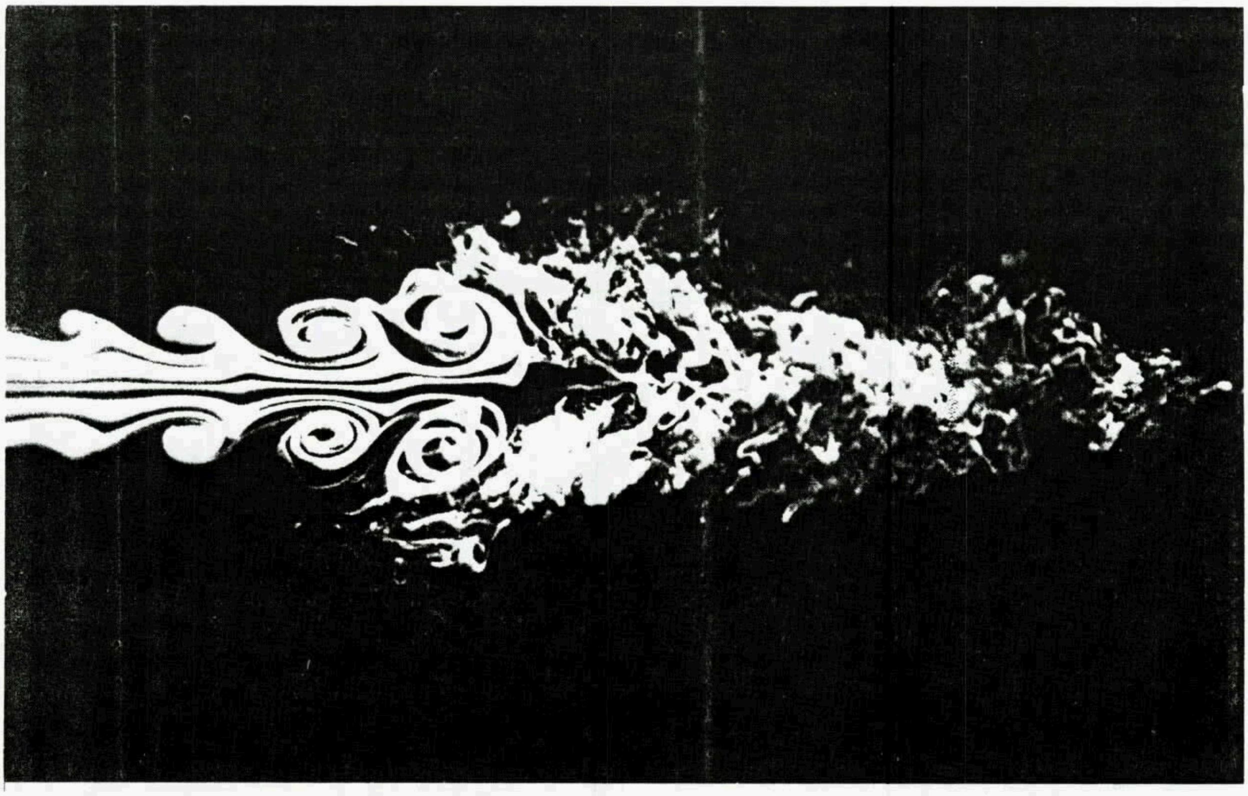
P. Heurre
Laboratoire d'Hydrodynamique
Ecole Polytechnique
91128 Palaiseau Cedex
France

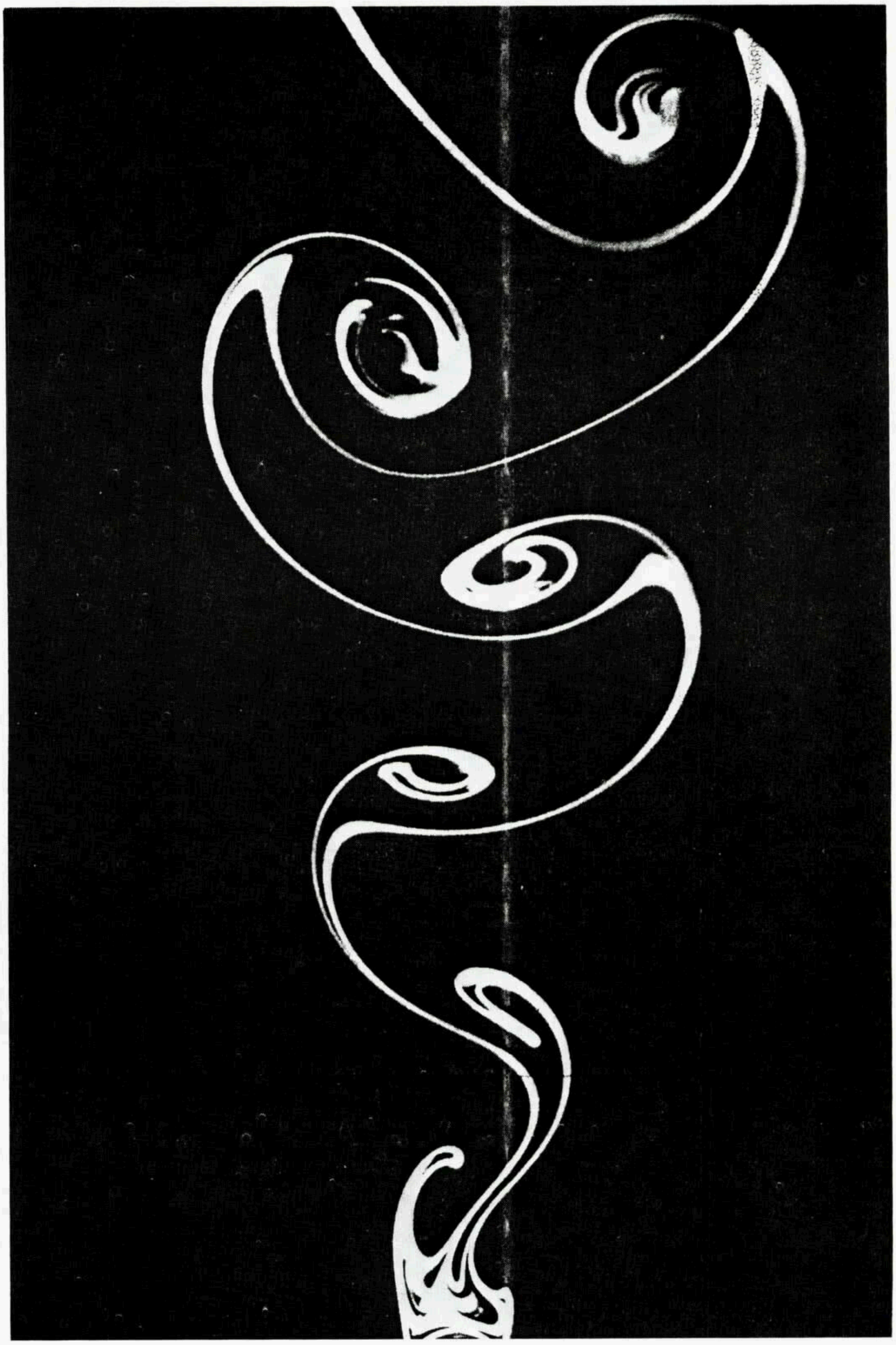
L.G. Redekopp
USC
Los Angeles
California
United States



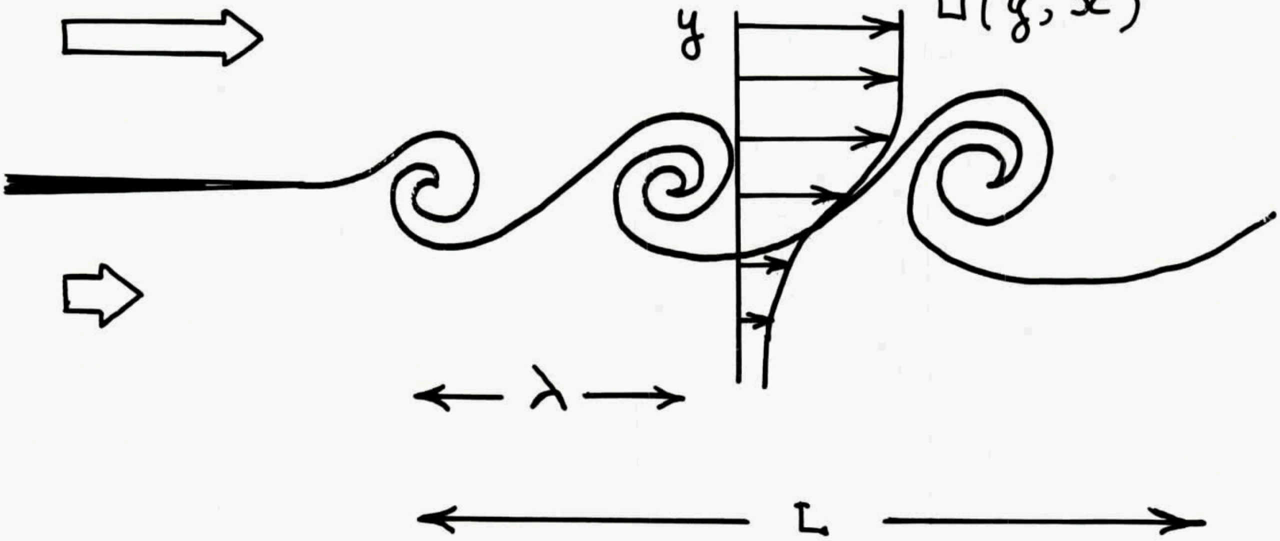



r, y
$\bar{U}(y;x) / \bar{U}(r;x)$
U_j





APPROCHE "DÉTERMINISTE" DES ÉCOULEMENTS TURBULENTS.



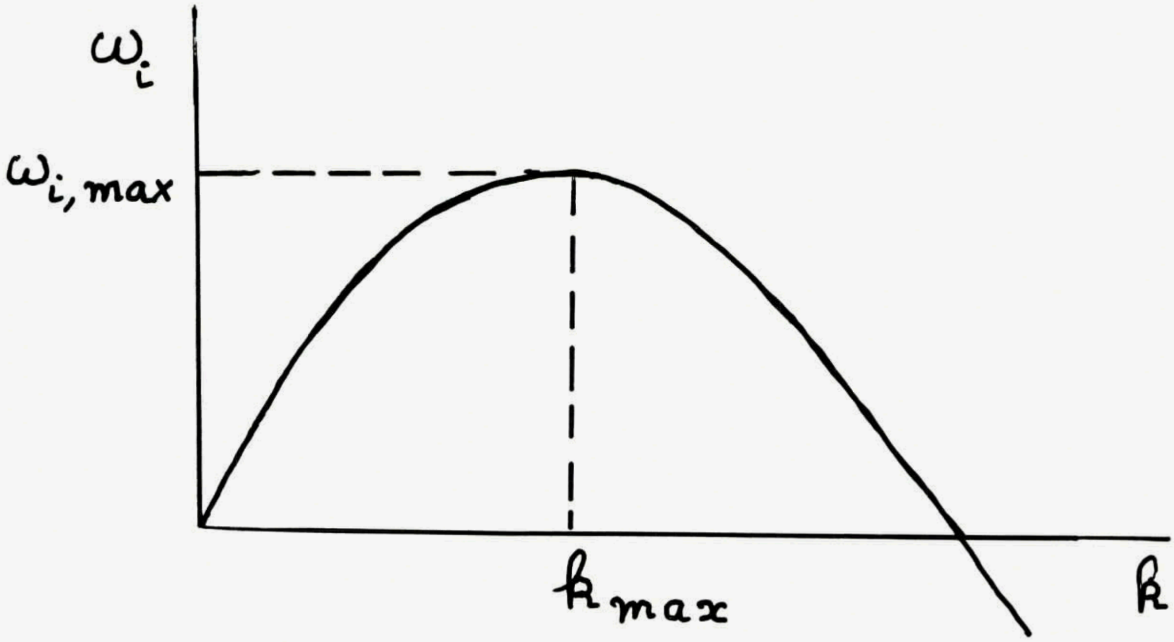
$$\boxed{\varepsilon = \frac{\lambda}{L} \ll 1}$$

Les structures cohérentes sont des ondes d'instabilité caractérisées par une relation de dispersion

$$D(k, \omega\ ;\ X, R, \dots) = 0$$

LOCAL CONCEPTS

$$D[k, \omega ; R] = 0$$



$$\frac{\partial \omega_i}{\partial k}(k_{max}) = 0$$

$\omega_{i,max}$	> 0	unstable
$\omega_{i,max}$	< 0	stable

Fig 1

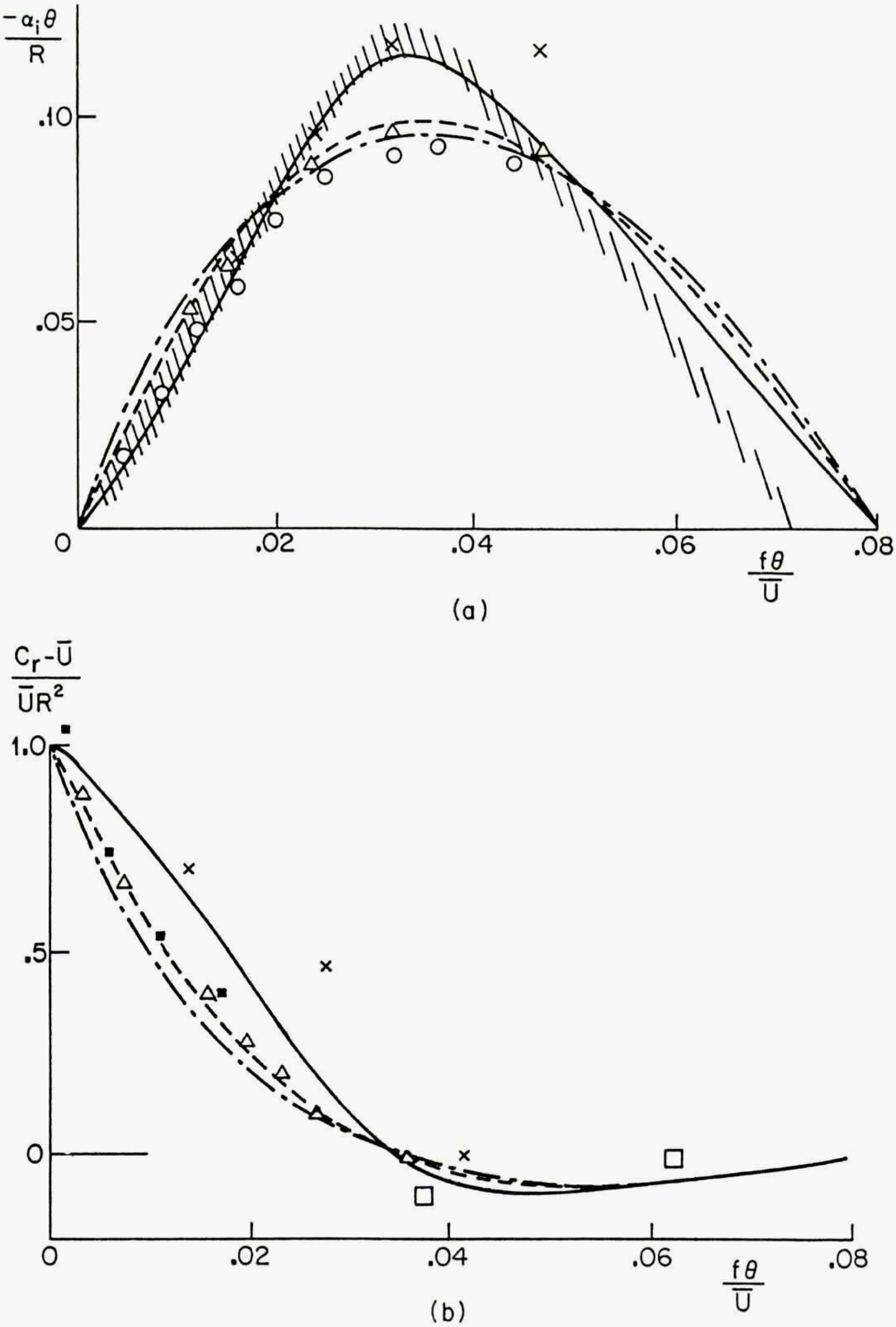


Fig 2 Variations of (*a*) normalized amplification rate and (*b*) normalized phase velocity with Strouhal number $f\theta/\bar{U}$. Linear stability theory (from Monkewitz & Huerre 1982): — $R = 1$; —·— $R = 0.5$; ——— $R \ll 1$. Experiments: □ $R = 1$ (Sato 1960); ○ $R = 1$ (Freymuth 1966); × $R = 0.72$ (Miksad 1972); \\\ $R = 1$ (Fiedler et al. 1981); △ $R = 0.31$ (Ho & Huang 1982); ■ $R = 1$ (Drubka 1981).

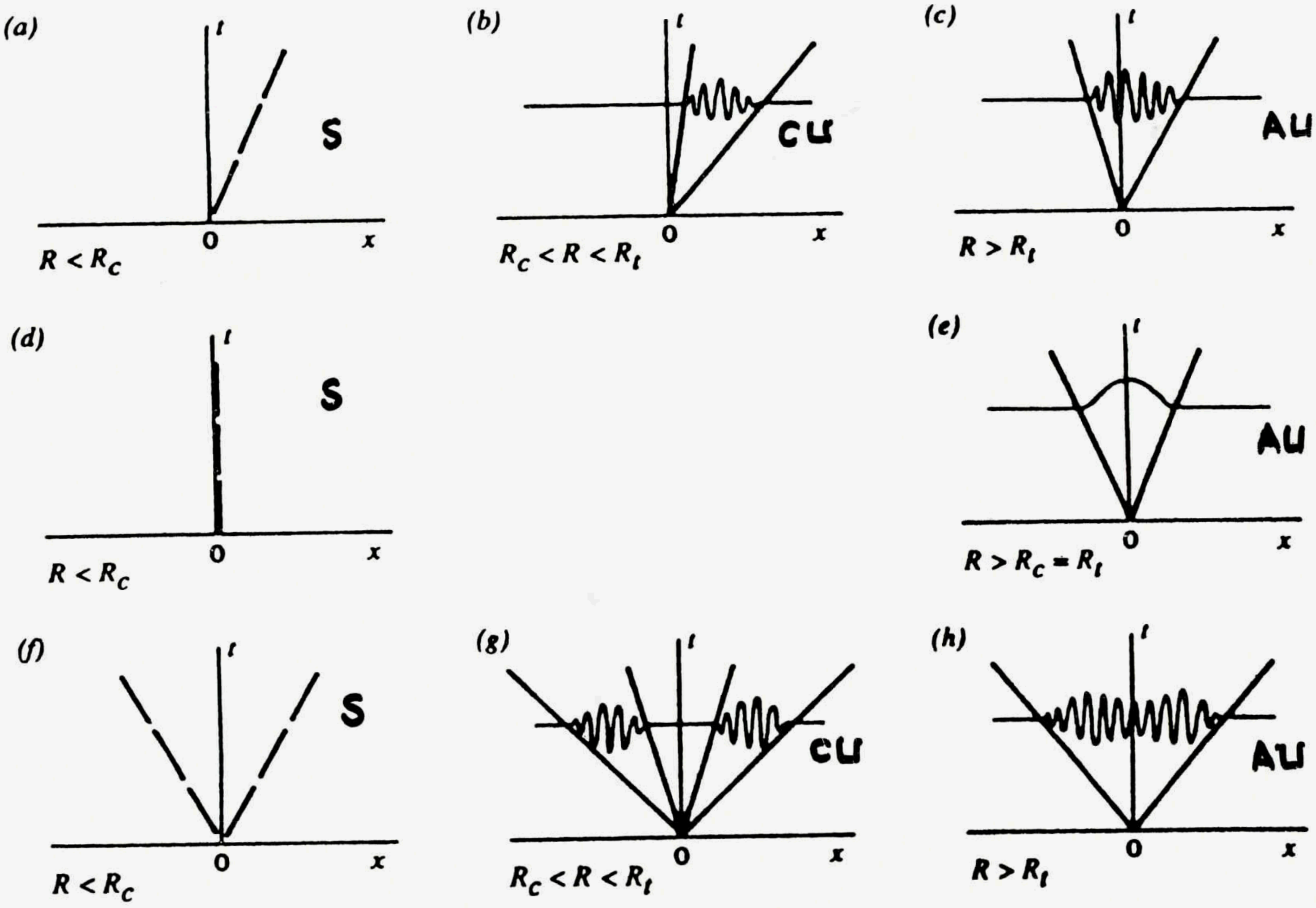


Figure 3 Sketches of typical impulse responses. Single traveling wave: (*a*) stable, (*b*) convectively unstable, (*c*) absolutely unstable. Stationary mode: (*d*) stable, (*e*) absolutely unstable. Counterpropagating traveling waves: (*f*) stable, (*g*) convectively unstable, (*h*) absolutely unstable.

Impulse response of a flat plate
boundary layer.
Tollmien-Schlichting instability

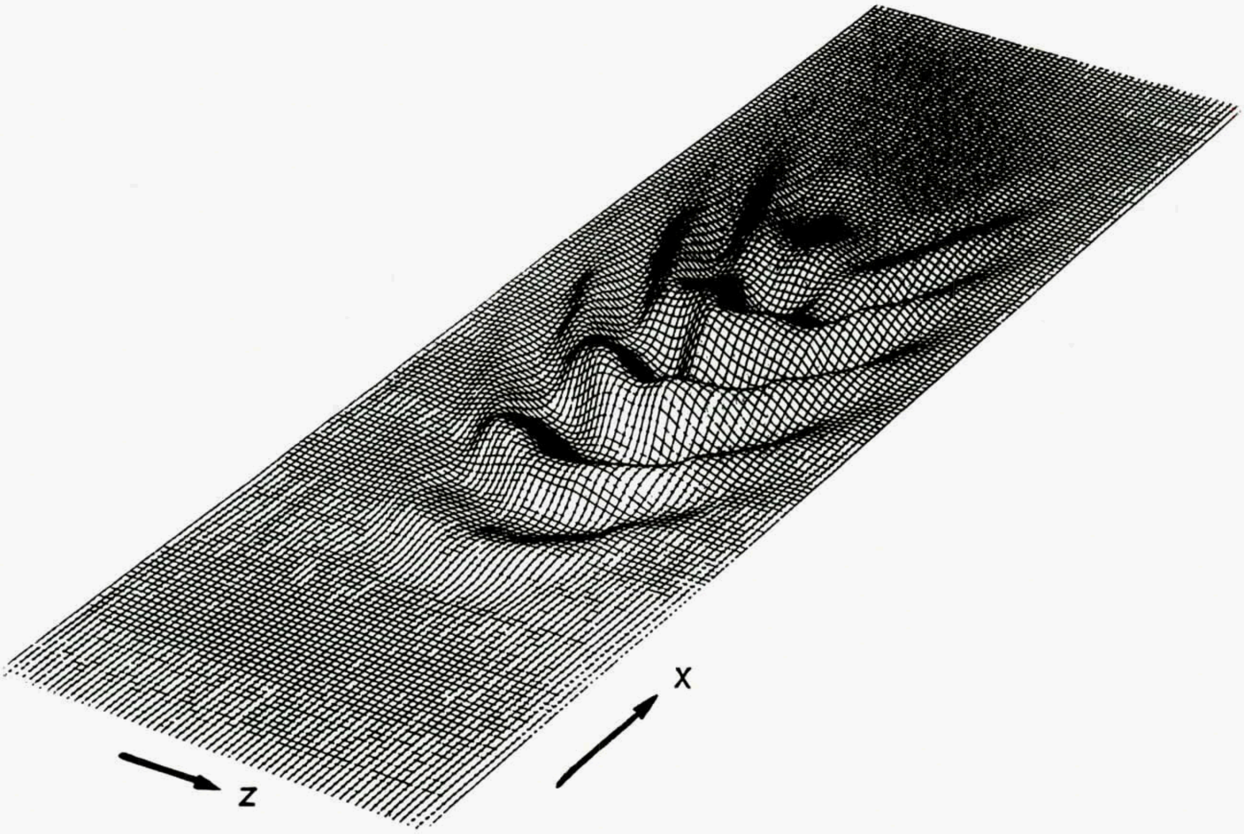
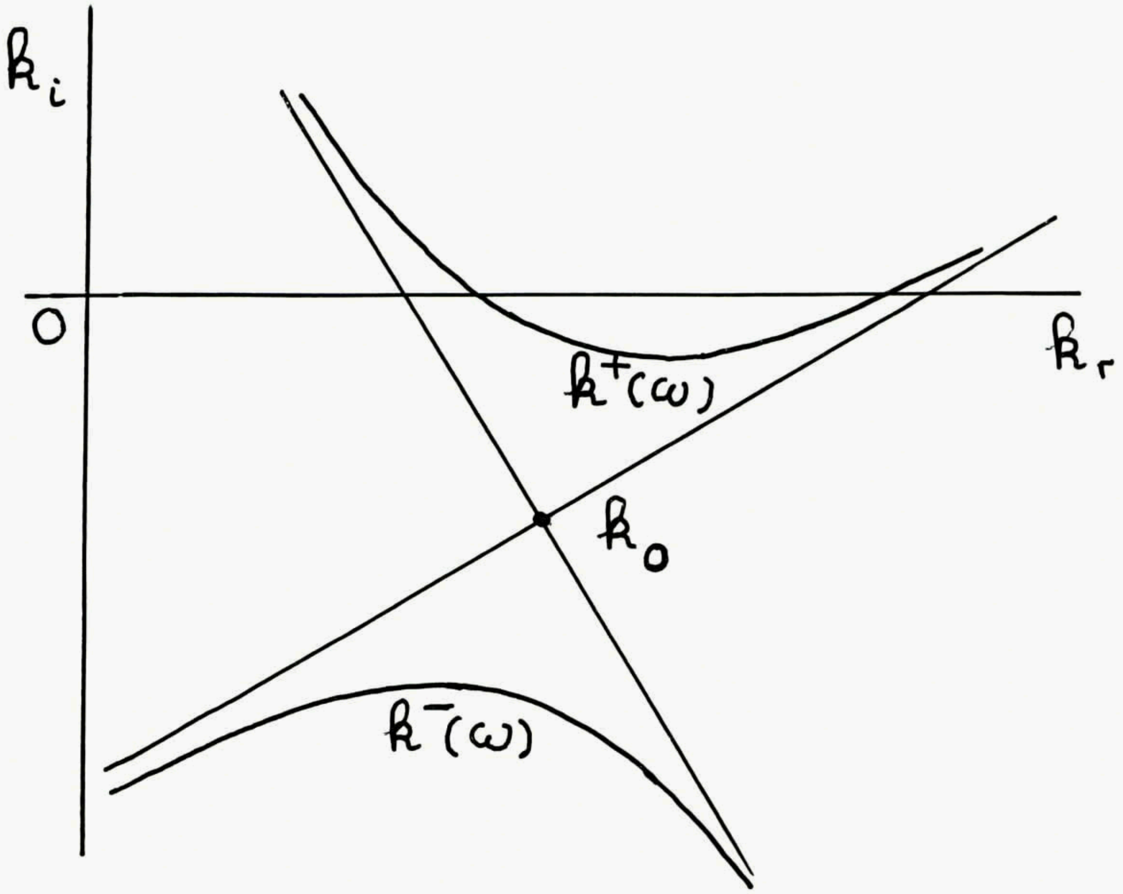


Figure 4

Absolute Growth Rate



$$\boxed{\frac{\partial \omega}{\partial k}(k_0) = 0}$$

$$\omega_0 = \omega(k_0)$$

$$\boxed{\begin{array}{ll} \omega_{0,i} > 0 & \text{absolutely unstable} \\ \omega_{0,i} < 0 & \text{convectively unstable} \end{array}}$$

Figure 5

CLASSES OF
SPATIALLY-DEVELOPING FLOWS

$$\omega_{i,max}(X)\ ;\ \omega_{o,i}(X)$$

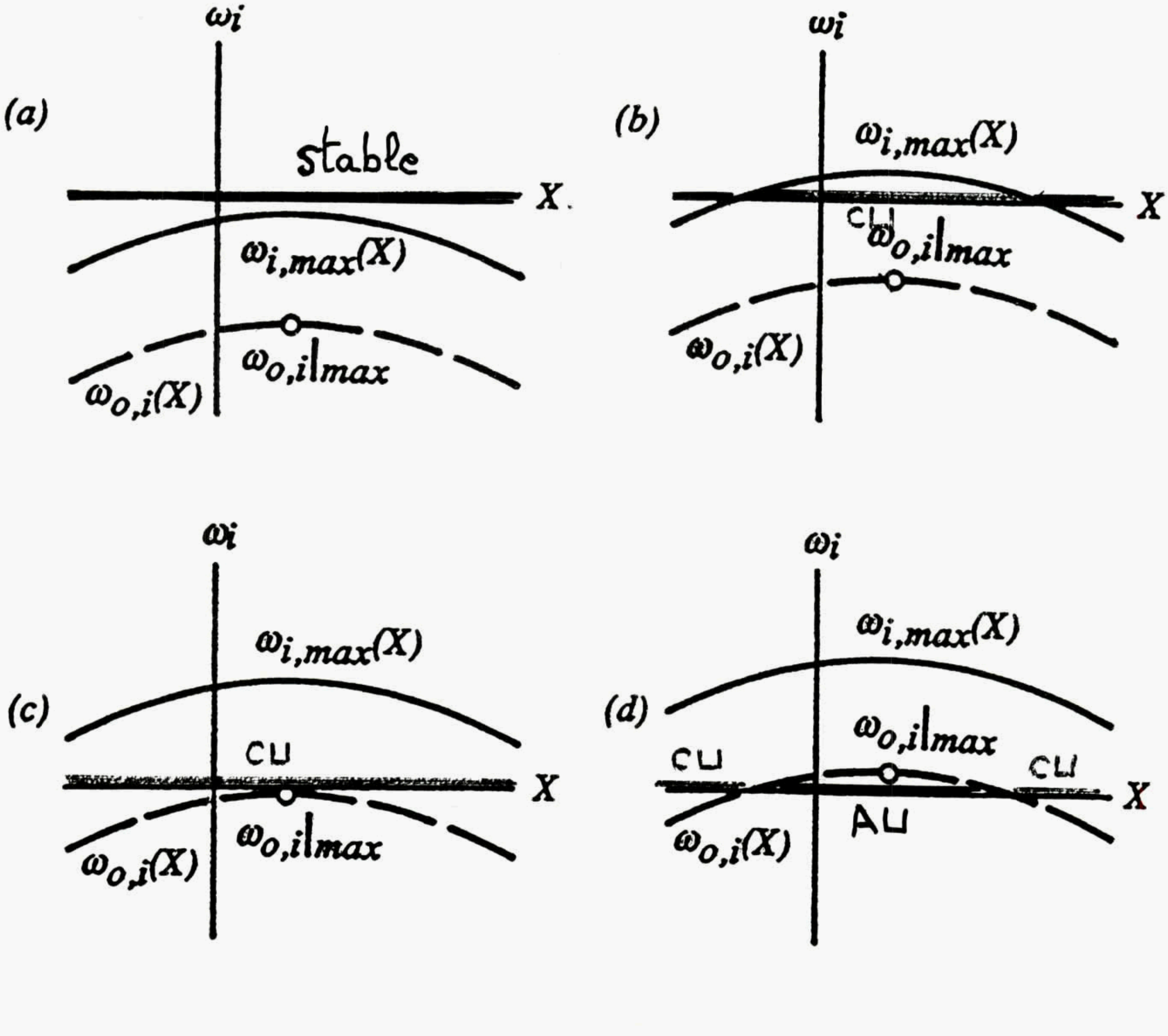


Figure 6

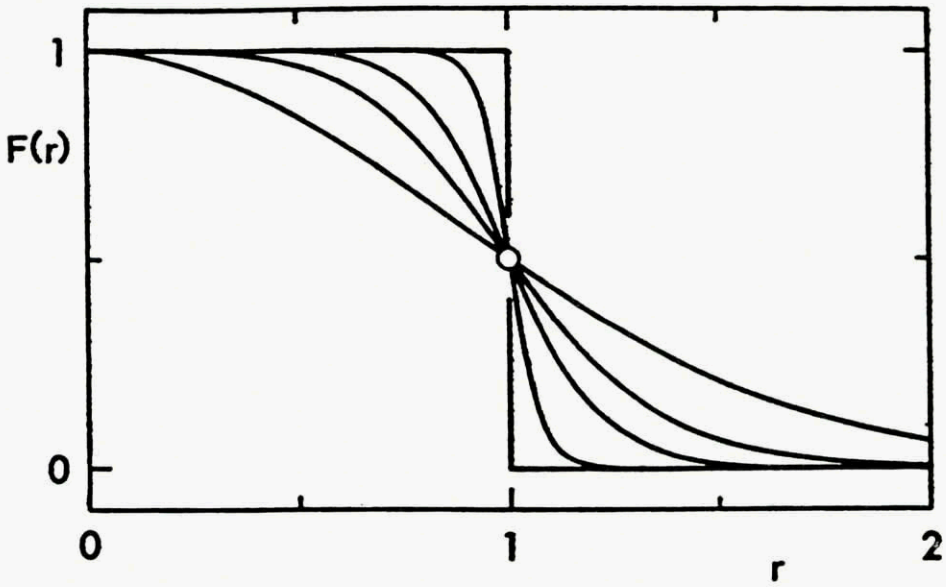


Fig. 7 **The normalized velocity profile *F(r)*, defined by Eq. (1), for $N^{-1} = 1, 0.5, 0.3, 0.1$, and 0.**

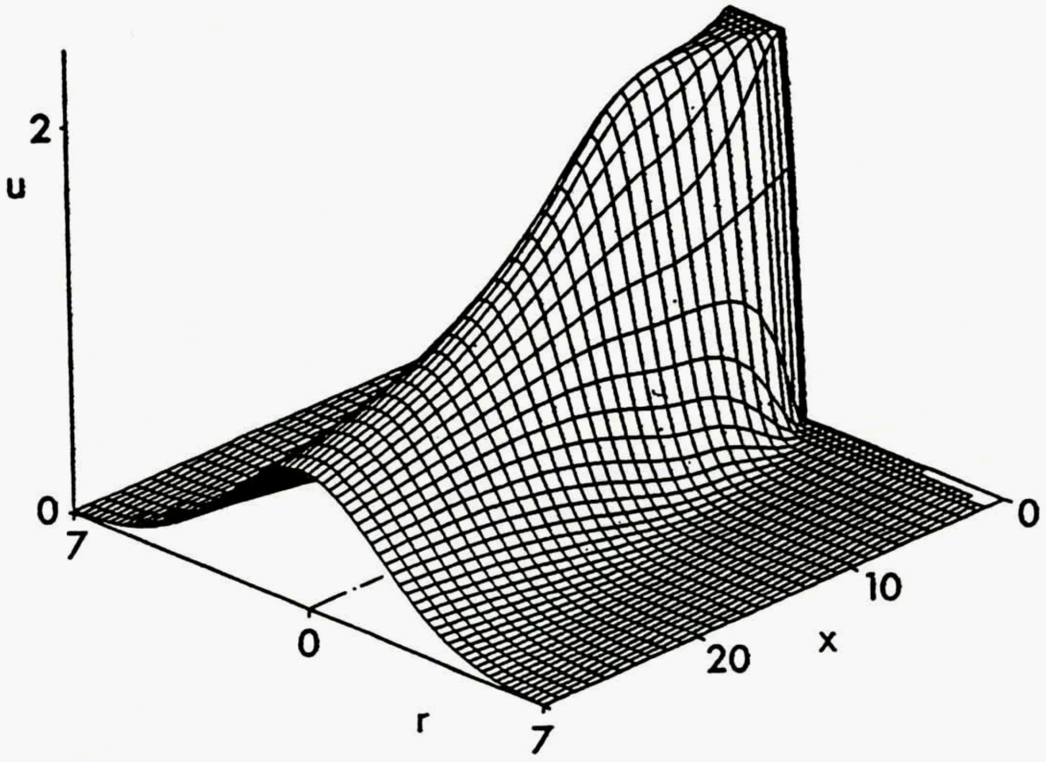


Fig. 8 **Example of the jet velocity vs *r* and *x* for $R_e = 1$, $S_e = 0.5$, and $M_e = 0$.**

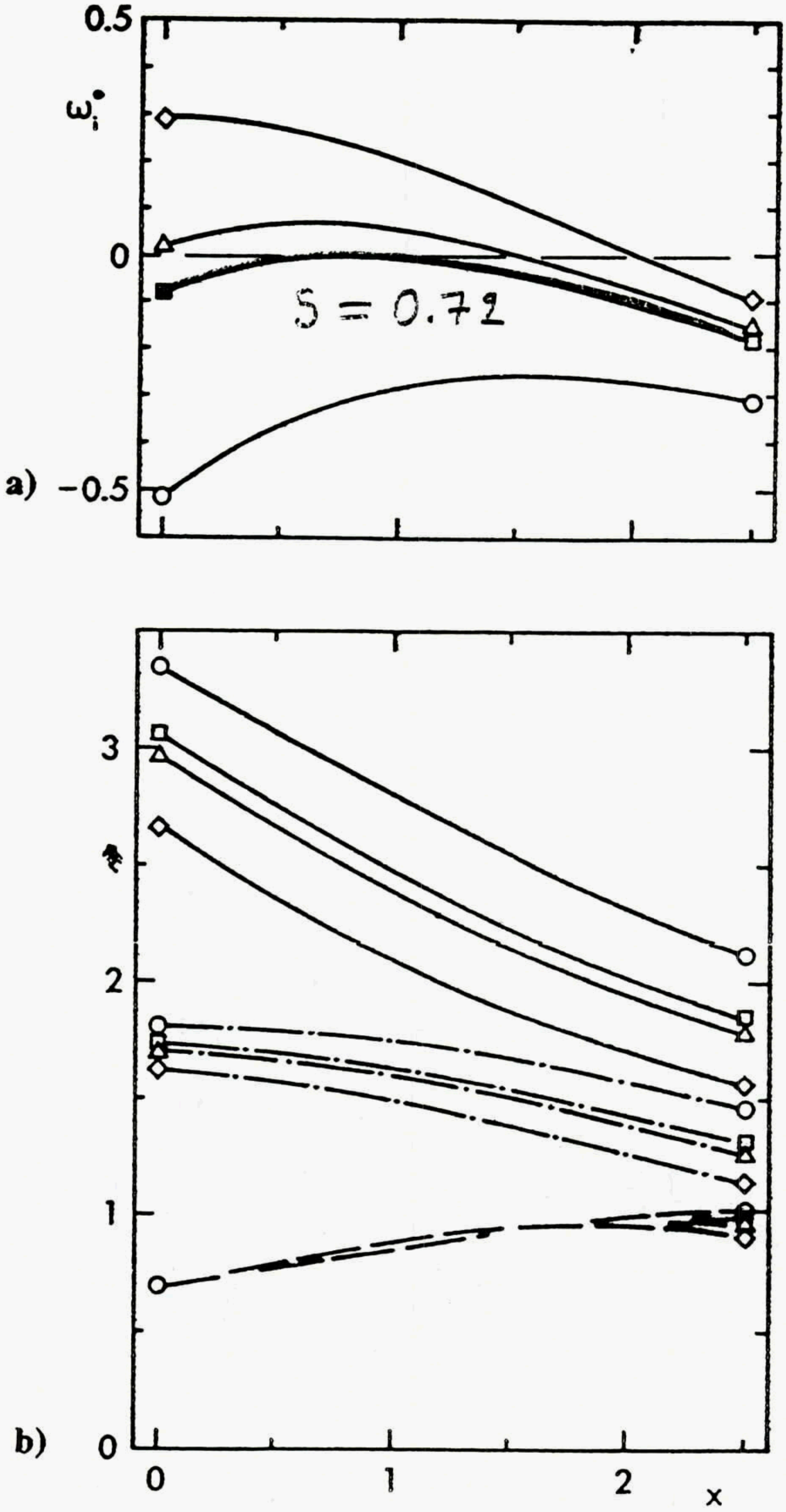


Fig. 9 a) absolute growth rates vs x for the $m=0$ mode and $R=1$, $\bar{M}=0$, and $S_e=1.0$ (○), $S_e=0.72$ (□), $S_e=0.66$ (△), and $S_e=0.5$ (◇). b) —, corresponding real parts ω_r^0; − − −, corresponding k_r^0; − · − · −, corresponding $-k_i^0$.

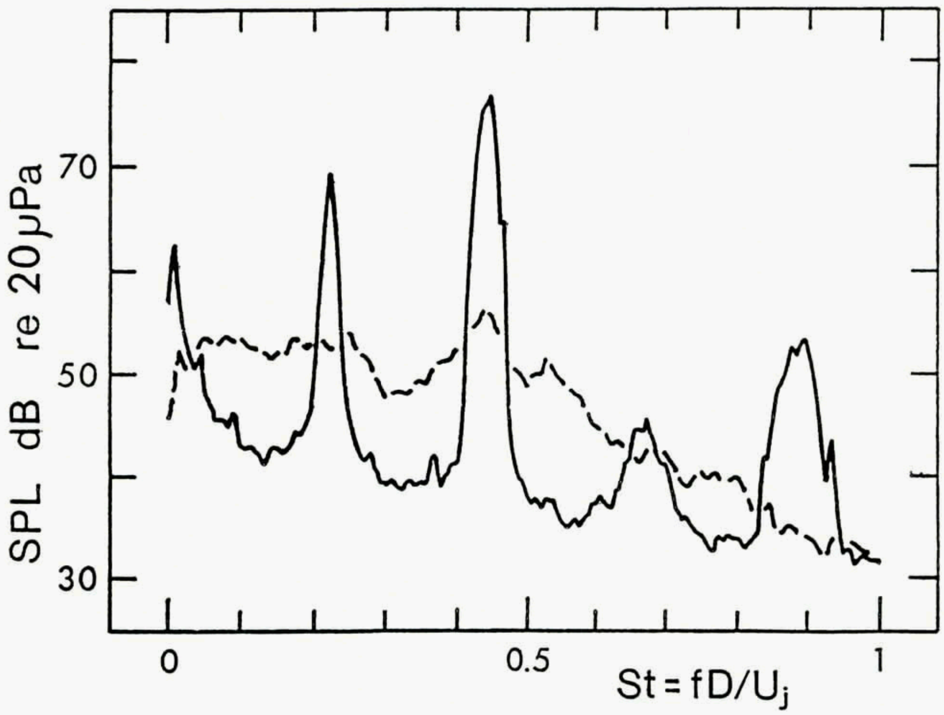


Figure 10

Near field pressure spectrum at $\frac{x}{D} = 0.2$ $\frac{r}{D} = 1$

— hot jet $S = \frac{\rho_j}{\rho_\infty} = 0.47$, $Re = 7500$, $U_j = 16.4$ m/s

--- cold jet $S = 1$ of equal dynamic head

Note St = 0.5 preferred mode of jet.

$$Fr = \frac{S U_j^2}{D(1-S)g} = 1.6 \times 10^3 \gg 1$$ buoyancy effects negligible

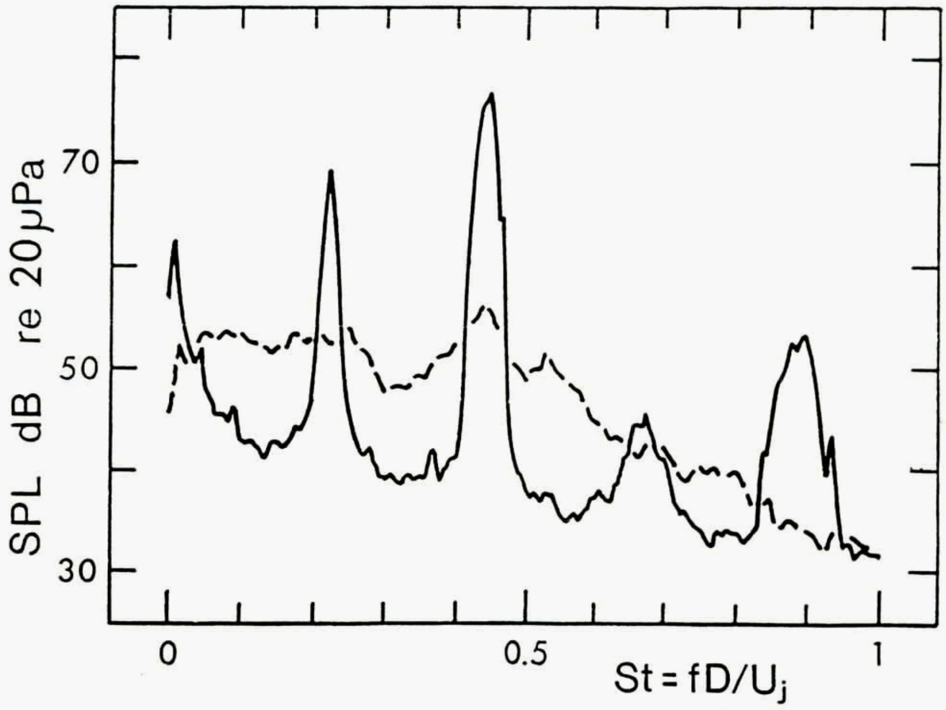


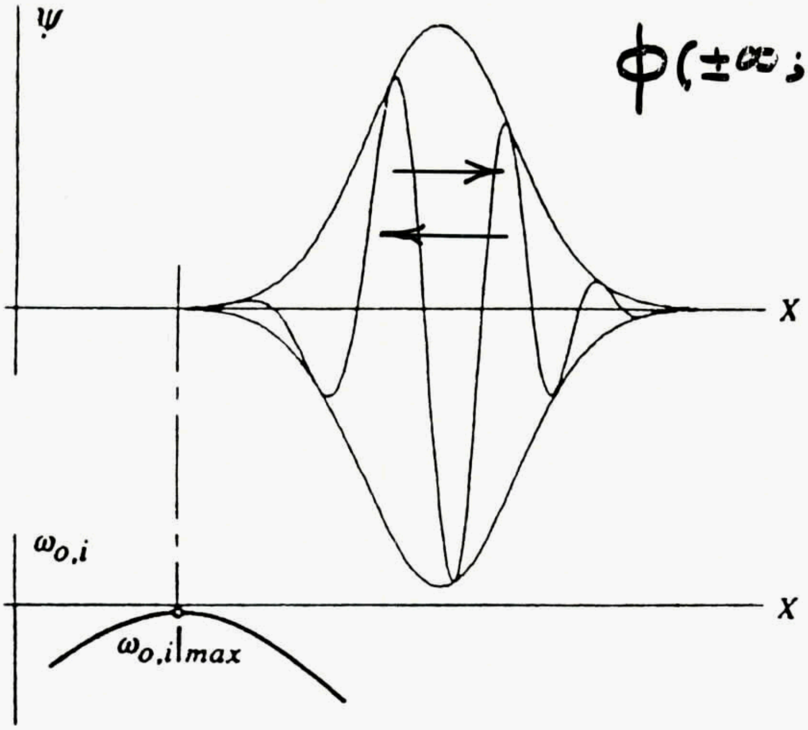
Figure 11

GLOBAL CONCEPTS :

Global Modes :

$$\psi(x,t;X)=\phi(x;X)e^{-i\omega_{G}}$$

$$\phi(\pm\infty;\pm\infty)=$$



$\omega_{G,i} > 0$ globally unstable

$\omega_{G,i} < 0$ globally stable

Is there a connection between local (absolu convective) properties and global properties ?

Figure 12

CHAOTIC TIME SERIES ANALYSIS

Jaroslav Stark
Long Range Research Laboratory,
Hirst Research Centre,
East Lane, Wembley,
Middlesex, HA9 7PP, UK.

SUMMARY

Perhaps the single most important lesson to be drawn from the study of non-linear dynamical systems over the last few decades is that even extremely simple deterministic dynamical systems can give rise to complex behaviour which from a statistical point of view can appear completely random. This phenomenon is frequently referred to as chaos. One of its obvious consequences is that it is possible to describe certain apparently complex signals using relatively simple non-linear models. This has led to the development of a variety of novel techniques for the manipulation of such "chaotic" time series. Initially, these methods concentrated on the characterization of chaotic signals using invariants such as fractal dimensions or Liapunov exponents. Later, attention focused on the possibility of predicting their future short term behaviour and this in turn has led to algorithms for noise reduction in time series having a chaotic component.

This paper begins with a brief description of the basic theoretical framework for analysing chaotic time series and then goes on to discuss recent work in this area in the Long Range Research Laboratory. As a particular example we shall show how the ability to make short term predictions of chaotic time series can be used to extract small signals from a complex deterministic background. Experiments with simulated data have shown that it is possible to recover signals to a reasonable accuracy in situations where the ratio of amplitudes of signal to chaotic background is as low as 10^{-5}:1.

1. INTRODUCTION

Until recently, the subject of time series analysis has been largely dominated by linear mathematics and the behaviour of time series has been modelled by linear equations. Unfortunately, many natural phenomena are governed by fundamentally non-linear laws and hence such linear models are often unable to describe or predict their behaviour adequately. As a result, conventional signal processing has for many years been hampered by its underlying assumption of linearity.

The main reason for the prevalence of linear approaches to time series has been that, in common with many other branches of mathematics, the passage from linear to non-linear is a difficult one. Nevertheless, the last few years have seen rapid progress in non-linear time series techniques. One of the main reasons for this has been the enormous advances made in the field of non-linear dynamical systems. Perhaps the most important of these has been the realization that even simple deterministic non-linear dynamical systems can give rise to complex behaviour. Such behaviour is often statistically indistinguishable from that produced by a completely random process and as a consequence has come to be referred to as *chaotic*. A simple example is given in Figure 1, which shows a time series $\{x_n\}$ obtained from the so called Hénon map

$$x_{n+1} = 1 - 1.4(x_n)^2 + y_n \tag{1a}$$

$$y_{n+1} = 0.3x_n \tag{1b}$$

Deterministic chaos of this kind occurs naturally in many different non-linear mechanical, electrical, and electronic systems. Its existence forces us to drastically rethink our notions of stability and predictability. As regards the former, within the traditional linear context stability implies bounded behaviour and vice versa. In particular, if we do not wish our system to diverge to infinity (which id likely to have unpleasant consequences), we are forced to design a linear system to be stable. By contrast, the behaviour shown in Fig. 1 is unstable (see below) but still bounded. Hence non-linearity may allow future system designers the freedom to drop the requirement of stability if all they are interested in is a bounded operating regime. Potentially, there could be substantial advantages in using the inherent instability of chaotic systems to obtain faster response to control parameters. We shall discuss this briefly in §7 below. Turning now to the issue of prediction, chaotic time series such as Fig. 1 typically have continuous broadband power spectra. Classically, they would thus be modelled by a linear stochastic process, such as, for example, an autoregressive moving average (ARMA) model:

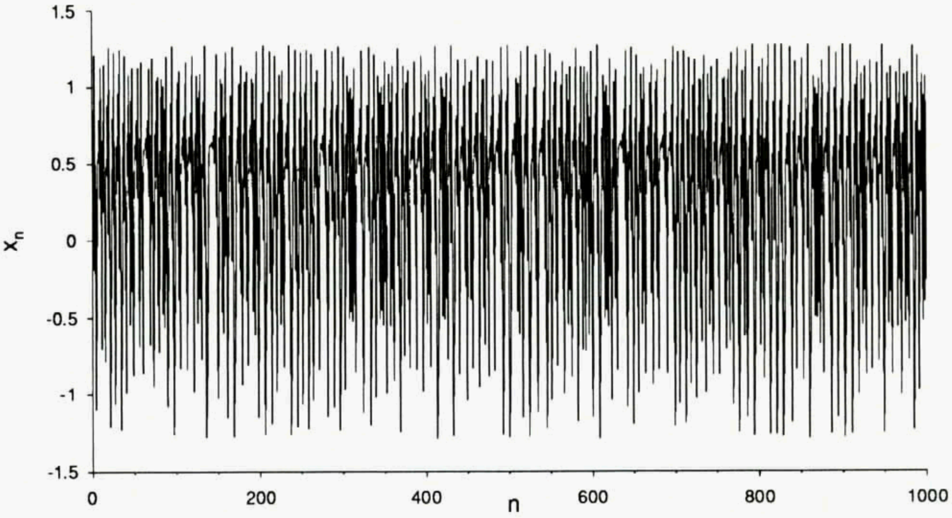


Figure 1. Sample chaotic time series $\{x_n\}$ derived from the Hénon map in equation 1.

$$x_n = a_0 + \sum_{j=1}^{k} a_j x_{n-j} + \sum_{j=0}^{1} b_j \varepsilon_{n-j} \quad (2)$$

Here ε_n is a sequence of uncorrelated random variables and the a_i and b_i are the parameters of the model. Such a model assumes that the complexity seen in Fig. 1 is due to the randomness represented by the ε_n and hence leads to rather poor estimates of the future behaviour of x_n. By contrast, as we shall see below, considerable progress has been made in the last few years in developing non-linear prediction schemes which are able to make extremely good short term forecasts of chaotic times series, even when the underlying dynamical equations (such as Eq. 1) are not known. Such prediction algorithms in turn can be used as the basis of novel techniques for signal processing including filtering, noise reduction and signal separation.

2. CHAOS AND INSTABILITY

Before we describe some of these algorithms, we wish to examine what exactly we mean when we call a system chaotic and what this implies about its stability. At first sight it might seem strange that a simple deterministic system, such as that given by Eq. 1, should lead to such apparently complex behaviour as that shown in Fig. 1. The principal mechanism behind this phenomenon is the extreme sensitivity of such systems to their initial conditions. To illustrate this we slightly perturbed the initial conditions (x_0,y_0) used to generate Fig. 1 to give (x'_0,y'_0) where $x'_0 = x_0 + 10^{-8}$ and $y'_0 = y_0 + 10^{-8}$. Figure 2 shows the resulting time series. Although the gross features of Fig. 1 and Fig. 2 are very similar, careful inspection will show that x_n and x'_n rapidly diverge. This is much more apparent in Figure 3 where we plot $\log|x_n - x'_n|$. We see that for about the initial 50 iterations $\log|x_n - x'_n|$ increases more or less linearly until there is essentially no correlation left between x_n and x'_n.

It turns out that for sufficiently small perturbations the distance between (x_n,y_n) and (x'_n,y'_n) always grows exponentially with n. This effect is known as *exponential divergence of trajectories* and is characteristic of chaotic systems. Indeed, it can be used to define chaos. In systems which exhibit such sensitive dependence on initial conditions any small perturbation to the state of the system is rapidly amplified and leads to a completely different future evolution of the system. Such systems are thus highly unstable.

This phenomenon also places strong limits on the long term predictability of such systems. This is because in practice we

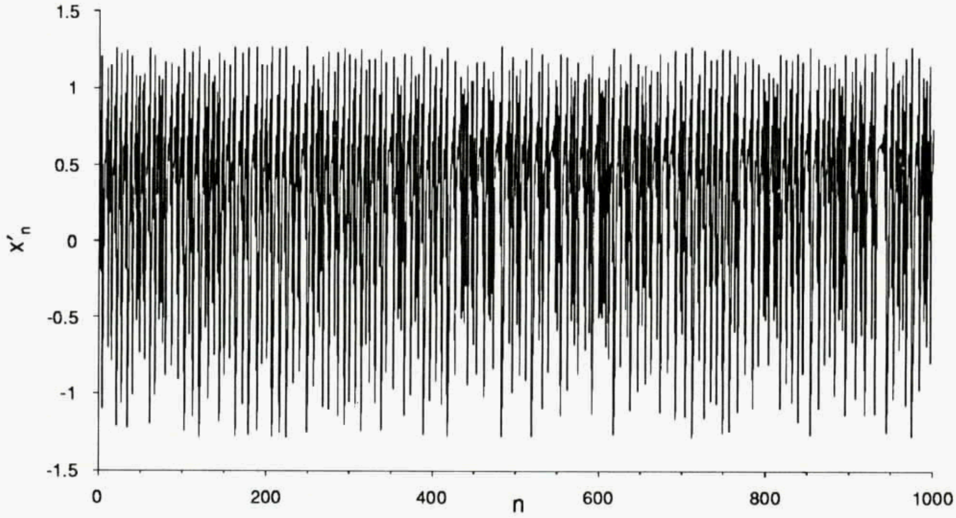


Figure 2. Time series derived from Eq. 1 and modified initial conditions (x'_0,y'_0).

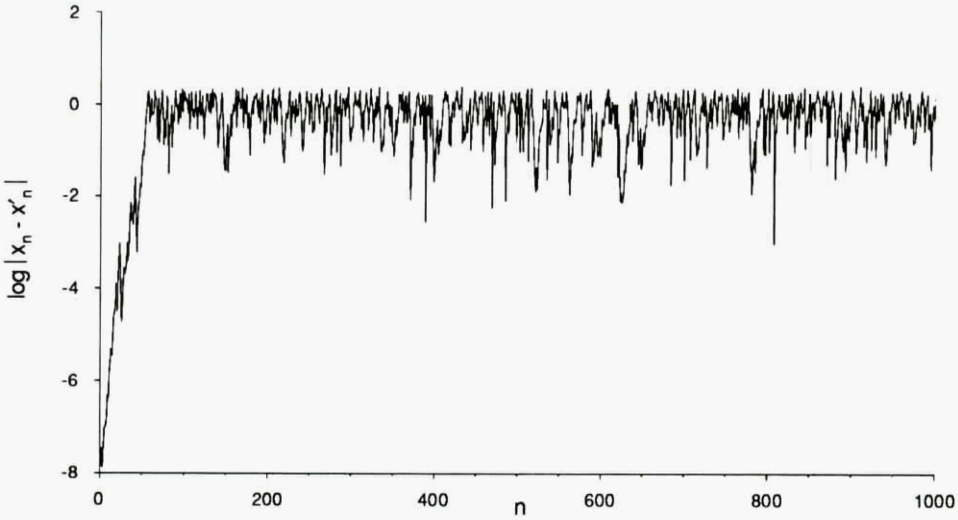


Figure 3. Behaviour of $\log|x_n - x'_n|$ where x_n and x'_n are as in Figures 1 and 2 respectively.

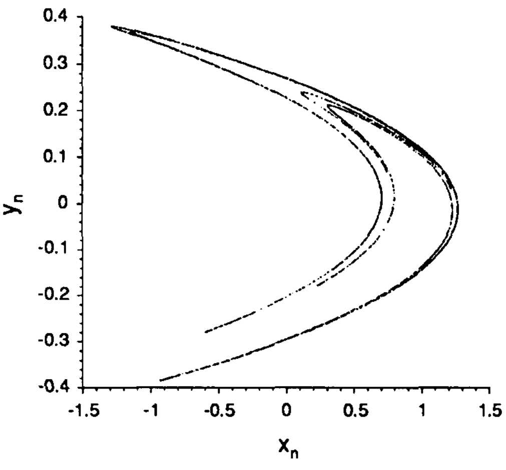


Figure 4. Plot of y_n against x_n for the Hénon map.

never know the initial state of the system to infinite precision. Any errors, no matter how small, in the determination of this state will grow as above and rapidly render our long term forecasts meaningless. We thus have the dichotomy that chaotic systems are highly predictable in the short term (due to their deterministic time evolution) but completely unpredictable in the long term (due to their sensitive dependence on initial conditions).

Finally, we mention that although locally highly unstable, chaotic systems often posses a large degree of global stability, as for instance illustrated by the overall similarity between Fig.1 and Fig. 2. Indeed, any quantity calculated by averaging over the whole time series will be identical for these two figures and for a wide range of other perturbations to (x_0,y_0). Furthermore the points (x_n,y_n) are severely constrained in where they may lie in the plane. A plot of x_n against y_n reveals the peculiar object, called a *strange attractor,* shown in Figure 4. All orbits in its neighbourhood rapidly converge to it, implying a global stability to the dynamics. A perturbation such as from (x_0,y_0) to (x'_0,y'_0) initially knocks a trajectory off the attractor, but it rapidly returns to it, albeit at a different point. Within the attractor itself orbits diverge as described above, leading to the separation of (x_n,y_n) and (x'_n,y'_n). To conclude, we thus see that chaotic systems exhibit an intriguing mixture of both stable and unstable behaviour.

3. EMBEDDING CHAOTIC TIME SERIES

For the Hénon map of Eq. 1 the time evolution of x_n depends on both x and y. Thus to generate Fig. 1, we had to compute both x_n and y_n for n = 1 ... 1000. Similarly, the behaviour of most practical systems will depend on more than one state variable. Thus, typically, the state of a dynamical system at time t will be described by a vector $\mathbf{u}_t \in \mathbb{R}^m$ whose time evolution will be given by a law such as $\mathbf{u}_t = \mathbf{f}^t(\mathbf{u}_0)$, where $\mathbf{u}_0$ is the initial state of the system. Here $\mathbf{f}^t$ may be the solution of a differential equation in which case time evolves continuously, or as in the case of the Hénon map may be obtained by the iteration of some single map **f** from $\mathbb{R}^m$ to $\mathbb{R}^m$. In the latter case, time will advance in discrete steps n = 0,1,2,... and we have $\mathbf{u}_{n+1} = \mathbf{f}(\mathbf{u}_n)$. Thus $\mathbf{f}^n$ is just **f** composed with itself n times, so that $\mathbf{f}^{n+1} = \mathbf{f}\circ\mathbf{f}^n$ with $\mathbf{f}^1 = \mathbf{f}$. For the Hénon map, **f** is given by $\mathbf{f}(u_1,u_2) = ((1-1.4(u_1)^2+u_2), 0.3u_1)$. For simplicity, we shall restrict ourselves for the remainder of this paper to discrete time systems. This will, in fact, be the relevant case in most practical applications, since one generally makes observations of the system's behaviour at regular time steps. Should one begin with a continuous time system, one can always reduce to the discrete time case by defining **f** to be $\mathbf{f}^\tau$ for some sampling interval τ and restricting time to $t = 0,\tau,2\tau, \ldots$.

In most practical situations we will have no direct knowledge of the state $\mathbf{u}_n$, or of the map **f**. Instead, we are confined to making measurements of one or more observables of the system such as velocity or temperature. Mathematically, such observables will just be given by a function $\varphi(\mathbf{u})$ of the state **u**. We shall only treat the case of a single observable, so that φ is real valued, but most of what follows is equally true for vector valued observations with only minor modifications. The sequence of repeated observations $\varphi(\mathbf{u}_n)$ that we make as the system evolves in time will then gives us a time series $x_n = \varphi(\mathbf{u}_n)$. Thus, in the example above the time series of Fig. 1 is obtained from the Hénon map using the rather simple observable $\varphi(u_1,u_2) = u_1$.

Given that we can neither observe $\mathbf{u}_n$ nor have any direct knowledge of **f**, it would appear at first sight that the fact that $\{x_n\}$ is generated by a deterministic system is of little use to us. In particular, it would seem that x_n, being one dimensional, cannot contain much information about the higher dimensional variable $\mathbf{u}_n$. Remarkably, this intuition turns out to be false due to a powerful result known as the *Takens Embedding Theorem.* This shows that for most systems **f** and most measurement functions φ, it is possible to reconstruct **u** and **f**, up to some nice co-ordinate change, just from knowledge of the time series $\{x_n\}$. More precisely, fix some integer d (called the *embedding dimension*) and consider the map Φ from $\mathbb{R}^m$ to $\mathbb{R}^d$ given by

$$\Phi(\mathbf{u}) \quad = \quad ((\varphi(\mathbf{u}), \varphi(\mathbf{f}(\mathbf{u})), \ldots, \varphi(\mathbf{f}^{d-1}(\mathbf{u}))) \qquad (3)$$

Then the Takens theorem [1,2] states that generically if $d \geq 2m + 1$, then Φ is an *embedding*. This means that is it is invertible on its image and both Φ and Φ^{-1} are smooth (continuously differentiable). The word "generically" here is a precise mathematical term which means "for all typical systems" somewhat in the same way that one might say that the roots of a typical polynomial are simple. From now on we shall always assume that we are in the typical case.

The embedding Φ can be thought of as a smooth change of co-ordinates between the dynamics given by **f** and the time evolution of the series $\{x_n\}$. To see this, define the map F on $\mathbb{R}^d$ by

$$F \quad = \quad \Phi\circ\mathbf{f}\circ\Phi^{-1} \qquad (4)$$

Thus F is just **f** under the co-ordinate change given by Φ. All of the co-ordinate independent properties of F and **f** will thus be identical. Then

$$\begin{aligned} F(x_{n-d}, x_{n-d+1}, \ldots, x_{n-1}) &= F(\Phi(\mathbf{u}_n)) \\ &= \Phi\circ\mathbf{f}\circ\Phi^{-1}\circ\Phi(\mathbf{u}_n) \\ &= \Phi\circ\mathbf{f}(\mathbf{u}_n) \\ &= \Phi(\mathbf{u}_{n+1}) \\ &= (x_{n-d+1}, x_{n-d+2}, \ldots, x_n) \qquad (5) \end{aligned}$$

Thus F just advances a block of elements of the time series $\{x_n\}$ forward by one time step. The importance of this is that whilst **f**, **u** and even Φ are inaccessible to us, the map F operates purely in terms of the time series $\{x_n\}$ and is thus in principle completely observable. In other words, despite the fact that x_n is one dimensional whilst $\mathbf{u}_n$ is m-dimensional, all of the important information about the behaviour of $\mathbf{u}_n$ is actually contained in x_n. It is this fundamental fact which lies behind all practical schemes for processing chaotic time series.

The first d-1 co-ordinates of F are trivial, since they just consist of copying the last d-1 co-ordinates of its argument. All the work is done by the last coordinate, which we shall denote by G. Thus

$$F(y_1, y_2, \ldots, y_d) \quad = \quad (y_2, y_3, \ldots, y_d, G(y_1, y_2, \ldots, y_d)) \qquad (6)$$

In terms of the time series $\{x_n\}$ we have

$$G(x_{n-d}, x_{n-d+1}, \ldots, x_{n-1}) \quad = \quad x_n \qquad (7)$$

G thus predicts x_n in terms of the previous d values of the time series. We can think of this equation as a non-linear autoregressive (AR) model for $\{x_n\}$. It is thus similar to Eq. 2, except that we have removed the stochastic terms due to the ε_n and replaced the linear combination $a_d x_{n-d} + a_{d-1} x_{n-d+1} + \ldots + a_1 x_{n-1}$ by the non-linear function G.

In general, it is not possible to give an explicit expression for G. However, due to the particular structure of the Hénon map, we can in fact write down G for the time series in Fig. 1 in closed form. Thus, from Eq. 1b, we have that $y_n = 0.3x_{n-1}$ and substituting this into Eq. 1a we get

$$x_{n+1} \quad = \quad 1 - 1.4(x_n)^2 + 0.3x_{n-1} \tag{8}$$

In this special case it is thus sufficient to take d = 2, so that we have d = m rather than d = 2m + 1 required in the general case.

The most serious difficulty in applying Takens's theorem is that although it gives an upper bound on the value of d required, this bound is in terms of m which will usually be unknown. In practice d therefore has to be determined essentially by trial and error. One possibility is to compute a trial G for each value of d and then use the smallest d which gives a good fit to Eq. 7. A better method is to look at some simple invariants of a dynamical system which characterize its complexity and also give a good estimate of the size of d required as a by-product

4. CHARACTERIZING CHAOS

Given an apparently complex time series $\{x_n\}$ such as Fig. 1, how do we tell whether it has come from a deterministic chaotic system or from some kind of stochastic process such as the ARMA model of Eq. 2? A variety of techniques for answering this question exist, all based upon the Takens embedding described above. Thus, from the scalar series $\{x_n\}$ we form the d-dimensional orbit $\{\mathbf{v}_n\}$ of F, where $\mathbf{v}_n = (x_{n-d}, x_{n-d+1}, \ldots, x_{n-1})$, so that $F(\mathbf{v}_n) = \mathbf{v}_{n+1}$. This process is called *embedding* $\{x_n\}$ and is similar to the classical time series method of delays. As mentioned above, the co-ordinate independent properties of $\{\mathbf{v}_n\}$ are the same as those of the original (unknown) state space orbit $\mathbf{u}_n$. As described in §2, the chaotic behaviour of such an orbit can be characterized by the exponential divergence of nearby trajectories. Thus if $\mathbf{u}'_n$ is a small perturbation of $\mathbf{u}_n$ we expect to see that

$$\|\mathbf{u}_n - \mathbf{u}'_n\| \quad \sim \quad e^{\lambda n} \tag{9}$$

for some constant λ. In fact, there will be m different values of λ, depending on the direction of the initial perturbation. These m values are known as the *Liapunov exponents* of **f** [3,4] and it is usual to describe **f** as chaotic if at least one is positive. It can be shown that the Liapunov exponents of F are the same as those of **f** and at least in principle an estimate of the exponents of F can be made from a sufficiently large sample of $\{\mathbf{v}_n\}$.

In practice, however, the Liapunov exponents are rather difficult to compute and cannot always distinguish between a deterministic time series and a stochastic one. It is thus preferable to measure a quantity called the *correlation dimension* D_C [3, 4]. This attempts to measure the dimension of the set in $\mathbb{R}^m$ on which the points $\{\mathbf{u}_n\}$ lie. In some sense it is thus a measure of the complexity or number of variables required to describe this set. If all the $\mathbf{u}_n$ are identical (so that all the x_n are constant) then D_C will be zero. If they lie on some curve then D_C will equal 1 and if they fill a plane it will equal 2. At the other extreme, if the $\mathbf{u}_n$ completely fill $\mathbb{R}^m$ then we will have $D_C = m$. Intriguingly, D_C need not be an integer (it is about 1.21... for the Hénon map) and is thus an example of a *fractal dimension* [3,4]. As with the Liapunov exponents, the correlation dimension is co-ordinate independent and hence is the same for **f** and F. It can be estimated from a finite sample $\{\mathbf{v}_1, \ldots, \mathbf{v}_N\}$ as follows. First form all the N^2 possible pairs $(\mathbf{v}_i, \mathbf{v}_j)$ of such points. Calculate the Euclidean distance $r_{ij} = \|\mathbf{v}_i - \mathbf{v}_j\|$ between each pair. Note that this can be defined in terms of the original time series $\{x_n\}$ by

$$r_{ij} \quad = \quad \left(\sum_{k=0}^{d-1} \left(x_{i-k} - x_{j-k} \right)^2 \right)^{\frac{1}{2}} \tag{10}$$

Now, for a given ε, let N(ε) be number of pairs such that $r_{ij} \leq \varepsilon$. Then $C(\varepsilon) = N(\varepsilon)/N^2$ is the proportion of pairs of points within a distance of ε of each other. A simple calculation shows that if all the points $\mathbf{v}_n$ lie randomly on some curve we have roughly $C(\varepsilon) \sim \varepsilon$ for large N and small ε. Similarly if the $\mathbf{v}_n$ lie on a surface we get $C(\varepsilon) \sim \varepsilon^2$. This suggests that C(ε) behaves exponentially for small ε with the exponent giving the dimension of the set on which the $\mathbf{v}_n$ lie. This motivates the definition of the correlation dimension

$$D_C \quad = \quad \lim_{\varepsilon \to 0} \frac{\log C(\varepsilon)}{\log \varepsilon} \tag{11}$$

The quantity C(ε) is called a *correlation integral* and several efficient numerical methods exist for calculating it, even for moderately large data sets (e.g. $N = 10^6$) [3,4]. To evaluate D_C we then plot log C(ε) against log ε and estimate the resulting slope.

So far, we have assumed that we know the size of d required to embed the time series $\{x_n\}$. When this is not the case, we proceed by trial and error. Thus, we calculate a correlation dimension $D_C(d)$ for each trial choice of d. When d is too small the set $\{\mathbf{v}_n\}$ will completely fill $\mathbb{R}^d$ and we will get $D_C(d) \cong d$ (we must of course always have $0 \leq D_C(d) \leq d$). Conversely once d is sufficiently large the computed value $D_C(d)$ should stabilize at approximately the correct correlation dimension of **f**. As an example, for the Hénon map we get $D_C(1) \cong 1$ and $D_C(2) \cong D_C(3) \cong D_C(4) \cong 1.21$.

When D_C is not an integer, as in this case, we say that the system contains a *strange attractor*. This is usually a sign of chaos (as indicated by a positive Liapunov exponent), although strange non-chaotic systems do exist (but are currently believed to be pathological).

It may of course happen that $D_C(d)$ continues to grow with d. This usually suggests that the time series $\{x_n\}$ was generated by a stochastic process rather than by a chaotic dynamical system. Thus, for example, a random process such as white noise will have $D_C = \infty$. Of course, we would also get this result if the original dynamical system **f** was genuinely infinite dimensional, but from many points of view such a system is indistinguishable from a random one.

In the above procedure, the smallest value of d at which $D_C(d)$ begins to stabilize yields the minimal embedding dimension required to adequately represent the dynamics of the system. Computing the correlation dimension in this way thus yields both a measure of a time series's complexity, and an estimate of the embedding dimension required for any further processing. It is thus usually the first step in analysing time series which we suspect might have been generated by a chaotic system. In practice, lack of data and numerical precision limit calculations to about $d \leq 10$ (and hence $D_C \leq 10$). From the point of view of this paper, therefore, any time series with D_C appreciably larger than 10 can be treated more or less as a truly random one.

5. PREDICTION OF CHAOTIC TIME SERIES

Having identified a particular time series $\{x_n\}$ as possibly arising from a deterministic system and obtained an estimate of d, the next step will usually be to construct an approximation of the function G. This in turn will enable us to predict the future behaviour of $\{x_n\}$, or perform more complex processing such as noise reduction. This is essentially a problem in multi-dimensional non-linear function fitting and any number of different techniques can be used, ranging from piecewise linear interpolation to neural networks (e.g. [5-9]). We shall describe one particular method, based upon *radial basis function interpolation* (e.g. [10]) which seems to work particularly well [8].

We are thus seeking to calculate an estimate $\tilde{G}$ of G, based upon a finite sample $\{x_1, \ldots, x_N\}$ of the time series. The basic idea behind the radial basis function approach is to choose a finite number of points $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(M)} \in \mathbb{R}^d$ called *radial basis centres* and look for an approximation $\tilde{G}$ of the form of

$$\tilde{G}(\mathbf{v}) \quad = \quad \sum_{i=1}^{M} \lambda_i \, \rho\left(\left\| \mathbf{v} - \mathbf{y}^{(i)} \right\|\right) \tag{12}$$

where ρ is the so called *basis function*, $\|\bullet\|$ is some norm on $\mathbb{R}^d$ (typically the Euclidean norm) and the λ_i are parameters which determine the function $\tilde{G}$.

In Powell's original approach the $\mathbf{y}^{(i)}$ are chosen from amongst the data points $\mathbf{v}_n$, so that $\mathbf{y}^{(i)} = \mathbf{v}_{n(i)}$ (where $\mathbf{v}_n = (x_{n-d}, x_{n-d+1}, \dots, x_{n-1})$ as before). Then the value G takes at the points $\mathbf{y}^{(i)}$ is known, and it is reasonable to require that $\tilde{G}(\mathbf{y}^{(i)}) = G(\mathbf{y}^{(i)})$. Substituting this into Eq. 12 we get the matrix equation

$$\mathbf{g} \quad = \quad \Psi\lambda \tag{13}$$

where $\mathbf{g} = (g_1, \dots, g_M)$ with $g_i = G(\mathbf{y}^{(i)}) = G(\mathbf{v}_{n(i)}) = x_{n(i)}$, $\lambda = (\lambda_1, \dots, \lambda_M)$ and Ψ is the M×M matrix with entries

$$\Psi_{ij} \quad = \quad \rho\left(\left\|\mathbf{y}^{(i)} - \mathbf{y}^{(j)}\right\|\right) \tag{14}$$

Note that, using Eq. 10, this is just $\rho(r_{n(i),n(j)})$. Thus both $\mathbf{g}$ and Ψ can be readily computed in terms of $\{x_n\}$. Remarkably, for a wide choice of basis functions ρ, the matrix Ψ is invertible if the $\mathbf{y}^{(i)}$ are disjoint. This immediately leads to the solution $\lambda = \Psi^{-1}\mathbf{g}$.

Figure 5 gives an example of the application of this algorithm to the Hénon time series from Fig. 1. The basis function used here was $\rho(r) = \sqrt{(r^2 + 10)}$, but similar results are obtained with a wide choice of other functions. We took $M = 70$, with $n(i) = i$ for $i = 1, \dots, M$. Note the artificially good predictions for the first 70 points. These are due to the fact that the λ_i are chosen to make $\tilde{G}$ agree with G exactly for these points. The observed error is thus simply due to the numerical errors arising from the numerical inversion of Ψ. The average of the prediction error $\Delta_n = |x_n - \tilde{G}(x_{n-2}, x_{n-1})|$ (ignoring these initial points) is about 1.2×10^{-5}. This appears to be largely due to a small number of very poor predictions so that the average of $\log \Delta_n$ is in fact -7.1.

Of course, the prediction error shown in Fig. 5 is the result of only predicting forward one time step. Thus for each n, we try to predict x_n from x_{n-2} and x_{n-1}. As we shall below, this is precisely what we need as the basis for noise reduction and signal extraction schemes. In other cases, however, we might wish to forecast $\{x_n\}$ further forward in time. As we have already remarked in §2, there are fundamental limits on how far forward one can do this, due to the exponential separation of trajectories in a chaotic system. This is illustrated in Figure 6, where we plot the prediction error $\Omega_n = |x_n - \tilde{x}_n|$ for multi-step predictions, always predicting from the same point (x_{70}, x_{71}). Here $\tilde{x}_n$ is defined by iterating $\tilde{G}$ forward from this point. Thus $\tilde{x}_n = \tilde{G}(\tilde{x}_{n-2}, \tilde{x}_{n-1})$ with initial conditions $(\tilde{x}_{70}, \tilde{x}_{71}) = (x_{70}, x_{71})$. Other details are as for Fig. 5. We see that the error rapidly rises so that it is impossible to make any kind of prediction beyond

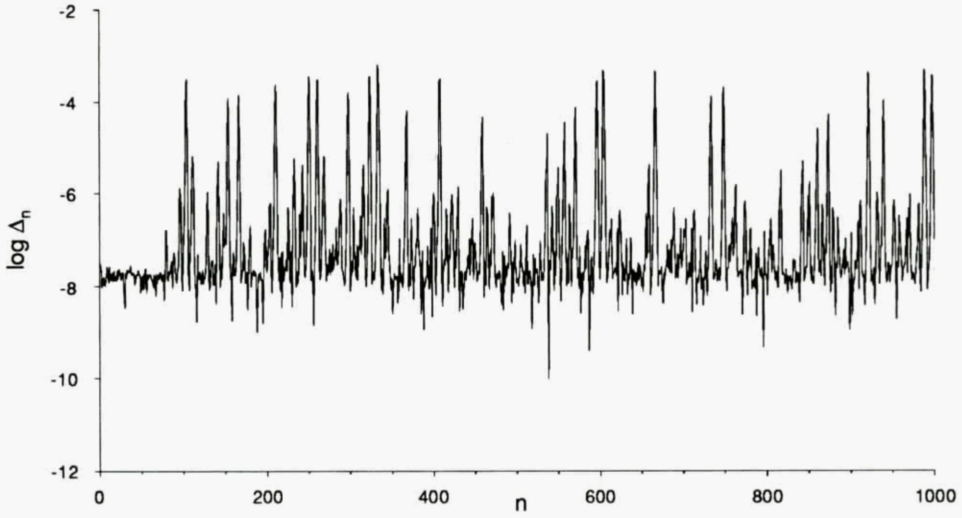


Figure 5. Logarithmic prediction error log Δ_n for Hénon time series using simple radial basis function scheme.

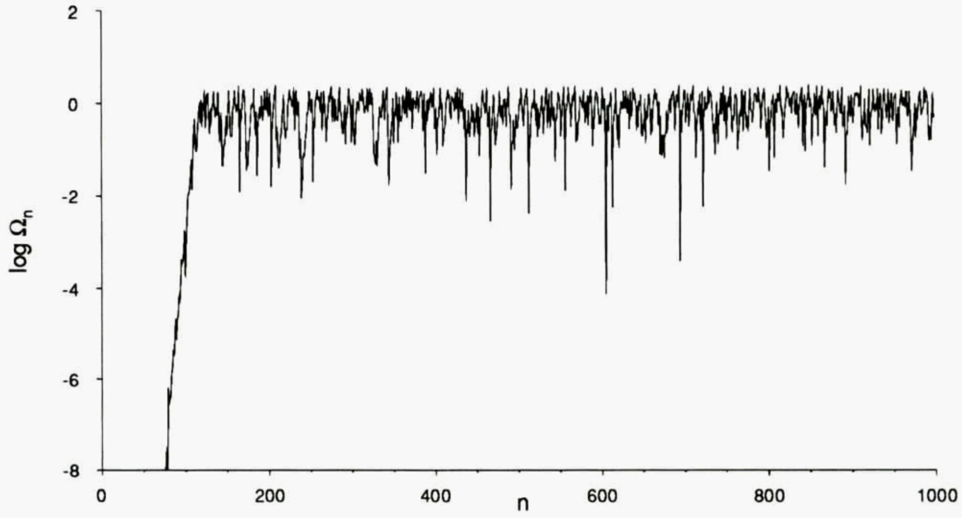


Figure 6. Multi-step logarithmic prediction error log Ω_n for Hénon time series.

about 30 time steps ahead. Notice the obvious similarity to Fig. 3.

The disadvantage of the above approach is that the number of centres M is always equal to the number of data points and the computation of $\tilde{G}$ requires the inversion of an M×M matrix. This severely limits the number of data points that we can use in the estimation of G. Although in the case of the Hénon time series it was sufficient to use $M = 70$, for which the inversion can be done reasonably rapidly, for more complex systems (particularly with larger embedding dimensions d) it is necessary to use much larger samples of data to obtain a sufficiently accurate estimate of G. Larger data sets are also required when $\{x_n\}$ is contaminated by noise or other error and we wish to average this out.

This can be achieved through a generalization due to Broomhead and Lowe [11]. Although they formulated it in terms of a particular neural network architecture, it is equally applicable to our case. Their idea is to take the sample of data points $\mathbf{v}_j$ for $j = 1, \ldots, N$, with $N \geq M$ and seek to minimize the difference between $\tilde{G}$ and G over these points. A natural choice is to minimize the least squares error

$$E \quad = \quad \sum_{j=1}^{N} \left[\tilde{G}(\mathbf{v}_j) - G(\mathbf{v}_j)\right]^2 \tag{15}$$

This is equivalent to finding a λ such that the vector $\mathbf{g} - \Psi\lambda$ has minimum Euclidean norm, where Ψ is now given by

$$\Psi_{ji} \quad = \quad \rho\left(\left\|\mathbf{v}_j - \mathbf{y}^{(i)}\right\|\right) \tag{16}$$

and $g_j = \tilde{G}(\mathbf{v}_j)$. This is a standard linear least squares problem [12] and a variety of practical algorithms exist for computing λ. Broomhead and Lowe [11] chose to use the singular value decomposition (SVD) of Ψ (e.g. see [13]) to obtain λ. Recall that the SVD of a matrix is a decomposition of the form $\Psi = \mathbf{U\Sigma V}^T$ where U is an N×M matrix whose columns are orthonormal, V is an M×M orthogonal matrix and Σ is an M×M diagonal matrix. Then the required λ is given by $\lambda = \mathbf{V\Sigma}^+\mathbf{U}^T\mathbf{g}$ where $\Sigma^+ = \mathrm{diag}(\sigma_1^+, \ldots, \sigma_M^+)$, with $\sigma_i^+ = \sigma_i^{-1}$ if $\sigma_i \neq 0$ and $\sigma_i^+ = 0$ if $\sigma_i = 0$ where $\sigma_1, \ldots, \sigma_M$ are the diagonal elements of Σ. The matrix $\mathbf{V\Sigma}^+\mathbf{U}^T$ is the so called Moore-Penrose pseudoinverse of Ψ (e.g. see [14]).

Figure 7 shows the results of this approach, using the same data as for Fig. 5. We again have $M = 70$ with $N = 500$ (so that the data points $\mathbf{v}_j$ were the first 500 pairs of the time series). As one might expect we get much more uniform errors than in Figure 4. Thus the mean error has dropped to 3.0×10^{-7} whilst the mean log error has slightly risen to -6.7. Note that the error for the first 500 points (which were used in the fitting procedure) is not markedly different from that for the second half of the time series. This suggests that a good overall fit to G has been obtained

The above approach works extremely well in many situations. Its one disadvantage is that it is very much a "batch" algorithm. It thus calculates an estimate $\tilde{G}$ of G once and for all using a predetermined block of observations, say $x_1, \ldots, x_N$. There is then no way of updating $\tilde{G}$ from further observations $x_{N+1}, x_{N+2}, \ldots$ as they are made. Should one decide to use a larger data sample to estimate G one has to discard the previous estimate and recalculate a new estimate from the beginning. This leads to several disadvantages

a) it limits the number of data points $x_1, \ldots, x_N$ that can be used in the estimation process. This is because for a given value of N, we have to form and manipulate the N×M matrix Ψ and hence the above algorithm has a memory requirement of at least MN. When M is of the order of 10^2 as above, or even larger, this rapidly becomes a serious restriction on the size of N.

b) no useful predictions can be made until all the observations $x_1, \ldots, x_N$ have been made and processed. In many applications it would be preferable to start making predictions (albeit rather bad ones) right from the start and have their quality improve as more and more data is assimilated.

c) in many situations the function G maynot be stationary but will vary slowly with time, or occasionally change suddenly. In such an environment a batch estimation scheme will be very unsatisfactory since it will repeatedly have to discard its previous estimate of G and compute a new one starting from scratch. Furthermore, as mentioned in b) during each such recalculation there will be a delay before the estimate based on new data becomes available.

It would thus be useful, particularly for real time signal processing applications to develop prediction algorithms which continuously update the estimate $\tilde{G}$ using new observations $x_{N+1}, x_{N+2}, \ldots$ as they are made. It turns out that this can be done using the framework of *recursive least squares estimation* [15]. Recursive least squares techniques are of course well known in linear signal processing and form the basis of most adaptive filter architectures [16]. Unfortunately the standard least squares algorithms used in such linear schemes are not sufficiently stable or accurate for application to chaotic time series and in proves necessary to use more sophisticated recursive approaches such as the Recursive Modified Gram-Schmidt (RMGS) algorithm of Ling et. al. [17]. This yields results comparable to those obtained from the SVD approach outlined above [18].

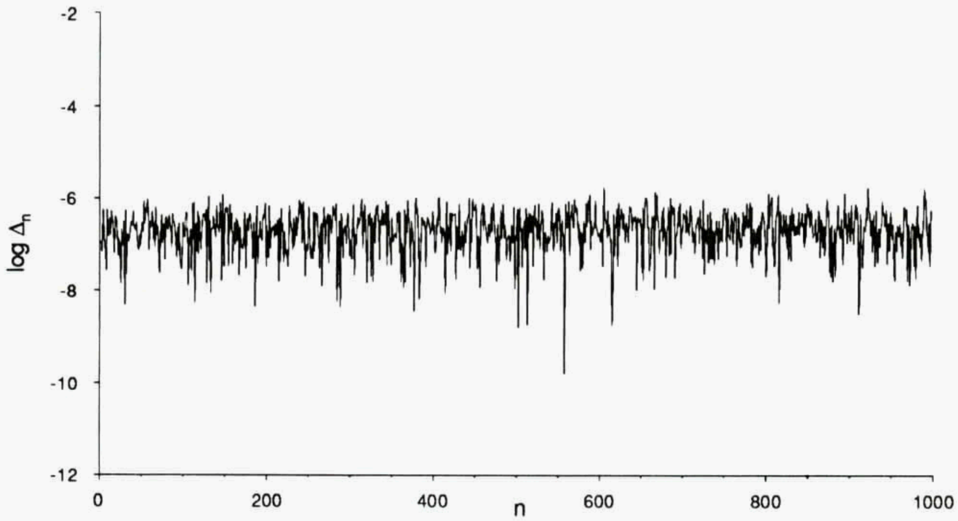


Figure 7. Logarithmic prediction error log Δ_n using SVD solution to least squares problem.

6. SIGNAL SEPARATION

So far we have considered the time series $\{x_n\}$ in isolation. In many applications, however, we are unlikely to be given such a pure chaotic signal. Instead, we will be asked to manipulate a mixture $z_n = x_n + s_n$ of a chaotic time series $\{x_n\}$ and some other signal $\{s_n\}$. The latter may represent noise, in which case we want to remove it from $\{z_n\}$, or it may be a signal that we wish to detect, in which case we want to extract it from $\{z_n\}$ and discard $\{x_n\}$. An example of the latter might be a faint speech signal $\{s_n\}$ masked by deterministic "noise" $\{x_n\}$ coming from some kind of vibrating machinery, such as an air conditioner [19]. In both cases, the mathematical problem that we face amounts to separating $\{z_n\}$ into its two components $\{x_n\}$ and $\{s_n\}$. Several schemes have been developed in the last few years for performing this task [19-22]. Here, we shall outline the overall framework behind all these approaches and describe a simple method recently developed in the Long Range Research Laboratory which is particularly appropriate when $\{s_n\}$ is a relatively slowly varying signal [22].

First observe that the decomposition of $\{z_n\}$ into $\{x_n\}$ and $\{s_n\}$ is not unique. In fact, we can choose $x'_1, \dots, x'_d$ arbitrarily, define the sequence $\{x'_n\}$ recursively by $x'_n = G(x'_{n-d}, x'_{n-d+1}, \dots, x'_{n-1})$ and obtain the trivial decomposition $z_n = x'_n + (z_n - x'_n)$. It is thus necessary to impose some additional constraints on $\{s_n\}$. The most common is to minimize the size of $\{s_n\}$ with respect to some appropriate norm [20]. In many cases, just requiring $\{s_n\}$ to be "small" for all n is sufficient to establish uniqueness. To see this, note that if $\{x_n\}$ is chaotic, then for most choices of $x'_1, \dots, x'_d$, the distance $|x_n - x'_n|$ will grow rapidly (if it is not large already). Hence if $\{s_n\}$ is small, $s'_n = z_n - x'_n = x_n - x'_n + s_n$ will be large for at least some values of n.

From now on we shall therefore restrict ourselves to the situation where $\{s_n\}$ is small in comparison to $\{x_n\}$. The problem of separating $\{z_n\}$ into its components then naturally falls into two parts:

a) Performing the decomposition when the function G is known.

b) Estimating the function G from the combined time series $\{z_n\}$ (rather than from $\{x_n\}$).

We already have all the tools required to solve the second problem. The basic idea is to apply the techniques of the last section to a large sample of $\{z_n\}$, in which case, with some luck, most of the effects of $\{s_n\}$ will average out and a reasonable estimate of G can be made. We can then proceed itera-

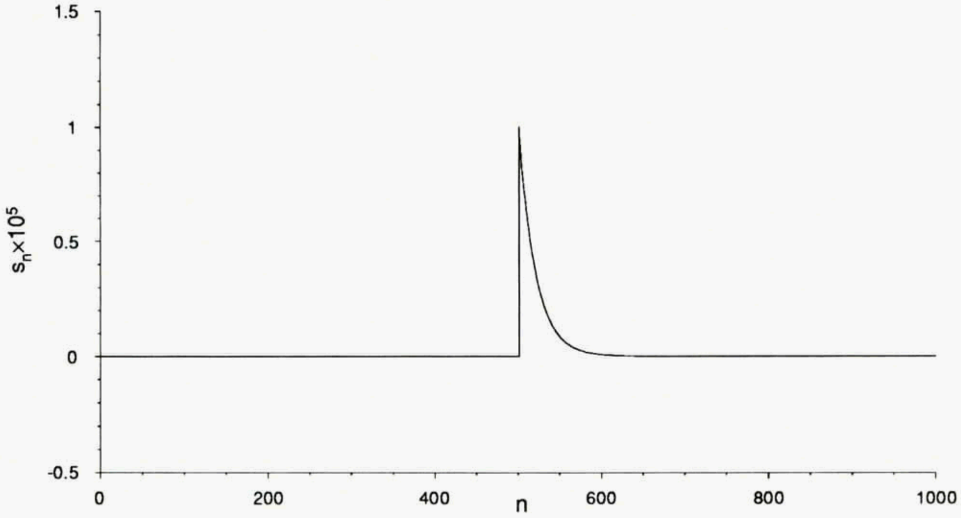


Figure 8. Sample signal s_n used for signal separation experiments.

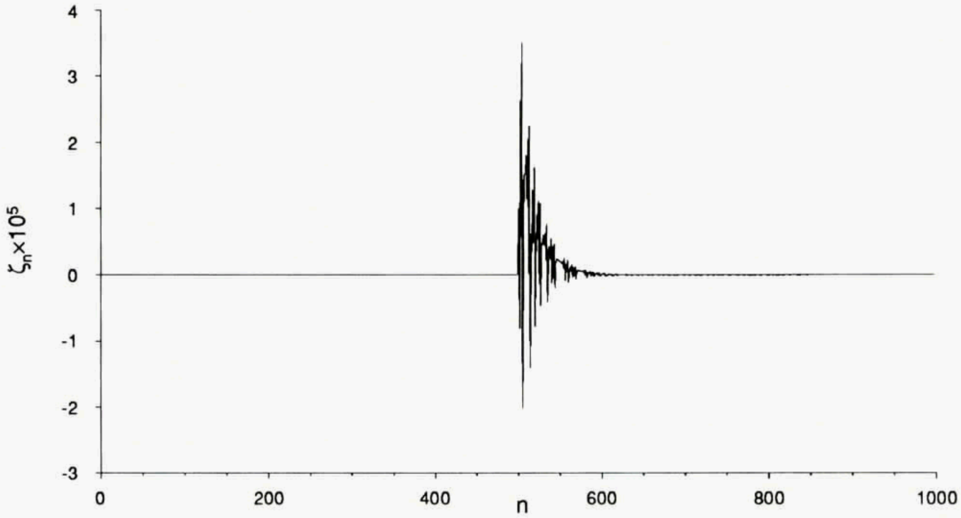


Figure 9. Prediction discrepancy ζ_n for signal from Fig. 8.

tively, using this estimate of G to obtain an estimate $\{\tilde{s}_n\}$ of $\{s_n\}$. Then $\tilde{x}_n = z_n - \tilde{s}_n$ should be much more deterministic than $\{z_n\}$ and hence we should be able to obtain a better estimate of G from it. This procedure can then be repeated as often as necessary.

Let us thus turn our attention to the first problem, namely that of performing the decomposition when G is known. The basic approach is to look for discrepancies between the observed value of z_n and that which is predicted by the deterministic dynamics $G(z_{n-d}, z_{n-d+1}, \dots, z_{n-1})$. If $s_{n-d} = s_{n-d+1} = \dots = s_n = 0$ then this discrepancy will be zero. A non-zero value of $\zeta_n = z_n - G(z_{n-d}, z_{n-d+1}, \dots, z_{n-1})$ will therefore indicate the presence of some non-trivial signal. This is illustrated by Figures 8 and 9. The signal from Fig. 8 was added to the Hénon time series from Fig. 1 and the resulting $\{\zeta_n\}$ is shown in Fig. 9. It is clear that the discrepancy $\{\zeta_n\}$ is able to detect the presence of the pulse $\{s_n\}$ and even to some extent extract its qualitative features, but does not yield much in the way of quantitative information.

To proceed further we expand ζ_n assuming that $\{s_n\}$ is small:

$$\begin{aligned} \zeta_n &= z_n - G(z_{n-d}, z_{n-d+1}, \dots, z_{n-1}) \\ &= x_n + s_n - G(x_{n-d} + s_{n-d}, \dots, x_{n-1} + s_{n-1}) \qquad (17) \\ &\cong s_n - \sum_{i=1}^{d} s_{n-i} \frac{\partial G(z_{n-d},\dots,z_{n-1})}{\partial z_{n-i}} \qquad (18) \end{aligned}$$

Allowing n to vary, Eq. 18 gives a set of simultaneous linear equations for $\{s_n\}$. These lie at the heart of most approaches to signal separation. Although several different techniques can be used to solve these equations, great care has to be taken when the dynamics of $\{x_n\}$ is chaotic, since in that case this set of equations becomes very badly conditioned. One possibility is to use the Singular Value Decomposition (see above), which is able to cope with very badly conditioned linear problems. This is essentially the technique used by Farmer and Sidorowich [20]. They, in fact, solve Eq. 18 repeatedly, regarding it as a Newton step in solving the full non-linear problem given by Eq. 17. They also impose additional equations designed to ensure that the final $\{s_n\}$ has minimal norm. A simpler alternative, which works well in practice [23] is to simply solve Eq. 18 using SVD.

Here we describe an alternative approach aimed at the situation where $\{s_n\}$ is slowly varying. This for instance is the case for the signal in Fig. 8, except in a small neighbourhood of the

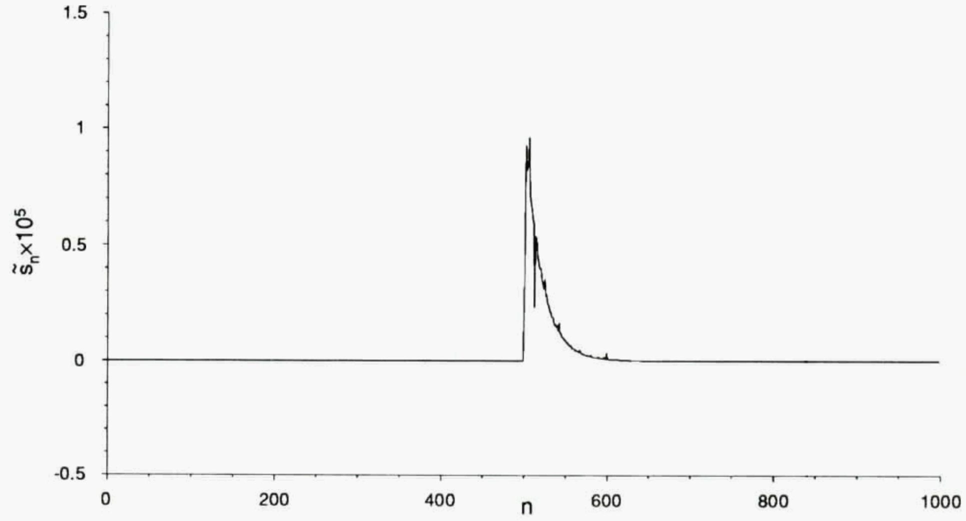


Figure 10. Extracted signal $\tilde{s}_n$ using Eq. 21, with same data as Fig. 9.

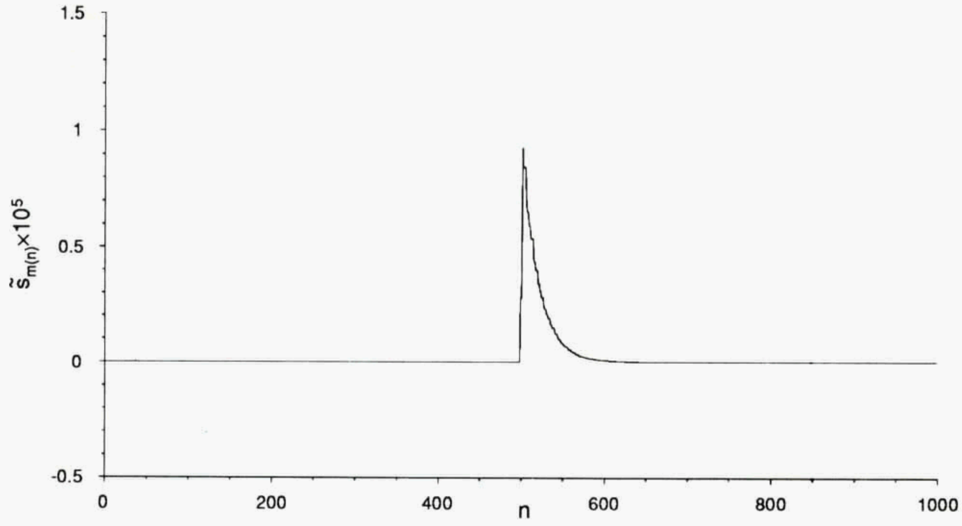


Figure 11. Effect of improved estimate using $\tilde{s}_{m(n)}$ instead of $\tilde{s}_n$.

transient (and as we shall see our scheme works reasonably well even there). We thus assume that $|s_n - s_{n-1}|$ is small in comparison to $|s_n|$. It is then reasonable to set $s_{n-d} \cong s_{n-d+1} \cong \ldots \cong s_n$ in Eq. 18, which gives

$$\zeta_n \cong s_n(1 - J_n) \quad (19)$$

where

$$J_n = \sum_{i=1}^{d} \frac{\partial G(z_{n-d}, \ldots, z_{n-1})}{\partial z_{n-i}} \quad (20)$$

This gives

$$\tilde{s}_n \cong \frac{\zeta_n}{1 - J_n} \quad (21)$$

as an estimate for s_n. Figure 10 shows the results of using Eq. 21 with the same data as Fig. 9. We get extremely good recovery of $\{s_n\}$ for almost all n, except for occasional values which can be quite wildly wrong. These are largely due to J_n coming close to 1, indeed the error in Eq. 21 is inversely proportional to $|1 - J_n|$. This can be overcome to a large extent by making the observation that under the assumption that $s_{n-d} \cong s_{n-d+1} \cong \ldots \cong s_n$, Eq. 21 is an equally valid estimate for any of $s_{n-d}, s_{n-d+1}, \ldots, s_n$. Thus, it is reasonable to use any of $\tilde{s}_n, \tilde{s}_{n+1}, \ldots, \tilde{s}_{n+d}$ as an estimate for s_n. The best estimate will be given by that $\tilde{s}_m$ for which the corresponding $|1 - J_m|$ is maximized. Let m(n) be this value, so that $n \leq m(n) \leq n+d$ and $|1 - J_{m(n)}| \geq |1 - J_k|$ for all $n \leq k \leq n+d$. Figure 11 plots $\tilde{s}_{m(n)}$ for the same data as Fig. 10, and we see that the occasional spurious values in Fig. 10 are removed without any adverse effects on the remainder of the signal.

This example shows that, at least in certain cases, it is possible to recover $\{s_n\}$ extremely accurately when G is known exactly. To demonstrate that this scheme is still useful when G has to be approximated from the combined signal $\{z_n\}$, we present the results of one last experiment. This time we take the random two-level signal of Figure 12 as $\{s_n\}$. This is added to Fig. 1 to give $\{z_n\}$ which is used to estimate G using the SVD as above. The same parameters were used as for Fig. 7, except that N was increased to 1000, so that the whole data sample was used in the computation of $\tilde{G}$. The above extraction algorithm was then applied using this $\tilde{G}$. The results are shown in Figure 13. We

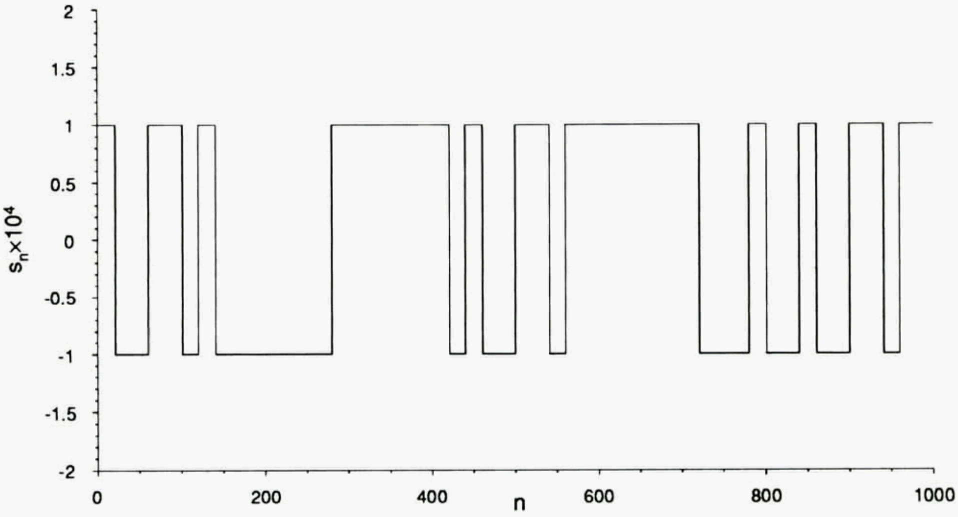


Figure 12. Sample signal s_n used for signal separation with unknown G.

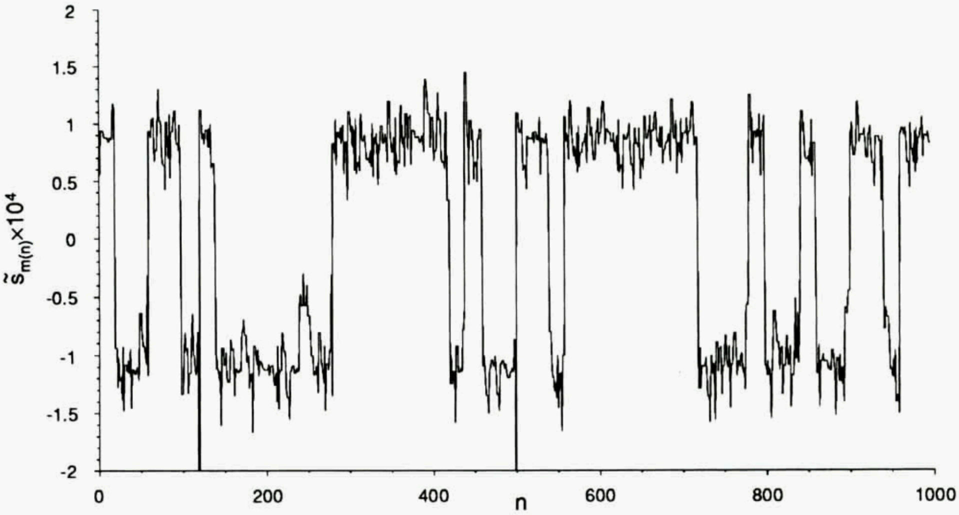


Figure 13. Extracted estimate of signal in Fig. 12.

can see that using an estimate of G, rather than G itself, leads to some loss of accuracy in recovering $\{s_n\}$, but nevertheless a useful signal can still be extracted.

Observe that in traditional signal processing terms, if we identify $\{x_n\}$ as the "noise" contaminating the signal $\{s_n\}$, we are able to recover the signal at a signal to noise ratio of -80dB. This is of course far in excess of what could be done using conventional linear filtering. In this context it should be pointed out that, in common with most such schemes, the above algorithm's performance increases as the amplitude of $\{s_n\}$ decreases, down to some lower limit set by numerical inaccuracy. This is in complete contrast to conventional signal processing techniques where signal extraction performance deteriorates with decreasing signal amplitude.

7. CONTROLING CHAOTIC SYSTEMS

So far, we have considered the presence or absence of chaos in a given system as something beyond our control. In many cases however, we would like to avoid chaotic behaviour, or at least modify it in such a way as to perform a useful function. One approach would be to make large and possibly costly alterations to the system to completely change its dynamics and ensure that the modified system is not capable of chaotic behaviour. This may, however, not be feasible for many practical reasons. An alternative, due to Ott et. al. [24] is to attempt to control the system using only small time dependent perturbations of an accessible system parameter. The key fact behind their idea is that a typical chaotic system, such as the Hénon map, will contain an infinite number of unstable periodic orbits. Normally, these will not be observed, precisely because they are unstable, but any typical trajectory of the system will come arbitrarily close to them infinitely often. This can be see in Fig. 1, where $\{x_n\}$ comes close to an unstable fixed point near to $x = 0.6$, at about $n = 150$ and again at $n = 640$ and $n = 690$.

The effects of the unstable periodic orbits can thus be seen in an observed time series $\{x_n\}$. Several groups have shown that an accurate estimate of their position and eigenvalues can be derived from the time series (e.g. [25,26]). Such estimates make use of many of the techniques for analysing chaotic time series that we have described in the previous sections. As usual, Takens's theorem is used to reconstruct the original dynamical system, whilst a local estimate to the function G around a periodic orbit can be used to calculate the orbit's eigenvalues.

The basic idea behind controlling a chaotic system is to choose one of these periodic orbits and attempt to stabilize it by small perturbations of a parameter. This is possible precisely because chaotic systems are so sensitive to small changes. Normally, such sensitivity simply leads to instability and complex behaviour of the kind seen in Fig. 1. However, if the perturbations are carefully chosen, they can push the system into the desired periodic regime and then keep it there. Ott et. al. first demonstrated the feasibility of their algorithm using numerical simulations [24], but since then it has been applied successfully to the control of a variety of real systems [27-29]. Once again, this algorithm relies on a local approximation of the function G around the periodic orbit.

One of the great advantages of this approach is that there is potentially a large number of different periodic orbits which we can stabilize. We can thus choose precisely that orbit which gives the best system performance in a given application. Furthermore, we can easily switch amongst the different orbits available, again using only small changes in the control parameter. In principle, it should thus be possible to obtain many substantially different classes of behaviour from the same chaotic system. This is in complete contrast to systems which lack a chaotic attractor and operate at a stable equilibrium or periodic orbit. For such systems, small parameter perturbations can only move the orbit by a small amount, but cannot generally lead to dramatically different behaviour. One is thus essentially restricted to whatever behaviour is given by the stable orbit and it is difficult to make substantial improvements in performance without major changes to the system.

A further extension to this idea is described in [30]. Here, rather than aiming to operate the system in a given periodic steady state, one is trying to direct the state of the system to a desired target state in as short a time as possible using only small perturbations of the control parameter. It turns out that essentially the same framework as above can be used to achieve this, and once again a chaotic system's extreme sensitivity to small perturbations can be used to our advantage.

REFERENCES

[1] T. Sauer, J. A. Yorke and M. Casdagli, 1991, Embedology, *J. Stat. Phys.*, **65**, 579-616.

[2] L. Noakes, 1992, The Takens Embedding Theorem, *Int. J. of Bifurcation and Chaos*, **1**, 867-872.

[3] J.-P. Eckmann and D. Ruelle, 1985, Ergodic Theory of Chaos and Strange Attractors, *Rev. Mod. Phys.*, **57**, 617-656.

[4] P. Grassberger, T. Schreiber and C. Schaffrath, 1992, Nonlinear Time Sequence Analysis, *Int. J. of Bifurcation and Chaos*, **1**, 521-547.

[5] J. D. Farmer and J. J. Sidorowich, 1987, Predicting Chaotic Time Series, *Phys. Rev. Lett.*, **59**, 845-848.

[6] P.S. Linsay, 1991, An Efficient Method of Forecasting Chaotic Time Series Using Linear Interpolation, *Phys. Lett. A*, **153**, 353-356.

[7] A. S. Lapedes and R. Farber, 1987, Nonlinear Signal Processing Using Neural Networks: Prediction and System Modelling, Technical Report, LA-UR-87-2662, Los Alamos National Laboratory.

[8] M. Casdagli, 1989, Nonlinear Prediction of Chaotic Time Series, *Physica D*, **35**, 335-356.

[9] M. Casdagli, 1992, Chaos and Deterministic versus Stochastic Non-linear Modelling, *J. Roy. Stat. Soc. B*, **54**, 303-328.

[10] M. J. D. Powell, 1987, Radial Basis Functions for Multivariable Interpolation : A Review, in *Algorithms for Approximation*, ed. J. C. Moxon and M. G. Cox, Clarendon Press, Oxford.

[11] D. S. Broomhead and D. Lowe, 1988, Multivariable Functional Interpolation and Adaptive Networks, *Complex Systems*, **2**, 321-355.

[12] C. L. Lawson and R. J. Hanson, 1974, *Solving Least Squares Problems*, Prentice-Hall, New Jersey.

[13] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, 1988, *Numerical Recipes in C, The Art of Scientific Computing*, Cambridge University Press, Cambridge.

[14] J. Stoer and R. Bulirsch, 1980, *An Introduction to Numerical Analysis*, Springer-Verlag, Berlin.

[15] P. Young, 1984, *Recursive Estimation and Time-Series Analysis, An Introduction*, Springer-Verlag, Berlin.

[16] S. T. Alexander, 1986, *Adaptive Signal Processing, Theory and Applications*, Springer-Verlag, Berlin.

[17] F. Ling, D. Manolakis and J. G. Proakis, A Recursive Modified Gram-Schmidt Algorithm for Least-Squares Estimation, *IEEE Trans. ASSP*, **34**, 829-835.

[18] J. Stark, 1992, Recursive Prediction of Chaotic Time Series, submitted to *J. of Nonlin. Sci.*

[19] W. W. Taylor, 1991, Quantifying Predictability for Applications in Signal Separation, to appear in *SPIE Proceedings*, **1565**.

[20] J. D. Farmer and J. J. Sidorowich, 1991, Optimal Shadowing and Noise Reduction, *Physica D*, **47**, 373-392.

[21] T. Schreiber and P. Grassberger, 1991, A Simple Noise-Reduction Method for Real Data, *Phys. Lett. A*, **160**, 411-418.

[22] J. Stark and B. Arumugam, 1992, Extracting Slowly Varying Signals from a Chaotic Background, to appear in *Int. J. of Bifurcation and Chaos*.

[23] J. Stark, 1992, Extracting Signals from Chaotic Time Series Using the Pseudo-Inverse, in preparation.

[24] E. Ott, C. Crebogi and J.A. Yorke, 1990, Controlling Chaos, *Phys. Rev. Lett.*, **64**, 1196-1199.

[25] D. Auerbach, P. Cvitanović, J.-P. Eckmann, G. Gunaratne and I. Procaccia, 1987, Exploring Chaotic Motion Through Periodic Orbits, *Phys. Rev. Lett.*, **58**, 2387-2389.

[25] D. P. Lathrop and E. J. Kostelich, 1989, Characterization of an Experimental Strange Attractor by Periodic Orbits, *Phys. Rev. A*, **40**, 4028-4031.

[27] W.L. Ditto, S.N. Rauseo and M.L. Spano, 1990, Experimental Control of Chaos, *Phys. Rev. Lett.*, **65**, 3211-3214.

[28] J. Singer, Y.-Z. Wang and H.H. Bau, 1991, Controlling a Chaotic System, *Phys. Rev. Lett.*, **66**, 1123-1125.

[29] E.R. Hunt, 1991, Stabilizing High-Period Orbits in a Chaotic System: the Diode Resonator, *Phys. Rev. Lett.*, **67**, 1953-1955.

[30] T. Shinbrot, E. Ott, C. Crebogi and J.A. Yorke, 1990, Using Chaos to Direct Trajectories to Targets, *Phys. Rev. Lett.*, **65**, 3215-3218.

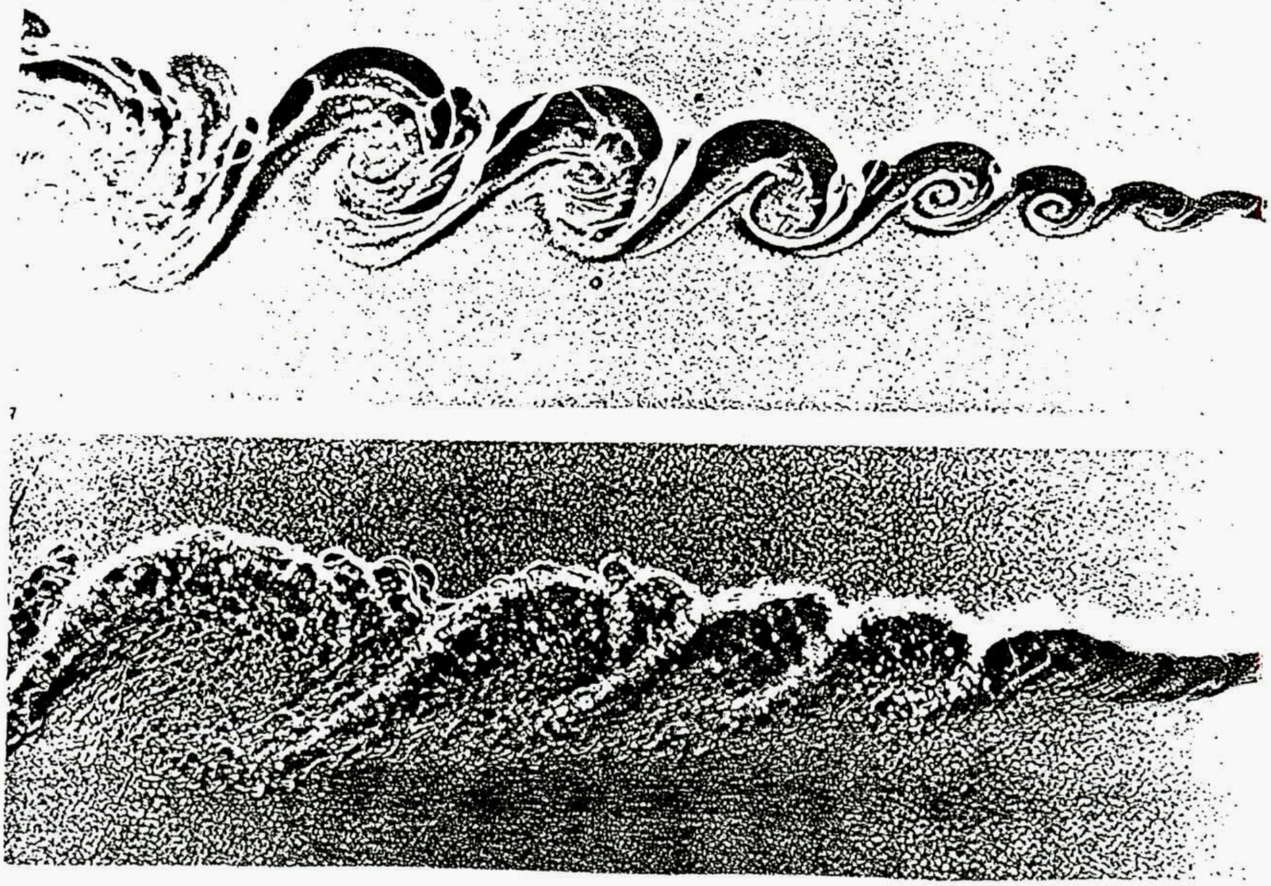
Exploring chaos: A toolkit and some ways to exploit it.

Manuel SAMUELIDES* **

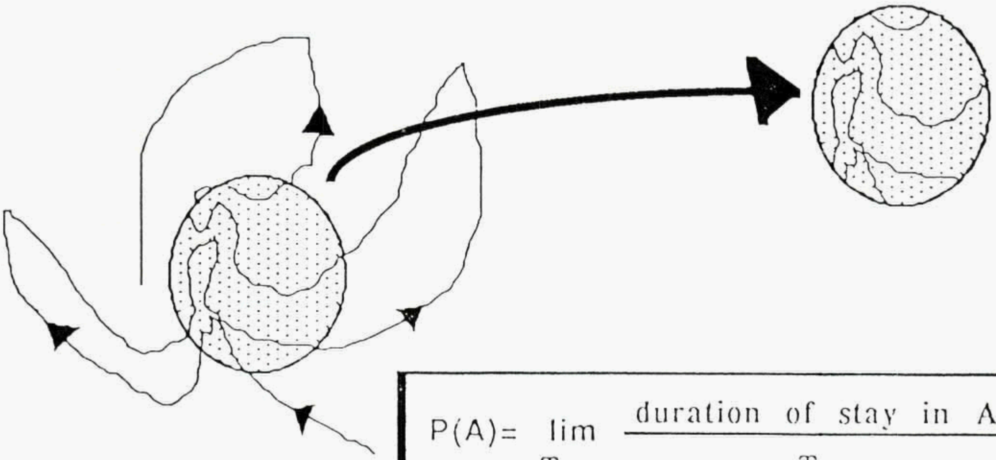
* Ecole Nationale de l'Aéronautique et de l'Espace, BP 4032, 31055 Toulouse Cedex FRANCE

** Groupement d'Intelligence Artificielle, Office National d'Etudes et de Recherches Aerospatiales, Centre d'Etudes et de Recherches de Toulouse.

1. First characteristics of Chaos
2. Attractors (Delay reconstruction, Poincare map)
3. Weak chaos and Smale's horsehoe
4. Strong chaos and Liapounov exponents
5. Controlling chaos



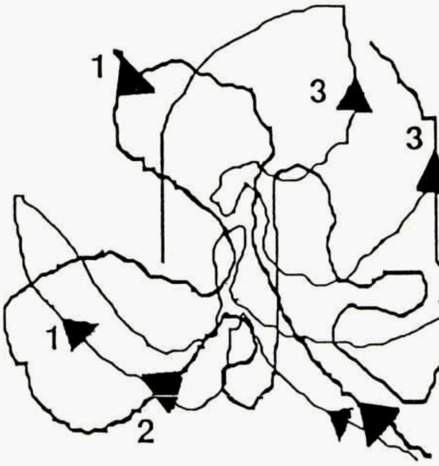
Turbulent flows (Experimental datas)
(From Brown & Roshko , Cal Tech)



$$P(A) = \lim_{T \to \infty} \frac{\text{duration of stay in } A}{T}$$

Inadequacy of deterministic modelisation/
Relevance of stochastic one modelisation

Unpredictibility:



Sensitivity to Initial Conditions

Vanishing of correlation function with time

$$C(u)=\frac{1}{T}\int_0^T x(t)\,x(t+u)\,dt$$

$$\lim_{u\to\infty} C(u) = 0$$

Continuity of Power Fourier Spectrum

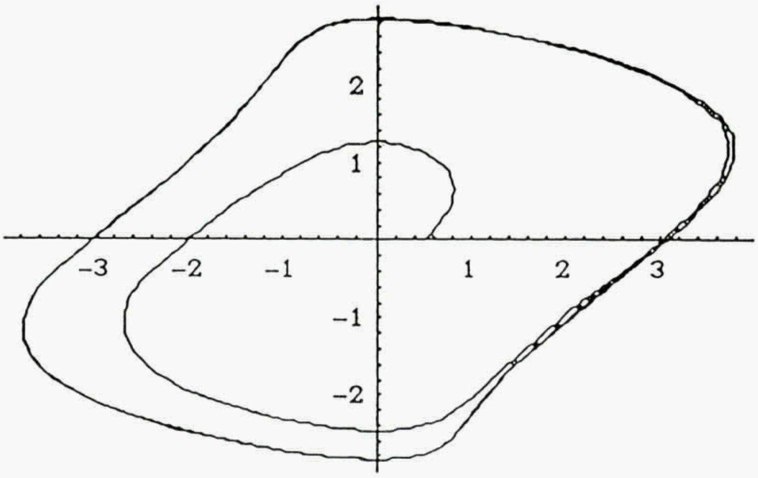
How is it possible for deterministic PDE to produce stationary solutions that behave like stochastic processes ?

The concept of ATTRACTOR

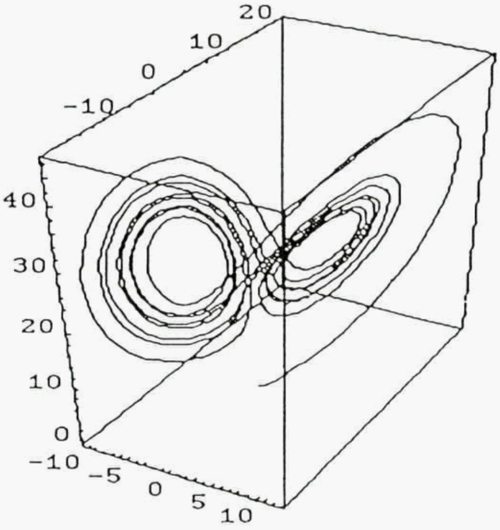
Attracting set: A closed invariant set which attracts any point in an open neighbourhhod

Examples:

- Asymptotically stable equilibrium
- Stable limit cycle: Van Der Pole Equation

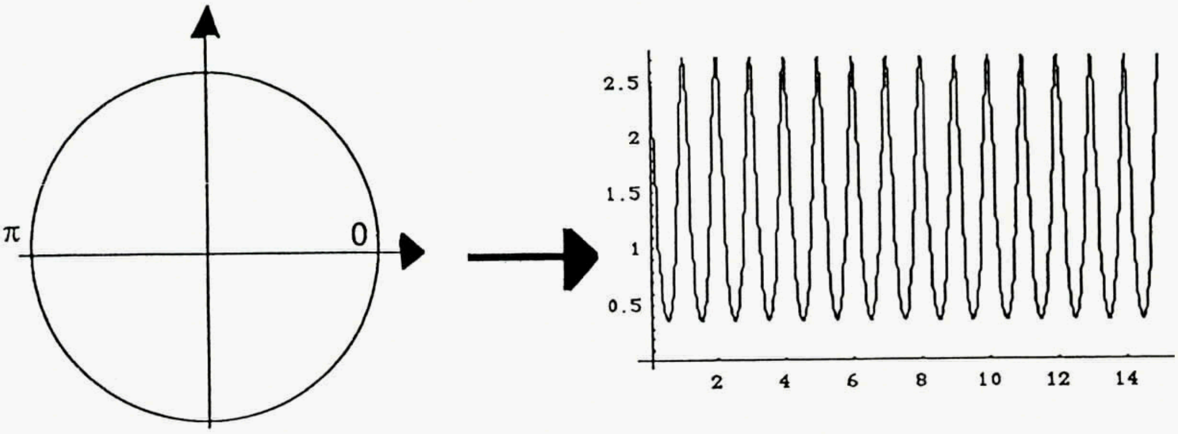


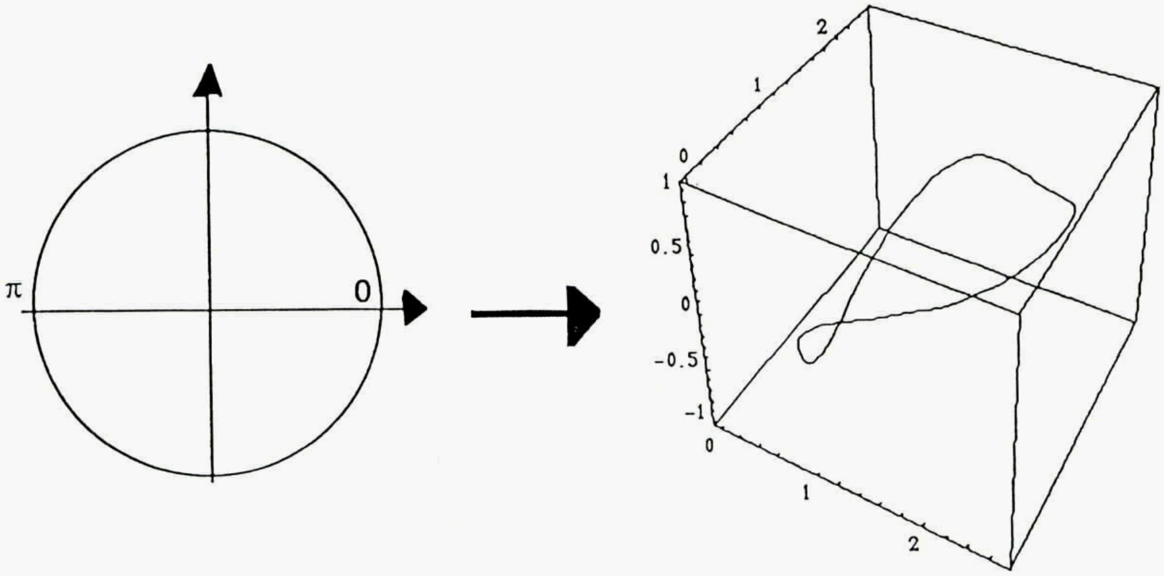
- Strange attractor: (Lorenz Equation)



Periodic orbit: $t \in [0, 15]$; $x(t) = \mathrm{Exp}[\mathrm{Cos}(2\pi t)]$,

$t \to e^{2i\pi t}$ $\omega \to e^{\mathrm{Cos}\omega}$





Periodic correlation function

Discrete spectrum: Dirac comb

(Multiple of the fundamental frequency)

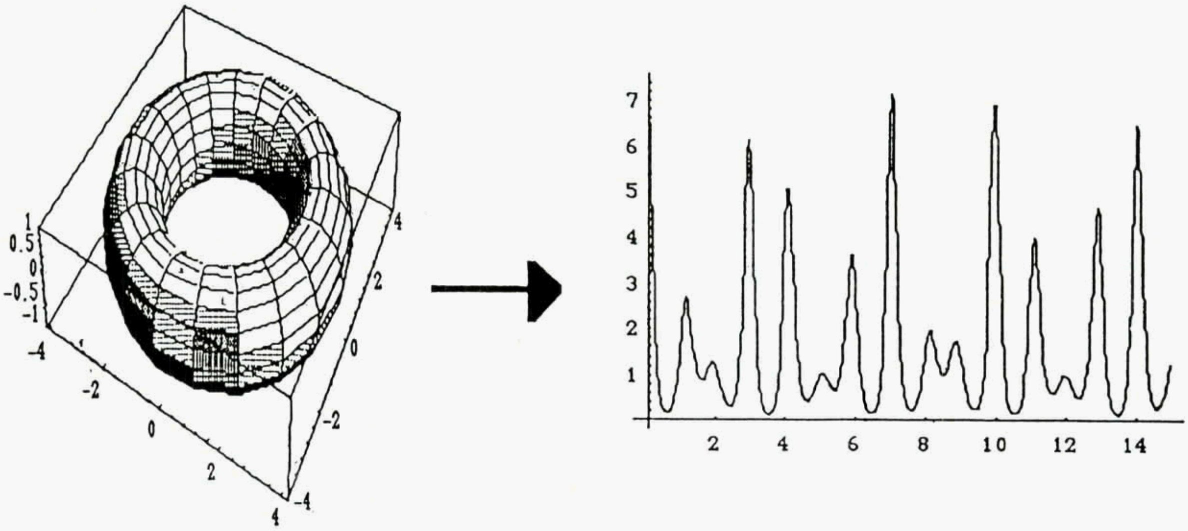
Almost-periodic orbit:

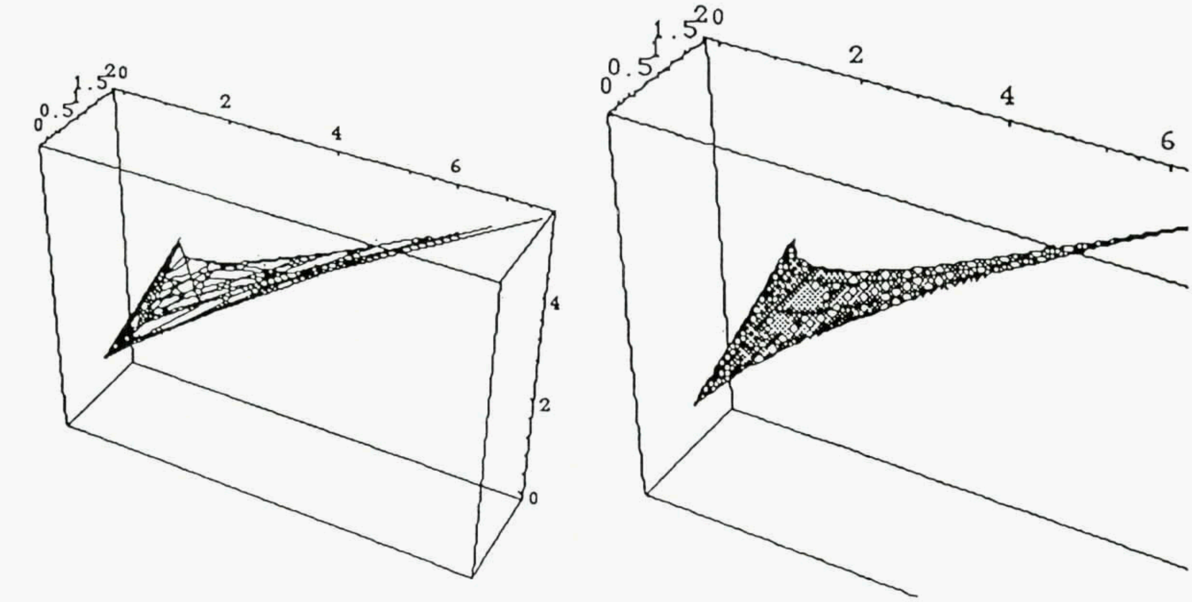
ergodic orbit on a n-dimensional torus

$t \in [0, 15]$; $x(t) = \mathrm{Exp}[\mathrm{Cos}(2\pi t)+\mathrm{Cos}(\sqrt{2}\pi t)]$

$t \to (e^{2i\pi t}, e^{\sqrt{2} i\pi t})$

$(\omega_1, \omega_2) \to e^{\mathrm{Cos}(\omega_1)+\mathrm{Cos}(\omega_2)}$





Almost-periodic correlation function

Point dense spectrum:

(Multiple of n fundamental frequencies:

High multiple are generally weak)

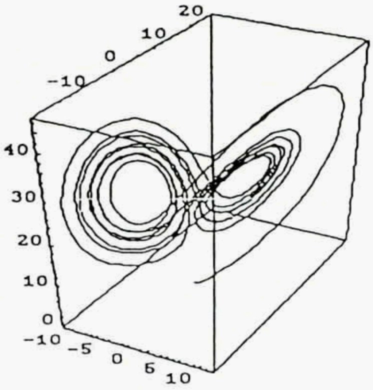
Landau interpretation of turbulence:

Attractor in the infinite dimensional phase space of fluid mechanics is a n-d torus with n big

(A likely explanation for continuous spectrum)

Lorenz discovery (1963)

Chaotic dynamics may occur in low dimensional manifolds



Lorenz system:

$x' = -3(x-y)$

$y' = 26{,}5\ x - y - xz$

$z' = xy - z$

Rössler Band:

$x' = -y - z$

$y' = x + ay$

$z' = 2 + z(x-4)$

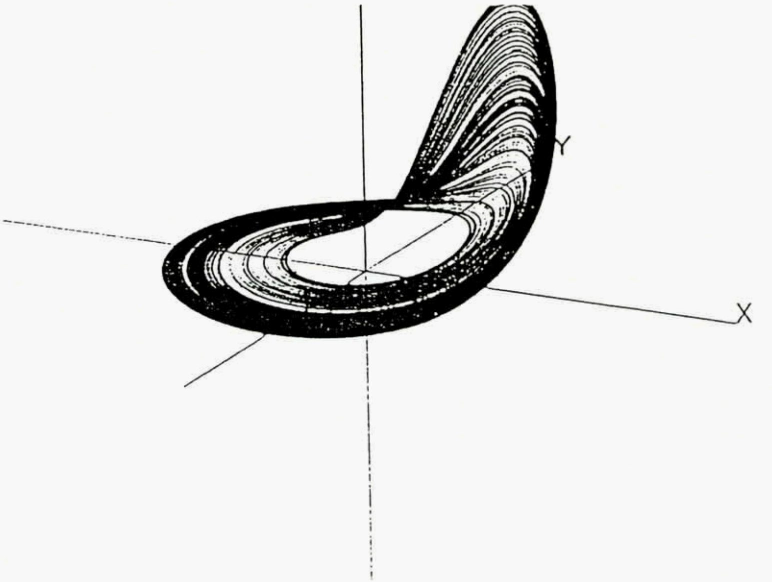


Figure 12.1 A post-transient trajectory of Rössler's equations (12.1) for the simply folded band attractor. Parameters are $a = 0.398$, $b = 2$, $c = 4$

From Thompson & Stewart (Wiley)

Reconstruction of the attractor

When an experimental signal is studied,
identify a low-dimensional attractor
into a high-dimensional phase space

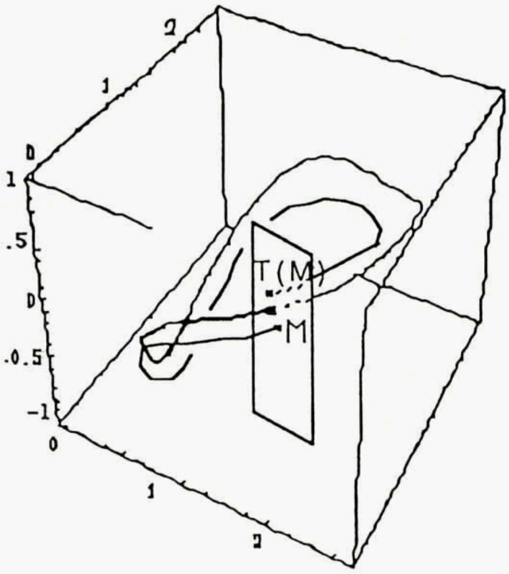
The Delay method:
Take an observable $y(t)=f[x(t)]$
Study trajectory of delay-coordinate map
$\{y(t),y(t-\Delta t),\ldots,y(t-n\Delta t)\}$

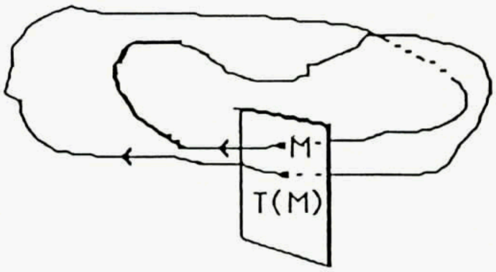
d dimension of the attractor & $n>2d$

$\Downarrow$

Embedding of the attractor into R^n

Poincare section and strange attractors





Continuous Flow ⇒ Transformation of the section

Limit-cycle ⇒ Fixed point

Doubling of period ⇒ Periodic point

ergodic orbit on 2d torus ⇒ Irrational translation on the circle

Weak chaos: Spreading and folding

(fractal dimension)

Smale's horsehoe

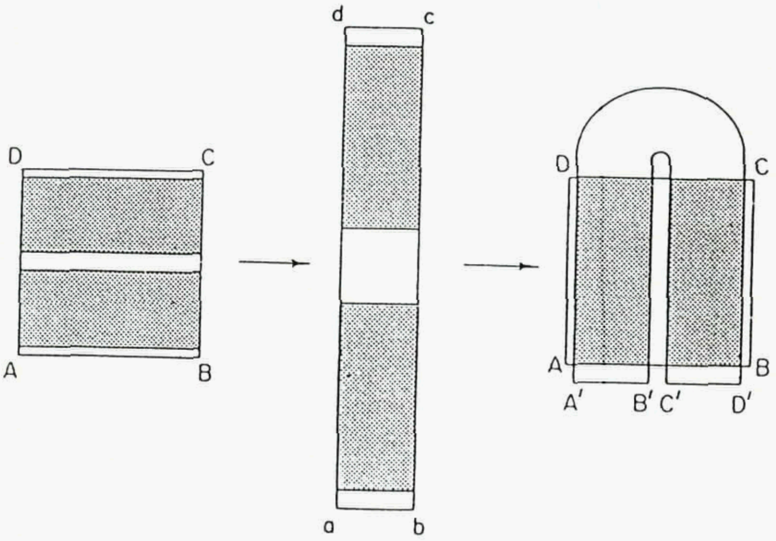


Figure 12.6 The horseshoe mapping of Smale, taking points in a square onto a U-shaped region overlapping the square

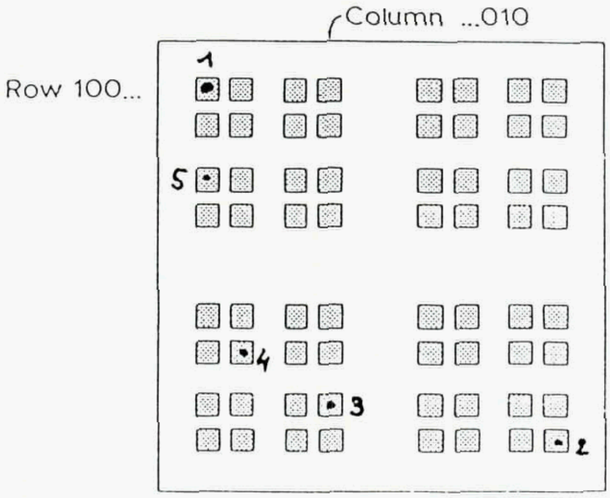


Figure 12.8 An approximate picture of the invariant set of the horseshoe mapping

Symbolic modelization of the Horsehoe: Chaotic dynamics of the shift

$$\Omega = \{0,1\}^{Z} \quad ; d(x,y) = \sum_{i \in Z} 2^{-|i|}(x_i - y_i)$$

Topological structure of Cantor set

$T(\ldots 0,1,0,0,1,1,0 \ldots) = (\ldots 1,0,1,0,0,1,1 \ldots)$

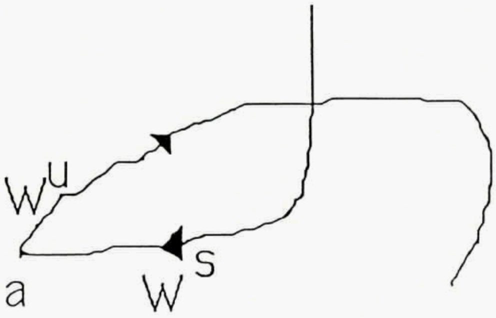
2 fixed points: (...0,0,0...) &(...1,1,1...)

Hyperbolic fixed point: a=(...0,0,0...)

Stable manifold: $\{x \,/ T^{n}(x) \rightarrow a\} = \{(\ldots 00 \ldots, 0\S\S\S \ldots)\}$

Unstable manifold: $\{x \,/ T^{-n}(x) \rightarrow a\} = \{(\ldots, \S\S\S 0, \ldots, 0, 0 \ldots)\}$

The stable and the unstable manifolds are intersecting transversally in an infinite number of points



Infinite number of unstable periodic orbits

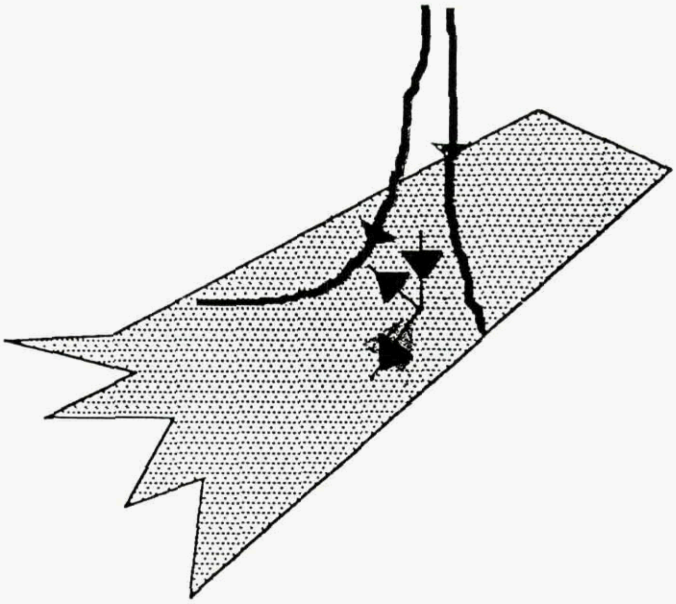
A weak chaotic dynamical system

contains a Smale's horseoe

Strong chaos:

Exponential divergence of trajectories

Liapounov exponents



Ergodic theorem of Ossedelets

Let $F_x = D_xT$,

$F_x^{(n)} = D_x(T^n) = F_{T^{n-1}(x)} \cdot \ldots \cdot F_x$

Then $\lim_{n\to\infty}[F_x^{(n)*}F_x^{(n)}]^{1/2n} = \Lambda_x$

exists and doesn't depend on x (a.e.)

Liapounov exponents are Log of the eigenvalues

Strong chaos: positive Liapounov exponent

Computation of Liapounov exponents is based on triangularization of F_x for each step:

- F_x is known then do $F_x = Q_x R_x$

Let $\mu_x^{(i)}$ be the i-th eigenvalue of R_x

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \mathrm{Log}\ \mu_{Tk(x)}^{(i)}$$

- F_x is not known

(experimental signal and/or reconstruction using delays)

Find other points y in the vicinity of x

such that the orbit of y stays in the

vicinity of the orbit of x for some iterations

Liapounov exponents are not smooth functions

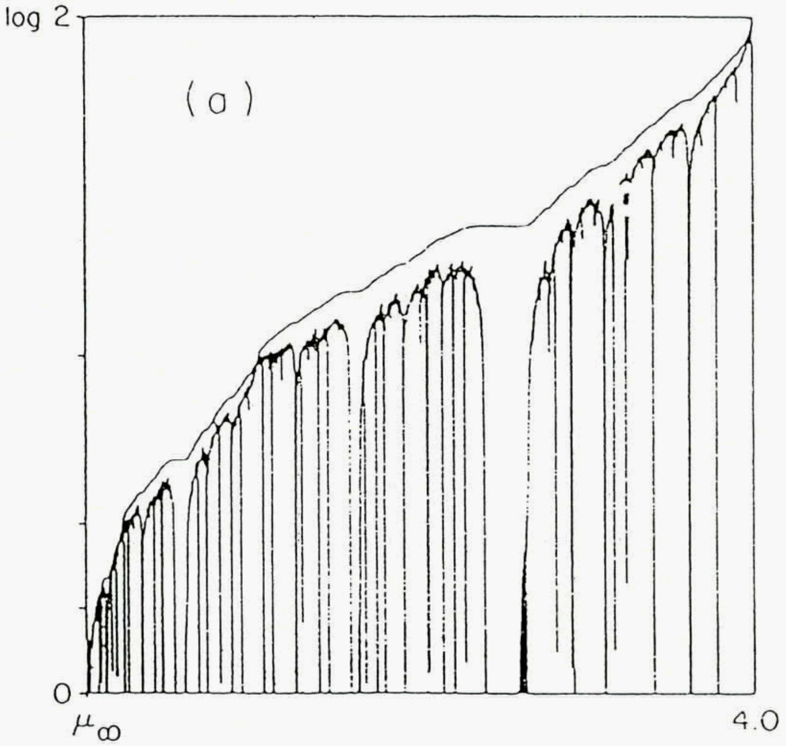


FIG. 12. (a) Topological entropy (upper curve) and characteristic exponent (lower curve) as a function of μ for the family $x \to \mu x(1-x)$. (Graph by J. Crutchfield.) Note the discontinuity of the lower curve.

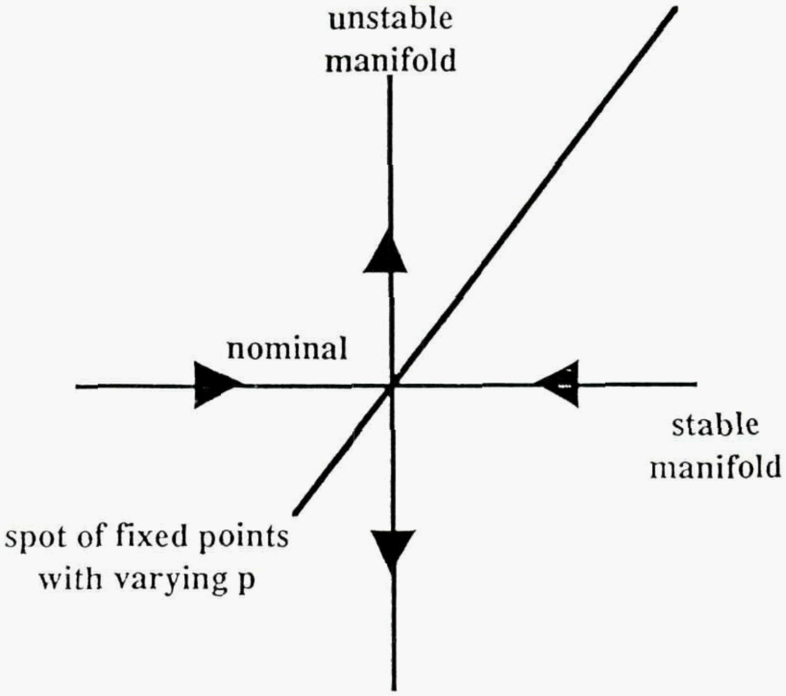
Controlling chaos

(from Ott,Grebogi,Yorke , *Phys.Rev.Letters* [1990])

Principle:

Stabilize an unstable periodic orbit by slight perturbation of the parameter.

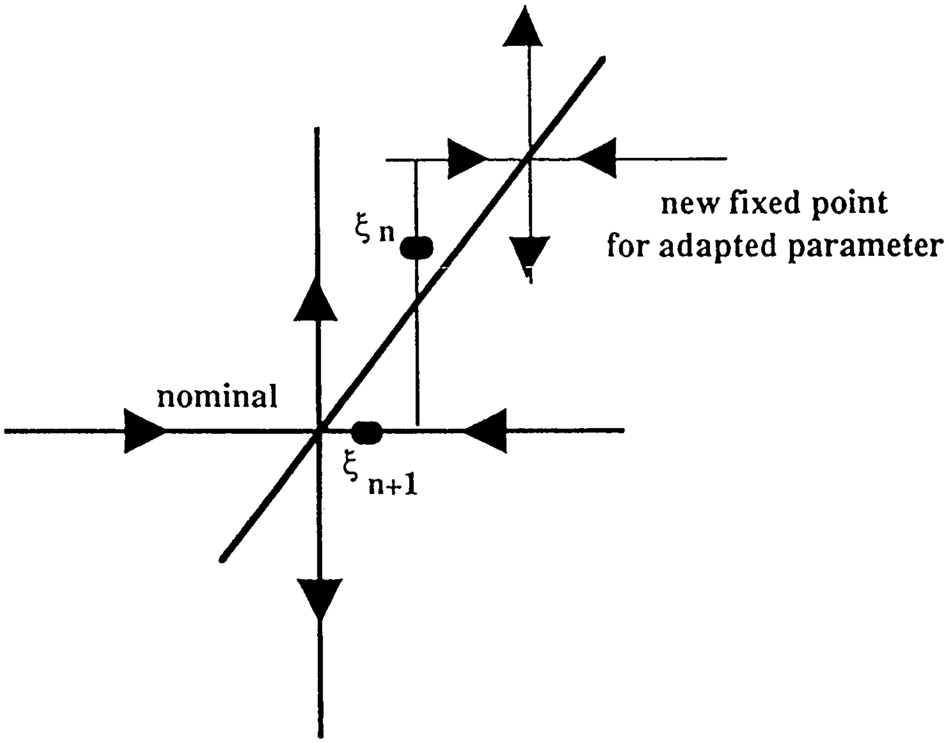
Close from linearization of controlled dynamical systems



Hypothesis:

Predetermination of the nominal orbit
Knolwedge of the characteristic exponents
Structural stability of the nominal orbit

Iterative adaptation of the parameter



Beware the effects of non-linearity and noise

Trajectory may escape out of controlled zone

Experimental realization

Cancellation of chaos with weak periodic perturbations

(from Braiman,Goldhirsch , *Phys.Rev.Letters* [1991])

Model: The parametric oscillator

$$\boxed{\theta'' + G\theta' + \sin\theta = I + A\sin\omega t + a\sin\beta\omega t}$$

2-periodic perturbation

widely studied for $\beta = 0$

Model of shunted Josephson junction

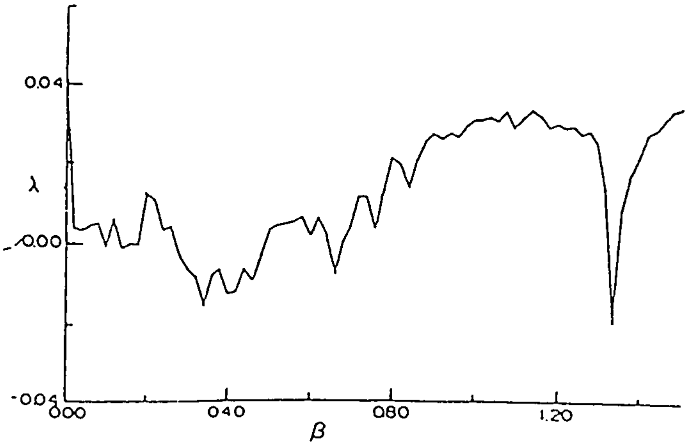


FIG. 1. Leading Liapunov exponent λ as a function of the parameter β (see text) with $G=0.7$, $A=0.4$, $I=0.905$, $w=2\pi/25.12$, and $a=0.0125$. The points represent actual results and they are connected by lines to guide the eye.

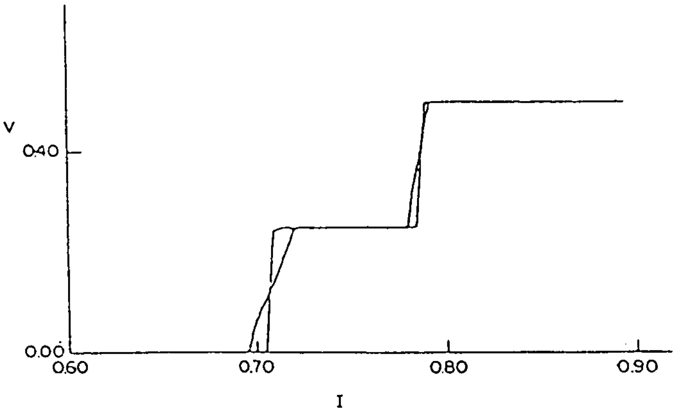


FIG 3. I-V characteristics corresponding to the parameters $G=0.7$, $A=0.4$, $w=2\pi/25.12$, $a=0.0125$, and $\beta=1.11803$. The smooth curve corresponds to $a=0$, i.e., no external perturbation. The dotted curve corresponds to $a=0.0125$ and $\beta=1.11803$.

Conclusions:

Chaotic dynamics is encountered in various situations (physics, biology, ...)

To controll it one can use some good properties:

Structural stability → Robustness

Exploration of a large zone of the state space

An important drawback

Unstability → difficult computation of characteristics

A research track:

Adaptive control and learning:

Let the system find its representation of the target by an unstable orbit

Favour this learning process

Stability Analysis and Aerospace Vehicle Dynamics

Pierre Y. Willems*
Université Catholique de Louvain
AUTO/CESAME
Place du Levant,2
B-1348 Louvain-la-Neuve
Belgium

1. SUMMARY

This paper presents, stability analysis results which can be useful to analyse the behaviour of aerospace vehicles.

A certain number of definitions are recalled. Liapunov stability criteria are given for autonomous systems described by ordinary differential equations and discrete-time equations.

A particular attention is paid to the stability of mechanical systems around equilibrium configurations. It is shown that the hamitonian function is a good Liapunov function for a rather large class of mechanical systems. When constraints are present and for particular dissipation interactions depending linearly on the variables, modified Liapunov functions are presented.

For general linear systems, stability criteria based on the characteristic polynomial are recalled. Extension permits to consider stability robustness and the introduction of isolated nonlinearities. Finally, some results on exact feedback linearization are presented.

Attitude stability analysis of orbiting satellites (and gyrostats) are presented as particular applications.

2. INTRODUCTION

The investigation of the stability is an important problem in dynamics and in particular for aerospace vehicle. When, the stability cannot be obtained passively, the possibility of an active stabilization by feedback can be considered. In both cases, the robustness of the stability (with respect to parameter changes and/or perturbations) should be guaranteed. The above defined problems have satisfactory solution for linear systems and some progresses have been made for nonlinear systems. Nonlinear considerations are clearly necessary when large motions or perturbations are involved, for instance during large manoeuvres or orientations changes in particular for helicopters or orbiting systems.

We will not provide a complete discussion of the stability problem but concentrate on the results which directly apply the aerospace vehicle dynamics.

The stability method considered here have also direct applications in other problems such in robotics or ground vehicle dynamics.

3. CONCEPT OF STABILITY

3.1 Continuous systems

These systems can be described by a set of n first-order differential equations written in vector form as :

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t), \tag{1}$$

where $\mathbf{x}$ is the n state-vector and the components of the function-vector $\mathbf{f}$ are such that the solution of the differential equations exist and is unique - i.e. $\mathbf{f}$ is assumed to be Lipschitz or of class C^1 which is (almost) always the case in mechanical applications. The solution of these equations then only depends on the initial condition $\mathbf{x}_0$ and the initial time t_0 and is written $\mathbf{x}(t; \mathbf{x}_0, t_0)$ or more simply $\mathbf{x}(t)$.

Various stability considerations are given by the following definitions [1].

Definition 3.1 – Without lost of generality, the point $\mathbf{x} = 0$ is said to be an *equilibrium* point of (1) if :

$$\mathbf{f}(0, t) = 0 \quad \forall t.$$

Definition 3.2 – The equilibrium point $\mathbf{x} = 0$ of (1) is said to be *Lagrange stable* or *bounded* if for all t_0 and some $\delta > 0$, the condition :

$$|\mathbf{x}_0| < \delta$$

implies the existence a finite scalar M such that :

$$|\mathbf{x}| < M.$$

Definition 3.3 – The equilibrium point $\mathbf{x} = 0$ is said to be *Liapunov stable* if for all initial time and all positive parameter ϵ there exist a positive parameter $\delta(\epsilon, t_0)$ such that :

$$|\mathbf{x}_0| < \delta \Rightarrow |\mathbf{x}(t)| < \epsilon \quad \text{for} \quad t \geq t_0;$$

---

*The research reported in this paper was partially supported by the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian State Science Policy Programming Office (Prime Minister's Office). The scientific responsibility is assumed by the author.

this point is said to *uniformly stable* if δ does not depend on the initial time.

Definition 3.4 – The equilibrium point $\mathbf{x} = 0$ is said to be *attractive* if for all initial time there exist a positive parameter $\delta\,(t_0)$ such that :

$$|\mathbf{x}_0| < \delta \;\Rightarrow\; \lim_{t \to \inf} |\mathbf{x}\,(t)| = 0;$$

this point is said to *uniformly convergent* if δ does not depend on the initial time.

Definition 3.5 – The equilibrium point $\mathbf{x} = 0$ is said to be (uniformly) *asymptotically stable* if it is (uniformly) stable and convergent.

Remark – These notion of stability can be extended to *global* stability considerations which are valid for all initial conditions (see [1]); further, for linear systems, these notions are equivalent.

In mechanical problems, there are generally various possible equilibrium states. Further one can be more interested in evaluating of the *domain of attraction* of these equilibria than in global stability considerations.

3.2 Discrete systems

These systems can be described by a set of n first-order difference equations written in vector form as :

$$\mathbf{x}_{k+1} = \mathbf{f}\,(\mathbf{x}_k, k)$$

where $\mathbf{x}_k$ is the n state-vector at instant t_k.

The point $\mathbf{x} = 0$ is said to be an *equilibrium* point if :

$$\mathbf{f}\,(0, k) = 0 \quad \forall k$$

and all the stability definitions for continuous systems can be adapted to these discrete systems by substituting the index sequence for the time variable.

4.LIAPUNOV METHOD

4.1 Stability theorems for autonomous systems

Definitions 4.1 – A scalar function $V(\mathbf{x})$ is said to be a *positive definite function* – or a *Liapunov function* – in some neighbourhood $\mathcal{U}$ of an equilibrium point $\mathbf{x} = 0$, if

- it is defined and differentiable on $\mathcal{U}$;
- $V(0) = O$;
- $V(\mathbf{x}) > 0 \quad \forall\ \mathbf{x}) \in (\mathcal{U} \backslash 0)$.

This function is said to be a *negative definite function* when $<$ is substitute for $>$ in the last item; this function is said to be positive (respectively negative) *semi-definite* when $\geq$ (respectively $\leq$) is substitute for $>$.

Definition 4.2 – The *time derivative* the function $V(\mathbf{x})$ along a solution of the equations $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ (or Lie derivative) is defined as :

$$\dot{V}(\mathbf{x}) = \mathcal{L}_{\mathbf{f}} V = [\operatorname{grad} V]^T \mathbf{f}(\mathbf{x}),$$

Liapunov stability theorem – The equilibrium state $\mathbf{x} = 0$ of the equation (1) is stable if – in some neighbourhood Ω of the origin – there exists a Liapunov function, $\mathrm{V}(\mathbf{x})$, whose Lie derivative, $\dot{V}(\mathbf{x})$ is a negative semi-define function (in Ω) - or identically equal to zero.

Liapunov asymptotic stability theorem – The equilibrium state considered in the previous theorem is asymptotically stable if $\dot{V}(\mathbf{x})$ is a negative definite function in Ω.

Lasalle asymptotic stability theorem – The equilibrium state considered in the Liapunov stability theorem is asymptotically stable if $\dot{V}(\mathbf{x})$ is a semi-negative definite function which does not vanish identically along any solution different from $\mathbf{x} \equiv 0$.

Remark These theorems can be extended in order to investigate the stability of non autonomous systems – and in particular of time periodic systems.

4.2 Instability theorem for autonomous systems

Chetayev theorem – If there exist a function $V(\mathbf{x})$ defined in some neighbourhood Ω of the origin and a sub-domain Ω_1 such that :

- V is differentiable in Ω,
- V and $\dot{V}$ are positive in Ω_1,
- $V = 0$ for $\mathbf{x} \in \partial\Omega_1$,
- $0 \in \partial\Omega_1$,

then the origin is unstable.

4.3 Liapunov linearization principle

Asymptotic stability theorem – If the solution $\mathbf{x} = 0$ of the linear system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{2}$$

is asymptotically stable and if $\mathbf{x} = 0$ is also a solution of the system :

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathrm{N}\,(\mathbf{x}) \tag{3}$$

with

$$\lim_{|\mathbf{x}| \to 0} \frac{|\mathbf{N}|}{|\mathbf{x}|} = 0,$$

then the zero solution of (3) is asymptotically stable.

Instability theorem – If the solution $\mathbf{x} = 0$ of (2) is unstable and the other conditions of the previous theorem are holding, the corresponding solution of (3) is also unstable.

4.4 Non autonomous systems

The theorems presented here for autonomous systems can be extended to non automonous systems.

5 LINEAR TIME INVARIANT SYSTEMS

Let us consider linear system of the form (2) where the matrix $\mathbf{A}$ is constant.

This system is asymptotically stable if all the eigenvalues of the matrix **A** – the roots of the characteristic polynomial :

$$|\mathbf{A} - s\mathbf{E}| = f(s) = 0$$

have negative real part.

The system is unstable if some eigenvalues have positive real parts.

When there are purely imaginary eigenvalues, the system is stable or unstable according to the form of the corresponding Jordan blocks. This case will not be considered here and we will only present some criteria which permit to guarantee asymptotic stability. Further reference for these problems can be found in [2],[3] and [4].

5.1 Real coefficient systems

First, let us note that if the order of the system is n, the characteristic polynomial is of degree n with real coefficients written :

$$f(s) = a_0 s^n + a_1 s^{n-1} + a_2 s^{n-2} + \dots + a_n. \quad (4)$$

If all the roots have negative real parts, $f(s)$ is clearly a product of term of the form :

$$s + a \quad \text{or} \quad (s+a+jb)(s+a-jb) = (s+a)^2 + b^2 \quad (5)$$

and consequently all the coefficients, a_i, have the sign of a_0. Without lost of generality, a_0 will be assumed to be positive and a *necessary* condition for stability is then :

$$a_i > 0 \qquad \text{for} \quad i = 0, 1, \dots, n.$$

5.1.1 Argument criterion

Writing

$$f(s) = a_0 \prod_{i=1}^{n} (s - s_i)$$

where from (5)

$$s_i = -a \quad \text{or} \quad s_i = -a \pm jb$$

it is easily seen that the argument of $f(s)$ for $s = j\omega$ increases monotically from 0 to $n\frac{\pi}{2}$ when ω increases from 0 to ∞, i.e. :

$$\Delta_0^{+\infty} arg\{f(j\omega\} = n\frac{\pi}{2};$$

this condition is also sufficient and is sometimes referred to as the *Mikhailov* or the *Leonhard criterion*.

Considering the symmetry of $f(s)$with respect to the real axis, this last relation can also be written :

$$\Delta_{-\infty}^{+\infty} arg\{f(j\omega\} = 2\,\Delta_0^{+\infty} arg\{f(j\omega\} n\pi \qquad (6)$$

Further $f(s)$ and can be written as :

$$f(s) = h(s^2) + s g(s^2)$$

$$\text{or} \quad f(j\omega) = h(-\omega^2) + j\omega g(-\omega^2) \quad \text{for } s = j\omega.$$

This implies the so called *Hermite-Biehler* result which states that the system is asymptotically stable if (and only if) the roots of $h(-\lambda)$ and $g(-\lambda)$ – where $\lambda = \omega^2$ – are distinct negative real and alternate, i.e. :

$$\lambda_1^h < \lambda_1^g < \lambda_2^h < \lambda_2^g < \dots,$$

and if and only if $a_1 > 0$ – has the sign of a_0.

5.1.2 Sturm's criterion

If the characteristic polynomial (4) has real coefficients, the relation (6) can be estimated from these coefficients of by

$$\frac{1}{\pi}\Delta_{-\infty}^{+\infty} arg\{f(j\omega\} = I_{-\infty}^{+\infty} \frac{a_1\omega^{n-1} - a_3\omega^{n-3} + \dots}{a_0\omega^n - a_2\omega^{n-2} + \dots}$$

or

$$\frac{1}{\pi}\Delta_{-\infty}^{+\infty} arg\{f(j\omega\} = I_{-\infty}^{+\infty} \frac{f_2(\omega)}{f_1(\omega)}$$

where $I_a^b f(\omega)$ is called the *Cauchy index* of the real rational function $f(\omega)$ between a and b and is equal to the number of jumps of $f(\omega)$ from $-\infty$ to $+\infty$ as ω increases from a to b.

This Cauchy index can be evaluated by constructing a Sturm sequence, i.e. defining a sequence of polynomials $f_1(\omega), f_2(\omega), f_3(\omega), \dots, f_m(\omega)$, with decreasing order and here $m = n+1$ (as the roots of the numerator and the denominator are distinct), by the relations :

$$f_1(\omega) = q_1(\omega) f_2(\omega) - f_3(\omega)$$

$$f_{i-1}(\omega) = q_{i-1}(\omega) f_i(\omega) - f_{i+1}(\omega)$$

$$f_m(\omega) = q_m(\omega) f_2(\omega)$$

It can be checked (*Sturm's theorem*) that :

$$I_{-\infty}^{+\infty} \frac{f_2(\omega)}{f_1(\omega)} = V(-\infty) - V(+\infty) \qquad (7)$$

where $V(a)$ is the number of sign variations in the Sturm sequence (when the polynomials are evaluated for the fixed value a).

Consequently the Hurwitz (asymptotic) stability is guaranteed if :

$$V(-\infty) - V(+\infty) = n$$

5.1.3 Routh' criterion

Further $V(\pm\infty)$ is evaluated from the coefficient of the higher order terms of the various polynomials of the sequence and it can be concluded that the Cauchy index is equal to n and consequently the system is (asymptotically) stable when all these coefficients, say c_i with $i = 1, \dots, m$, have the same sign (i.e. the sign of a_0), i.e. are positive.

5.1.4 Hurwitz's criterion

The computation of the Routhian coefficients, c_i, is rather cumbersome, but it can be checked – by rather simple algebraic manipulations – that they are obtained from the principal minors, Δ_i, of the $n \times n$ Hurwitz matrix,

$$H = \begin{bmatrix} a_1 & a_3 & a_5 & \dots & \dots \\ a_0 & a_2 & a_4 & \dots & \dots \\ 0 & a_1 & a_3 & \dots & \dots \\ 0 & a_0 & a_2 & \dots & \dots \\ 0 & 0 & a_1 & \dots & \dots \\ 0 & 0 & a_0 & \dots & \dots \\ . & . & . & \dots & \dots \\ . & . & . & \dots & \dots \end{bmatrix},$$

as $c_i = \dfrac{\Delta_i}{\Delta_{i-1}}$.

The stability criterion then simply reads :

$$\Delta_i > 0 \quad \text{for} \quad i = 1, \dots n.$$

5.2 Complex coefficient systems

The results obtained for real coefficients systems can be extended to systems with complex coefficient in a rather straightforward manner. Indeed, after having divided the characteristic polynomial by the (complex) coefficient a_0, we obtained the equivalent polynomial :

$$f(s) = [s^n + a_1 s^{n-1} + \dots + a_n] + j[b_1 s^{n-1} + \dots + b_n]$$

In this case the argument variation has to be taken from $-\infty$ to $+\infty$ has the symmetry with respect to the real axis is not anymore guaranteed. The following stability condition can then be verified :

$$\frac{1}{\pi}\Delta_{-\infty}^{+\infty} arg\{f(j\omega\} = I_{-\infty}^{+\infty}\frac{f_2(\omega)}{f_1(\omega)} = n$$

where

$$\begin{aligned} f_1 &= \omega^n - b_1\omega^{n-1} - a_2\omega^{n-2} + b_3\omega^{n-3} + \dots, \\ f_2 &= a_1\omega^{n-1} - b_2\omega^{n-2} - a_3\omega^{n-3} + b_4\omega^{n-4} + \dots. \end{aligned}$$

These function permits to construct a Sturm sequence and the Sturm's theorem (7) can then be applied.

In thus case the Hurwitz matrix has to be modified as follows :

$$H^* = \begin{bmatrix} a_1 & -b_2 & -a_3 & b_4 & a_5 & \dots & \dots \\ 1 & -b_1 & -a_2 & b_3 & a_4 & \dots & \dots \\ 0 & a_1 & -b_2 & -a_3 & b_4 & \dots & \dots \\ 0 & 1 & -b_1 & -a_2 & b_3 & \dots & \dots \\ 0 & 0 & a_1 & -b_2 & -a_3 & \dots & \dots \\ 0 & 0 & 1 & -b_1 & -a_2 & \dots & \dots \\ . & . & . & . & . & \dots & \dots \\ . & . & . & . & . & \dots & \dots \end{bmatrix},$$

5.3 Discrete systems

A linear discrete system is written under the form :

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$$

and its characteristic polynomial is given by :

$$f(z) = |\mathbf{A} - z\mathbf{E}| = a_0 z^n + a_1^{n-1} + \dots + a_n$$

In order to have stability, all the roots should have a norm smaller than one – should be located in the unit circle centered at the origin of the complex plane. A necessary condition is then that the product of all the roots has a norm smaller than one; this implies that :

$$|\frac{a_n}{a_0}| < 1$$

The principle of the argument is also valid here and it is seen that for stable systems $f(e^{j\theta})$ increases monotonically from 0 to $n\pi$ when θ increases from 0 to π; here too this condition is also sufficient.

Writing $f(z)$ under the form of a sum of a symmetrical and an antisymmetrical part as :

$$f(z) = \frac{1}{2}\left[f(z) + z^n f(\frac{1}{z})\right] + \frac{1}{2}\left[f(z) - z^n f(\frac{1}{z})\right]$$

$$\text{or} \quad f(s) = h(z) + g(z)$$

it is seen that $h(z)$ and $g(z)$ have alternating roots on the unit circle. This permits to obtain results equivalent to the Routh-Hurwitz criteria for continuous systems [2].

5.4 Feedback Stabilization

If the open loop transfer function (of a single input single output system) is given by :

$$x(s) = G(s)\,e(s) \quad \text{with} \quad G(s) = \frac{p(s)}{f(s)}$$

with a constant feedback gain, k, i.e. with

$$e = u + y \quad \text{with} \quad y = k\,x,$$

the closed loop transfer function is written :

$$\mathbf{x}(s) = \frac{G(s)}{1 + kG(s)}u(s) = \frac{q(s)}{f(s) + k\,q(s)}u(s)$$

The stability of closed loop system can be investigated by using the closed loop characteristic polynomial : $f(s) + k\,q(s)$

The Popov's analysis permit to extent these results to particular nonlinearities described by :

$$y = \phi(x) \quad \text{with} \quad \frac{\phi(x)}{x} < k \quad \forall x.$$

Some extension exist for multivariable systems.

5.5 Stability robustness

Stability robustness and simultaneous stability analysis are current research topics of real practical importance.

Let us just mention the Kharitonov results concerning the stability of systems described by the characteristic polynomial (4) for which the various parameters can vary between independent limits or :

$$\underline{a}_i \leq a_i \leq \overline{a}_i.$$

The system will be stable for all these parameters, if the four following polynomial guarantee the stability.

$$f_1(s) = \overline{a}_n + \underline{a}_{n-1}s + \underline{a}_{n-2}s^2 + \overline{a}_{n-3}s^3 + \overline{a}_{n-4}s^4 + \dots,$$

$$f_2(s) = \overline{a}_n + \overline{a}_{n-1}s + \underline{a}_{n-2}s^2 + \underline{a}_{n-3}s^3 + \overline{a}_{n-4}s^4 + \dots,$$

$$f_3(s) = \underline{a}_n + \overline{a}_{n-1}s + \overline{a}_{n-2}s^2 + \underline{a}_{n-3}s^3 + \underline{a}_{n-4}s^4 + \dots,$$

$$f_4(s) = \underline{a}_n + \underline{a}_{n-1}s + \overline{a}_{n-2}s^2 + \overline{a}_{n-3}s^3 + \underline{a}_{n-4}s^4 + \dots.$$

Such results have extension for system for which independent physical parameters are varying between bound. Results and references for these topics can also be found in [2].

6 EXACT LINEARIZATION BY FEEDBACK

The system of equation :

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sum_{i=1}^{m} \mathbf{g}_i\, u_i = \mathbf{f}(\mathbf{x}) + \mathbf{g}\,\mathbf{u} \qquad (8)$$

is said to be feedback linearizable if there exist a diffeomorphism from a neighbourhood of the origin to $\mathcal{R}^n$ and controls given by :

- $\Phi : U_{x_0} \longrightarrow \mathcal{R}^n \quad x \longrightarrow \mathbf{z} = \phi(\mathbf{x})$
- $\mathbf{u} = \alpha(\mathbf{x}) + \beta(\mathbf{x})\,\mathbf{v}$

such that the system is transformed by Φ into the linear system

$$\dot{\mathbf{z}} = \mathbf{A}\,\mathbf{z} + \mathbf{B}\,\mathbf{v}.$$

6.1 Distributions

Let us consider d smooth vector fields, $\mathbf{f}_i$, $i = 1, \dots, d$, defined in an open set U. For each point $\mathbf{x}$ the corresponding vectors $\mathbf{f}_i(\mathbf{x})$ span a vector space, $\Delta(\mathbf{x})$ given by :

$$\Delta(\mathbf{x}) = span\{\mathbf{f}_i(\mathbf{x})\} \quad \text{or simply} \quad \Delta = span\{\mathbf{f}_i\}.$$

The assignment to each point of U of the corresponding linear subspace $\Delta(\mathbf{x})$ is called a (smooth) *distribution.*

For any vector fields $\mathbf{f}_1$ and $\mathbf{f}_2$, the *Lie bracket*, $[\mathbf{f}_1, \mathbf{f}_2]$, of these vector fields is a new vector field defined as :

$$[\mathbf{f}_1, \mathbf{f}_2] \overset{\text{def}}{=} \frac{\partial \mathbf{f}_2}{\partial \mathbf{x}}\mathbf{f}_1 - \frac{\partial \mathbf{f}_1}{\partial \mathbf{x}}\mathbf{f}_2.$$

Further, a distribution $\mathcal{D} = span\{\,\mathbf{f}_i\,\}$ is said to be *involutive* if $[\mathbf{f}_i, \mathbf{f}_j] \in \mathcal{D}$.

6.2 Conditions for exact linearization

The vectors fields $\mathbf{f}$ and $\mathbf{g} = \{\,\mathbf{g}_i\,\}$, defined in (8), permits to construct a sequence of distributions :

$\mathcal{D}_1 = span\{\,\mathbf{g}\,\} = span\{\,ad^0_{\mathbf{f}}\,\mathbf{g}\,\}$
$\mathcal{D}_2 = span\{\,\mathbf{g}, [\mathbf{f}, \mathbf{g}]\} = span\{\,\mathbf{g}, ad^1_{\mathbf{f}}\,\mathbf{g}\,\}$
...
$\mathcal{D}_n = span\{\,\mathbf{g}, ad^1_{\mathbf{f}}\mathbf{g}, \dots, ad^{n-1}_{\mathbf{f}}\,\mathbf{g}\,\}$
or
$\mathcal{D}_i = span\{\,ad^k_{\mathbf{f}}\,\mathbf{g} \,:\, k = 1, \dots, i-1\,\}$;

this sequence is said to be *nested* as clearly : $\mathcal{D}_1 \subset \mathcal{D}_2 \subset \dots \subset \mathcal{D}_n$

The exact linearization of the system (8) is possible, see [5], if :

- $dim\{\,\mathbf{g}\,\} = m$
- $dim\mathcal{D}_i \quad i = 1, \dots, n$ constant in U_{x_0}
- $dim\mathcal{D}_n = n$
- $\mathcal{D}_j$ involutive $j = 1, \dots, n-1$

It should be noted that in the previous conditions $\mathcal{D}_n$ should be replaced by $\mathcal{D}_k$ if $dim\mathcal{D}_k = n$.

The above mentioned conditions also implies the following properties :

$\exists\,\{r_1, \dots, r_m\}(r_i < n)$ and $\{\lambda_i(x),\ i = 1, \dots, m\}$

such that :

- $L_{\mathbf{g}_i} L^k_{\mathbf{f}}\, \lambda_j(x) = 0 \quad \forall i, j$ and $k = 0, 1, \dots, r_{j-2}$;
- the matrix :

$$\mathbf{M}(\mathbf{x}) = \left[L_{\mathbf{g}_j} L^{r_i - 1}_{\mathbf{f}}\, \lambda_i\right] \ \text{is nonsingular;}$$

- $\displaystyle\sum_i^m r_i = n$;
- $\mathbf{u} = \alpha(\mathbf{x}) + \beta(\mathbf{x})\,\mathbf{v}$ with :

$$\beta = \mathbf{M}^{-1}(\mathbf{x}) \qquad \alpha = \mathbf{M}^{-1}\left[L^{r_i}_{\mathbf{f}}\,\lambda_i\right];$$

- $\mathbf{z} = col\{L^{k_i - 1}_{\mathbf{f}}\,\lambda_i \,:\, i = 1, \dots, m \quad k_i = 1, \dots, r_i\}$

$$\dot{\mathbf{z}} = \mathbf{A}\,\mathbf{z} + \mathbf{B}\,\mathbf{v} \quad \text{with } \mathbf{A} = diag\{\mathbf{A}_{r_1}, \dots, \mathbf{A}_{r_m}\}$$

$$\text{and} \quad \mathbf{B} = diag\{\mathbf{B}_{r_1}, \dots, \mathbf{B}_{r_m}\}$$

where

$$\mathbf{A}_{r_i} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \qquad \mathbf{B}_{r_i} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

6.3 Construction of the new variables

The functions λ_i which permits to define the change of variables can be constructed from the following considerations :

$L_{\mathbf{g}_i} L^k_{\mathbf{f}}\, \lambda_j(x) = 0 \quad \forall i, j$ and $k = 0, 1, \dots, r_{j-2}$

$\equiv :\ \langle d\lambda_i, ad^k_{\mathbf{f}}\mathbf{g}_j\rangle = 0 \quad \forall i, j$ and $k = 0, 1, \dots, r_{j-2}$

$\equiv :\ d\lambda_i \in \mathcal{D}^{\perp}_{r_i - 1}$

and also that $span\{d\lambda_i, dL_{\mathbf{f}}\lambda_i\} \subset \mathcal{D}^{\perp}_{r_i - 2}$, as solution of system of (partial) differential equations.

7 MECHANICAL SYSTEMS

The n variables $q_1, \ldots, q_n$ are called *generalized coordinates* if the configuration of a material system – i.e. all the position vectors of the material elements with respect to inertial space – can be described by purely algebraic functions of these variables and the time. The *n generalized coordinates vector* is then given by :

$$\mathbf{q} = [q_1 \ \ldots \ q_n]^T$$

The position of a material element $\mathcal{X} \in \mathcal{M}$, where $\mathcal{M}$ is the considered material system, is then written as :

$$\underset{\sim}{\mathbf{x}} = \underset{\sim}{\mathbf{x}}(q_1, \ldots, q_n, t) = \underset{\sim}{\mathbf{x}}(q_\alpha, t)$$

and the configuration is :

$$\mathcal{V} = \{\underset{\sim}{\mathbf{x}}(q_\alpha, t)\}.$$

The velocities are then linear functions of the generalized velocities, $\dot{q}_\alpha$, i.e.:

$$\dot{\underset{\sim}{\mathbf{x}}} = \sum_{\alpha=1}^{n} \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial t} = \dot{\underset{\sim}{\mathbf{x}}}(q_\alpha, \dot{q}_\alpha, t).$$

The *kinetic energy* is then quadratic in the generalized velocities and can be written as :

$$T = \int_{\mathcal{M}} \dot{\underset{\sim}{\mathbf{x}}} \bullet \dot{\underset{\sim}{\mathbf{x}}} \, dm = \frac{1}{2} \dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}} + \boldsymbol{\Gamma}^T \dot{\mathbf{q}} + T_0,$$

where the matrix $\mathbf{M}$, the vector $\boldsymbol{\Gamma}$ and the scalar T_0 are functions of the generalized coordinates and time.

7.1 Potential Power Principle

The local equation for any material system (without internal momentum distribution) are given, for any material element $(\mathcal{X})$, by :

$$\ddot{\underset{\sim}{\mathbf{x}}} = \underset{\sim}{\mathbf{f}} \qquad (9)$$

where $\underset{\sim}{\mathbf{f}}$ is the local force density.

The scalar product of (9) by any vector field $\underset{\sim}{\mathbf{a}}$ integrated on the considered material system provides the relation :

$$\int_{\mathcal{M}} (\ddot{\underset{\sim}{\mathbf{x}}} - \underset{\sim}{\mathbf{f}}) \bullet \underset{\sim}{\mathbf{a}} \, dm = 0. \qquad (10)$$

It should be noted that if $\underset{\sim}{\mathbf{a}}$ is the (actual) velocity field ($\underset{\sim}{\mathbf{a}} = \dot{\underset{\sim}{\mathbf{x}}}$) the previous relation is equivalent to the mechanical energy relation :

$$\dot{T} = \int_{\mathcal{M}} \ddot{\underset{\sim}{\mathbf{x}}} \bullet \dot{\underset{\sim}{\mathbf{x}}} \, dm = \int_{\mathcal{M}} \underset{\sim}{\mathbf{f}} \bullet \dot{\underset{\sim}{\mathbf{x}}} \, dm = P,$$

where P is the power of all the interactions acting onto the considered system.

One can also define a *potential velocity* field compatible with a choice of independent generalized coordinates as :

$$\Delta \dot{\underset{\sim}{\mathbf{x}}} \overset{\text{def}}{=} \dot{\underset{\sim}{\mathbf{x}}}(q_\alpha, \dot{q}_\alpha + \Delta \dot{q}_\alpha, t) - \dot{\underset{\sim}{\mathbf{x}}}(q_\alpha, \dot{q}_\alpha, t) = \sum_{\alpha=1}^{n} \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial q_\alpha} \Delta \dot{q}_\alpha$$

where $\Delta \dot{q}_\alpha$ are arbitrary, independent, finite changes of the generalized velocities. It should be noted that the potential velocity field is linear in these potential changes.

For such a velocity field, the relation (10) becomes :

$$\int_{\mathcal{M}} \ddot{\underset{\sim}{\mathbf{x}}} \bullet \Delta \dot{\underset{\sim}{\mathbf{x}}} \, dm = \int_{\mathcal{M}} \underset{\sim}{\mathbf{f}} \bullet \Delta \dot{\underset{\sim}{\mathbf{x}}} \, dm. \qquad (11)$$

The left hand side of (11)can be expressed as :

$$\int_{\mathcal{M}} \ddot{\underset{\sim}{\mathbf{x}}} \bullet \Delta \dot{\underset{\sim}{\mathbf{x}}} \, dm = \sum_{\alpha=1}^{n} \left(\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_\alpha} - \frac{\partial T}{\partial q_\alpha} \right) \Delta \dot{q}_\alpha,$$

even in practise, this expression can be computed in a much more efficient way. This form is given here in order to show the structure of the obtained equation of motion.

7.2 Lagrangian mechanics

7.2.1 Conservative systems

A mechanical system is said to be *lagrangian* – or *conservative from an hamiltonian point of vue* – if the potential power of all the interactions acting onto the system can be expressed as :

$$\Delta P = \int_{\mathcal{M}} \underset{\sim}{\mathbf{f}} \bullet \Delta \dot{\underset{\sim}{\mathbf{x}}} \, dm = -\sum_{\alpha=1}^{n} \frac{\partial U}{\partial q_\alpha} \Delta \dot{q}_\alpha$$

where $U = U(\dot{q}_\alpha, t)$ is a function of the generalized coordinates and time an is called the *potential energy* of the system.
The *Lagrangian* of the system is then defined as : $L = T - U$.

Further, the $\Delta \dot{q}_\alpha$ having been assumed to be independent, the equations of motion have the form :

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{\mathbf{q}}} - \frac{\partial T}{\partial \mathbf{q}} + \frac{\partial U}{\partial \mathbf{q}} = 0$$

or

$$\frac{d}{dt} (\mathbf{M} \dot{\mathbf{q}} + \boldsymbol{\Gamma}) - \frac{\partial L}{\partial \mathbf{q}} = 0$$

The *Hamiltonian* function is then defined as :

$$H = H(\mathbf{p}, \dot{\mathbf{q}}, t) \overset{\text{def}}{=} \mathbf{p}^T \dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}, t)$$

where $\dot{\mathbf{q}}$ is expressed as a function of the *generalized momemtum vector*, $\mathbf{p}$ defined as :

$$\mathbf{p} \overset{\text{def}}{=} \frac{\partial L}{\partial \dot{\mathbf{q}}} = \mathbf{M} \dot{\mathbf{q}} + \boldsymbol{\Gamma}.$$

It can be easily checked that for conservative systems, the time derivative of the Hamiltonian along a trajectory is given by :

$$\dot{H} = \mathcal{L}_{\mathbf{f}} H = \frac{\partial H}{\partial t};$$

7.2.2 Dissipative systems

A mechanical system is said to be *dissipative from a Rayleigh point of vue* if the potential power of the interactions which do not "derive" from a potential can be expressed as :

$$\Delta P_d = \int_{\mathcal{M}} \underset{\sim}{\mathbf{f}}_d \bullet \Delta\underset{\sim}{\dot{\mathbf{x}}}\, dm = -\sum_{\alpha=1}^{n} \frac{\partial R}{\partial \dot{q}_\alpha}\Delta \dot{q}_\alpha$$

where R is a quadratic positive (semi) definite function of the generalized velocities.

For such systems the equations of motion have the form :

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{q}}} - \frac{\partial T}{\partial \mathbf{q}} + \frac{\partial U}{\partial \dot{\mathbf{q}}} + \frac{\partial R}{\partial \dot{\mathbf{q}}} = 0$$

Further the Lie derivative of the Hamiltonian is here :

$$\dot{H} = \frac{\partial H}{\partial t} - \mathbf{q}^T \frac{\partial R}{\partial \dot{\mathbf{q}}}.$$

7.2.3 Autonomous systems

A system is said to be *autonomous* if the corresponding Lagrangian does not explicitly depend on time, i.e. if the matrix $\mathbf{M}$, the vector $\boldsymbol{\Gamma}$ and the scalar $T_0 - U$ are not explicit functions of time.

It should be noted that for an autonomous system, the position vector of the various point can be described by explicit functions of time. This is a classical situation in rotational dynamics.

For such systems, the equilibrium configurations are given by the following system of algebraic equations :

$$\frac{\partial}{\partial \mathbf{q}}(U - T_0) = 0 \qquad (12)$$

Further, in this case, the time derivative of the Hamiltonian along a trajectory is given for conservative and dissipative systems respectively by :

$$\dot{H} = 0 \quad \text{and} \quad \dot{H} = -\dot{\mathbf{q}}^T \frac{\partial R}{\partial \dot{\mathbf{q}}} \leq 0.$$

Consequently, the hamiltonian is a "good" candidate as a Liapunov function for such systems.

The dissipation is said to be *pervasive* along a trajectory if it is never identically equal to zero along this trajectory (except at the equilibrium).

7.2.4 Linearized autonomous system

Around a given equilibrium – given by (12) – the quadratic forms (obtained by appropriate series expansions) of L and R are given respectively by :

$$L = \frac{1}{2}\dot{\mathbf{q}}^T\mathbf{M}\dot{\mathbf{q}} + \mathbf{q}^T\mathbf{N}^T\dot{\mathbf{q}} + \frac{1}{2}\mathbf{q}^T\mathbf{K}\mathbf{q}$$

$$\text{and} \quad R = \frac{1}{2}\dot{\mathbf{q}}^T\mathbf{C}\dot{\mathbf{q}}$$

where $\mathbf{M}$,$\mathbf{N}$,$\mathbf{K}$ and $\mathbf{C}$ are constant matrices. With $\mathbf{G} \overset{\text{def}}{=} \mathbf{N} - \mathbf{N}^T = \mathbf{G}^T$, the corresponding equations of motion are given by :

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{G}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = -\mathbf{C}\dot{\mathbf{q}}.$$

The corresponding hamiltonian and its time derivative (expressed in terms of the generalized coordinates are the respectively given by :

$$H = \frac{1}{2}\dot{\mathbf{q}}^T\mathbf{M}\dot{\mathbf{q}} + \frac{1}{2}\mathbf{q}^T\mathbf{K}\mathbf{q} \quad \text{and} \quad \dot{H} = -\dot{\mathbf{q}}^T\mathbf{C}\dot{\mathbf{q}}.$$

The hamiltonian is an appropriate Liapunov function.

If the damping is pervasive, using the abovementioned theorems (Lasalle and Chetaev), the corresponding equilibrium is asymtotically stable or unstable when the Hamiltonian is, respectively, positive definite or sign undefinite [8].

Systems with integrals of motion or constraints

When the constraints are independent algebraic functions of the generalized coordinates, some of these variables can be eliminated and a reduced order system is obtained, the corresponding Hamiltonian can then be used as a Liapunov function.

More interesting are (independent) non holonomous constraints and integrals of motion which are linear in the velocities. They both can be written as m linear relations between the generalized velocities, under the form of the system :

$$\mathbf{A}\dot{\mathbf{q}} + \mathbf{b} = 0,$$

If $\mathbf{A}_0$ is a regular $m \times m$ submatrix of $\mathbf{A}$ (possibly obtained after having reordered the variables), this system can be rewritten as :

$$\mathbf{A}_0\dot{\mathbf{q}}_0 + \mathbf{A}_1\dot{\mathbf{q}}_i + \mathbf{b} = 0 \qquad (13)$$

if the generalized coordinate vector has been partitioned as :

$$\mathbf{q} = [\mathbf{q}_0^T \ \mathbf{q}_1^T]^T.$$

One can use the Lagrange multiplier technique to find the equations of motion.

On the other hand, the stability can be directly analysed by eliminating $\dot{\mathbf{q}}_0$ – using (13) from the hamiltonian or by using of a modified Hamiltonian of the form :

$$V = H + (\mathbf{A}\dot{\mathbf{q}} + \mathbf{b})^T\mathbf{Q}(\mathbf{A}\dot{\mathbf{q}} + \mathbf{b})$$

whose time derivative is clearly equal to $\dot{H}$ and where $\mathbf{Q}$ can be selected in order to eliminate generalized velocities which do not appear in $\dot{H}$. The stability with respect to the remaining variables can then be analysed in a straightforward manner.

Constraint damping

In many aerospace applications (such as rotating systems) the dissipative interactions (in particular the aerodynamical interactions or the interactions in bearings) do not derive from a Rayleigh function. In fact, the corresponding generalized forces are then functions of the generalized velocities and of the generalized variables and can be written in linear approximation as :

$$\mathbf{Q}_d = -(\mathbf{C}\dot{\mathbf{q}} + \mathbf{F}\mathbf{q})$$

and the equations of motion are then given by :

$$\mathbf{M}\ddot{\mathbf{q}} + (\mathbf{G}+\mathbf{C})\dot{\mathbf{q}} + (\mathbf{K}-\mathbf{F})\mathbf{q} = 0.$$

It is still possible to obtain a simple Liapunov function, when one can write $\mathbf{F} = \mathbf{CS}$ and if $\mathbf{S}$, $\mathbf{MS}$, $\mathbf{KS}$ are skew symmetric matrices and $\mathbf{GS}$ is a symmetric matrix [9].

In this case, a modified Hamiltonian is given by :

$$V = \frac{1}{2}\dot{\mathbf{q}}^T\mathbf{M}\dot{\mathbf{q}} - \mathbf{q}^T\mathbf{SA}\dot{\mathbf{q}} + \frac{1}{2}\mathbf{q}^T(\mathbf{K}-\mathbf{GS})\mathbf{q}, \qquad (14)$$

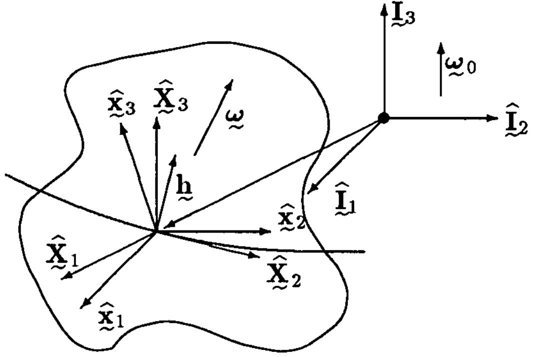
the time derivative of this function along a trajectory being given by the negative (semi) definite function :

$$\dot{V} = -(\dot{\mathbf{q}}+\mathbf{Sq})^T\mathbf{C}(\dot{\mathbf{q}}+\mathbf{Sq}).$$

In other case, the search for an appropriate Liapunov function is more difficult and it is generally easier to use the techniques presented in section **5**.

8. APPLICATION TO GYROSTATS

We will show now, as an example, how the above presented method can be used and we will consider a deformable orbiting satellite with internal rotating parts. This can be a dual spin system or a system which can be controlled by flywheels.



Orbiting Gyrostat

The orbit will be assumed to be circular and the corresponding angular velocity vector is :

$$\underset{\sim}{\omega}_0 = \omega_0\hat{\underset{\sim}{\mathbf{X}}}_3 = [\hat{\underset{\sim}{\mathbf{X}}}]^T\begin{bmatrix} 0 \\ 0 \\ \omega_0 \end{bmatrix}$$

where ω_0 is constant. The orbital base is the base $\{\hat{\underset{\sim}{\mathbf{X}}}_\alpha\}$, with $\hat{\underset{\sim}{\mathbf{X}}}_1$ aligned with the ascendent local vertical, $\hat{\underset{\sim}{\mathbf{X}}}_2$ tangent to the orbit and $\hat{\underset{\sim}{\mathbf{X}}}_3$ aligned with the normal to the orbit.

The body frame, $\{\hat{\underset{\sim}{\mathbf{x}}}_\alpha\}$, has a relative orientation with respect to the orbital base which is described by the orientation matrix $[A]$, i.e. :

$$[\hat{\underset{\sim}{\mathbf{x}}}] = [A][\hat{\underset{\sim}{\mathbf{X}}}]$$

and its rotational velocity with respect to inertial space is the vector $\underset{\sim}{\omega}$ with components ω_1, ω_2, ω_3 in the body frame or

$$\underset{\sim}{\omega} = [\hat{\underset{\sim}{\mathbf{x}}}]^T[\omega]$$

where $[\omega]$ is the corresponding component matrix.

The relative internal momentum of the rotors – assumed to be symmetrical with respect to the rotation axis – is represented by the vector $\underset{\sim}{\mathbf{h}}$. Each rotor can possibly be controlled by torques (along there rotation axis); the corresponding control vector normalized by the respective moments of inertia is the vector $\mathbf{u}$.

Deformation will be considered in order to be able to introduce internal damping. These deformations will be described by deformation variables, represented by the vector $\boldsymbol{\beta}$; the corresponding displacements will be assumed to be algebraic function of these variables; the inertia tensor and the internal angular momentum are clearly functions of these deformation variables.

For low altitude orbits, the gravity-gradient torque has to be considered. This torque can be expressed in terms or the actual inertia tensor of the system $\underset{\sim}{\mathbf{J}}$ (expressed as a function of the deformation variables) as :

$$\underset{\sim}{\mathbf{L}} = \omega_0^2\hat{\underset{\sim}{\mathbf{X}}}_1 \times \underset{\sim}{\mathbf{J}}\cdot\hat{\underset{\sim}{\mathbf{X}}}_1.$$

8.1 Equations of motion

The general equations of rotation can be written :

$$\dot{\underset{\sim}{\mathbf{H}}} = \underset{\sim}{\mathbf{L}}$$

$$\text{with}\quad \underset{\sim}{\mathbf{H}} = \underset{\sim}{\mathbf{J}}\cdot\underset{\sim}{\omega} + \underset{\sim}{\mathbf{h}} \qquad \underset{\sim}{\mathbf{L}} = 3c\omega_0^2\hat{\underset{\sim}{\mathbf{X}}}_1 \times \underset{\sim}{\mathbf{I}}\cdot\hat{\underset{\sim}{\mathbf{X}}}_1,$$

where $c = 1$ if the gravity torque has to considered and $c = 0$ otherwise.

The rotors equations around their axis of rotation will be written :

$$\dot{p}_i = M_i, \qquad i = 1, 2, 3$$

where p_i are the moment of momentum of the rotors about their axes and M_i are the corresponding applied torques.

We also have to consider the kinematical relation which relate the matrix $[A]$ to the rotational velocities of the orbital and body bases, i.e. :

$$[\dot{A}] = -[\omega]^{\sim}[A] + [A][\omega_0]^{\sim}$$

where *tilde* matrices are are expressed in terms of the corresponding vector by the relation :

$$[\omega] = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \qquad [\omega]^{\sim} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

and the equations of deformation.

8.2 Equilibrium configuration

At equilibrium, the matrices $[A]$ and $[h]$ are constant. Further, the rotational velocity is given by :

$$\underset{\sim}{\omega} = \underset{\sim}{\omega}_0 = \omega_0\hat{\underset{\sim}{\mathbf{X}}}_3$$

$$\underset{\sim}{\mathbf{H}} = \underset{\sim}{\mathbf{I}}\cdot\underset{\sim}{\omega}_0 + \underset{\sim}{\mathbf{h}}_0 = \underset{\sim}{\mathbf{H}}_0$$

with

$$\underset{\sim}{\mathbf{h}}_o \overset{\text{def}}{=} \omega_0 [\hat{\underset{\sim}{\mathbf{x}}}]^T \begin{bmatrix} J_1 \\ J_2 \\ J_3 \end{bmatrix}.$$

This permits to write the following equilibrium conditions

$$\omega_0 \hat{\underset{\sim}{\mathbf{X}}}_3 \times \underset{\sim}{\mathbf{H}}_0 = 3\omega_0^2 \hat{\underset{\sim}{\mathbf{X}}}_1 \times \underset{\sim}{\mathbf{I}} \bullet \hat{\underset{\sim}{\mathbf{X}}}_1. \qquad (15)$$

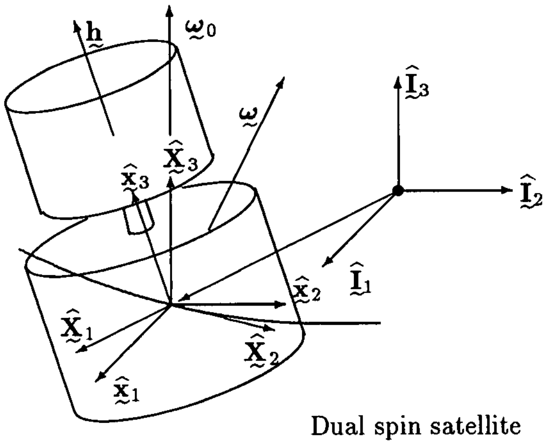
which implies that at equilibrium the inertia matrix has the form :

$$[I] = \begin{bmatrix} I_1 & 0 & -\alpha J_1 \\ 0 & I_2 & -J_2 \\ -\alpha J_1 & -J_2 & I_3 \end{bmatrix},$$

where $\alpha = \frac{1}{4}$ if the gravitational torque is considered and $\alpha = 1$ in free space. Additional conditions concerning the deformation variables have to be satisfied. We will assume that they correspond (possibly after an appropriate change of definition) to $\boldsymbol{\beta} = 0$.

8.3 Dual spin satellite

Such systems are designed to have $J_1 = J_2 = 0$. At equilibrium, the inertia matrix is thus a diagonal matrix. In other words the principal axes are aligned with the orbital axes or with a nominal rotating frame.



Dual spin satellite

For small angular deviation around this equilibrium, we will describe the matrix $[A]$ by a sequence of rotations around successive 1-, 2- and 3 axis respectively; the corresponding rotation angles θ_1, θ_2 and θ_3 are called Tait-Bryan angles.

The linearized kinematical relation provide the following relations between these angles and the components of the rotational velocity vector :

$$\omega_1 = \dot{\theta}_1 - \omega_0\theta_2, \quad \omega_2 = \dot{\theta}_2 + \omega_0\theta_1 \quad \text{and} \quad \omega_3 = \omega_0 + \dot{\theta}_3.$$

The rotors will be suppose to have constant relative angular velocities. In order to obtain exact linearized equations in the deformation variables, it is necessary to used a second order description of the displacement field with respect to the deformation variables. Failure to do so eliminates the so-called stiffening terms due to the rotation [10].

Further, we will, a priori, assume that damping terms are linear in the deformation velocities.

The corresponding linearized equations (around the nominal equilibrium) have the expected form :

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{G}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = -\mathbf{C}\dot{\mathbf{q}},$$

where here

$$\mathbf{q} = [\theta_1 \; \theta_2 \; \theta_3 \; \boldsymbol{\beta}^T]^T$$

and the matrices $\mathbf{M}$, $\mathbf{G}$, $\mathbf{K}$ and $\mathbf{C}$ are given by the relations :

$$\mathbf{M} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{m} \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_\beta \end{bmatrix}$$

$$\mathbf{G} = \omega_0 \begin{bmatrix} 0 & -G_1 & 0 & \mathbf{A}_1 \\ G_1 & 0 & 0 & \mathbf{A}_2 \\ 0 & 0 & 0 & \mathbf{A}_3 \\ -\mathbf{A}_1{}^T & -\mathbf{A}_2{}^T & -\mathbf{A}_3{}^T & \mathbf{0} \end{bmatrix}$$

with $G_1 = (I_1 + I_2 - I_3 - J_3)$ and

$$\mathbf{K} = \omega_0^2 \begin{bmatrix} K_1 & 0 & 0 & -\mathbf{A}_2 \\ 0 & K_2 & 0 & \mathbf{A}_1{}' \\ 0 & 0 & K_3 & \mathbf{A}_4 \\ -\mathbf{A}_2{}^T & \mathbf{A}_1{}'^T & \mathbf{A}_4{}^T & \boldsymbol{\Pi} \end{bmatrix}$$

where

$$K_1 = I_3 - I_2 + J_3 \quad K_2 = a(I_3 - I_1) + J_3 \quad K_3 = b(I_2 - I_1)$$

and the matrices $\mathbf{A}_i$ and $\boldsymbol{\Pi}$ are functions of the deformation shapes associated with the deformation variables. In these expressions, for freely spinning systems $a = 1$ and $b = 0$ and for gravity stabilized systems $a = 4$ and $b = 3$. Further the expression of $\boldsymbol{\Pi}$ differs in these two cases.

8.3.1 Gravity stabilized satellite

For gravity stabilized gyrostats, the hamiltonian can be taken as Liapunov function and further it can be checked that the damping is pervasive. The corresponding asymptotic stability conditions can be written :

$$I_3 - I_2 + J_3 > 0, \quad 4(I_3 - I_1) + J_3 > 0, \quad I_2 - I_1 > 0$$

together with conditions on the generalized stiffnesses. In the above-mentioned conditions J_3 can be expressed as

$$J_3 = I' \frac{\Omega}{\omega_0}$$

where Ω is the relative angular velocity of the rotor and I' is the corresponding moment of inertia.

8.3.2 Spin stabilized free system

For spin stabilized systems, the damping is clearly not complete as the system can be in equilibrium when

$$\boldsymbol{\beta} = 0 \quad \text{and} \quad \omega_1 = \omega_2 = 0,$$

i.e., when θ_1 and θ_2 are solution of the system

$$\dot{\theta}_1 - \omega_0\theta_2 = 0 \qquad \dot{\theta}_2 + \omega_0\theta_1 = 0.$$

For such a system, there is no external torques. This means that the total angular momentum is a constant vector. In particular the components of this vector along the axes $\hat{\underset{\sim}{\mathbf{X}}}_1$ and $\hat{\underset{\sim}{\mathbf{X}}}_2$ are identically equal to zero, or

$$\underset{\sim}{\mathbf{H}}\bullet\hat{\underset{\sim}{\mathbf{X}}}_1 = \underset{\sim}{\mathbf{H}}\bullet\hat{\underset{\sim}{\mathbf{X}}}_2 = 0.$$

Theses relations are linear in $\dot{\theta}_1$ and $\dot{\theta}_2$ and these variables can be eliminated from the Hamiltonian which will give stability conditions for the system without external perturbations.

8.3.3 Shaft constraint damping

A simple example of constraint damping is provide by the damping in the shaft assembly connecting the rotor to the main body.

The deformation of the shaft can be described by to rotation angles say β_1 and β_2. In first approximation the dissipation in the body part of the bearing-shaft assembly can be described by the following Rayleigh function :

$$R_c = \frac{1}{2}c(\dot{\beta}_1^2 + \dot{\beta}_2^2),$$

where c is the corresponding linear damping coefficient.

Energy dissipation takes place in the rotor part of this assembly (in particular in the rotating shaft) even when the deformation angles remain constant; the corresponding deformation rates being then equal to $\Omega\beta_2$ and $-\Omega\beta_1$ respectively. The corresponding Rayleigh function will be written under the form

$$R_{c'} = \frac{1}{2}c'\left[(\dot{\beta}_1 + \Omega\beta_2)^2 + (\dot{\beta}_2 - \Omega\beta_1)^2\right]$$

where c' is the corresponding linear damping coefficient.

In general, one cannot find a simple modified hamiltonian for this system and classical method should be used.

Nevertheless, when both bodies (i.e. the platform and the rotor are symmetrical, the conditions for taking the modified hamiltonian (14) are satisfied. In particular, if the platform has no nominal angular velocity, the "main" stability condition is given by the relation :

$$I' - \frac{c'}{c + c'}I > 0,$$

where I is the global transverse moment of inertia ($I_1 = I_2 = I$). It should be noted that this condition is then always satisfied when there is no dissipation associated with the rotor.

For this symmetrical system, one can also define the following complex variables :

$$\theta^* = \theta_1 + j\theta_2 \quad \text{and} \quad \beta^* = \beta_1 + j\beta_2$$

and obtain a complex system of equations (whose characteristic polynomial has complex coefficients) and use the corresponding modified Hurwitz stability criterion. The results – obviously – are equivalent but can be extended to more complicated systems such as multi-rotor systems.

8.4 Rigid gyrostat

For gravity stabilized rigid gyrostats, the variables of the non linear system are the components of the rotational velocity, the rotor momenta, and the elements of the orientation matrix $[A]$, together with the orthogonality constraints associated with this matrix. It is interesting to mention that the corresponding system of equations is quadratic in the corresponding variables – what somewhat simplifies the analysis.

It can then be shown, [11], that the exact linearization conditions are satisfied – except for isolated configurations which are known to be non controllable. This conclusion can be extended for freely spinning systems without external perturbations.

The variables of the corresponding linear system include the three projections of the total moment of momemtum on the orbital axes, $\underset{\sim}{\mathbf{H}}\bullet\hat{\underset{\sim}{\mathbf{X}}}_i \quad i = 1,2,3$, the three deviations from the equilibrium conditions (15), the various constraints – further, the constraints define slices in the state manifold and consequently are decreasing the dimension requirements for the corresponding distributions – of as well as some combination of the orientation parameters.

This linearization permits to define optimal procedures for large manoeuvres such as detumbling, capture and seasonal reversals.

REFERENCES

[1]Willems, J.L., *Stability theory of dynamical Systems*, Nelson, London, 1970.

[2] Mansour, M. *Robust stability in systems described by rational functions*, Report ETH Zurich, 1991.

[3] Porter, B., *Stability criteria for linear dynamical systems* Academic Press, 1968.

[4] Gantmacher, F.R., *The theory of matrices* (vol I & II) (translation) Chelsea Publ. Co., New York,1959.

[5] Isidori, A., *Nonlinear control systems* Springer Verlag, 1989.

[6] Barnett, S., A new formulation of the theorems of Hurwitz, Routh and Sturm, *J. of the Institute of mathematics and its applications*, Vol 8, n 2, 1971.

[7] Willems, P.Y, *Introduction à la mécanique*, Masson éd., Paris, 1979.

[8] Zajac, E.E., The Kelvin-Tait-Chetaev theorem and extensions, *J. of Astronautical Sciences* vol 11, 1964.

[9] Mingori, D.L., Stability of linear systems with constraint damping and integrals of motion, *Astronautica Acta*, vol 16, n 5, 1971.

[10] Willems, P.Y.,Attitude stability of deformable satellites, *Proc. CNES int. symp.* Paris, 1968.

[11] Atanossov, V., and P.Y. Willems, On the nonlinear control of orbiting gyrostats, *Proc. IUTAM congr. on Nonlinear Optimal control*, Moscow, mai 1992.

ON NON-LINEAR LONGITUDINAL STABILITY OF AN AIRCRAFT IN A DIVE IN THE PRESENCE OF ATMOSPHERIC DISTURBANCES

by

L.M.B.C. Campos[1,2] and **A.A. Fonseca**[2]
[1] Alexander von Humboldt Scholar, at Max-Planck-Institut für Aeronomie
Postfach 20, Max-Planck Strasse 20
3411 Katlenburg-Lindau, Germany on leave from:
[2] Instituto Superior Técnico, 1096 Lisabon Coden, Portugal

SUMMARY

We consider the non-linear longitudinal stability problem (§1) of aircraft starting a dive from an initial velocity far removed from the steady dive speeds: the aim is to find a pitch control law, which will keep the aircraft in a constant glide slope, compensating for the phugoid mode (§2). The problem is extended to account for the presence of arbitrary atmospheric winds, e.g. windshears (§3). The theoretical stability curves are compared with flight test data (§4), obtained using the BAFR (Basic aircraft for flight research) in portugal (§5). The model includes a number of effects, and has also some restrictons, which do not affect its suitability for the present application: (i) it includes compensation of the phugoid, but not short-period, mode i.e. neglects rotational inertia; (ii) it accounts for boundary-layer and induced drag, but not wave drag, i.e. applies at low Mach numbers, typical of approach to land; (iii) it neglects lateral motion, but allows for non-linear effects on longitudinal motion; (iv) it leads to a free-flying control law, in stall free conditions, and in the absence of autopilot or active control. These additional effects would be relevant to other applications, and would require extension or modification of the present model.

INDEX

LIST OF SYMBOLS

a, b, d	parameters of non-linear mathematical model of longitudinal stability (9a, b,c)
a_x, a_y, a_z	longitudinal, lateral and vertical acceleration
c	airfoil chord or mean wing chord
f	longitudinal stability coefficient (30)
f_o, f_1, f_2	parameters of engine thrust versus velocity relation (7)
g	acceleration of gravity
h	error in airspeed measurement
k	coefficient of induced drag in (5)
l	lengthscale of windshear
m	aircraft mass
n	parameter in normalization of dive velocity (42b)
p	fluid pressure
t	time
u	longitudinal wind velocity
v	perturbation of groundspeed due to wind
w	vertical wind velocity
x	distance along flight path
z	altitude
A	magnitude of longitudinal windshear (37)
C_D	total drag coefficient
C_{Df}	boundary layer drag coefficient in (5)
C_L	lift coefficient at arbitrary angle-of-attack
D_{LO}	lift coefficient at zero angle-of-attack
$C_{L\alpha}$	lift slope $C_{L\alpha} \equiv \partial C_L/\partial\alpha$
D	drag force
F	total dimensionless force along flight path (8)
J	groundspeed U normalized to minimum drag speed U_{md}
J_o	normalized groundspeed at time $t = 0$ or start of test run $t = t_o$
$J_\pm$	normalized steady dive speed
J_1	normalized ground at end of test run $t = t_1$
M	Mach number
P	perturbation of groundspeed divided by groundspeed (38a)
Q	perturbation of airspeed divided by groundspeed (38c)
S	reference surface area for aerodynamic forces
T	thrust
T_{min}	thrust for flight at minimum drag speed
$\bar{T}$	thrust corrected for weight along flight path
U	groundspeed along flight path
U_o	initial groundspeed along flight path
U_+	stable steady dive groundspeed
U_-	unstable steady dive groundspeed
U_{md}	minimum drag speed
V	true air speed
V_s	stalling speed
X	distance x along flight path normalized to lengthscale
Y	function in stability response (35b)
α	angle-of-attack
γ	glide slope angle
ϵ	angle of thrust-line with flight path
ϵ_o	angle of thrust-line with aircraft longitudinal datum
$\bar{\epsilon}$	average value of ϵ
λ	coefficient in non-parabolic lift-drag polar (5)
μ	aircraft parameter for longitudinal response (38b)
ν_o, ν_1	thresholds for glide slope at start and end of dive
ρ	mass density of atmospheric air
τ	aerodynamic time scale of longitudinal response (30)
θ	pitch angle
$\bar{\theta}$	average pitch angle
ζ	elevator deflection angle

§1– INTRODUCTION

The type of stability analysis most often performed concerns small perturbations from a constant or slowly evolving mean state; it leads to systems of linear ordinary differential equations with constant coefficients, which can be analysed using Laplace transforms [3–5] or the usual methods of control theory [6–7]. When the deviations from the reference state are large, one is lead to non-linear differential equations [8]; the non-linear problems have received much recent attention, e.g. in connection with flight at high-angles of attack and spins[9]. A distinct, but also non-elementary, stability problem arises for small perturbations, when the reference state is itself varying rapidly, i.e. on time scales comparable to those of the perturbations[10]; an example is the excitation of the phugoid or short period mode, by atmospheric disturbances of comparable time scale. This kind of problem leads to linear differential equation, with coefficients varying with time [11]; it was a topic of considerable research in the past, and although it has gone a bit out of fashion, it has not lost its relevance. We address both types of non-elementary stability problem [12–14], viz. non-linear and unsteady, in connection with aircraft longitudinal response in a dive.

§2 NON-LINEAR LONGITUDINAL STABILITY IN A DIVE

For the present section (§2) we consider flight in still air, so that we need not distinguish true airspeed V from groundspeed U, and will reason in terms of the latter. We consider and aircraft which starts a dive at an arbitrary initial groundspeed $U_o = U(0)$, which may be far removed from any of the two steady dive speeds $U_\pm$. It will be shown later that $U_+ > U_-$, and the upper steady dive speed U_+ is stable, and the lower U_- unstable; thus (i) after a long time $t \to \infty$, the aircraft Either tends to the stable dive speed $U(t) \to U_+$, or (ii) it diverges from the unstable dive speed ($U_- - U(t)$ increases, towards a stall $U(t_s) \to U_s$ at a finite time $t = t_s$). Either in the stable (i) or unstable (ii) case, the problem is non-linear if the groundspeed $U(t)$ at some time differs considerably from both steady dive speeds, i.e. neither of the conditions $\mid U^2 - U_+^2 \mid \ll U_+^2$ or $\mid U^2 - u_-^2 \mid \ll U_-^2$ is met. From example, the problem would be non-linear from time $t = 0$, if the aircraft started the dive at a ground far removed from both steady dive speeds, i.e. none of two conditions $\mid U_o - U_\pm \mid \ll U_\pm^2$ is met.

In operations, there could be several reasons why an aircraft might start a dive from a cruise or loiter speed far removed from the steady dive speeds, e.g.: (a) it could start an emergency approach to a landing strip seen late due to poor visibility; (b) it could start tracking a target of opportunity spotted at the last moment. In such cases the pilot would try to keep on a constant gilde slope, either by adjusting incidence, or by throttling the engine, or both, so that the trajectory lies on an inclined plane; the lateral control inputs, e.g. yaw or roll, if they remain small, i.e. linear, do not couple to the longitudinal motion, which may be non-linear. In this case, of small heading changes, the trajectory is approximately a straight line, and our non-linear stability problem has two-degrees-of-freedom, associated with the phugoid and short-period mode. Since we are concerned with longer time scales, we may omit the short-period mode, by neglecting the rotational intertie of the aircraft, i.e. assuming that a desired value of the angle-of-attack can be obtained almost instantaneously, i.e. on a time short compared with the period of the phugoid. Thus we are left with a one-dimensional non-linear stability problem, in a sense inverse to Lanchester's [15] phugoid, and which may be stated as follows: if an aircraft starts a dive from an arbitrary initial ground $U_o \equiv U(0)$, how should the groundspeed $U(t)$ (or angle-of-attack ($\alpha(t)$) vary with time t, so that it keeps on a constant slide slope?; an alternative, and equivalent, statement is: what is the pitch control law

$\theta(t)$, to be flown by the pilot, or programmed into an automatic control system, which will exactly compensate the phugoid, and deep the aircraft on a constant slide slope angle γ? This problem is almost as old [16] as Lanchester's phugoid [17], and could be called the "inverse phugoid problem " [18] it was not settled as definitively [19–21] as the phugoid problem, for the inverse problem has been the subject of later research [22–23].

Our one-dimensional non-linear longitudinal stability problem is represented in Figure 1: the aircraft flies along a constant glide slope γ with groundspeed $U = dx/dt$, where x is the distance along the flight path; it is acted upon by four forces, namely, the weight W along the vertical downwards, the lift L transverse to the flight path upwards in a vertical plane, the drag D opposite to the motion, and the thrust T along the motion (in some cases the thrust may make an angle ϵ with the flight path, viz. $\epsilon = \theta + \epsilon_o$, where θ is the pitch angle and ϵ_o is the angle of the thrust line with the aircraft's longitudinal datum). The mathematical model is specified by the balance of transverse and longitudinal forces. The former, transverse force balance, is static, and states that the lift:

$$\frac{1}{2}C_L(\alpha)\rho S U^2 = L = W\cos\gamma - T\sin\epsilon, \tag{1}$$

is balanced by the component of weight transverse to the flight path (minus the traverse component of thrust, if latter does not lie along the flight path); the latter, longitudinal force balance, states that the inertial force, equal to mass $m = W/g$ times acceleration:

$$(W/g)dU/dt = T\cos\epsilon - D - W\sin\gamma, \tag{2}$$

balance the (longitudinal component of) thrust T, minus drag D and longitudinal component of weight. The drag is specified by a formula similar to the lift (1):

$$D = \frac{1}{2}C_D(\alpha)\rho S U^2, \tag{3}$$

replacing lift C_L by drag C_D coefficient. We can make (2) dimensionless by dividing by the weight, and using (1) and (3), we conclude that the longitudinal acceleration, made dimensionless by dividing by the acceleration of gravity:

$$g^{-1}dU/dt = -\sin\gamma - (T/W)\cos\epsilon + \cos\gamma(C_D/C_L)/\{1 + (T/L)\sin\epsilon\}, \tag{4}$$

is determined by the longitudinal component of weight, the thrust-to-weight ratio T/W projected on the flight path, and the drag-to-lift ratio C_D/C_L modified by glide slope $\cos\gamma$. The last term in (4) can be omitted $(T/L)\sin\epsilon \ll 1$, if either thrust makes a small angle with path $\sin\epsilon \ll 1$, or lift-to-drag ratio is high $T/L \sim D/L \ll 1$, or product of these two is negligible compared to unity.

Since our problem is one-dimensional, the r.h.s. or (4) should be expressible in terms of groundspeed alone. To do this, we note that for low Mach number flight, i.e. incompressible flow $M \lesssim 0.3$, we can neglect wave drag, and the drag coefficient consists of three other terms:(i) boundary layer drag, which is due to skin friction, and thus independent of lift [24]; (ii) induced drag, which is proportional to the square of the lift coefficient [25]; (iii) the parabolic lift-drag polar resulting from (i) and (ii), may have a non-parabolic correction, proportional to lift:

$$C_D(\alpha) = C_{Df} + k\{C_L 9\alpha)\}^2 + \lambda C_L(\alpha). \tag{5}$$

Thust last term in (4) can be written, after omission of the factor in curly brackets:

$$\begin{aligned}\cos\gamma\{C_D(\alpha)/C_L(\alpha)\} &= \cos\gamma C_{Df}/C_L(\alpha) + k\cos\gamma C_L(\alpha) + \lambda\cos\gamma = \\ &= (\rho S/2W)C_{Df}U^2 + (2W/\rho S)kU^{-2}\cos^2\gamma + \lambda\cos\gamma\end{aligned} \tag{6}$$

where in the last formula we substituted $C_L(\alpha)$ from (1), ommiting the term - $T\sin\epsilon/W$
$\cos\gamma \approx (D/L)\sin \ll 1$, as assumed before. We assume that the thrust-to-weight ratio has a dependence on groundspeed similar to (6), viz.:

$$T/W = f_o - f_1U^2 - f_2/U^2, \tag{7}$$

where f_o, f_1, f_2 are constants for a given engine type.

Substituting (6) and (7) into (4) we obtain a single equation for the groundspeed:

$$g^{-1}dU/dt = a - bU^2 - d/U^2 \equiv F(U), \tag{8}$$

where the total force per unit mass along the flight path, involves three coefficients:

$$a \equiv f_o\cos\bar{\epsilon} - \sin\gamma - \lambda\cos\gamma \tag{9a}$$

$$b \equiv f_1\cos\bar{\epsilon} + C_{Df}(\rho S/2W), \tag{9b}$$

$$d \equiv f_2\cos\bar{\epsilon} + k\cos^2\gamma(2W/\rho S); \tag{9c}$$

with two exceptions, these coefficients are constant, since they depend on aircraft parameters W, S, aerodynamic coefficients C_{Df}, k, λ, flight parameters γ, and propulsion characteristics f_o, f_1, f_2 which are all constant. The first exception is the angle of the thrust-line with the aircraftś flight path ϵ, for which we take an average value $\bar{\epsilon} = \epsilon_o + \bar{\theta}$ involving the constant angle to the aircraftś longitudinal datum ϵ_o, and an average pitch value $\bar{\theta}$; small deviations from the latter $\theta - \bar{\theta}$, would introduce a negligible correction $f(\cos\theta - \cos\bar{\theta}) \simeq f(\theta^2 - \bar{\theta}^2)/2 \ll f$ with $f \equiv f_o, f_1, f_2$. The second variable parameter is the air density, which would not be constant in a long dive, involving significant altitude change, Since in the following integration we treak the parameters (9a,b,c) as constant, our model applies to the followings parctical situation, corresponding to test flights detailed later (§4). A low-speed aircraft ($U \le 200kt, M \le 0.3$), performing a not very long dive (altidude loss $\Delta z \lesssim 700m$) along a constant glide slope. Another application, to which the present model would not apply, is the case of a modern high-speed airliner, with cruise speed optimized by supercritical airfoil design, starting a long dive from an altitude around the tropopause; in such a case wave drag would be important, and furthermore, density change with altitude could not be neglected.

Our mathematical model of longitudinal stability is specified by equation (7), which may be written:

$$g^{-1}dU/dt = -b(U^2 - U_-^2)(U^2 - U_+^2)/U^2, \tag{10}$$

where $U_\pm$ denote the two steady dive speeds:

$$U_\pm = \{a \pm \sqrt{a^2 - 4bd}\}/2b, \tag{11}$$

for which the total force along the flight path $F(U_\pm) = 0$ is zero, and hence there is no longitudinal acceleration $dU/dt = 0$ for $U = U_+$ or $U = U_-$. The existence of two steady dive speeds can be inferred from Figure 2, where we plot versus groundspeed U, the drag-to-weight ratio D/W (thick line) and the thrust-to-weight ratio T/W shifted by $T = T\cos\epsilon - W\sin\gamma$ (two thin lines).

The intersection of the lines specifies the two steady dive speeds $U_{\pm}$, for which the inertia force vanishes. The steady dive speeds would coincide if the thrust was the minimum for steady flight, i.e. at the minimum drag speed:

$$U_{md} = a/2b = (f_o \cos\bar{\epsilon} - \sin\gamma - \lambda\cos\gamma)/f_i \cos\bar{\epsilon} + C_{Df}\rho S/2W); \tag{12}$$

steady flight is possible only above this velocity $U \geq U_{md}$ i.e. if the condition $a^2 \geq 4bd$ in (11) is satisfied by (9a,b,c), so that the minimum thrust for steady dive is given by:

$$\begin{aligned} &(f_o \cos\bar{\epsilon} - \sin\gamma - \lambda\cos\gamma)^2 \geq \\ &\geq 4(f_1 \cos\bar{\epsilon} + C_{Df}\rho S/2W)(f_2 \cos\bar{\epsilon} + k\cos^2\gamma 2W/\rho S). \end{aligned} \tag{13}$$

The simplest particular case of (12,13) is level flight $\gamma = 0$, of an aircraft with symmetrical lift-drag polar $\lambda = 0$, with constant thrust $f_1 = 0 = f_2$ along the flight path $\bar{\epsilon} = 0$, for which we obtain well-known results [26–27]for the minimum drag speed $U_{md} = 2Wf_o/C_{Df}\rho S = 2T/C_{Df}\rho S$ and minimum thrust for steady flight $T/W = f_o \geq C_{Df}/k = T_{min}/W$.

For any groundspeed distinct from the steady dive speeds $U \neq U_{\pm}$, there is an acceleration $dU/dt \neq 0$ along the flight path, whose sign can be deduced either from equation (10) or figure 2, viz.: (i) between the steady dive speeds $U_- < U < U_+$ there is an excess of thrust (the $(T/W)\cos\epsilon + \sin\gamma$ lies above D/W in Figure 2), and thus acceleration $dU/dt > 0$ in (10); (ii) otherwise $U > U_+$ or $U < U_-$, there is an excess of drag (the thick line lies above the thin lines in Figure 2) , and thus decceleration $dU/dt < 0$ in (10). It follows that: (a) the upper steady dive speed U_+ is stable., because acceleration in its vicinity points towards it; (b) the lower steady dive speed U_- is unstable, because the acceleration in its vicinity points away from it. The evolution of the groundspeed as a function of time $U(t)$ is specified by integrating (10) from an initial velocity $U_o = U(0)$ at time $t = 0$, leading to:

$$\begin{aligned} e^{-t/\tau} = \{(U(t) - U_+)(U + U_+)/ \\ /(U - U_-)(U - U(t))\}^{1/(1-U_-/U_+)} \times \\ \times\{(U - U_-)(U(t) + U_-)/ \\ /(U(t) - U_-)(U + U_-)\}^{1/(U_+/U_- -1)}, \end{aligned} \tag{19}$$

where τ is an aerodynamic time scale, determined by the characteristics of the aircraft:

$$1/\tau \equiv 2bg(U_+ + U_-). \tag{20}$$

It is clear that as the elapsed time significantly exceeds the aerodynamic time scale $t \gg \tau$ then $e^{-t/\tau} \to 0$, and the groundspeed either approaches the upper steady dive speed $U(t \gg \tau) \to U_+$, or it diverges from the lower steady dive speed, i.e. $U(t \gg \tau) \to U_-$ becomes large. This confirms that the former U_+ is stable and the latter U_- unstable.

For short time compared to the aerodynamic time scale $t \ll \tau$, such that the groundspeed has not deviated much from the initial value:

$$\{U(t \ll \tau) - U_o\}^2 \ll (U_{\pm} - U_o)^2, \tag{21}$$

the equation (19) may be linearized with respect to groundspeed;

$$\begin{aligned} U(t) = U_o + \{(U_+^2/U_o^2 - 1)(1 - U_-^2/U_o^2) \\ (U_+ + U_-)/2\}(1 - e^{-t/\tau}), \end{aligned} \tag{22}$$

leading to an exponential dependence on time, with time scale τ. For $t = 0$ it follows from (22) that $U(0) = U_o$, but for $T \gg \tau$ we do not have $U(t \gg \tau) = U_+$ because (22) is not valid for long times. For long times we have to use the non-linear relation(19), which expresses time $t(U)$ as function of groundspeed, and cannot be readily inverted for $U(t)$. Thus we plot the stability curves in Figure 3, in dimensionless form, with time divided by aerodynamic time scale t/τ as abscissa and groundspeed divided by minimum drag speed $J(t) \equiv U/U_{md}$ as ordinate. We have one curve for each initial velocity $J_o \equiv U_o/U_{md}$. The two steady dive speeds $U_{\pm}$ (11) are related to the minimum drag speed U_{md} (12) by:

$$U_+^2 + U_-^2 = a/b = 2U_{md}^2, \qquad J_{\pm}^2 + J_-^2 = 2, \tag{23a,b}$$

so that of the two parameters $J_+ \equiv U_{\pm}/U_{md}$ only one is independent.

Thus, if we choose an aircraft for which the stable steady dive speed is 1.2 times the minimum drag speed $J_+ = U_+/U_{md} = 1.2$, it follows that the unstable steady dive speed is $J_- = U_-/U_{md} = 0.748$ times the minimum drag speed. This is the case illustrated in Figure 3, where we see that: (i) for initial groundspeed below the unstable dive speed $J_o < J_- = 0.748$, the drag is too high, and speed drops rapidly towards a stall; (ii) for initial groundspeed between the unstable and stable steady dive speeds $0.748 = J_- < J_o < J_+ = 1.2$, there is excess thrust, and the aircraft accelerates gradually to the upper, stable dive speed ; (iii) if $J_o > J_+ = 1.2$, there is excess drag, and the aircraft decelerates gradually to the stable upper steady dive speed.

§ 3 - RESPONSE TO ATMOSPHERIC WIND DISTURBANCES

In the preceding account (§2) we have assumed flight in still air, but the stability theory could be extended to include the effect of atmospheric disturbances [28], of which the windshear [29] has received greatest attention. The atmospheric wind conditions can be measured using meterological methods [30–31], Doppler radars [32–33], airborne lasers [34], or comparison of INS with non-inertial velocity measurements [32–34] . A windshear (Figure 4) is a toroidal vortex near the groundspeed, causing a downflow through its core, which is deflected by the groundspeed into a radial outflow. The effects of the downflow combined with headwind changes on flight [35–36], are of particular concern during approach to land [37–40]. The safety aspect [41] is related to aircraft response and control in such conditions [42–44], which, as the analysis of specific events [45–46] shows, can be quite hazardous. Since there are many possible combinations of aircraft characteristics and windshear locations relative to the flight path, the flight simulators, though useful, can hardly train the pilot for all possible events. We extend the preceding non-linear longitudinal stability problem, to the case of flight path directly through a windshear,or any other condition with wind in a vertical plane through the constant glide slope of the aircraft; this is equivalent to determining the air or groundspeed (or angle-of-attack or pitch) as a function of time, so that the aircraft keeps on a constant glide slope, compensating for the phugoid mode induced by the wind [47–48].

In the presence of wind we should distinguish between the groundspeed and true air speed, e.g. for the non-linear longitudinal stability model (8), the total force depends on the true air speed V, and the acceleration in an inertial frame is the time derivative of the groundspeed:

$$g^{-1}d(U + v)/dt = F(V); \tag{24}$$

we denote the groundspeed in the presence of wind by $U + v$, to distinguish from the groundspeed U in the absence of wind (8), i.e. v is the perturbation in groundspeed due to the wind:

$$g^{-1}dv/dt = F(V) - F(U). \qquad (25\ )$$

We allow (Figure 1) for the presence of an horizontal u and vertical w wind, which are given functions of distance x along the flight path, and hence, alternatively, a function of time. Bearing in mind that the angle of the flight path with the horizontal is the glide slope, the true air speed V is related to groundspeed $U + v$ and wind components u, w, by:

$$V^2 = (U + v + u\cos\gamma - w\sin\gamma)^2 + (u\sin\gamma + \cos\gamma)^2. \qquad (26\ )$$

We know the groundspeed U in the absence of wind from the previous (§2) solution of (8), and calculate the airspeed by (26), and then integration of (25) gives us the groundspeed perturbation v, or groundspeed in the presence of wind $U + v$.

For arbitrary strong winds (25) is a non-linear differential equation with variable coefficients, namely, the groundspeed in the absence of wind U and the wind components u, w,; since the latter are most conveniently given as functions of position $u(x), w(x)$, we replace time by space derivatives:

$$dv/dt = (dv/dx)dx/dt = (U + v)dv/dx, \qquad (27\ )$$

in (25), which specifies, in general, a non-linear wind perturbation of a non-linear background state:

$$g^{-1}(U - v)dv/dx = F(V) - F(U); \qquad (28\ )$$

thus (28) will apply even if the wind components u, w were comparable to the groundspeed U. The latter can be written as function of position, if we re-write the mean state equation (10) as:

$$(U/g)dU/dx = -b(U^2 - U_-^2)(U^2 - U_+^2)/U^2, \qquad (29\ )$$

and integrate it from an initial groundspeed $U_o \equiv U(0)$ at position $x = 0$, viz.:

$$\begin{aligned}\exp(-2bgx) = (U(x)^2 - U_+^2)/(U_o^2 - U_+^2)^{1/(1-U_-^2/U_+^2)} \times \\ \times(U_o^2 - U_-^2)/(U(x)^2 - U_-^2)^{1/(U_+^2/U_-^2-1)}. \qquad (30\ )\end{aligned}$$

In conclusion, for the general problem: (i) the groundspeed for the non-linear mean state is given as a function along flight path x by (30) (or as a function of time t by (19,20)); (ii) the horizontal $u(x)$ and vertical $w(x)$ wind components, which could be comparable to the grondspeed for non-linear wind perturbations, lead (26) to the airspeed V; (iii) the difference of the respective total forces (8;9a,b,c) then leads to a non-linear differential equation (28) with variable coefficients, for the perturbation in groundspeed v; in general, this will have to be integrated numericaly, given the complex dependence of the total force (8) on v through (26).

We can use analytical integration in the case, which is almost always true in real flight conditions, of wind speed small relative to groundspeed $u^2, w^2 \ll U^2$; for example, if the windspeed does not exceed $u, w \leq 20$m/s, and the groundspeed exceeds $U \geq 60$ m/s, we have $(u/U)^2, (w/U)^2 \ll 0.3 = 0.09 \ll 1$. It will be shown subsequently that the perturbation in groundspeed is of the same order of magnitude as the wind $v \sim u, w$, so that the linearization with regard to the wind is consistent with the linearization with regard to the perturbation in groundspeed $v^2 \ll U^2$. Thus the problem of most practical interest is of a linear wind perturbation, with variable coefficients, from a non-linear mean state, which may be varying on comparable time or lenghthscales. We apply the linearization first to the airspeed (26), by noting that since the glide slope γ is small $\gamma^2 \ll 1$, we have $\cos\gamma = 1 - \gamma^2/2 \simeq 1$, and since $u/U, w/U$ are also small, we can neglect $(u, w)\sin\gamma \sim (u, w)\gamma \ll U$, so that:

$$V^2 = (U + v + u)^2 + w^2 = U^2 + 2(u + v)U, \qquad (31a)$$

$$V = U + u + v, \qquad (31b)$$

where we have used $u^2, v^2, w^2 \ll U^2$. The equation (28) for the wind perturbation may also be linearized:

$$(U/g)dv/dx = F(U + u + v) - F(U) = (u + v)(dF/dU), \qquad (32\ )$$

so that all effects of the non-linear mean state appear in the coefficient:

$$f(U) \equiv (g/U)(dF/dU) = 2g(d/U^4 - b). \qquad (33\ )$$

where (8) was used, and the coefficients b, d are given by (9b,c).

The groundspeed perturbation (32) satisfies:

$$dv/dx - f(U(x))v = f(U(x))u(x), \qquad (34\ )$$

which may be integrated for any given mean state $U(x)$, e.g. (30) and wind $u(x)$, viz.:

$$v(x) = v_oY(x) + Y(x)\int_0^x \{u(\zeta)f(\zeta)/Y(\zeta)\}d\zeta, \qquad (35a)$$

where

$$Y(x) \equiv \exp\{\int_0^x f(\zeta)d\zeta\}. \qquad (35b)$$

In conclusion, we start from the non-linear mean state (30) in the absence of wind, and calculate the force influence factor (33), which appears in (35b); this together with the horizontal wind $u(x)$, specifies the linear groundspeed perturbation (35a). Note that the linearization procedure has implied (27a) that we may take the horizontal wind along the glide slope, and the vertical wind across it; also, the vertical wind w does not affect the airspeed (27b), although it introduces a small, but non-negligible change w/U in the angle of attack. A further simplification occurs if the non-linear mean state varies on lengthscales much longer than those ℓ of wind, i.e. $dU/dx \ll U/\ell$; in this case we can calculate (37) for a constant background velocity $f_o \equiv f(U_o)$, so that (35a,b) lead to a groundspeed perturbation:

$$v(x) = v_oe^{f_ox} + e^{f_ox}f_o\int_0^x e^{-f_o\zeta}u(\zeta)d\zeta, \qquad (36\ )$$

due to a wind $u(\zeta)$.

As an example, we consider a typical windshear profile (Figure 5), with a headwind followed by a tailwind, on a lengthscale ℓ:

$$u(x) = AU_o\sin(2\pi x/\ell), \qquad (37\ )$$

with maximum magnitude $\pm AU_o$, such that $(u/U_o) \leq A^2 \ll 1$. Substituting (37) into (35) and assuming that there is no initial groundspeed perturbation $v_o = 0$, we conclude that the wind (34) induces a groundspeed perturbation:

$$P \equiv v(x)/U_o = (A/(1+\mu^2))\{\mu(1-\cos(2\pi x/\ell)) - \sin(2\pi x/\ell)\}, \qquad (38a)$$

where the characteristics of the aircraft appear only through the parameter:

$$\mu \equiv 2\pi/f_o\ell = \pi/\{g\ell(d/U_o^4 - b)\}. \qquad (38b)$$

The perturbation of airspeed (27b) is given by:

$$\begin{aligned} Q \equiv V(x)/U_o - 1 &= \{u(x)+v(x)\}/U_o = \\ &= (A/1+\mu^2))\{\mu(1-\cos(2\pi x/\ell)) + \mu^2\sin(2\pi x/\ell)\}. \end{aligned} \qquad (38c)$$

The groundspeed and airspeed perturbations vanish both at the start of the headwind $u(0) = 0 = v(0)$ and at the end of the tailwind $u(\ell) = 0 = v(\ell)$, and they coincide, but are non-zero, at the change from head to tailwind $u(\ell/2) = 0, v(\ell/2) = 2A\mu/(1+\mu^2)$. The perturbation of the groundspeed has a maximum:

$$v_{max} = v((\ell/2\pi)\,arccot\,\mu) = AU_o(\mu + \sqrt{1+\mu^2})/(1+\mu^2), \qquad (39a)$$

whereas the perturbation of the airspeed has maxima and minima:

$$\begin{aligned} (u+v)_{max,min} &= (u+v)_{x=\ell/2\pm(\ell/2)\,arctan\,\mu} = \\ &= U_oA(2\mu + \mu^3 \pm \mu^4)/(1+\mu^2)^2, \end{aligned} \qquad (39b)$$

both dependent on the aircraft parameter.

We plot in Figure 6 the groundspeed (top) and airspeed (bottom) perturbation, normalized to the groundspeed, versus $X \equiv x/\ell$ distance x divided by windscale ℓ. We give curves for five values of the aircraft parameter $\mu = 1, 2, 4, 7, 10$ whose mining we discuss next. This parameter is given by (38b), where for flight at constant thrust $f_1 = 0 = f_2$, (9b,c) reduce to the second terms; here, for reference conditions $W\cos\gamma = L = \frac{1}{2}C_{LO}\rho S U_o^2$, we have $2W/\rho S = C_{LO}U_o^2/\cos\gamma$, so:

$$\mu = (\pi U_o^2/g\ell\cos\gamma)/\{C_{LO}k - C_{Df}/C_{LO}\} \qquad (40\ )$$

vanishes at the minimum drag speed, and changes of sign across it. We consider as reference condition an approach to land at $U_o = 30 - 60$m/s, with the lower speed corresponding to a light general aviation aircraft and the higher value to a heavy jet airliner or high-performance fighter; a glide slope $\gamma \sim 3^o$ for convention, or $\gamma \sim 6^o$ for STOL, has little effect on the acceleratio of gravity projected along the flight path $\bar{g} = g\cos\gamma$, which retains the value $g = 9.8$ m/s. Thus for a wind lengthscale $\ell = 1000 - 2000$m, the factor in curved brackets in (40) is $\pi U_o^2/g\ell = 0.14 - 1.54$. For an aircraft in the landing configuration, the induced drag is much larger than the boundary layer drag $kC_{LO}^2 \geq C_{Df}$, so that the term in curly brackets reduces to kC_{LO}; the lift coefficient is about unity at landing $C_{LO} \sim 1 - 2$, and the induced drag factor is given [24-25] by $k = (1+\delta)/\pi\Lambda$ where Λ is the aspect ratio and δ measures the deviation from elliptic loading or uniform downwash. Taking $\delta = 0.25$, we have $k = 0.40/\Lambda$, and for an aspect ration $\Lambda = 2 - 8$ ranging from a high-performance fighter to a subsonic transport, it follows that $C_{LO}k = 0.1 - 0.2$. Thus the aircraft parameter is lowest for a lightplane $\mu \sim 0.14/0.2 \sim 1$, intermediate for a heavy jet transport $\mu \sim 1.54/0.2 \sim 7$, and highest for a fighter $\mu \sim 1.54/0.1 \sim 15$. We see from Figure 6 that as μ increases, the groundspeed perturbation is smaller, and the airspeed perturbation becomes similar to the longitudinal wind; for small μ, the glide slope will be maintained, if the groundspeed increases in the tailwind phase, so that the airspeed increases significantly in the headwind phase, but does not fall much below the initial value during the tailwind phase.

§4 - COMPARISON WITH FLIGHT TEST RESULTS

The predictions of the non-linear theory of longitudinal stability (§2), can be checked by performing flight tests. One can take as the reference, for the theory, figure 3, showing how the groundspeed U (normalized to the minimum drag speed) evolves with time, for different values of the initial groundspeed U_o. For the purpose of flight testing, the region $U_o < U_-$ below the unstable steady dive speed U_- is unsafe, due to rapid divergence towards the stall. The region $U_o > U_+$ above the stable steady dive speed U_+ is not very interesting, since an aircraft would seldom be flying horizontally at a speed higher than is sought at the end of the dive. Thus the region of greater interest is that between the steady dive speeds $U_+ > U_o > U_-$, which is represented in Figure 9, in terms of $J \equiv U/U_{md}$ groundspeed U normalized to minimum drag speed U_{md}: several initial normalized velocities $J_o \equiv U_\pm/U_{md}$ should be tried an flight, for comparison with the curves in Figure 9. In order for the comparison with flight test data to be significant, we should check compliance with the main assumptions of the theory, namely: (i) that the effects of the short period mode can be neglected; (ii) that there is no coupling of longitudinal to lateral motion; (iii) that the atmospheric wind do not cause significant deviations from still air results.

In order to comply with these conditions, the pilot was instructed to start the dive from steady, straight and level flight, and then to keep as closely as possible a constant glide slope, using ILS for glide slope indication, if it could be received in the altitude range of the test; horizontal flight was recovered only after the dive speed has stabilized, to make sure the steady dive speed was attained. We can use the flight data records to check whether these objectives were met, e.g. we show in Figure 7 the following data for Flight 1: (top left) the altitude z plot shows that an approximately constant glide slope was maintained, between entry into, and recovery from, the dive; (top right) the tas V, which apart from wind effects, should coincide with groundspeed U, increased steadily, but non-linearly, with time, between constant values before and after; (bottom left) the maintenance of an approximate constant glide slope at rapidly varying tas, implied significant excursions of angle-of-attack α, obtained through pitch control; (bottom right) this control activity is testified by the vertical acceleration, which, as expected, was mostly negative during entry, and positive during exit from the dive, with smaller variation in between.

The theory (§2) is non-linear but concerns only longitudinal motion, i.e. no lateral coupling. Thus the pilot was instructed to fly on a constant glide slope, using pitch control alone, and was given no other task, such as heading or localizer catch, which might induce activity in the heading or roll axis. Since it is important to check this assumption (ii) of the theory, that the lateral control activity is minimal, we reproduce, in Figure 8, additional data records for Flight 1, showing that: (top left) the transverse acceleration a_y, in the range -0.03 g to $+0.015g$ was small compared with the normal acceleration, in the range $-0.45g$ to $+0.52g$; (top left) the longitudinal acceleration a_x, in the range $-0.17g$ to $-0.01g$, was not so small, but reflected groundspeed changes along the flight path $a_x = dU/dx$; (center left) the heading χ changed rather little, by 1^o between 312.9^o and 311.9^o; (center right) the roll angle ϕ remained small, i.e. between $+1.5^o$ and -0.5^o, during most of the test run $t \leq 19s$; (bottom left) the pitch angle θ varied significantly, between $+5.5^o$ and -3.7^o, demonstrating the control activity required to maintain a constant glide slope; (bottom right) this control activity was exercised through the elevator, whose deflection angle ζ took values between -16.2^o and -21.2^o. The only indication of significant lateral control activity, was the roll angle exceeding $\phi \geq 1.5^o$ for time $t \geq 19s$ (center right of figure 9); however, by this time the airspeed has already stabilized (top right of figure 7), so that this does not change the conclusion that assumption (ii) of the theory was met.

It is clear that it is possible to check automatically compliance with the theory, as regards assumption (ii), by putting bounds on certain parameters (lateral acceleration a_y, heading χ change, roll angle ϕ), whileleaving others (longitudinal acceleration a_x. pitch angle θ and elevator deflection ζ) unrestricted. We proceed to show that it is also possible to check automatically compliance with assumption (i) of the theory, which concerns neglect of the effects of the short period mode. The latter would appear as an oscillation in pitch, and although there is some evidence of this (at bottom left of Figure 8), it is not significant compared with the overall pitch excursion. Another way to assess compliance with assumption (i), is to note that it is associated with neglect of the rotational inertia of the aircraft; the latter implies that the angle-of-attack α (or pitch angle θ) required to keep a constant gilde slope cannot be achieved instantaneously. The associated excursions of the aircraft velocity vector, would cause deviations from a constant glide slope; although glide slope deviations can be seen (at top left of Figure 7), they are clearly small compared with the altitude change. Thus compliance with assumption (i) can also be checked automatically, by putting bounds on short period oscillations of the pitch angle θ, or on deviations of the glide slope from the mean.

The measurement of glide slope $\gamma \equiv dz/dx$ as a function of time $\gamma(t)$, allows automation of another significant task: finding where, in the recorded signal, is the time series with the phenomenon of interest. This is usually done by the flight test engineer, using his knowledge and experience to interpret the data. We have already given one example, viz. the larger roll angle values $\phi \geq 1.5^o$ (in center right of Figure 8), are no concern to lateral-conditional coupling, because they occur outside the time range of interest $t_o = 5s \ll t \ll 19s \equiv t_1$ for our study of longitudinal stability. We could determine this time range (t_o, t_2), by measuring the glide slope $\gamma(t)$ as a function of time, and averaging it over the data record $\bar{\gamma}$. The start of the dive is then taken as the time t_o a fraction ν_o of the mean slope is first attained $\gamma(t > t_o) \gg \nu_o\bar{\gamma}$, i.e. the aircraft is already accelerating in the dive condition; the end of the dive corresponds to the glide slope be coming small again $\gamma(t > t_1) \leq \nu_1\bar{\gamma}$, i.e. acceleration is small and as close to steady dive speed. The values chosen for $\nu_o, \nu_1 \ll 1$, determine the precise start and end of the test run; we can take "reasonable" initial values, e.g. $\nu_o = 0.1 = \nu_1$ and subject them to adjustment or learning process, as for the bounds used to check for compliance with assumptions (i) and (ii) of the theory.

In order to be able to compare flight test data with the mathematical model, it is necessary to use the same scaling, to for independent (time or distance) and dependent (speed, altitude, ect.) variables. In the case of Figure 3, the time t is normalized to the aerodynamic time scale τ which is specified (20) by the acceleration of gravity, steady dive speeds $U_\pm$ and parameter b of the mathematical model (9b); the latter involves further data, such as the form drag coefficient C_{Df} and thrust coefficient f_1 and wing loading W/S . If not all this data is known accurately , then instead of calculating τ from (20), we may infer it from the flight test records. The mathematical model in Figure 3 shows that the dimensionless time for stabilization when speed is close to the steady value is $t/\tau = 8$; thus we choose as time scale $\tau = m(t_1 - t_o)/8$, where t_o is the time at the start and t_1 the time at the end of the dive, and this is equivalent to using a dimentionaless time $(t - t_o)/\tau = 8(t - t_o)/(t_1 - t_o)$. The dependent variable is also re-scaled linearly, i.e. we change from the groundspeed U in kt, to the dimensionless velocity $J \equiv U/U_{md}$ normalized to the minimum drag speed U_{md}, viz.:

$$J = J_o + n(U - U_1); \qquad (41)$$

we need not know the aerodynamic and propulsion data which goes into the minimum drag speed (12). It is sufficient to know that the ratio $J_+ = U_+/U_{md}$ from the stable steady dive speed J_+, since this is only parameter in our non-linear longitudinal stability model. The tas at the start $U_o = U(t_o)$ and end $U_1 = U(t_1)$ of the dive then specify the parameters:

$$J_o = J_+U_o/U_1, \qquad n = (J_+ - J_o)/(U_1 - U_o), \qquad (42a,b)$$

in the linear relation (41), with initial value (J_o, U_o) in (42a) and slope n between (J_oU_o) and $(J_1, U_1) \equiv (J_+, U_+)$ given by (42b).

We can now calculate any number of experimental points from the fight data records, e.g. every 2s, and mark them in Figure 9, in comparison with the theoretical stability curves. Each est of points, corresponds to one flight, starting at a different initial normalized velocity $J_o \equiv U_o/U_{md}$. An exact fit of the measurement to the theory would correspond to the successive test points, for the same flight, all lying on the theoretical curve passing through the initial point J_o. Such an exact fit should not occur due to measurement errors, and the theory is deemed to be stratified to within experimental accuracy h, if the deviation between the theoretical curve passing through J_o, and the remaining flight data points, never exceeds h. This can be checked by applying the bar shown in Figure 9, to every flight data point, of the 6 flights considered; alternatively as shown for Flight F5, two lines $\pm h$ apart from the theoretical curve can be drawn, and the flight data points should lie between them. The accuracy of measurement of airspeed depends on the location and calibration of the Pilot tube; for the flight test aircraft used (see §5) the error does not exceed $h_1 = 4kt$. It should be borne in mind that the Pilot tube measures airspeed and for the theory we need groundspeed, and these coincide only under the assumption (iii) of absence of wind. The presence of a longitudinal wind (§3) of amplitude $u_{max} = UA$, causes deviation between ground and airspeed which is given, at most, by (39b) e.g. $(2\mu + \mu^3 + \mu^4)/(1 + \mu^2)^2 = 328/289 = 1.14$ times the wind, for a light transport $\mu \sim 4$. Thus the strongest wind found 8 kt would cause a $9kt$ error, which is larger than the calibration $h_1 = 4kt$. It is therefore essential to subtract the wind from the airspeed, to obtain the ground speed. The wind speed can be obtained comparing airspeed with the inertial reference, introducing an additional error $h_2 = 1kt$. Thus the total error in groundspeed is $h = h_1 + h_2 = 5kt$ or a maximum 5%. We could instead take groundspeed directly from the inertial reference, for an error $h_2 = 1kt$; however, the groundspeed in the absence and presence of wind $u_{max} = 8kt$; differ at most (39a) by $(\mu + \sqrt{1 + \mu^2})/(1 + \mu^2) = 0.5u_{max} \simeq 4kt \sim h_3$, leading to an error $h_1 + h_3 = 5kt$, much as before $h = 5kt$. The flight test data points are all within $4kt$ of the theoretical curves.

§5 - THE BASIC AIRCRAFT FOR FLIGHT RESEARCH (BAFR)

It may be worthwhile at this stage to mention some features of the aircraft used for the preceding test flights. This is [49–51] CASA 212 Aviocar twin turboprop, fitted with an instrumentation system described below (§5), which we designate BAFR (Basic Aircraft for Flight Research), and is the flying component of LNEV (National Flight Test Laboratory). Most Aeronautical research or development programmes start with the letter "A" for advanced ; we start the designation of our research aircraft with "B" for basic. Our more modest aspiration are justified by the fact that the BAFR was developed using equipment offered by foreign aeronautical research institutions, which also provided technical training and supervision of programme. The design of the flight test instrumentation system and its installation, involving 7000 man-hours of work, 1000 pages of documentation and 4 km of cabling, were entirely performed in Portugal. Also the aircraft is not dedicated fulltime to flight testing. and must be converted in one day to the operational configuration. All these constraints make the BAFR perhaps worth describing as a concept of affordable flight test aircraft. The ultimate aim of a flight test facility is to be able develop instrumentation packages tailored for different applications in various types of aircraft. For example, the test facility has developed instrumentation packages for application as diverse testing helicopter hover performance over a rough sea (typical bad weather search and rescue mission), or measuring loads on a wing due to an external store (typical new store - existing aircraft integration task). In order to gain this flexible capability, an appropriate first step would be to develop a relatively 'comprehensive' flight test instrumentation system in a dedicated aircraft. By "comprehensive" we do not mean a larger number of identical sensors, but rather a wide variety of sensors, say for performance, aerodynamics, control, propulsion, structural, navigation and systems data. Such a "comprehensive" system gives the breath of experience in types of sensors, signals and equipment, upon which dedicated of "at hoc" applications can be bsed. This approach rejects the attitude of starting flight testing on small packages, dependent on piecemeal aquisition of further equipment or knowledge, in a chicken feed situation. It aims instead at the smallest self-consistent "quantum jump", by independent, flight test capability is dependent, on the availability of wide range of instrumentation, in the present case offered by NLR. The list of parameters selected for measurement on BAFR (figure 10) includes a variety of sensors and signals, besides a few related to the aircraft configuration and test conduct. The air data sensors for pressure, temperature and angle-of-attack and side-slip should be mounted on a long boom (described later) if they are to achieve a greater accuracy than aircraft instruments. Linear and angular accelerations are measured by a platform under the floor, near the c.g., which gives results which are independent from, and can be compared with, those of the Inertial Navigation System. Other navigation data, is extracted from the Dopplelr radar, radio altimeter, ILS, etc..., with special provisions to avoid corrupting the original signal. The positions of all control and lift surfaces are measured, as well as essential propulsion parameters; strain gauges on the engine mountings are used to give in-flight measurement of thrust. Although the number of sensors is not large by modern standards, their variety is such as to represent most of what will be needed, perhaps in larger quantities, in other applications.

The aircraft allocated to the BAFR is not dedicated fulltime to flight testing; in fact the latter occupies only up to 30 flight hours per year, and the aircraft is mostly operated in aerial photography, earth resources, search and rescue, and other miscellaneous duties. Thus the flight test instrumentation system must interfere as little as possible with the operational missions, or else be quickly removable; the conversion between operational and flight test configuration should not take more than one day. The concepts of "quickchange" and "non-interference" with operations tend to be antagonistic: (i) leaving most equipment permanently in the aircraft would provide for quick-change, but would interfere with operations; (ii) conversely, removing most equipment to avoid interference with operations, but would imply tedious re-installation. The key to a compromise solution lies in leaving permanently installed in the aircraft all that does not interfere with the operational mission. Also avoiding that the list of removable includes too many sensors requiring calibration after re-installation; what needs to be removed should be packages as much as possible in a quick-change configuration.

The preceding considerations suggest that most sensors should be left permanently installed in the aircraft to avoid lenghty calibrations after each removal or re-installation, provided they can by fitted in such a way as not to interfere with aircraft operations. Such an installation is possible for many sensors, e.g. gyros, rate gyros and accelerometers can be put in unobtrusive locations, and synchros measuring the position of control and high-lift surfaces have to be carefully placed to ensure non-interference with linkages, and breakability in case of jamming as well as accurate data. Some sensors like the flight data boom interfere with aircraft operations, and have to be removable; the same applies to the drum in the roof of the fuselage from with is unrealed the cable to the trailing cone, deployed from its tail housing. The cabling and connectors, which involve many man-hours on installation work, should be left in place, and additionally require a minimum amount of rewring should new sensors be added. This is best done (Figure 11) by having "regional" connector panels for sensors in each section of the aircraft (nose, tail, centre fuselage, right and left wing), and a "central" connector panel near the data aquisition system. The later system which, due to weight and volume, must be removable, should be easy to connect and disconnect.

The Data Aquisition System is housed in a single rack, which can be installed and removed using the load rails in the cabin. It contains the "central" connector panel, to which are connected the "regional" connector panels, e.g. the "nose" connector panel takes signals from the strain bridge mounted on the control column which measures pilot stick forces, and from the strain gauges measuring rudder pedal forces. The signal from "regional" connector panels e.g., the under-deck connector panel (UDCP) in Figure 12 are collected at the main rack connector panel (MRCP), and routed to the appropriate signal conditioning unit (SCU) or synchro-to-digital conversion unit (SDCU) and digital conditioning unit (DCU); the signals are pulse code modulated (PCM) prior to on-board tape recording or telemetry transmission to the ground. The main rack also includes the power distribution unit (PDU). The choice of the CASA 212 Aviocar twin-turboprop aircraft as the basic of BAFR was dictated by several considerations including: (i) relatively low operating costs per flight hour, a desirable feature for a basic test aircraft; (ii) large cabin volume, allowing easy installation and removal of the main rack, and work around it; (iii) relative insensivity to external modifications, the most significant of which is the removable air data boom.

The design of the air data boom (Figure 13) is a good example of the compromise between measurement accuracy and structural ridigity, and the constraints on location and removability. The location above the fuselage was decided by exclusion of other possibilities: (i) on the nose, the boom support structure would interfere with the weather or Doppler radar radiation pattern; (iii) on the fuselage side, interference with the propeller slipstream would be excessive; (iv) on the wing outboard of the propeller disc, misalignement due to wing bending would be excessive. The air data boom over the fuselage needs to be relatively long to avoid excessive aerodynamic interference from the nose; a long boom would have its resonances excited by atmospheric turbulence, needing an inverted-V support on the fuselage nose.

This support has little effect on aircraft aerodynamics or crew visibility . It is removed together with the boom, using three attachment points, two on the sides of the nose and one on the top of the fuselage. The air boom has pressure and temperature sensors and wind vanes for angle of attack and sideslip.

§6 - CONCLUSIONS

The analysis of stability of an aircraft must be based [18,4] on a mathematic model (§2). Its verification by flight test (§4) requires the measurement of two kinds of parameters: (i) those whose temporal (or spatial) evolution is to be compared with prediction; (ii) those which should remain within certain bounds, to ensure that effects omitted in the theory are indeed negligible. Both process can be readily automated, and, of course, the highest achievable accuracy is important for (i); this depends on the quality of the instrumentation (§5) and the way tests are conducted. The occurence of atmospheric disturbances during the tests, may effect the accuracy, unless these effects are corrected for, which requires further modification of the theory. This may also be automated, leaving one more task: the identification of where the event of interest lies, in the recorded time series. We have given a particular, glide slope criterion, for the specific case of longitudinal stability in a dive. In general, for other kinds of flight testes, the automated location of an "atmospheric event", or an "aircraft manouever", in a data record, can be made by using a disturbance intensity indicator. This is the topic to which we turn our attention elsewhere [52–54]

References

1. Bryan, G.H., *Stability in Aviation*, MacMillan, 1911.
2. Duncan, W.J., *Principles of control and stability of aircraft*, Campridge University Press, 1952.
3. Lepage, W.R., *Complex variable and Laplace Transform for Engineers*, McGraw-Hill 1961.
4. Churchill, R.V., *Operational mathematics*, McGraw-Hill, 3rd ed. 1972.
5. Sneddon, i.n., *The use of integral transforms*, McGraw-Hill, 1972.
6. Babister, A.W., *Aircraft dynamic response and stability*, Pergamon Press, 1980.
7. McRuer, J. & Askhenas, S., *Aircraft stability and control*, McGraw-Hill, 1986.
8. Forsyth, A.R. *Treatise of differential equations*, MacMillan, 6th erd 1929.
9. Gvicheteau, P., Application de la theorie des bifurcations á la perte de controle sur les avions de combat, *AGARD Conference Proceedings CP. 319*, paper 17 (1981); also *Rechearche Aerospatiale* p. 61-73.
10. Brockhaus, R., *Flugregelung*, Holdenburg, 1977.
11. Ince, E.L., *Linear differential equations*, Dover 1956.
12. Routh, E.J., *Dynamics of a system of rigid bodies*, Cambridge University Press, 1905.
13. Painlevé, P., *Cours de Mécanique*, Gauthier-villars, 1930-6.
14. Fuller, A.T., *Stability of motion*, Taylor & Francis, 1975.
15. Lanchester, F.W., *Aerodonetics*, Conctable 1908.
16. Painlevé, P., Étude sur le régime normal d'an avion, *Technique Aeronautique* 1 (1910), 3-11.
17. Mises, R. von, *Theory of Flight*, Dover, 1945.
18. Campos, L.M.B.C. & Aguiar, A.J.M.N., On the inverse phugoid problem as an instance of non-linear stability in pitch, *Aeronautical Journal*, paper 1559/1 (1989), p. 241-253.
19. Perkins, C.D. & Hage, R.E., *Airplane performance, stability and control*, Wiley, 1950.
20. Lecomte, P., *Mécanique du vol*, Dunod, 1962.
21. Etkin, B., *Dynamics of flight stability and control*, Wiley, 1974.
22. Neumark, S., Prolems of longitudinal stability below the minimum draf speed, and the theory of stability under constraint, *Aeronautical Research Council report* 2983, 1953.
23. Pinsker, W.J.G., Glide path stability of aircraft under speed constant, *Aeronautical Research Council report* 3705, 1971.
24. Milne-Thomson, L.M., *Theoretical Aerodynamics*, MacMillan, 1966.
55. Prandtl, L. & Tjietjens, O., *Hydro-und Aeromechanik*, Springer, 1929-1944.
26. Houghton, E.L., & Brock, A.E., *Aerodynamics for engineering students*, Edward Arnold, 1960.
27. Lighthill, M.J., *An informal introduction to fluid mechanics*, Oxford University Press, 1986.
28. Etkin, B.,AIAAA Wright Brothers Lecture, 1980: The turbulent wind and its effect on flight, *University of Toronto Aerospace Department*, Review 44.
29. Schlickenmeier, H., FAA integrated wind shear program, *2nd International Conference on Aviation Safety*, Touluse 1986, Ed. Cepadues.
30. Clodman, J., Müller, F.B. & Morrisey, B.G., Wind regime in the lowest one hundred meters, as relateld to aircraft takoffs land landings, *World Health Organization Conference*, London 1986, p. 28-43.
31. Glazunov, V.G. & Guerava, V.Z., A model of windshear in the lower 50 meter section of the glide path, from data of low inertia measurments. *Mathematical Lending Library of Science and Technology*, NLL-M23036, 1973.
32. Sweezy, W.B., Moninger, W.R. & Strauch, R.G., Simulation of radar-measured Doppler velocity profiles in low-level windshear, *FAA RD-78-46*.
33. Zrnic, D.S., Doviak, L.J., Lee, J.T. & Eilts, M.D., Weather phenomena that affect aviation. Proceedings *2nd International Conference on Aviation Safety*, Touluse, ed. Cepadues.
34. Woodfield, A.A. & Vaughan, J.M., Using an airborne CO_2 SW-laser fot free stream airspeed and windshear measurements, *AGARD Conference Proceedings CP-373*, paper 22, 1984.
35. Woodfield, A.A. & Woods, J.F., Worldwide experience with windshear 1981-2, *AGARD CP-347*,Paper 11.
36. Vorsmann, P., An on-line realization for precise wind vector measurements on board the Do 28 research aircraft, *International Congress of Aerospace Sciences*, Paper 84-5. 10.1.
37. Zhu, S. & Etkin, B., A fluid dynamic model of a downburst, *University of Toronto Aerospace Department*, Report 271, 1983.
38. Jones, J.G., Modelling of gust and windshear for aircraft assessment and certification, *RAE*, 1976.
39. Bray, R.S.A., A method of three-dimensional modelling of windshear environments for flight simulator applications, *NASA TM-85969*, 1984.
40. Bochis, V., Dynamics of an aircraft under an windshear of arbitrary direction, *AIAA Journal of Guidance* **7**, 615-619.
41. Frost, W., Cosby, B. & Camp, D.W., Flight through thunderstorm outflow, AIAA *Journal of Aircraft* 16, 11-15.
42. Leurs, J.K. & Reeves, J.B., Effect of shear on aircraft landing, *NASA CR-2287*, 1973.
43. Cavalcanti, S.G., Critical conditions of the automatic control of landing from decision height in variable winds, *University of Toronto Aerospace Department*, Report 284, 1984.
44. Schanzer, G., The effect of gusts and windshears for automatic STOL approach and landing, *AGARD CP-140*, 1973.
45. Reid, L.D., Markov, A.B. & Graf, W.O., The application of ltechniques for predicting STOL aircraft response to windshear and turbulence during landing approach, *University of Toronto Aerospace Department* report 215, 1977.
46. Schanzer, G., Infuence of windshear on flight safety, *AGARD CP-347*, paper 12.

47. Diederich, F.W., Response of an airplane to random altmospheric disturbances, *NASA TN - 3910*, 1957.

48. Brockhaus, R. & Wuest, P., Open-loop compensation of windshear effects on low level flight, *AGARD CP- 240*, Paper 19, 1978.

49. Van der Waart, J.C., Aircraft response to windshears and down draughts, *AGARD CP-260*, Paper 16, 1979.

50. Woodfield, A.A. & Woods, J.F., Windshear from headwind measurements on British Airways B747-236 aircraft, *RAE TM-409*, 1981.

51. Fujita, T.T., Microburst windshear at New Orleans internations airport, Kenner, Louisiana on 9 Luly 1982, *Univerity of Chicago*, SMRP research paper 199, 1983.

52. Campos, L.M.B.C., On the influence of atmospheric disturbances on aircraft aerodynamics, *Aeronautical Journal*, paper 1085 (1984), p. 257-264.

53. Campos, L.M.B.C., On aircraft flight performance in a perturbed atmosphere, *Aeronautical Journal*, paper 1305 (1986), p. 302-312.

54. Campos, L.M.B.C., On the disturbance intensity as an indicator of aircraft performance degradation in a perturbed atmosphere, *2nd International Symposium on Aviation Safety*, 18-21 november 1986, Toulouse, proceedings Editions Cépadues p. 175-190.

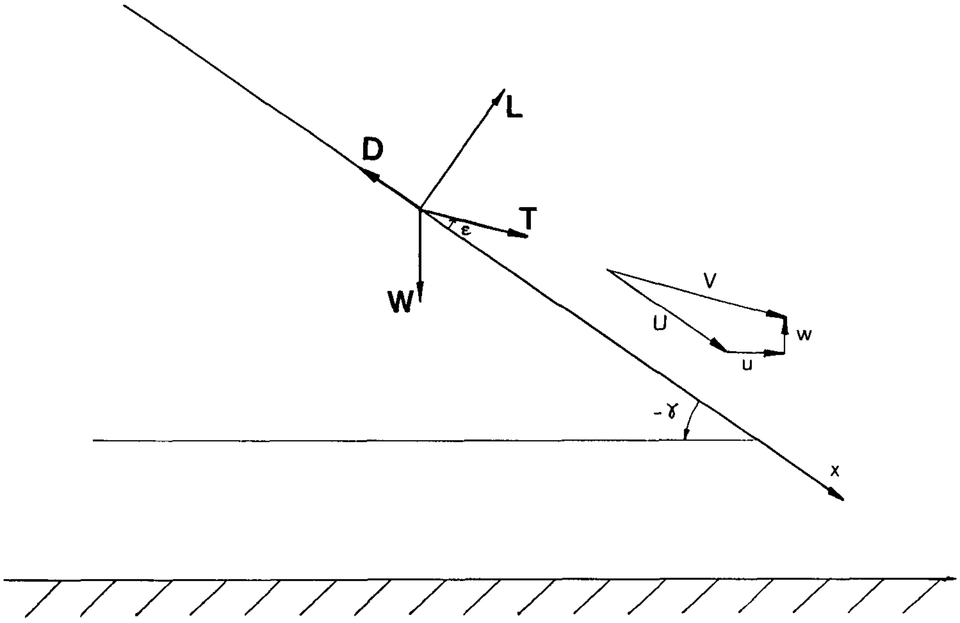


Figure 1 - Aircraft flying on a constant glide slope γ, with weight W vertically downwards, lift L perpendicular to flight path, drag D opposite to motion and thrust T at an angle ϵ to it. The airspeed V coincides with groundspeed U, only in the absence of horizontal u and vertical w wind.

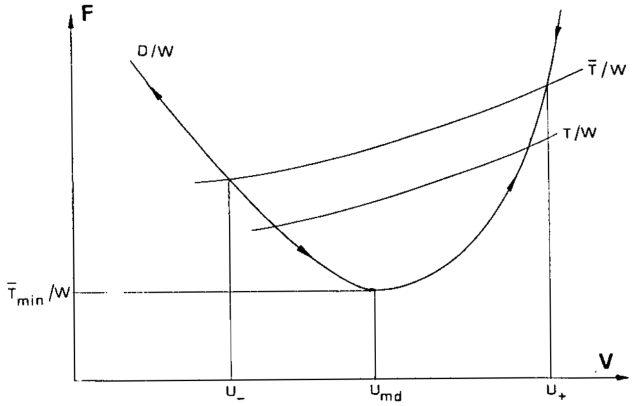


Figure 2 - Plot of forces F versus airspeed V. The dimensionless forces are the ratios to weight W, of drag D, thrust T and thrust plus weight projected along flight path $\bar{T}$. In the absence of wind the airspeed V coincides with groundspeed U, and we indicate the minimum drag speed U_{md} corresponding to the minimum thrust $\bar{T}_{min}$ for steady dive. Below $\bar{T} < \bar{T}_{min}$ steady dive is not possible, and above $\bar{T} > \bar{T}_{min}$, there are two steady dive speeds $U_{\pm}$; the lower U_- is unstable, in the sense that acceleration is away from it, and the upper U_+ is stable, in the sense that the acceleration is towards it.

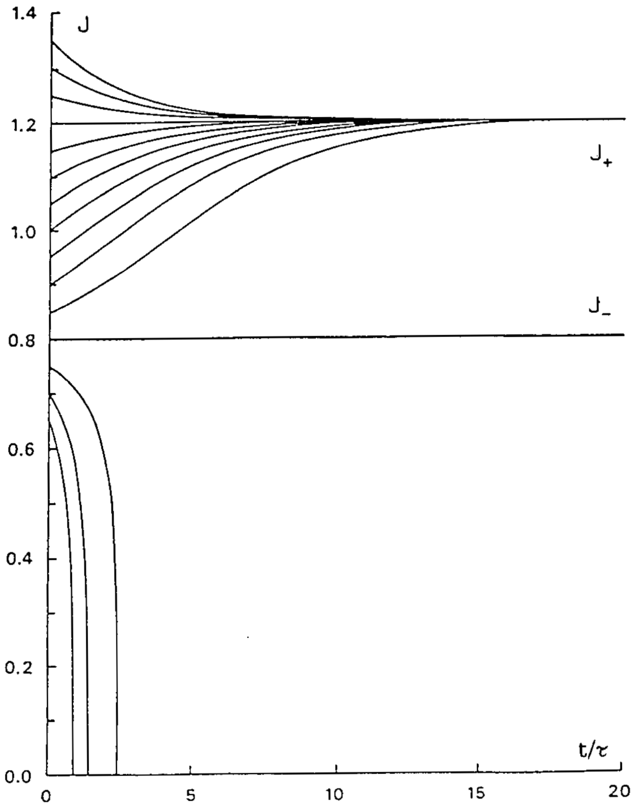


Figure 3 - Groundspeed U normalized $J \equiv U/U_{md}$ to minimum drag speed U_{md}, plotted versus time t normalized to aerodynamic time scale τ, for 15 values of initial groundspeed at time $t = 0$. The stability curves are given for an aircraft with $J_+ = 1.2$ stable dive speed 20% above minimum drag speed $U_+ = 1.20 U_{md}$; since the non-linear longitudinal stability model has only one free parameter, the unstable steady dive speed is not independent, viz. $J_+^2 + J_-^2 = 2$ implies $J_- = 0.75$.

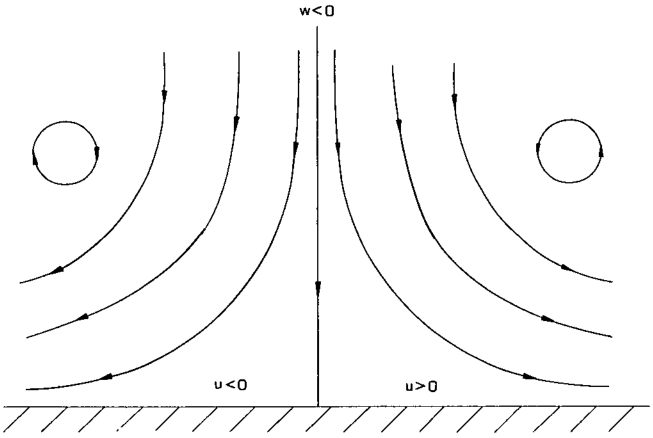


Figure 4 - The simplest model of a windshear is a toroidal vortex, causing a downflow $w < 0$ through its core, which becomes a radial outward wind near the ground; a section through the axis of the torus, shows two opposite vortices, a downflow $w < 0$ between, and a headwind $u > 0$ or tailwind $u < 0$ below, with changeover at mid point.

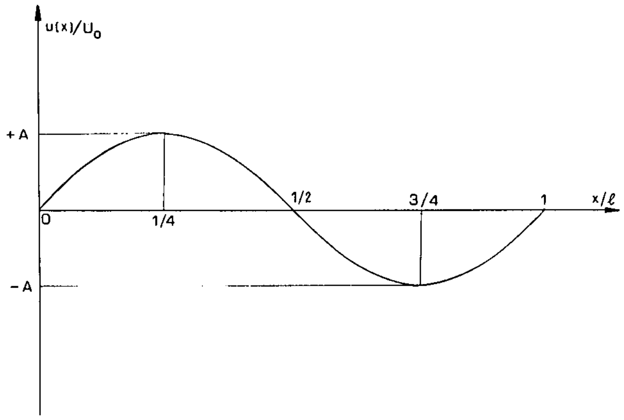


Figure 5 - The longitudinal wind model due to a windshear is a plot of wind normalized to groundspeed versus distance normalized to lengthscale. The simplest model is a sinusoid, with peak headwind equal to peak tailwind, equally spaced from start of headwind, change from head-to-tailwind, and end of tailwind.

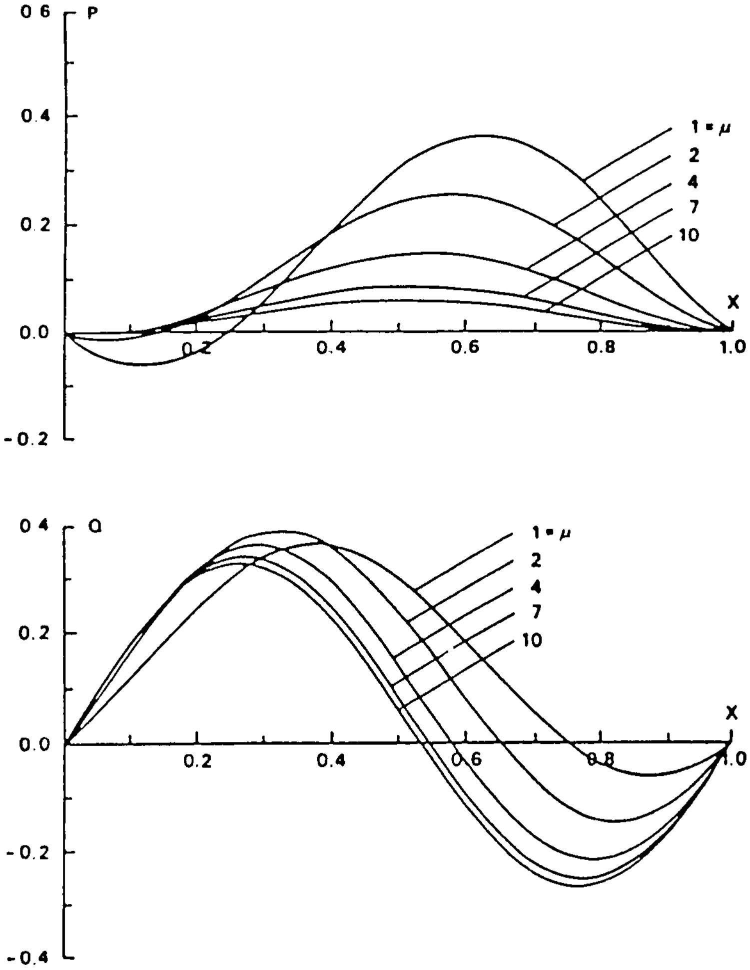


Figure 6 - The perturbation of groundspeed P and airspeed Q, normalized to mean state groundspeed, are plotted versus $X \equiv x/\ell$ distance x divided by windscale ℓ, for five values of the aircraft parameter μ ranging from a lightplane $\mu \sim 1$, through a heavy jet transport $\mu \sim 6$, to a high-performance fighter $\mu \sim 10$.

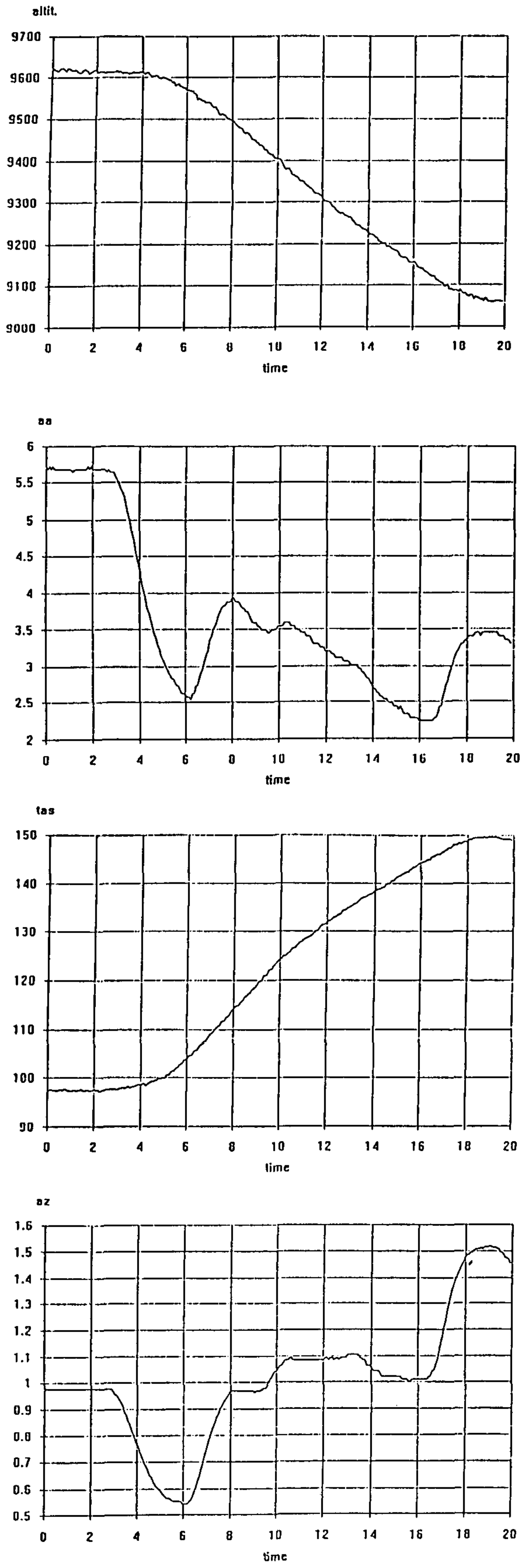


Figure 7 - The comparison of the non-linear theory of longitudinal stability with flight test data is based mainly on plots, versus time, of altitude z (top left) and airspeed V (top right), with the angle-of-attack α (bottom left) and vertical acceleration a_z (bottom right) testifying to the control activity needed to maintain a constant glide slope.

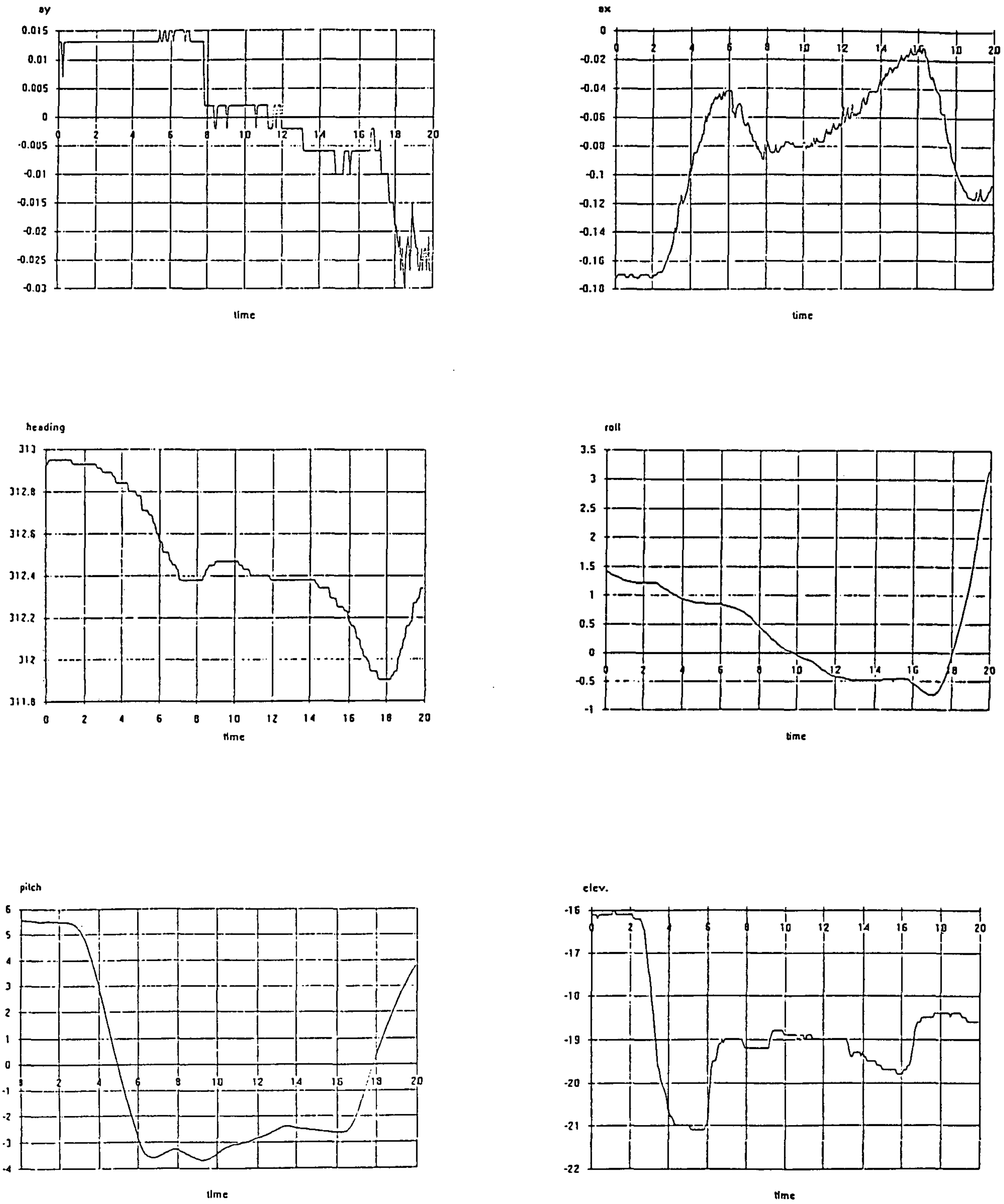


Figure 8 - To satisfy the assumption made in the non-linear theory of longitudinal stability, the lateral acceleration a_y (top left), heading χ (center left) and roll angle ϕ (center right) should be small, but the longitudinal acceleration a_x (top right), pitch angle θ (bottom left) and elevator deflection ζ (bottom right) need not be so restricted.

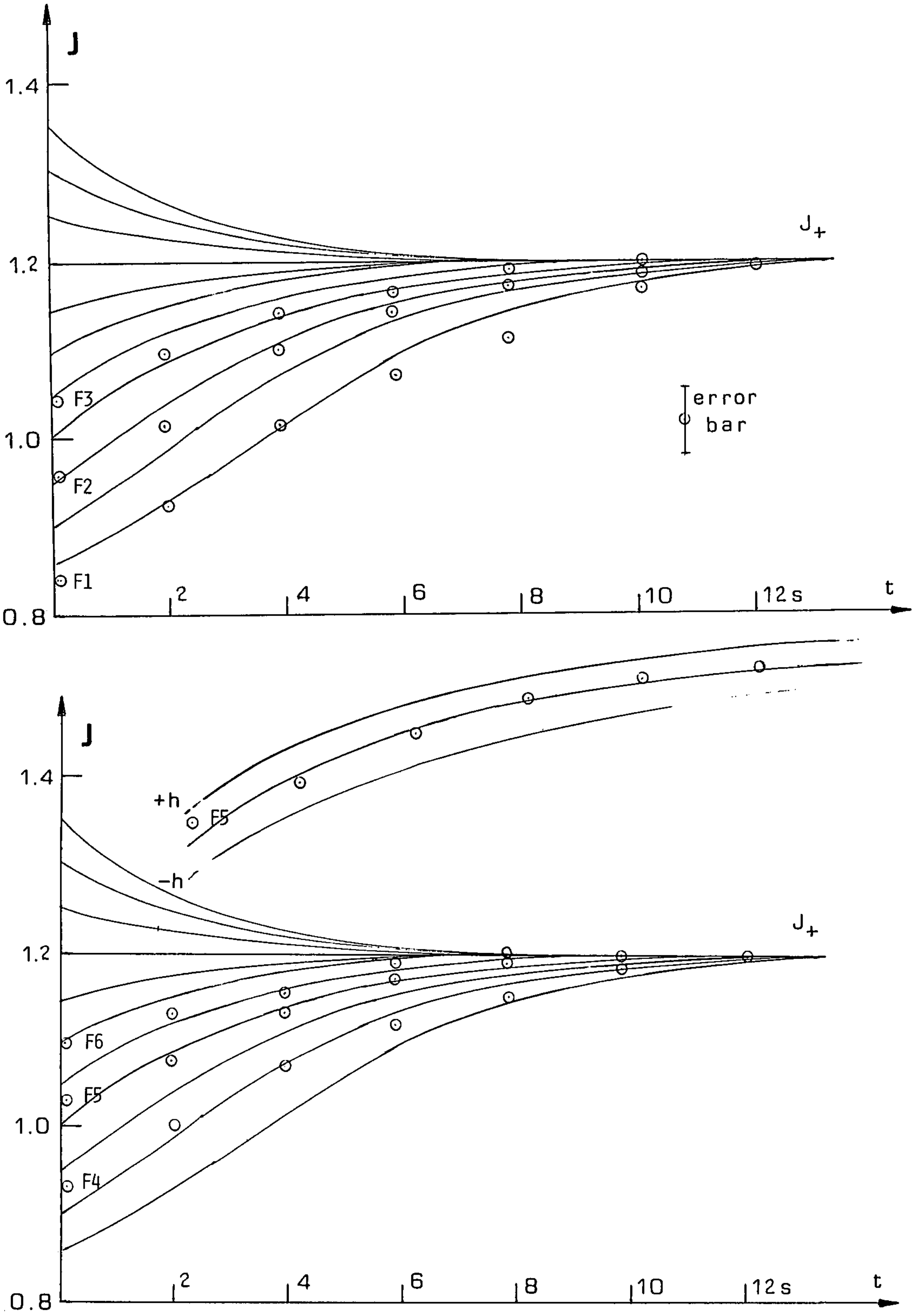


Figure 9 - If there would be no measurement errors, and the non-linear longitudinal stability theory were exactly satisfied, the data points (for each Flight F1 to F6) would lie on one stability curve. The error bar due to measurement errors (5 kt maximum) is larger than the maximum deviation (4 kt) of data points, as shown for Flight F5.

Category	Ref	Parameter	Code	Range
General	01	test signal	TS	
	02	time base 1	TB1	
	03	time base 2	TB2	
	04	run counter	RC	1/99
Air Data	05	diferential pressure	PD	0/10 kPa
	06	static pressure	PS	30/105 kPa
	07	calibration pressure	PC	0/15 mb
	08	total air temperature	TAT	-50/+50 °C
Configuration	09	ground/flight switch	GFS	on/off
	10	wing flap position	DF	0/45 deg
Control	11	elevator deflection	DE	-30/+20 deg
	12	LT aileron deflection	DA1	-20/+20 deg
	13	RH aileron deflection	DA2	-20/+20 deg
	14	rudder deflection	DR	-25/+25 deg
	15	elevator force	FE	-450/+450 N
	16	aileron force	FA	-300/+300 N
	17	rudder strain A	FR1	
	18	rudder strain B	FR2	
	19	rate of pitch	RP	-20/+20 deg/s
	20	rate of roll	RR	-60/+60 deg/s
	21	rate of yaw	RY	-20/+20 deg/s
	22	acceleration (X-dir)	AX	-1/+1 g
	23	acceleration (Y-dir)	AY	-1/+1 g
	24	acceleration (Z-dir)	AZ	-2.5/+2.5 g
	25	angle of attack	AA	-35/+35 deg
	26	angle of side-slip	AS	-35/+35 deg
	27	angle of pitch	AP	-90/+90 deg
	28	angle of roll	AR	-90/+90 deg
	29	INS valid signal	INSf	on/off
Propulsion	30	engine speed L	N1	0/41730 rpm
	31	engine speed R	N2	0/41730 rpm
	32	fuel flow L	FF1	0/1050 lb/h
	33	fuel flow R	FF2	0/1050 lb/h
	34	turbine gas temperature L	TGT1	0/930 °C
	35	turbine gas temperature R	TGT2	0/930 °C
	36	torque pressure L	TP1	0/65 psi
	37	torque pressure R	TP2	0/65 psi
Auto-Flight	38	autopilot engaged*	AE	on/off
	39	fight director mode	FDM	1/5
Navigation	40	true heading	HDG	0/360 deg
	41	HDG valid signal	HDGf	on/off
	42	radio altitude	RA	0/2500 ft
	43	RA valid signal	RAf	on/off
	44	localizer deviation	LLD	-90/+90 deg
	45	LLD valid signal	LLDf	on/off
	46	glide slope deviation	GSD	-80/+80 deg
	47	GSD valid signal	GSDf	on/off
	48	drift angle	DFT	-180/+180 deg
	49	DFT valid signal	DFTf	on/off
Thrust	50	A mount strain (L engine)	FA1	
	51	B mount strain (L engine)	FB1	
	52	C mount strain (L engine)	FC1	
	53	D mount strain (L engine)	FD1	
	54	E mount strain (L engine)	FE1	
	55	F mount strain (L engine)	FF1	
	56	G mount strain (L engine)	FG1	
	57	H mount strain (L engine)	FH1	
	58	A mount strain (R engine)	FA2	
	59	B mount strain (R engine)	FB2	
	60	C mount strain (R engine)	FC2	
	61	D mount strain (R engine)	FD2	
	62	E mount strain (R engine)	FE2	
	63	F mount strain (R engine)	FF2	
	64	G mount strain (R engine)	FG2	
	65	H mount strain (R engine)	FH2	

Figure 10 - The list of parameters measured in the BAFR is not very numerous, but includes a relatively wide variety of sensors and signals.

sensor location

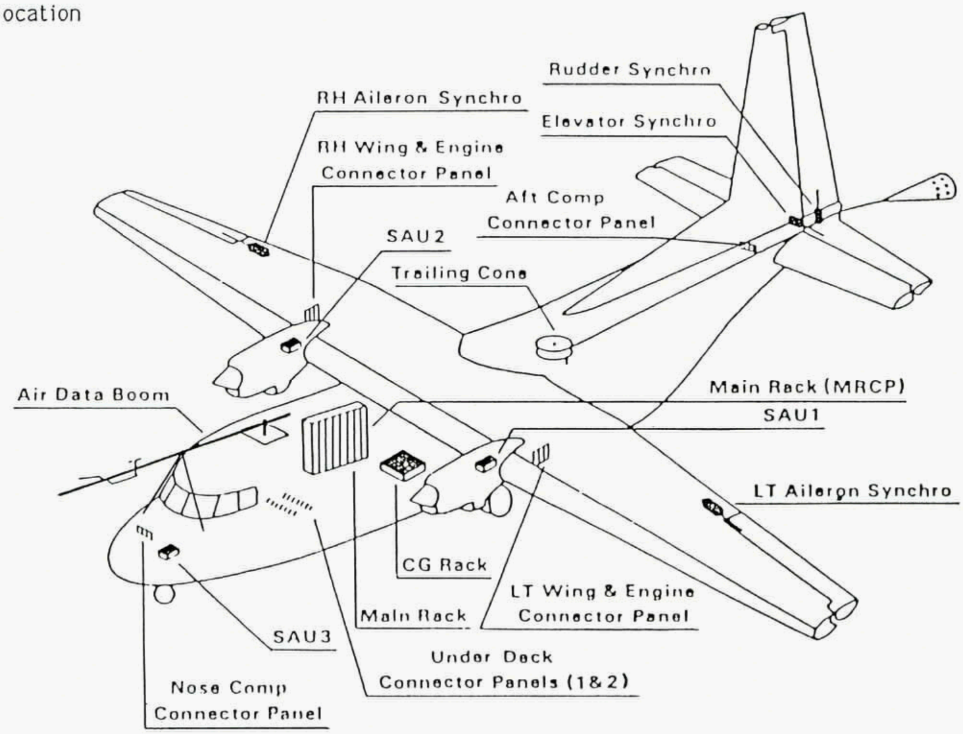


Figure 11 - The signals from sensors located in each part of the aircraft (nose, tail, fuselage underfloor, right and left wing) are collected in panels connected to a central one in the main rack.

DAS block diagram

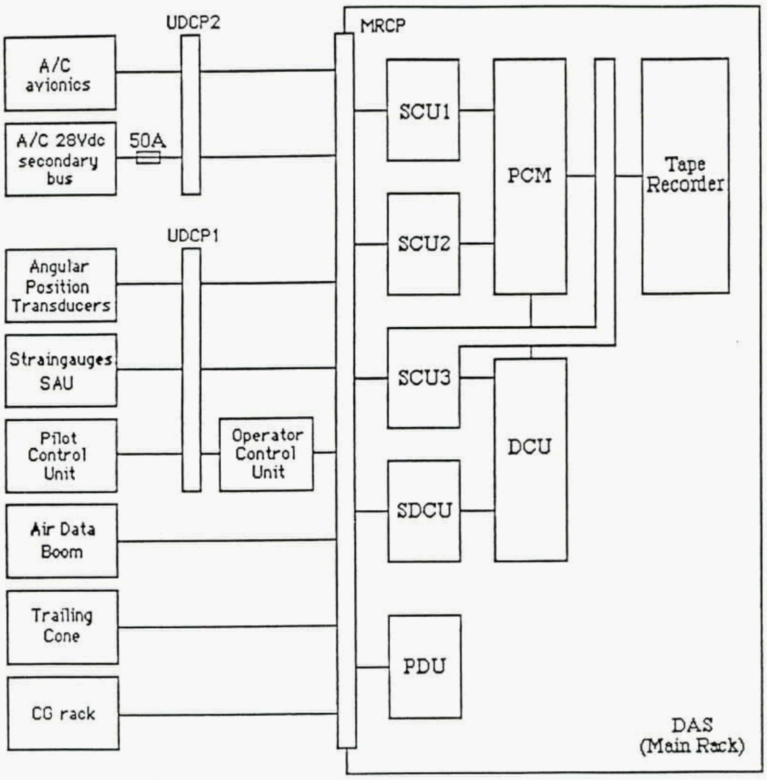
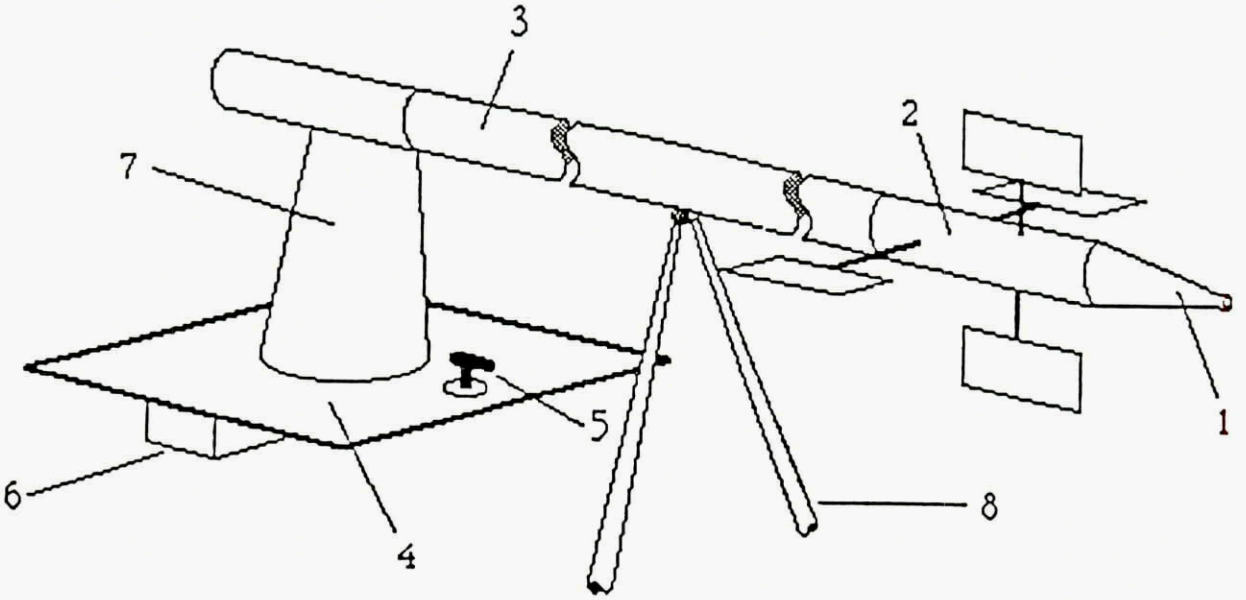


Figure 12 - The main rack is removable, and includes the central connector panel, and the data aquisition system, including the power distribution unit.



notes:
1 - pitot tube and static port
2 - two section cilindrical vane
3 - cilindrical body
4 - upper emergency exit
5 - temperature sensor
6 - connector panel
7 - main support
8 - supports

Figure 13 - Another removable item is the air data boom above the fuselage, with inverted - *V* support on the nose, containing pressure and temperature sensors and vanes for angle-of attack and sideslip.

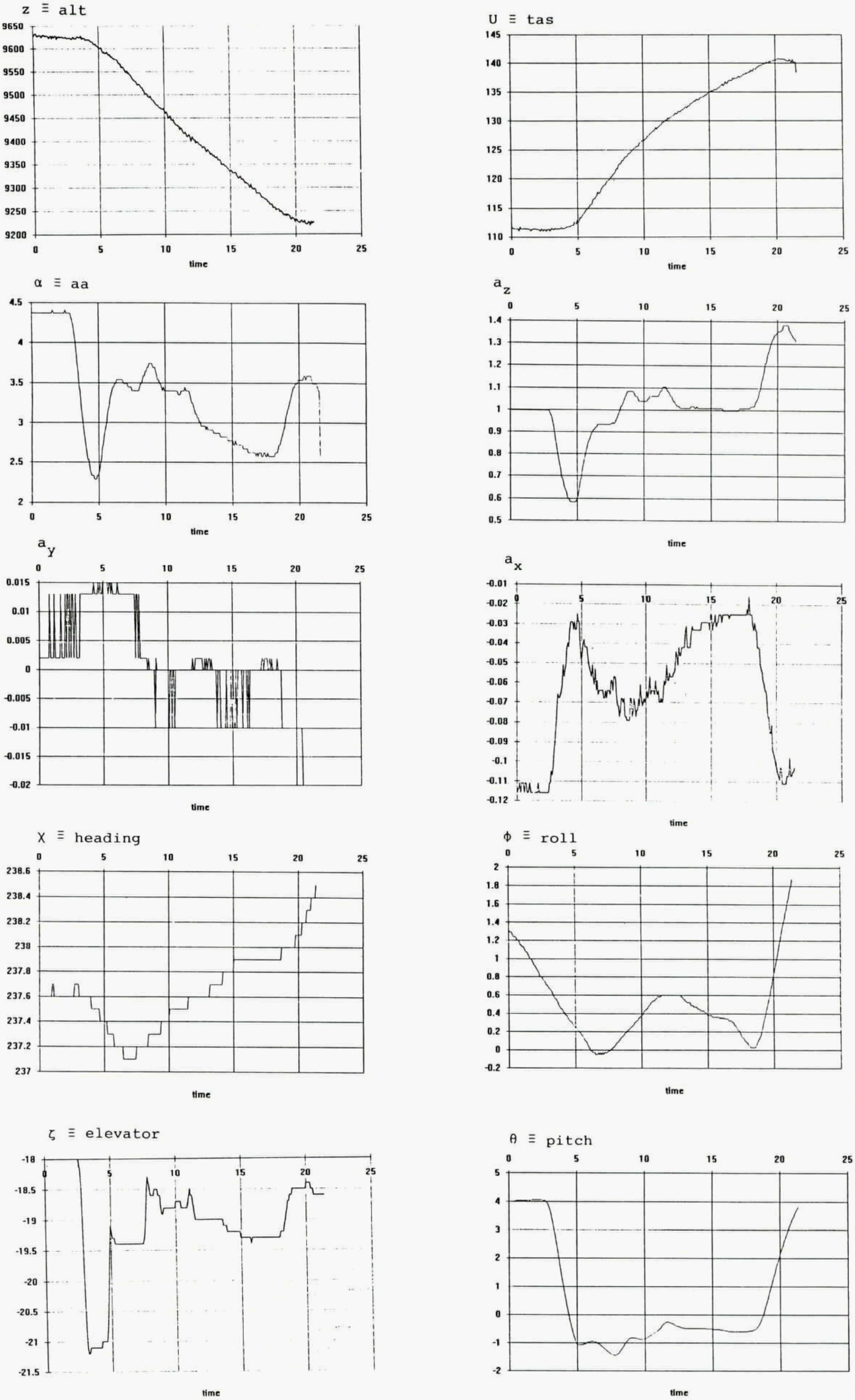


Figure 14 - Sample of data recorded for Flight F2, containing data corresponding to the combination of Figures 7 and 8 for Flight F1.

Discussion

SESSION I - QUESTIONS & ANSWERS (PAPERS 1, 2, 6)

PAPER 1: P. HEURRE

Question:

You said that:
- If a flow is convective unstable, we use real Ω, complex k.
- Reverse for absolutely unstable.

What about a convectively unstable flow with absolutely unstable pockets?

Answer:

Then we have to use a and k complex. This is at the root of the global mode idea which is detailed in Chance, Heurre & Redekopp (1990), Studies in Applied Mathematics

PAPER 2: J STARK

Question:

Most of your results, e.g., the existence of the reconstruction Function G, apply to a 'typical' system. What is a 'typical' system? How do you define it? How do you know that a system is 'typical'?

Answer:

A 'typical system' in this context is one which satisfies the Takens Embedding Theorem. Unfortunately there is no known set of sufficient conditions which will ensure that a particular system falls in this class. In practical applications one thus has to proceed heuristically, trying the techniques described in the paper and evaluating how well they perform.

It is important, however, to stress that systems which are not 'typical' are very rare. In particular (see Reference G3 of the paper) if one picks a system at random the probability that it fails to satisfy the Takens Theorem is zero. Furthermore, arbitrary small perturbations of any system will satisfy the Theorem.

One would thus be very surprised if a given system failed to be 'typical'. When this occurs it is usually due to special features of the system such as symmetries or degeneracies. In such cases one may often be able to take account of such features and still obtain useful algorithms.

PAPER 6: LMBC CAMPOS

Question:

What exactly did you assume the pilot has control over? (groundspeet 3, pitch)

Answer:

The pilot had control over pitch, and used it to keep on a constant glide slope. Velocity (TAS or groundspeed) was a consequence of this.

Question:

You assumed that you started from a horizontal flight to a dive. You are in a transient stage when you enter the microburst, do you assume that the speed is steady or that it is still varying?

Answer:

There are two cases:
- if the length scale of aircraft stability is comparable to the microburst, we perform one kind of integration, taking this into account;
- the integration is simplified if the aircraft states changes on much longer scales than the microbursts.

Question:

When you enter a microburst, you distinguished the vertical and horizontal component. I could not evaluate which component gives maximum effect (it probably varies during the crossing of the microburst). Is it acceptable to compute separately the effect of the horizontal and the vertical components?

Answer:

The two effects can be separated, because the equations can be linearized with regard to the wind (if windspeed $\leq 1/3$ aircraft speed, then square of windspeed is negligible compared to square of aircraft speed). One can distinguish the effects of head-/tailwind from downflow. One can find an 'equivalence': the downflow which produces the same effect as a tailwind.

STABILITY OF VISCOELASTIC FLOW PHYSICAL AND NUMERICAL CONSIDERATIONS

by

M.J. Crochet and **Ch. Bodart**
Mécanique Appliquée
Université Catholique de Louvain
Place du Levant 2
B1348 Louvain la Neuve
Belgium

Summary

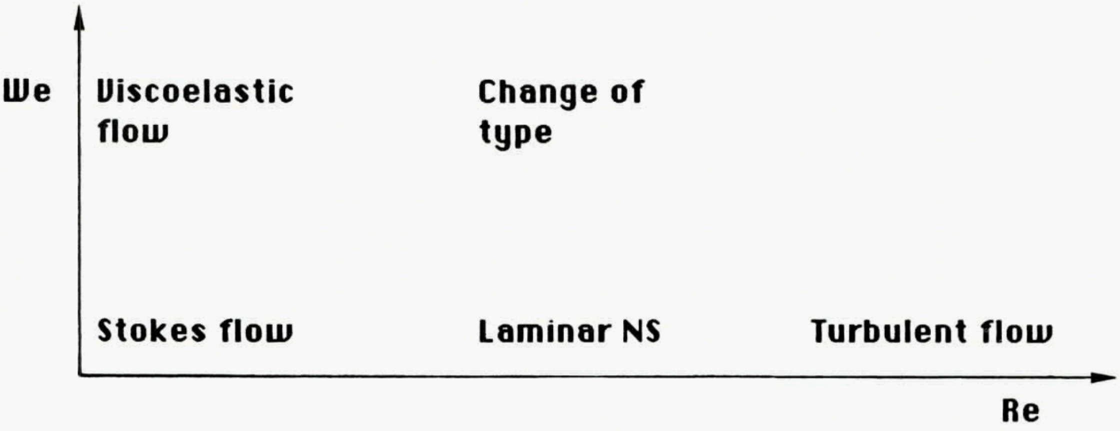
- Motivation
- Viscoelastic flow
- Numerical difficulties: the high Weissenberg number problem
- Efficient algorithms for smooth problems
- Efficient algorithms for singular problems
- Experimental evidence of instability
- Numerical procedure for verifying stability
- Numerical results
- Conclusions and outstanding problems

The notion of Weissenberg or Deborah number

Natural time of the fluid : λ

Characteristic time of the flow : $\tau = \dot{\gamma}^{-1}$, $\tau = L/V$

We or De = λ / τ, Re = $\rho\, V\, L / \eta$



We

Viscoelastic flow

Change of type

Stokes flow

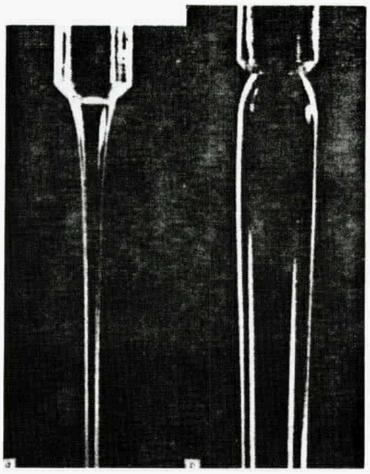
Laminar NS

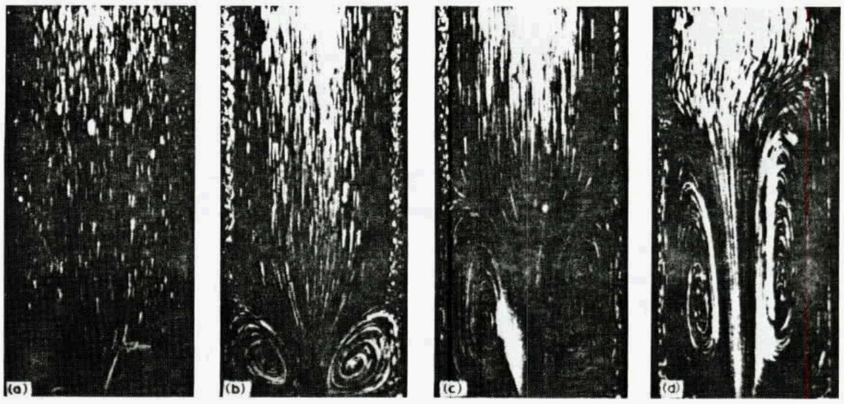
Turbulent flow

Re

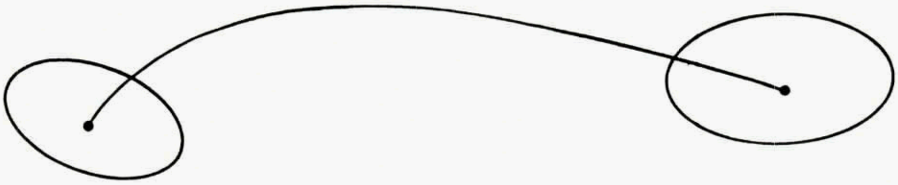
Viscoelastic effects

Rod climbing, extrudate swelling, drag reduction, generation of vortices, spurt, sharkskin and melt fracture, draw resonance...





Constitutive equations for viscoelastic fluids



General functional form: $\sigma = -p\,I + T; \quad T = \mathcal{T}\{\,C_t(t-s)\,;\, s>0\,\}$

Differential form: $T = \Sigma\, T_i\,; \quad g_i\, T_i + \lambda \overset{\square}{T_i} = 2\,\eta_i\, d$

Integral form: $T = \int_0^\infty m(s)\, h(I_1, I_2)\, C_t^{-1}(t-s)\, ds$

In addition: **incompressibility,** $\nabla . v = 0$

linear momentum, $-\nabla p + \nabla T + f = \rho a.$

Example : plane flow of a Maxwell fluid with one relaxation time set ; *implicit, non-linear* partial differential equations

$$Txx + \lambda(Txx,t + Txx,xu + Txx,yv - 2Txxu,x - 2Txyu,y) = 2\eta u,x$$

$$Tyy + \lambda(Tyy,t + Tyy,xu + Tyy,yv - 2Txyv,x - 2Tyyv,y) = 2\eta v,y$$

$$Txy + \lambda(Txy,t + Txy,yu + Txy,yv - Txxv,x - Tyyu,y) = \eta(u,y + v,x)$$

$$-p,x + Txx,x + Tyx,y + fx = \rho(u,t + u,xu + u,yv)$$

$$-p,y + Txy,x + Tyy,y + fy = \rho(v,t + v,xu + v,yv)$$

$$u,x + v,y = 0$$

The high Weissenberg number problem

Early developments: late 70's with	finite differences finite elements
Typical problems:	⇒flow of a Maxwell fluid through a four to one contraction ⇒extrudate swelling of a Maxwell fluid
Difficulties:	● lack of convergence beyond We ≅ 1 ● lack of convergence with mesh refinements ● limit points
Proposed reasons:	● true limit points, wrong fluid, numerical noise
True reason:	● numerical errors generate ill-posed problems
Cure:	● appropriate mixed f.e. representation (T, v, p) or other methods ● correct treatment of hyperbolic problem

Efficient algorithms for viscoelastic flow calculations

Mixed finite elements	4x4 sub-el,	streamline upwind (SU)
	4x4 sub-el,	streamline upwind Petrov-Galerkin (SUPG)
	EVSS	SU
	EVSS	SUPG

EEME (explicitely elliptic momentum equations)

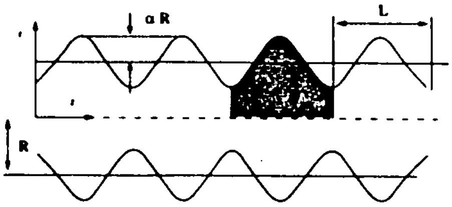
Spectral methods

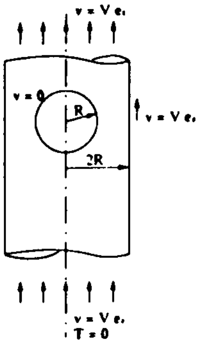
Efficiency:
- smooth problems
- problems with singularities

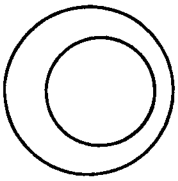
Note on SU :
- extremely stable at high We
- at best of order h !

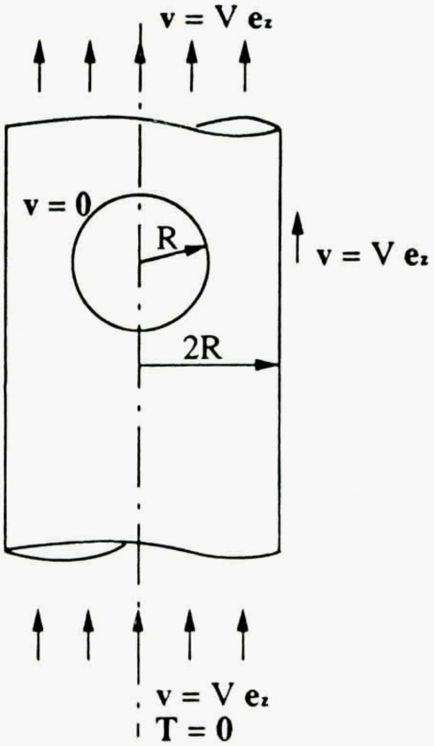
Smooth test problems

Wavy tube	Sphere in tube	Journal bearing
$fRe = 2\pi\, \delta P\, R^4 / L \eta Q$	$K = D / 6\pi\eta V R$	

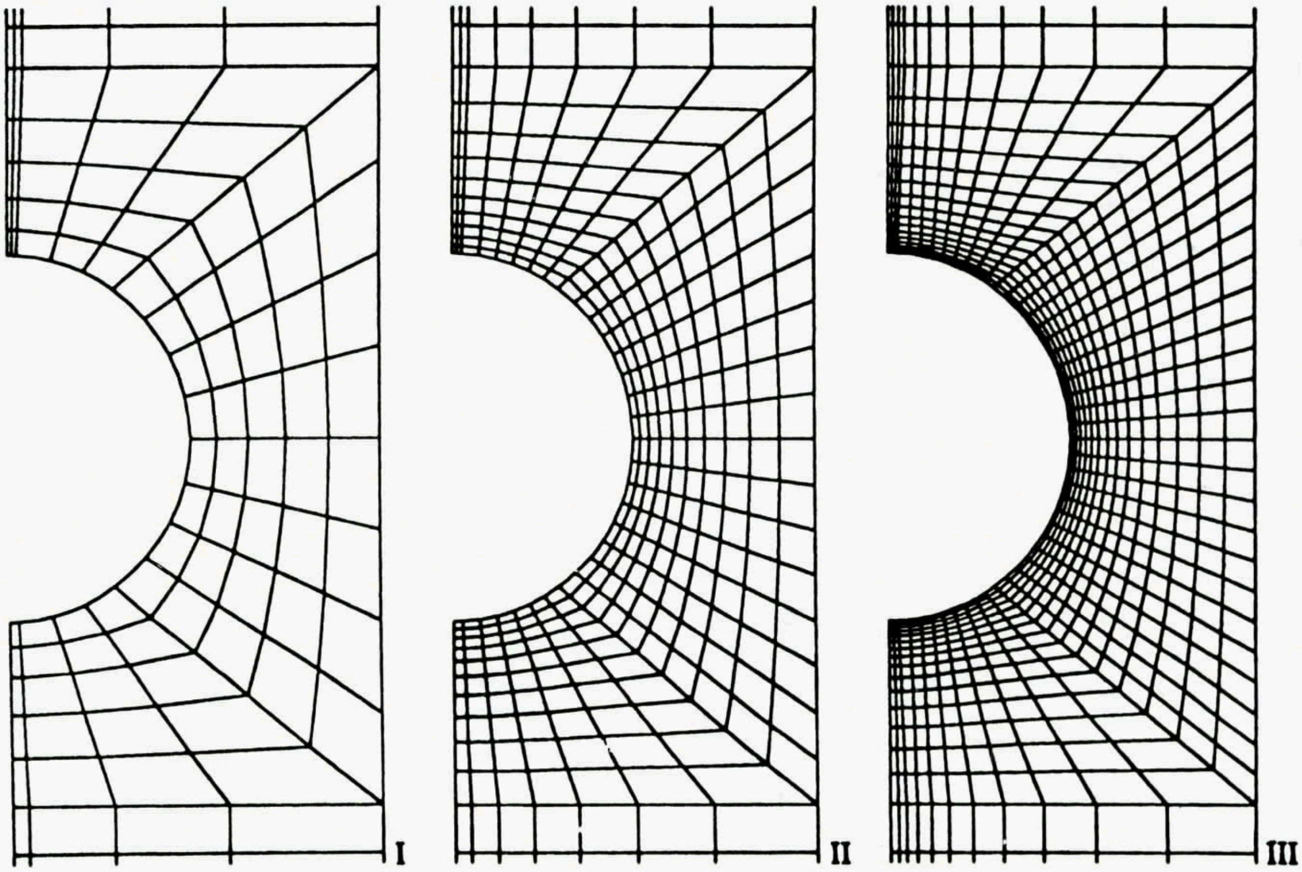








Geometry of the problem and boundary conditions.



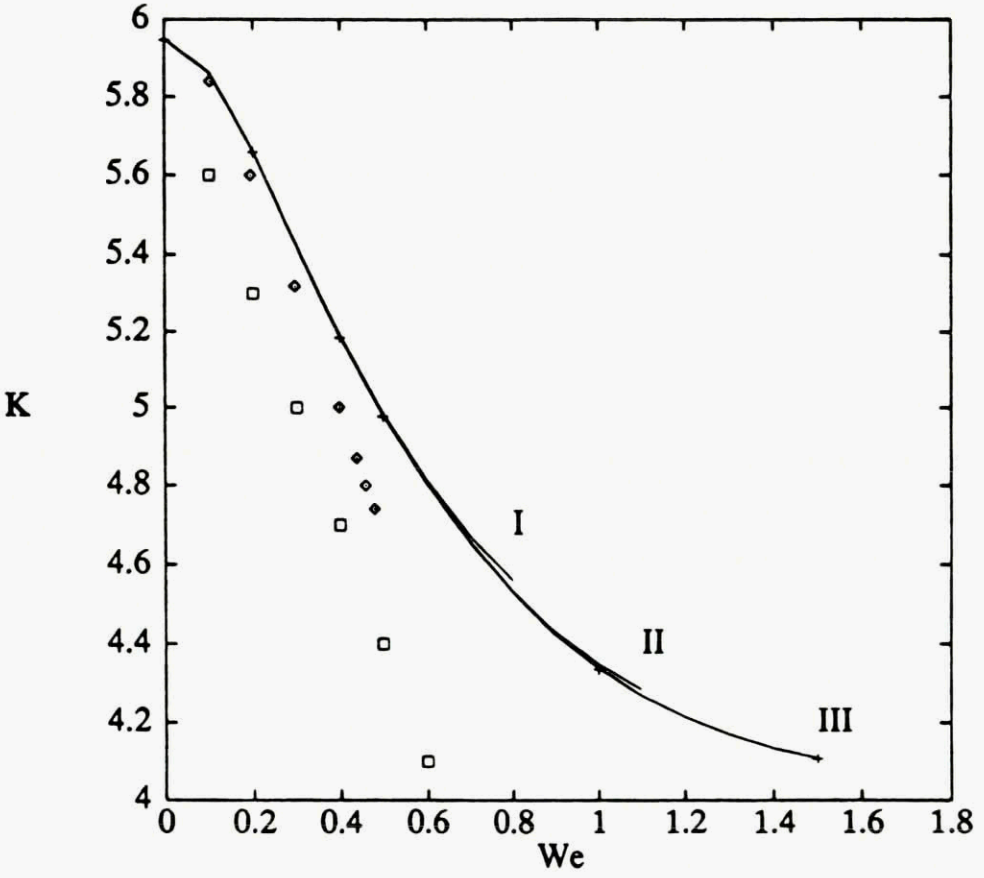
Central portion of finite element meshes used for solving the sphere problem.

Table I Characteristic dimensions of finite element meshes for the sphere problem

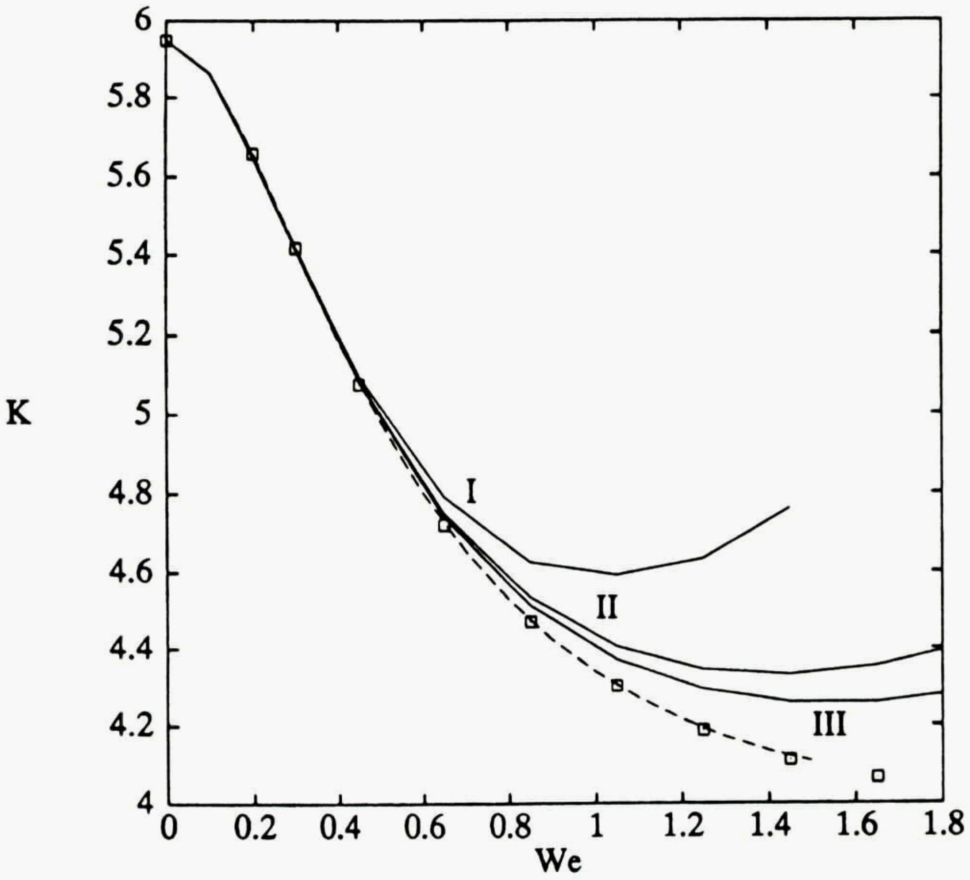
Mesh	Elements	Velocity nodes	Degrees of freedom
I	207	925	16126
II	510	2175	38644
III	1009	4213	75512

Table II. Drag correction factor for the flow of a Maxwell fluid around a sphere in a tube; SUPG 4x4 method.

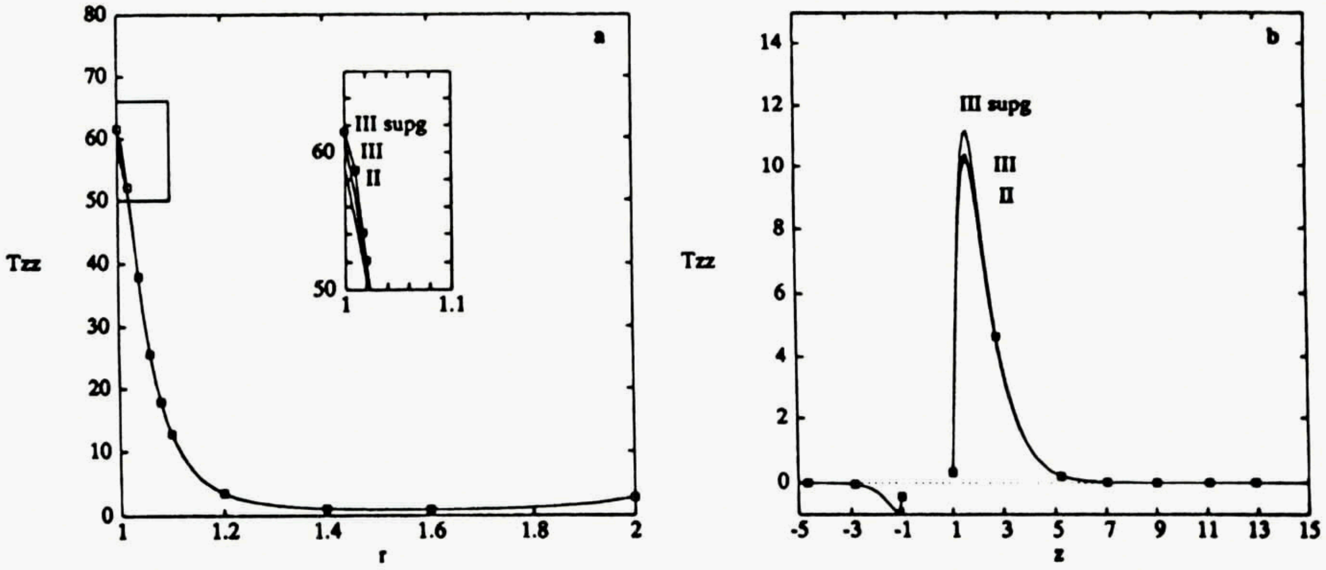
We	Mesh I	Mesh II	Mesh III	Ref [11]
0.	5.9490	5.9476	5.9475	5.9465
0.1	5.8638	5.8620	5.8621	
0.2	5.6624	5.6598	5.6597	5.6583
0.3	5.4235	5.4197	5.4195	
0.4	5.1920	5.1872	5.1868	5.1845
0.5	4.9867	4.9801	4.9795	4.9766
0.6	4.8125	4.8031	4.8021	
0.7	4.6697	4.6548	4.6533	
0.8	4.5593	4.5320	4.5299	
0.9	Diverges	4.4314	4.4244	
1.0		4.3502	4.3405	4.3364
1.1		4.2862	4.2697	
1.2		Diverges	4.2159	
1.3			4.1712	
1.4			4.1352	
1.5			4.1089	4.1072
			Diverges	Diverges



Drag correction factor for the flow around a sphere in a tube filled with a Maxwell fluid; plain lines : SUPG4x4; +:EEME [11]; ◊ : Zheng et al [14]; ▫ : Sugeng and Tanner [13].



Drag correction factor for the flow around a sphere in a tube; plain lines correspond to SU4x4 with the meshes of Fig.2; the dotted line to Mesh III with SUPG4x4; the symbols to the Richardson extrapolation.



Extra-stress component T_{zz} in the equatorial plane (a) and along the axis of symmetry (b) obtained with the SU4x4 method for meshes II and III; the symbols indicate SUPG4x4 results with Mesh III.

Problems with singularities

Examples:
- abrupt contractions, corners
- stick-slip flows
- extrudate swelling

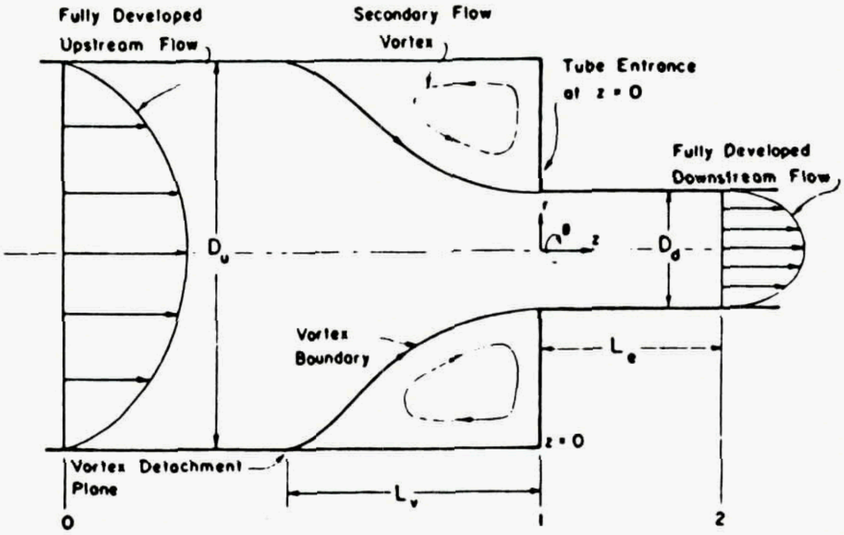
Essential difficulty:
- the nature of the singularity is not known
- stresses might be non-integrable

Present approaches:
- modify the fluid near the singularity
- modify the constitutive equations (MUCM)
- use a "robust" numerical method
  (which may also locally modify the fluid)

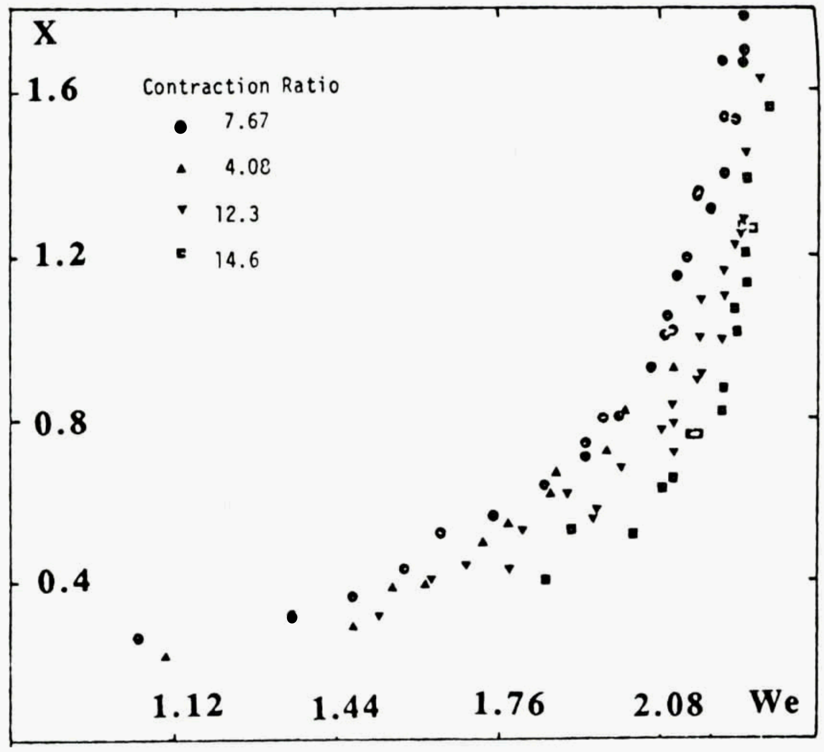
Example of problem with singularity:

Boger fluid through a four-to-one contraction

Vortex size: $X = L_v / D_u$, Weissenberg number: $We = \lambda \dot{\gamma}_w$

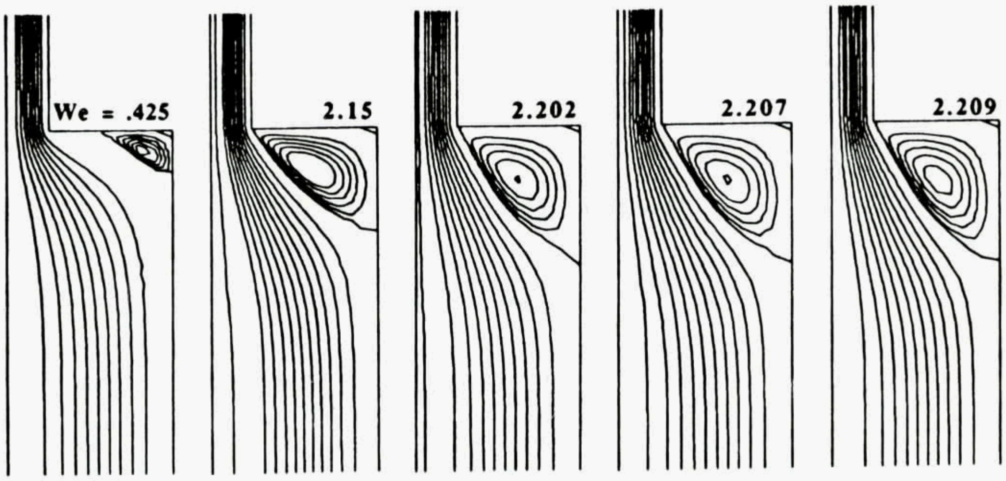


Geometry of the flow through an abrupt circular contraction

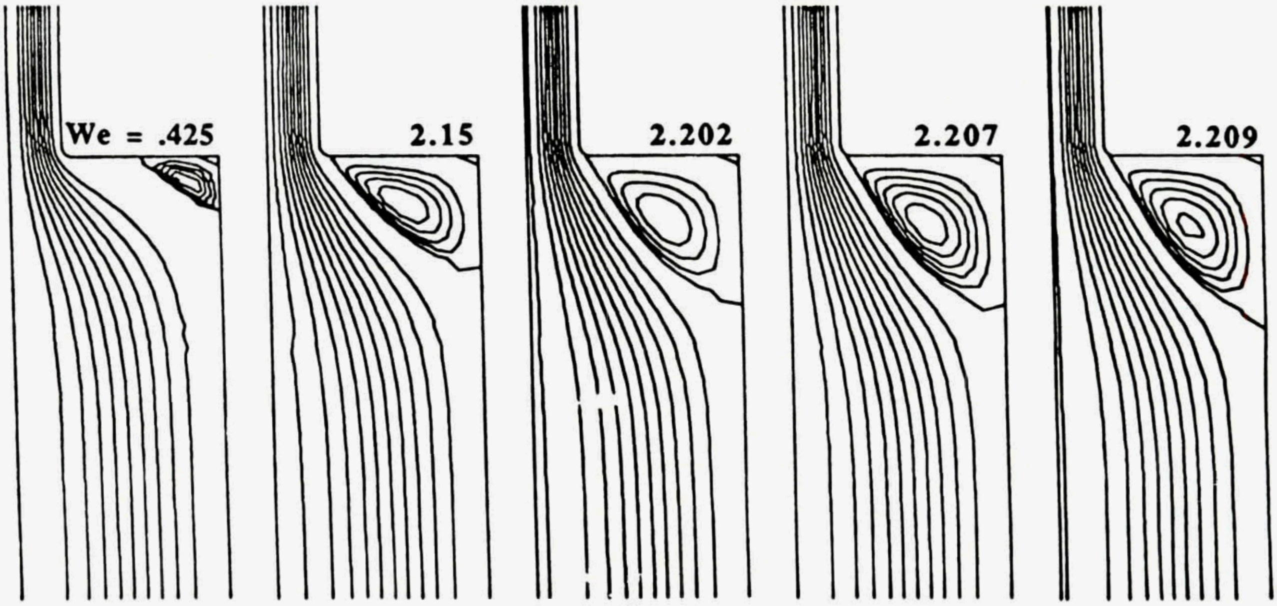


Dimensionless vortex length X as a function of We for fluid E1.

Other approach: White-Metzner fluid for the numerical simulation



The Oldroyd-B and the White-Metzner fluid
show a "saturation" of the viscoelastic effects at high values of We:
the vortex stops growing
(note: independent of corner and numerical method)



Outstanding question

We obtain steady state solutions. Are we on stable branches ?

Experimental evidence:

- **measured and calculated drag for Boger fluid (extensional viscosity ?)**
- **melt fracture (experiments by Piau)**
- **periodic flow through 4:1 contraction**
- **LDV observations by Lawler et al.**

Journal of Non-Newtonian Fluid Mechanics, 20 (1986) 51–92
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

LASER DOPPLER VELOCIMETRY MEASUREMENTS OF VELOCITY FIELDS AND TRANSITIONS IN VISCOELASTIC FLUIDS

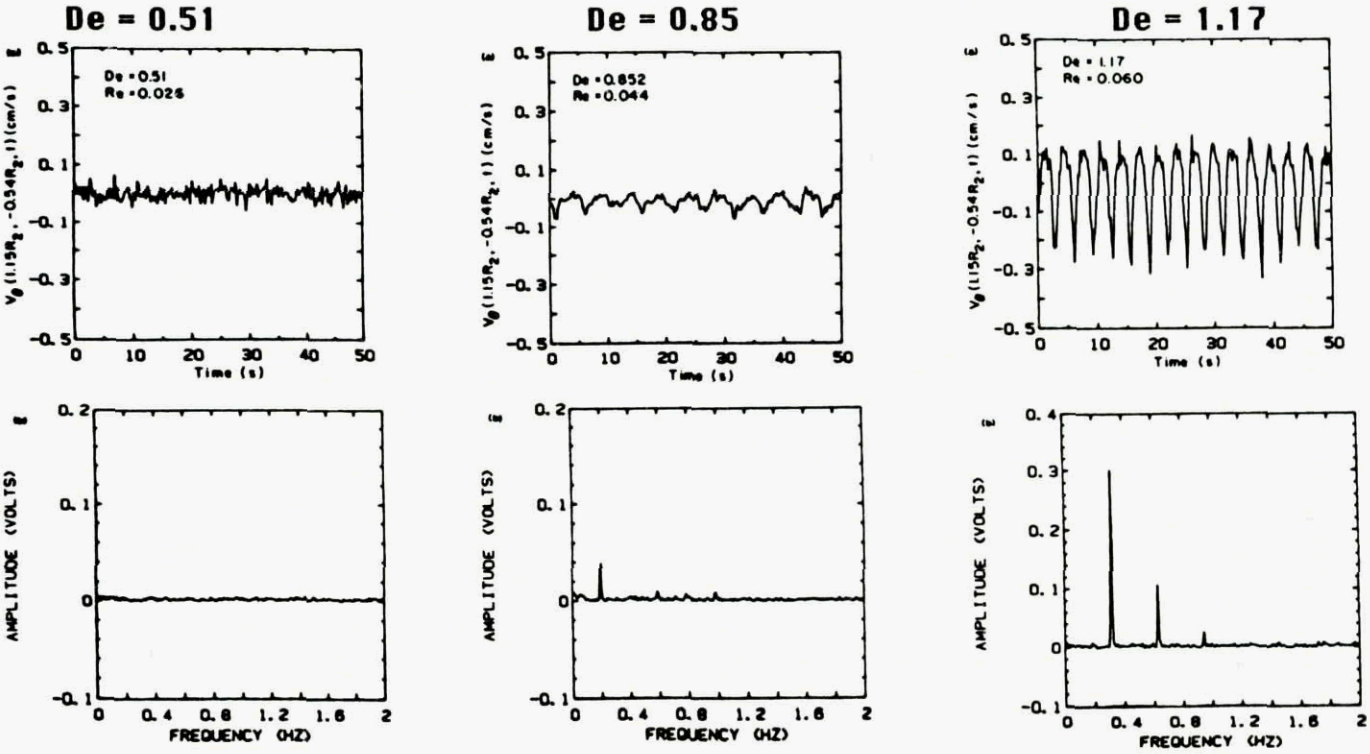
J.V. LAWLER, S.J. MULLER, R.A. BROWN and R.C. ARMSTRONG

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 (U.S.A.)

(Received October 9, 1985)

Observations by Lawler, Muller, Brown and Armstrong

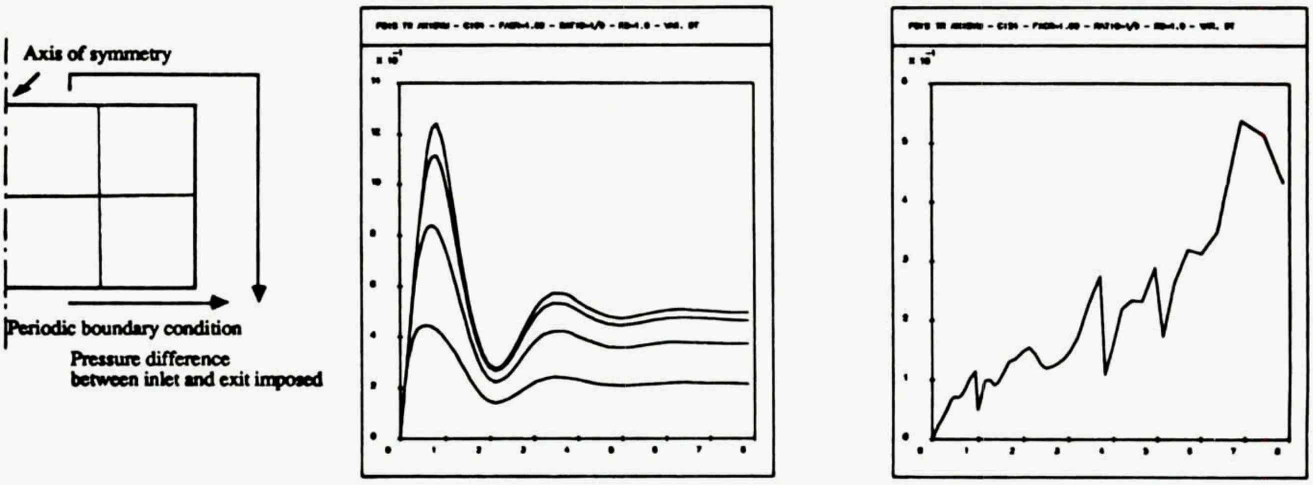
Time dependence of azimuthal velocity at one point near contraction



Numerical procedure

- **Fully coupled problem:** constitutive equations
  momentum equations
  incompressibility
- **Fully implicit time marching scheme**
- **Automatic calculation of time step based on predictor-corrector estimate**
- **Swirling flow included in axisymmetric description**
- **Test of accuracy: time-dependent Poiseuille flow of Maxwell fluid; comparison with analytical results**

Time-dependent Poiseuille flow at De = 1



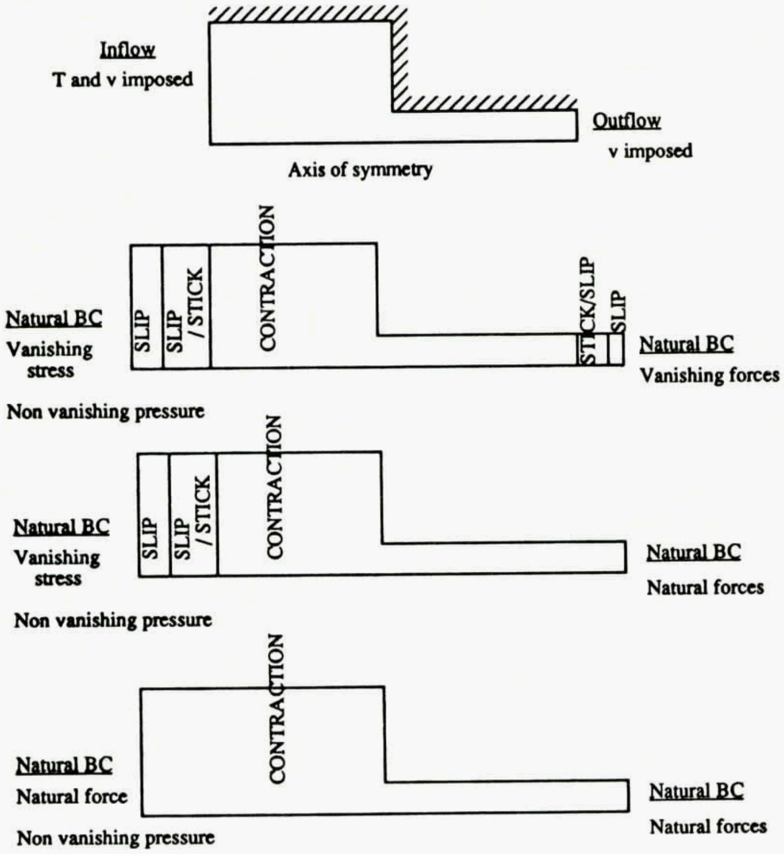
Geometry

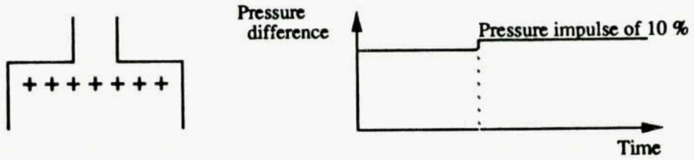
Axial velocity

Time step vs. time

Plane flow through a four to one contraction

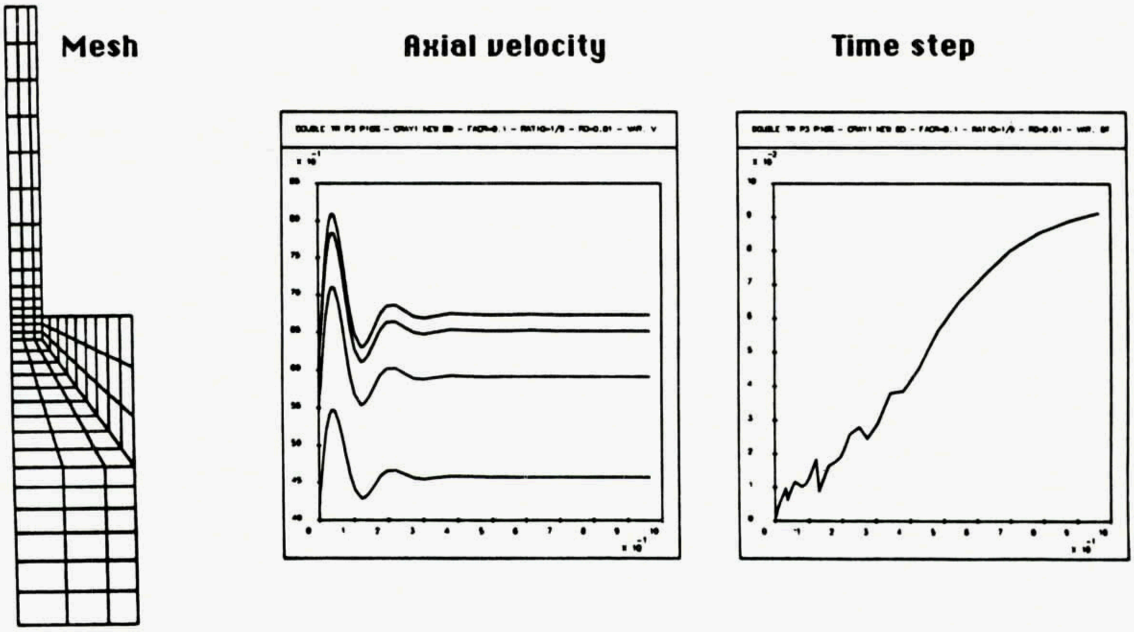
Steady state :





Plane flow through a four to one contraction

Galerkin method, De = 2



Mesh

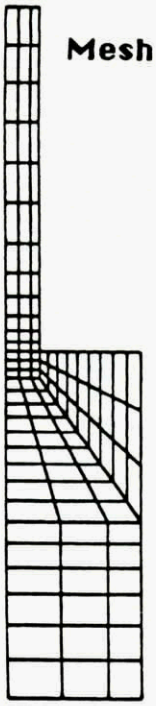
Axial velocity

Time step

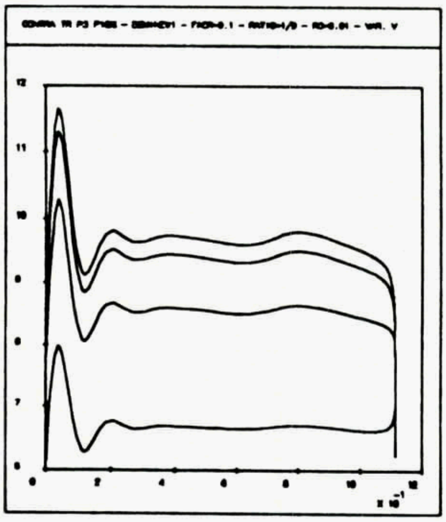
Plane flow through a four to one contraction

Galerkin method, De = 3

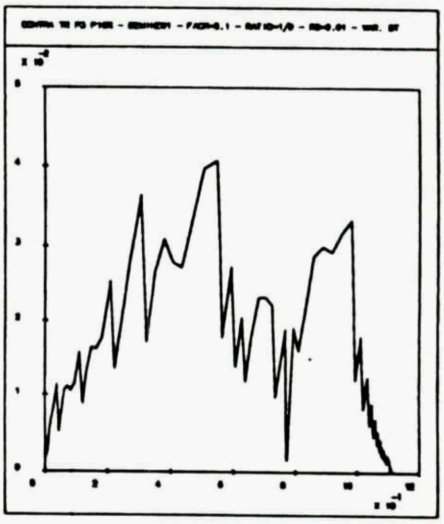
Mesh



Axial velocity



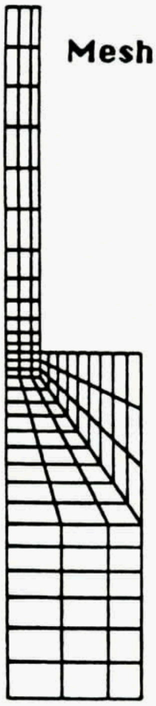
Time step



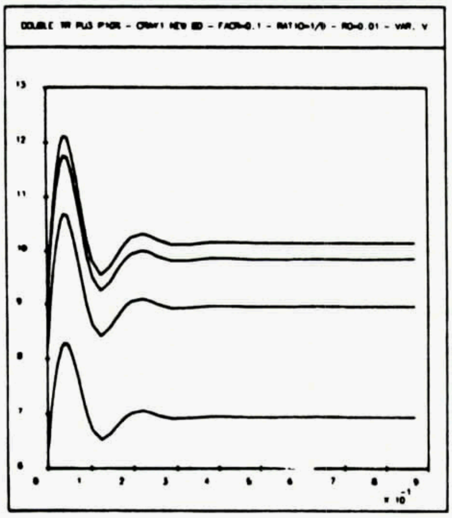
Plane flow through a four to one contraction

Streamline-upwind method, De = 3

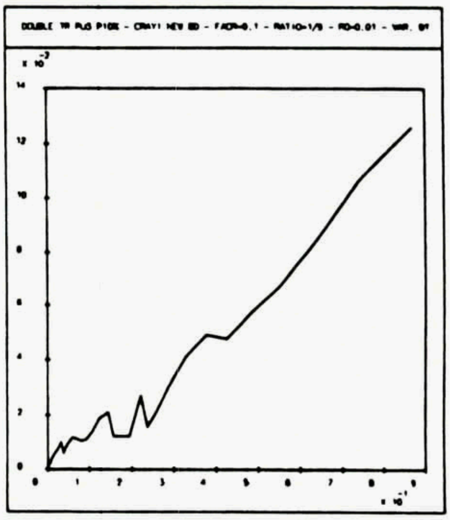
Mesh



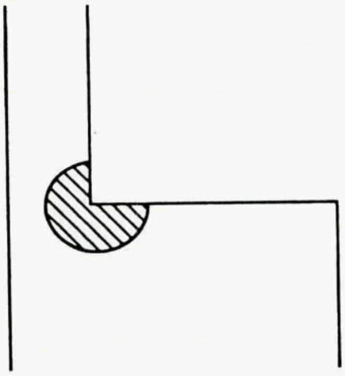
Axial velocity

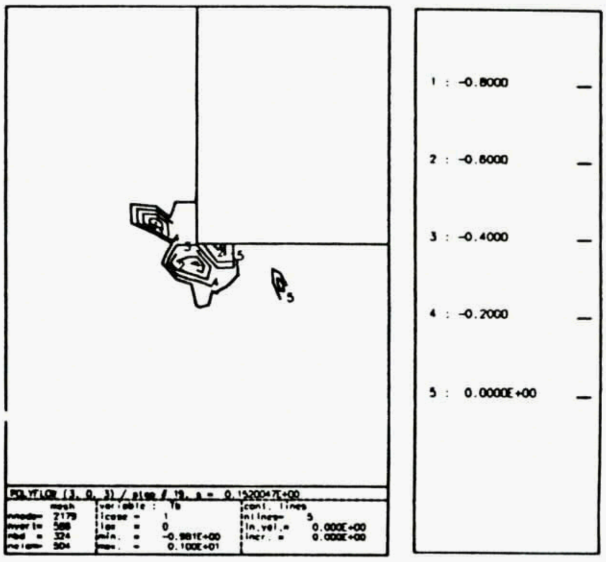


Time step



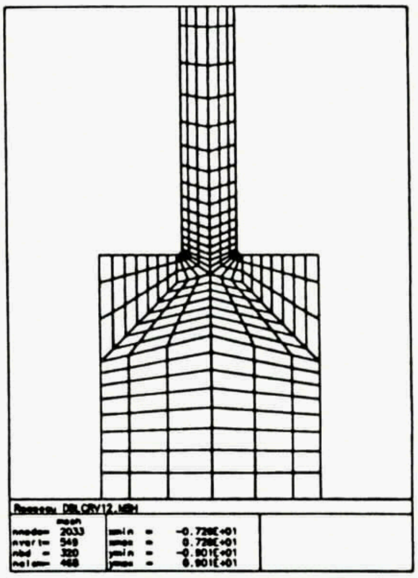
- **Previous results based on Galerkin method**
- **Not optimal for solving hyperbolic problems**
- **Are we calculating numerical errors ?**
- **Simple test: the tensor $T_A = T + \lambda/\eta\, I$ must be positive definite**
- **Results at De = 4 : lack of positive-definiteness**





Summary of numerical results

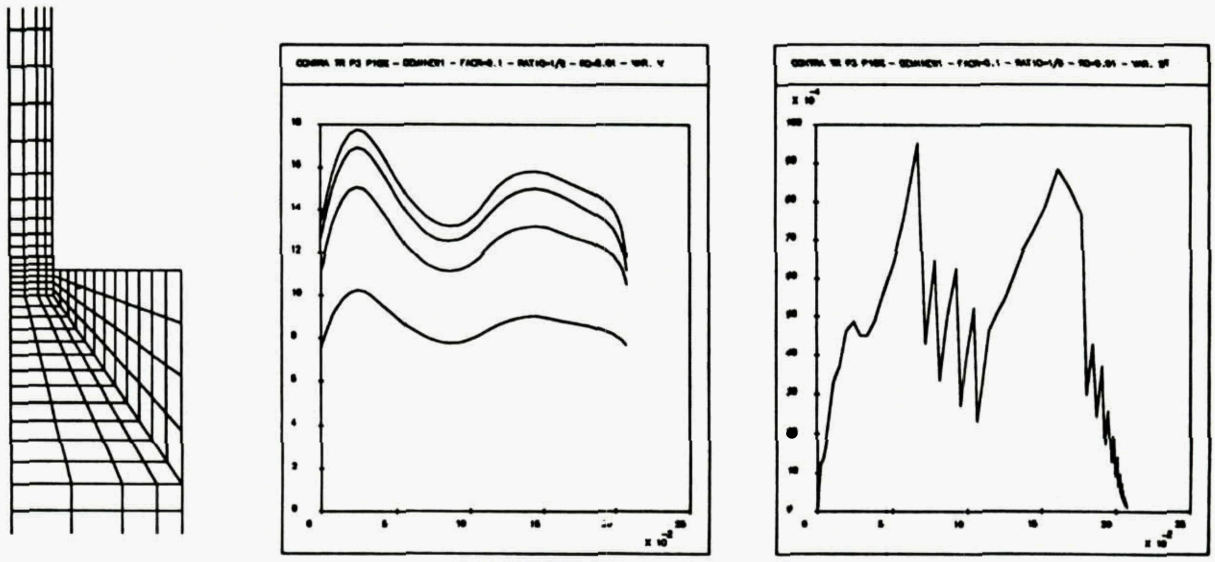
Numerical tests for the flow through an abrupt contraction						
Deborah number		1.0	2.0	3.0	4.0	5.0
Galerkin	MESH I			Stab $T_{A,B} < 0$	Stab $T_{A,B} < 0$	Stab $T_{A,B} < 0$
Galerkin	MESH II	Stab $T_{A,B} > 0$	Stab	± Unst $T_{A,B} < 0$	Unst $T_{A,B} < 0$	
Galerkin	MESH III		Stab $T_{A,B} < 0$	Unst $T_{A,B} < 0$		
SU	MESH I			Stab $T_{A,B} > 0$		
SU	MESH II	Stab $T_{A,B} > 0$	Stab	Stab	Stab $T_{A,B} > 0$	Stab
SU	MESH III		Stab	Stab		



For a smooth contraction: all numerical results are stable

Axisymmetric contraction with swirling capability

Loss of stability at De = 7. No swirling component.



Mesh **Axial velocity** **Time step**

Conclusions

- The numerical simulation of viscoelastic flow is well understood
- Converged results are available with various techniques for smooth problems
- Problems with singularities are much more difficult: the nature of the singularity is unknown (non-integrable forces ?)
- Stability problems are essential. However, are we observing a numerical instability due to numerical error or to physical causes ?
- Same is true for the calculation of eigenvalues
- Outstanding problems for future years

DYNAMICS AND CONTROL OF COHERENT STRUCTURES IN THE TURBULENT WALL LAYER - AN OVERVIEW[*]

Gal Berkooz, Philip Holmes, John Lumley
Cornell University
Ithaca, NY 14853, USA

ABSTRACT

We expand the velocity field in the vicinity of the wall in empirical eigenfunctions obtained from experiment. Truncating our system, and using Galerkin projection, we obtain a closed set of non-linear ordinary differential equations with ten degrees of freedom. We find a rich dynamical behavior, including in particular a heteroclinic attracting orbit giving rise to intermittency. The intermittent jump from one attracting point to the other resembles in many respects the bursts observed in experiments. Specifically, the time between jumps, and the duration of the jumps, is approximately that observed in a burst; the jump begins with the formation of a narrowed and intensified updraft, like the ejection phase of a burst, and is followed by a gentle, diffuse downdraft, like the sweep phase of a burst. The magnitude of the Reynolds stress spike produced during a burst is limited by our truncation. The behavior is quite robust, much of it being due to the symmetries present (Aubry's group has examined dimensions up to 128 with persistence of the global behavior). We have examined eigenvalues and coefficients obtained from experiment, and from exact simulation, which differ in magnitude. Similar behavior is obtained in both cases; in the latter case, the heteroclinic orbits connect limit cycles instead of fixed points, corresponding to cross-stream waving of the streamwise rolls. The bifurcation diagram remains structurally similar, but somewhat distorted. The role of the pressure term is made clear - it triggers the intermittent jumps, which otherwise would occur at longer and longer intervals, as the system trajectory is attracted closer and closer to the heteroclinic cycle. The pressure term results in the jumps occurring at essentially random times, and the magnitude of the signal determines the average timing. Stretching of the wall region shows that the model is consistent with observations of polymer drag reduction. Change of the third order coefficients, corresponding to acceleration or deceleration of the mean flow, changes the heteroclinic cycles from attracting to repelling, increasing or decreasing the stability, in agreement with observations. The existence of fixed points is an artifact introduced by the projection; however, a decoupled model still displays the rich dynamics. Numerous assumptions made in Aubry *et al.* (1988) can now be proved exactly. Feeding back eigenfunctions with the proper phase can delay the bursting, (the heteroclinic jump to the other fixed point), decreasing the drag. It is also possible to speed up the bursting, increasing mixing to control separation. Our approach is optimal for short time tracking in control.

THE PROPER ORTHOGONAL DECOMPOSITION

Lumley (1967) proposed a method of identification of coherent structures in a random turbulent flow. This uses what Loève (1955) called the Proper Orthogonal Decomposition, and which is often called the Karhunen-Loève expansion. An advantage of the method is its objectivity and lack of bias. Given a realization of an inhomogeneous, energy integrable velocity field, it consists of projecting the random field on a candidate structure, and selecting the structure which maximizes the projection in quadratic mean. The calculus of variations reduces this problem to a Fredholm integral equation of the first kind whose symmetric kernel is the autocorrelation matrix. The properties of this integral equation are given by Hilbert Schmidt theory. There is a denumerable set of eigenfunctions (structures). The eigenfunctions form a complete orthogonal set, which means that the random field can be reconstructed. The coefficients are uncorrelated and their mean square values are the eigenvalues themselves. The Kernel can be expanded in a uniformly and absolutely convergent series of the eigenfunctions and the turbulent kinetic energy is the sum of the eigenvalues.

The most significant point of the decomposition is perhaps the fact that the convergence of the representation is optimally fast since the coefficients of the expansion have been maximized in a mean square sense. Berkooz *et al.* (1990) have shown that the n terms of this decomposition contain at least as much energy as n terms of any other decomposition.

Application of the Proper Orthogonal Decomposition to the Shear Flow of the Wall Region

The flow of interest here is three dimensional, approximately homogeneous in the streamwise direction (x_1) and spanwise direction (x_3), approximately stationary in time (t), inhomogeneous and of integrable energy in the normal direction (x_2). In the homogeneous directions the spectrum of the eigenvalues becomes continuous, and the eigenfunctions become Fourier modes, so that the proper orthogonal decomposition reduces to the harmonic orthogonal decomposition in those directions. See Lumley (1967, 1970, 1981) for more details.

We want a three dimensional decomposition which can be substituted in the Navier-Stokes equations in order to recover the phase information carried by the coefficients. We measure the two velocities at the same time and determine $<u_i(x_1,x_2,x_3,t)\, u_j(x'_1,x'_2,x'_3,t)> = R_{ij}$. From R_{ij} we will determine the eigenfunctions. Since the flow is quasistationary, R_{ij} does not depend on time, nor do the eigenvalues and eigenfunctions. The information in time is carried by the coefficients $a^{(n)}$ which are still "stochastic", but now evolve under the constraint of the equations of motion. We also change the Fourier integral into a Fourier series, assuming that the flow is periodic in the x_1 and x_3 directions. The periods L_1, L_3 are determined by the first non-zero wave numbers chosen. Finally, each component of the velocity field can be expanded as the triple sum

$$(1)\quad u_i(x_3,x_2,x_3,t) =$$

$$\frac{1}{\sqrt{L_1 L_3}} \sum_{k_1 k_3 n} e^{2\pi i(k_1 x_1 + k_3 x_3)} a^{(n)}_{k_1 k_3}(t)\phi_{i\,(k_1 k_3)}^{(n)}$$

where

$$(2)\qquad \int \Phi_{ij}(x_2, x'_2)\phi_j^{(n)}(x'_2)dx'_2 = \lambda^{(n)}\phi_i^{(n)}(x_2),$$

and we have to solve equation (2) for each pair of wave numbers (k_1,k_3). Φ_{ij} now denotes the Fourier transform of R_{ij} in the x_1, x_3 directions.

EXPERIMENTAL RESULTS

The candidate flow we are investigating is the wall region (which reaches $x_2^+ = 40$; x_2^+ is the distance from the wall normalized by kinematic viscosity and friction velocity) of a pipe flow with almost pure glycerine (98%) as the working fluid Herzog (1986). From this data the autocorrelation tensor R_{ij} was obtained and the spatial eigenfunctions were extracted by numerical solution of the eigenvalue problem. The results show that approximately 60% of the total kinetic energy is contained in the first eigenmode (figure 1) and that the first three eigenmodes capture essentially the entire flow field as far as these statistics are concerned.

THE DYNAMICAL EQUATIONS

We decompose the velocity—or the pressure—into the mean (defined using a spatial average) and fluctuation in the usual way. We substitute this decomposition into the Navier-Stokes equations. Taking the spatial average of these equations we obtain, in the quasi stationary

---

[*] Prepared for presentation at NATO/AGARD Workshop on *Stability in Aerospace Systems*: Toulouse, France 23-26 June 1992. Supported in part by: the U. S. Air Force Office of Scientific Research, The U. S. Office of Naval Research (Mechanics Branch and Physical Oceanography Program), The U. S. National Science Foundation (programs in Applied Mathematics, Fluid Mechanics, Meteorology and Mechanics, Structures & Materials) and the NASA Langley Research Center.

case, an approximate relation between the divergence of the Reynolds stress and the mean pressure and velocity.

$$\text{(3)} \quad \langle u_{i,j} u_j \rangle = - 1/\rho \, P_{,i} + \nu \, U_{,jj} \delta_{i1}.$$

(where $u_{i,j}$ indicates the derivative with respect to x_j of u_i, and similarly for the other terms; repeated indices are summed). Equation (4) may be solved to give the mean velocity in terms of the Reynolds stress in a parallel flow. This reduces the slope of the mean velocity as the structures become stronger, stabilizing the system. (This depends on the sign of the Reynolds stress, which is certainly positive for the first structure, though not necessarily for the higher modes).

After taking the Fourier transform of the Navier Stokes equations and introducing the truncated expansion, we apply Galerkin projection by multiplying the equations by each successive eigenfunction in turn, and integrating over the domain.

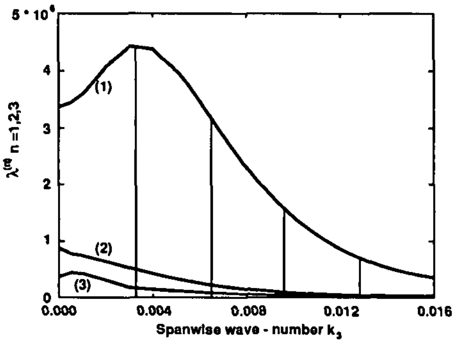


Figure 1. Convergence of the proper orthogonal decomposition in the near-wall region ($x_2^+ = 40$) of a pipe flow according to experimental data. Turbulent kinetic energy in the first three eigenmodes. $\lambda^{(n)}$ (n = 1, 2, 3) function of the spanwise wavenumber (from Herzog, 1986).

By use of the continuity equation and the boundary conditions (vanishing of the normal component at the wall, and at infinity) it can be seen by integration by parts that the pressure term would disappear if the domain of integration covered the entire flow volume. Since this is not the case (rather the domain is limited to $X_2^+ = 40$, where X_2^+ indicates the value of x_2^+ at the upper edge of the integration domain), there remains the value of the pressure term at X_2^+, which represents an external perturbation coming from the outer flow.

ENERGY TRANSFER MODEL.

The exact form of the equations obtained from the decomposition, truncated at some cut-off point (k_{1c}, k_{3c}, n_c), does not account for the energy transfer between the resolved (included) modes and the unresolved smaller scales. The influence of the missing scales will be parameterized by a simple generalization of the Heisenberg spectral model in homogeneous turbulence. Such a model is fairly crude, but we feel that its details will have little influence on the behavior of the energy-containing scales, just as the details of a sub-grid scale model have relatively little influence on the behavior of the resolved scales in a large eddy simulation. This is a sort of St. Venant's principle, admittedly unproved here, but amply demonstrated experimentally by the universal nature of the energy containing scales in turbulence in diverse media having different fine structures and dissipation mechanisms (see Lumley (1972) for a fuller discussion). The only important parameter is the amount of energy absorbed.

We will refer to α_1 as a Heisenberg parameter. We will adjust α_1 upward and downward to simulate greater and smaller energy loss to the unresolved modes, corresponding to the presence of a greater or smaller intensity of smaller scale turbulence in the neighborhood of the wall. This might correspond, for example, to the environment just before or just after a bursting event, which produces a large burst of small scale turbulence, which is then diffused to the outer part of the layer.

A term representing the energy fluctuation in the unresolved field due to the resolved field appears in the equation for the resolved field, and can be combined with the pressure term. We assume that the deviation (on the resolved scale) in the kinetic energy of the unresolved scales is proportional to the rate of loss of energy by the resolved scales to the unresolved scales. This term gives some quadratic feed-back. For generality we call this parameter α_2, although in all work presented in this paper, we have set $\alpha_1 = \alpha_2$.

Thus the Heisenberg model introduces two parameters in the system of equations, one, α_1, in the linear term, the other one, α_2, in the quadratic term. The equations therefore have the following form:

$$\text{(4)} \quad d a_{k_1 k_3}^{(n)} / dt = L + (\nu + \alpha_1 \nu_T) L' + Q + \alpha_2 Q' + C$$

where L and L' represent the linear terms, Q the direct quadratic terms, Q' the quadratic pseudo-pressure term and C the cubic terms arising from the Reynolds stress.

IMPLICATIONS FOR THE FLOW IN THE WALL REGION

Numerical integrations of 3, 4, 5 and 6 mode models have been carried out, but we shall only report in detail on the 6 mode (5 active mode) simulations here. Note that the (0, 0) mode is uncoupled and inactive.

There is a rich dynamical behavior, but we focus here on the behavior for $1.37 < \alpha < 1.61$, when a family of globally attracting double homoclinic cycles G exists, connecting pairs of saddle points which are π out of phase with respect to their second (x_2, y_2) components. The system spirals away from one saddle (the laminar phase) until it is far enough to leap to the other, and then repeats the process, to return to the first. The existence of the cycles G implies that, after a relatively brief and possibly chaotic transient, almost all solutions enter a tubular neighborhood of G and thereafter follow it more and more closely. As they approach G, the duration of the "laminar" phase of behavior increases while the bursts remain short. In an ideal, unperturbed system, the laminar duration would grow without bound, but small numerical perturbations, such as truncation errors, prevent this occurring in our numerical simulations. More significantly, the pressure perturbation will limit the growth of the laminar periods. Thus there is an effective maximum duration of events, which is reduced as α is decreased from the critical value $\alpha_b \sim 1.61$.

In Figure 2 we show the time histories of the modal coefficients for $\alpha = 1.45$. A description of the motion of the eddies during a burst is given in Figure 3 for $\alpha = 1.4$ by plotting u_2 and u_3 at 14 different times during one of the transitions shown in figure 2. Before and after the event, two pairs of streamwise vortices are present in the periodic box. However, pictures 1 and 14 are shifted in the spanwise direction by π. Moreover it is possible to adjust the value of the Heisenburg parameter ($\alpha \sim 1.5$) so that the bursting period is 100 wall units as experimentally observed (Kline *et al.*, 1967). It is found that, in this case, the "burst" lasts 10 wall units which is also the right order of magnitude. During one of these events there is a sudden increase in Reynolds stress, though smaller than observed. An event consists of a sudden intensification and sharpening of the updraft between eddies (5, 6 & 7, fig.3), followed by a drawing apart of the eddies, and the establishment of a gentle downdraft between them (9, 10 & 11, fig. 3): these are similar respectively to the ejection and sweep events that are observed.

PHYSICAL INTERPRETATION

Keith Moffatt points out that non-trivial solutions to the Navier Stokes equations, having no streamwise variation and driven by a streamwise (mean) velocity dependent only upon distance from the wall, should ultimately decay. This is easily seen from a simplified model with a single cross-stream Fourier mode for each velocity component and a fixed linear mean velocity profile. The streamwise velocity component (u_1) is fed from the mean velocity gradient by the component normal to the wall (u_2). However, neither u_2 nor u_3 has a source of energy. Both u_2 and u_3 decay exponentially from their initial values, with u_1 at first rising, but ultimately decaying exponentially also. The ratio of the Reynolds stress to the energy at first rises, but ultimately decays to zero algebraically.

In our ten-dimensional model, however, the ratio of Reynolds stress to energy does not decay, but is bounded away from zero, as is easily proved (Berkooz *et al.*, 1990), providing the energy source which makes the non-trivial fixed points and heteroclinic cycles possible. Berkooz has also shown (op cit) that, since the contributions of some of the higher modes to the Reynolds stress are of opposite sign to that of the first mode, the Reynolds stress for higher-order approximations will not be bounded away from zero. Thus we expect an "accurate" model lacking streamwise variations, but including many spanwise modes and several eigenfunctions, to exhibit the appropriate decay properties, the trivial solution u = 0 being a stable fixed point.

The proximal cause for the non-zero Reynolds stress/energy ratio when only the first eigenfunction is included, therefore, is the fact that the vector eigenfunctions have scalar coefficients. Hence, the u_1 and u_2 components in each mode are held in a non-evolving ratio. The eddies which occur in the real boundary layer, of course, have streamwise variation, and temporal variation. They each go through a life cycle, growing to a maximum and decaying. Only in a statistical sense is the ensemble stationary. The stationary behavior of the model reflects the stationary behavior of the ensemble, rather than the non-stationary behavior of the members. The Reynolds stress of the model (relative to the energy) is endowed by the empirical eigenfunctions with the value measured in the real boundary layer. In this way the cross-stream velocity components can extract energy from the mean flow. Hence, the empirical eigenfunctions are, in a sense, a closure approximation that embodies the effects of streamwise structure and unsteadiness in the value of the Reynolds stress represented by the relative sizes of their components. In this sense the model only appears to belong to the subspace of fields without streamwise variation.

In the present context, the vital question is whether the complex and apparently physically significant dynamical behavior of the ten-dimensional model is an artifact of the projection, like the fixed points. Happily we can give strong assurance that this is not so. We have constructed a decoupled model (Berkooz *et al.*, 1990) in which the streamwise component and those normal to the streamwise direction have separate coefficients. Solutions of this model decay properly, as described in the first paragraph (figure 4). The Reynolds stress (relative to the energy) decays to zero. The "fixed points" now drift slowly toward the origin. They are still connected by "ghosts" of heteroclinic cycles, so that the same bursting phenomenon occurs, but the bursts are now modulated by the slow decay. The bursts only occur while the cross-stream components are non-zero. There is a relatively long period after the cross-stream components have decayed during which only the streamwise component remains, no bursting occurs, and the streamwise component decays slowly to zero. We feel that this is probably the explanation for the common observation that the sublayer consists primarily of "streaks" - the streamwise remnants of eddies whose cross-stream components have decayed. The fraction of time during which there is cross stream activity (u_2, u_3 and bursting) is relatively short, and most of the time the scene would be dominated by the streak left behind.

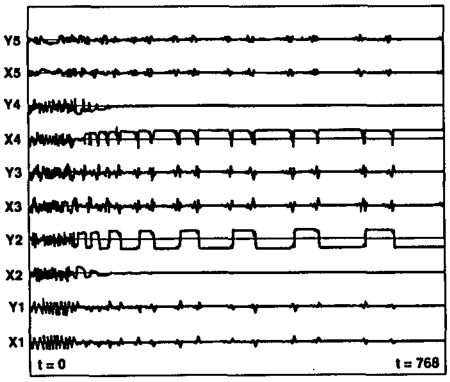


Figure 2. Time histories of the real (x_i) and imaginary (y_i) parts of the coefficients for a value of the Heisenberg parameter of $\alpha = 1.45$.

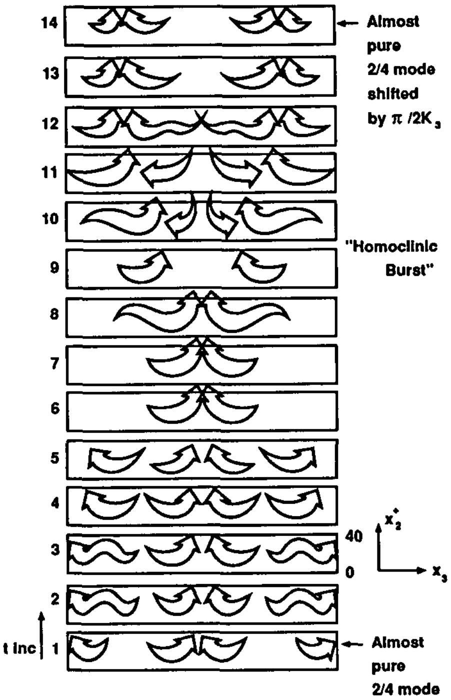


Figure 3. Intermittent solution (corresponding to an Heisenberg parameter a = 1.4) during a burst, times equally speced from the bottom. Each snapshot is a cross section of the flow (normal to the streamwise firection) from the wall (bottom) to $x_2^+ = 40$ (top), of width $Dx_3^+ = 333$.

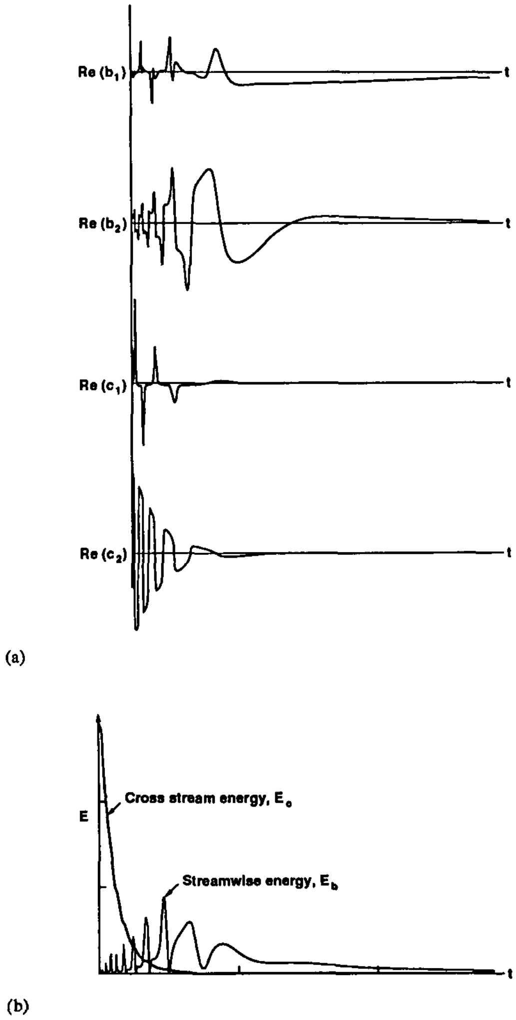


Figure 4. (a) Evolution of modal components for the uncoupled model: b_j are the streamwise and c_j the cross stream components respectively. (b) Evolution of the energy for the uncoupled model.

Holmes *et al.* (1990) have investigated in some depth the subspace of no streamwise variation. If we let P( ) be a projection operator, which is equivalent to a streamwise average, we can split the field into resolved modes $\mathbf{r} \in R$ and unresolved modes $s \in S$, so that $\mathbf{u} = \mathbf{r} + \mathbf{s}$, and $P(s) = 0$, $P(\mathbf{u}) = \mathbf{r}$, then Berkooz (Berkooz *et al.* 1990) has shown the correspondence P(SOLUTION OF NS) to the SOLUTION P(NS). That is, if we streamwise average the Navier Stokes equations, how does the solution of the averaged equations correspond to the streamwise average of the solution of the full equations. He has shown in addition: that the Leonard stresses (the cross stresses between the resolved and unresolved modes) vanish on the average; that the perturbation Reynolds stresses can only transfer energy from R to S; and that the energy loss from R to S can be represented by an eddy viscosity. Many of these were assumed in Aubry *et al.* (1988).

We have truncated in our ten-dimensional system the mechanism that represents the production of higher wavenumber energy when an intense updraft is formed, presumably as a result of a secondary instability. Thus, although our eddies are capable of exhibiting the basic bursting and ejection process, the labor is in vain. A contribution is made only to the low wavenumber part of the streamwise fluctuating velocity and the Reynolds stress. Recently however, Aubry & Sanghi (1989) have extended the model to include 1, 2 and 3

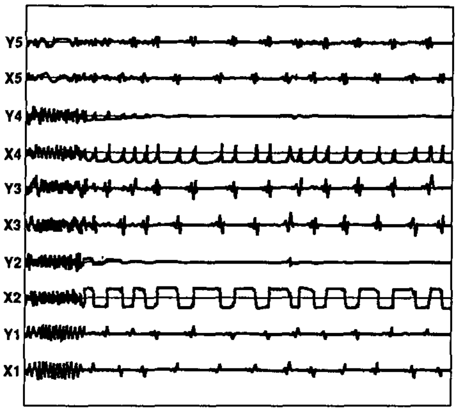


Figure 5. Similar to figure 2, but with the pressure term operative. Note that the inter-burst period is randomized, and on average, stabilized.

streamwise Fourier components, going up to studies of 38 complex (78 real) differential equations (very recently extended to 128 real). Addition of the streamwise components does not change the basic behavior of the system. For the most part the streamwise components are relatively quiescent; following a burst, however, they are excited, contributing to the Reynolds stress.

Initially we did not exercise the pressure term, which appeared due to the finite domain of integration. The order of magnitude that we estimated for this term was small, and for that reason we at first neglected it. It has, however, an important effect, while not changing the qualitative nature of the solution.

The term has the form of a random function of time, with a small amplitude. This slightly perturbs the solution trajectory constantly; away from the fixed points this has little effect, but when the solution trajectory is very close to these points, the perturbation has the effect of throwing the solution away from the fixed point, so that it need not wait long to spiral outward. This results in a thorough randomization of the transition time from one solution to the other, while having little effect on the structure of the solution during a burst. While in the absence of the pressure term (and round-off error), the interburst time tends to lengthen as the solution trajectory is attracted closer and closer to the heteroclinic cycle, with the pressure term, the mean time stabilizes.

One of the important findings of this work is the suggestion of the etiology of the bursting phenomenon. That is, presuming that the abrupt transitions from one fixed point to the other can be identified with a burst, these bursts appear to be produced autonomously by the wall region, but to be triggered by pressure signals from the outer layer. Whether the bursting period scales with inner or outer variables has been a controversy in the turbulence literature for a number of years. The matter has been obscured by the fact that the experimental evidence has been measured in boundary layers with fairly low Reynolds numbers lying in a narrow range, so that it is not really possible to distinguish between the two types of scaling. The turbulent polymer drag reduction literature is particularly instructive, however, since the sizes of the large eddies, and the bursting period, all change scale with the introduction of the polymer (Kubo & Lumley,1980; Lumley & Kubo, 1984). The present work indicates clearly that the wall region is capable of producing bursts autonomously, but the timing is determined by trigger signals from the outer layer. This suggests that events during a burst should scale unambiguously with wall variables. Time between bursts will have a more complex scaling, since it is dependent on the first occurrence of a *large* enough pressure signal *long* enough after a previous burst; "long enough" is determined by wall variables, but the pressure signal should scale with outer variables.

FURTHER CONSEQUENCES

We are, of course, concerned about the robustness of our findings. We have tried eigenfunctions generated from exact numerical simulations of channel flow, by Moser and Moin at the Center for Turbulence Research (Stanford/NASA Ames). These eigenfunctions are superficially similar to those from Herzog's data, but result in changes of the order of 20% in the values of the coefficients in the equations. The bifurcation diagram is similar, but the fixed points are replaced by limit cycles. Physically, this means that the eddies are wiggling from side to side instead of sitting still. This makes no essential difference, and is even more realistic physically. The intermittent behavior remains.

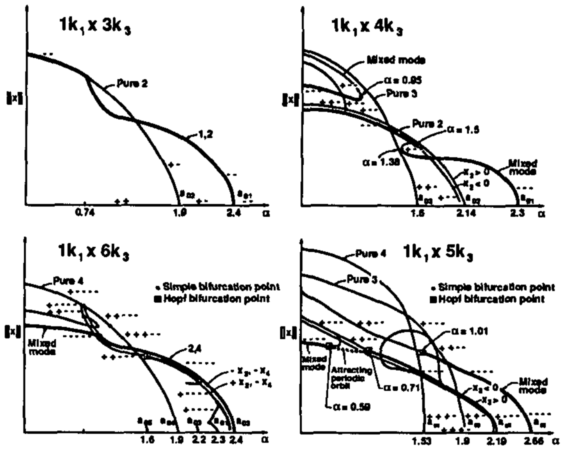


Figure 6. Bifurcation diagrams for models with varying numbers of cross-stream modes (from Stone, 1989). Note in the six-mode model that modes 3 and 5 have been suppressed for clarity. Note the similarity of the basic structure.

In her thesis, Stone (Stone & Holmes, 1990) investigated models with various numbers of cross-stream modes: 3, 4, 5 and 6. She found that the bifurcation diagrams had a backbone common to all of these systems, and were all structurally similar. In particular, the intermittent behavior was common to all. This is illustrated in figure 6.

In addition, Stone (Stone & Holmes, 1990) found that a small change in the value of the coefficients of the third order terms could change the heteroclinic cycles from attracting to repelling. This is illustrated in Figure 7 (lower), where one can see that the system begins on a traveling wave, but is gradually attracted to the heteroclinic cycle. In Figure 7 (upper) we see the opposite - the system starts on the heteroclinic cycle, but is repelled by it, and ends on a traveling wave. (We show only the values of the first two transverse Fourier modes - the others are quiescent). This would be a dynamical systems curiosity, if we could not relate it to the physics. However, if we consider the case $\partial_t U_1 = kU_1$ (an exponential increase or decrease of the mean velocity) we find that this results in a change in the real part of the cubic terms for $k_1 = 0$. When k changes sign the addition to the real part of the cubic term changes sign. This phenomenon is related to the destabilization and stabilization known to be induced by deceleration and acceleration of the flow (as by an adverse or favorable pressure gradient). Although we have discussed here the effect of temporal acceleration and deceleration, the same qualitative effect is obtained from a spatial acceleration and deceleration. Making the heteroclinic cycle more attractive would increase the time between bursts, stabilizing the flow, and vice versa.

Stone (Stone & Holmes, 1990) also predicted and measured histograms of the bursting period (Figure 8a). These look reasonably similar to measurements of the same by Kline *et al.* (1967), (Figure 8b).

Bloch and Marsden (1989) have shown that it is possible to stabilize this system by feedback, in the absence of noise. That is, if an eigenfunction is fed back with the proper phase, the system can be held in the vicinity of a fixed point for all time. In the presence of noise, however, (such as the pressure perturbation from the outer layer) the system cannot be stabilized completely; however, it can be held in a neighborhood of the fixed point for a longer time. When the system finally wanders so far from the fixed point as to make it uneconomical to recapture it, it is allowed to leave. The same procedure is carried out at the other fixed point. The effect is to increase the mean time between bursts, and hence to reduce the drag. Of course, the system can be made to work the other way, also, kicking the system away from the fixed point whenever it comes too close, resulting in a decrease in the mean time between bursts, and an increase in drag. This would be useful in avoiding separation or improving mixing, for instance. In recent work, Berkooz (1990) introduced the notion of short term tracking time, which measures the time over which a dynamical system model tracks the true dynamics accurately; for control, it must be of the order of the wall region time scales. Berkooz (1990) showed that dynamical systems based on the Proper Orthogonal Decomposition have, on the average, the best short term tracking time for a given number of modes.

In drag reduction by polymer additives, one of the accepted mechanisms (Kubo & Lumley, 1980; Lumley & Kubo, 1984) is the stabilization of the large eddies in the turbulent part of the flow, allowing the eddies to grow bigger and farther apart, as observed. Aubry *et al* (1989) tried stretching the eddy structure in the wall region, producing drag reduction, and found the bifurcation diagrams morphologically unchanged, except that the bifurcations occurred for larger and larger values of the Heisenberg parameter. This suggests that the motions giving rise to the bifurcations are more and more unstable, the more the region is stretched, requiring a larger and larger value of the Heisenberg parameter to stabilize them. Now, the Heisenberg parameter represents the loss of energy to the unresolved modes. However, the crudeness of the model is such, that it cannot distinguish between loss to the unresolved modes and loss to any other dissipation mechanism, such as viscosity or extensional viscosity. As far as the large scales are concerned, all losses are the same. Hence the findings of Aubry *et al* (1989) are completely consistent with the idea of the larger eddies being less stable, and able to grow to this larger, less stable size due to the stabilizing effect of the polymer.

Finally, Bloch and Marsden (1989) have shown that, within the assumptions of the scenario above, polymer drag reduction is equivalent to control of the wall region. That, is they showed that an increase of the Heisenberg parameter was equivalent to control by feeding back eigenfunctions, and would lead to a reduction in the bursting rate, and hence to a decrease in the drag. According to the scenario above, this would lead to a stabilization, and result in a growth of the eigenfunctions. This has the important consequence that a controlled boundary layer would be very similar to a polymer drag-reduced boundary layer. From experience with the polymer-drag-reduced boundary layer, we know that it would be a robust layer, still turbulent though with a reduced bursting rate, relatively insensitive to roughness and external disturbance. This is important from the standpoint of applications. Other drag reduction schemes connected with stabilization of the laminar layer are not robust in this sense, and are very sensitive to external disturbances and surface roughness.

From a practical point of view, the feed back could be implemented by an array of hot film sensors to detect the presence, location and strength of an eigenfunction. Piezoelectrically raised welts could produce a negative eigenfunction (see figure 9) by overturning the vorticity in the boundary layer.

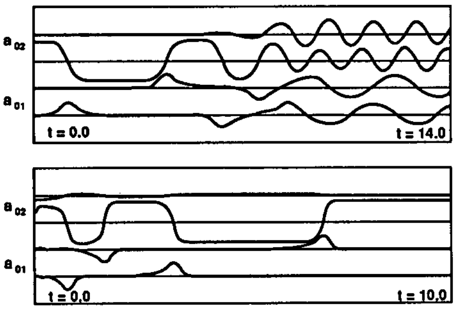


Figure 7. A model system with two streamwise modes (one active) and three cross-stream modes (two active). We show only the 01 and 02 modes (the others are unexcited). In the upper figure one of the third order coefficients is -2.69, and the heteroclinic cycle is repelling; in the lower, it is -3.00, and it is attracting (Stone, 1989).

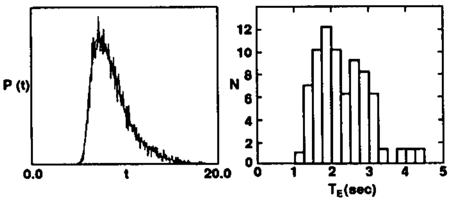


Figure 8. (a) The predicted and measured histogram of the bursting period from our model (Stone, 1989). (b) The measured histogram of the bursting period in the turbulent boundary layer, from Kline *et al.* (1967).

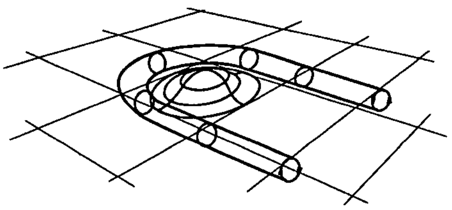


Figure 9. Schematic of a piezoelectrically produced welt on the surface, overturning the vorticity in the boundary layer and producing a negative eigenfunction.

BIBLIOGRAPHY

Armbruster, D., Guckenheimer, J. and Holmes, P. 1987. Heteroclinic cycles and modulated traveling waves in systems with O(2) symmetry. *Physica D* (to appear).

Aubry, N., Holmes, P., Lumley, J. L. and Stone, E. 1988. The dynamics of coherent structures in the wall region of a turbulent boundary layer. *J. Fluid Mech.* **192**: 115-173.

Aubry, N., Lumley and Holmes, P. , J. L. 1989 The effect of drag reduction on the wall region. *Theoretical and Computational Fluid Dynamics.* **1**, 229-248

Aubry, N. and Sanghi, S. 1989. Streamwise and cross-stream dynamics of the turbulent wall layer. Proceedings, July meeting of ASME, New York. ed Ghia.

Berkooz, G. (1990). In preparation.

Berkooz, G., Holmes, P. J. and Lumley, J. L. (1990). Intermittent dynamics in simple models of the turbulent wall layer. *J. Fluid Mech.* (submitted for publication)

Bloch, A. M. and Marsden, J. E. 1989. Controlling Homoclinic Orbits. *Theoretical and Computational Fluid Dynamics.* **1**(3):179-190.

Corino, E. R., and Brodkey, R.S. 1969. A visual investigation of the wall region in turbulent flow. *J. Fluid Mech.* **37**(1):1-30.

Golubitsky, M. & Guckenheimer, J. 1986. (eds). *Multiparameter Bifurcation Theory.* A.M.S. Contemporary Mathematics Series, No. 56. American Mathematical Society, Providence, R.I.

Herzog, S. 1986. *The large scale structure in the near-wall region of turbulent pipe flow.* Ph.D. thesis, Cornell University.

Holmes, P. J., Berkooz, G., and Lumley, J. L. 1990. Turbulence, dynamical systems and the unreasonable effectiveness of empirical eigenfunctions. *ICM '90 Proceedings*, Kyoto, Japan. In Press.

Kline, S.J., Reynolds, W.C., Schraub, F.A. and Rundstadler, P.W. 1967. The structure of turbulent boundary layers. *J. Fluid Mech.* 30(4): 741-773.

Kubo, I. and Lumley, J. L. 1980. *A study to assess the potential for using long chain polymers dissolved in water to study turbulence.* Annual Report, NASA-Ames Grant No. NSG-2382. Ithaca, NY: Cornell.

Loève, M. 1955. *Probability Theory*. New York: Van Nostrand

Lumley, J. L. 1971. Some Comments on the energy method. In *Developments in Mechanics 6*, eds. L. H. N. Lee and A. H. Szewczyk, pp. 63-88. Notre Dame IN: Notre Dame Press.

Lumley, J. L. and Kubo, I. 1984. Turbulent drag reduction by polymer additives: a survey. In *The Influence of Polymer Additives on Velocity and Temperature Fields.* IUTAM Symposium Essen 1984. Ed. B. Gampert. pp. 3-21. Berlin/Heidelberg: Springer.

Lumley, J.L. 1967. The structure of inhomogeneous turbulent flows. In *Atmospheric Turbulence and Radio Wave Propagation*, A.M. Yaglom and V.I. Tatarski:, eds.: 166-178. Moscow: Nauka.

Lumley, J.L. 1970. *Stochastic tools in turbulence.* Academic Press, New York.

Lumley, J.L. 1981. Coherent structures in turbulence. *Transition and Turbulence*, edited by R.E. Meyer, Academic Press, New York: 215-242.

Moffat, H. K. 1989 Fixed points of turbulent dynamical systems and suppression of non-linearity. In *Whither Turbulence*, ed. J. L. Lumley. Heidelberg: Springer. In press.

Moin, P. 1984. Probing turbulence via large eddy simulation. AIAA 22nd Aerospace Sciences Meeting.

Smith, C.R. and Schwarz, S.P. 1983. Observation of streamwise rotation in the near-wall region of a turbulent boundary layer. *Phys. Fluids* **26**(3) 641-652.

Stone, E. and Holmes, P. J. 1990. Random perturbations of heteroclinic attractors. *SIAM J. Appl. Math.* **50**: 726-743.

Tennekes, H. and Lumley, J.L. 1972. *A first course in turbulence.* Cambridge, MA: M.I.T. Press.

CHAOS, ENTROPY AND REVERSIBILITY: A SIMPLE EXAMPLE

C. Marchal
Office National d'Etudes et de Recherches Aérospatiales – D.E.S.
BP 72
92320 Châtillon, France

SUMMARY

The contradiction between mathematical reversibility and physical irreversibility has traditionally led to the suspicion of "hidden correlations" that will allow, in some cases, the decrease of the entropy of isolated systems

This idea seems wrong, as shown by a very simplified model.

The true reason of the physical irreversibility is certainly the very large number of parameters of irreversible systems.

LIST OF SYMBOLS

In this paper with many very large numbers we will use the notation "by figures and sizes" with the letter p for "positive power of ten" and the letter n for "negative power of ten".

Hence for instance:
6.02p23 = Avogadro number = 6.02×10^{23}
1.66n24 = inverse of Avogadro number $=1.66\times10^{-24}$.

Let us recall that the "figure", left to the p or n (*i.e.* here 6.02 and 1.66) is always between 1 and 10. This gives an unambiguous definition to the "size" (here p23 and n24). The size is the main element of very large and very small quantities, it is even very often their only known element.

INTRODUCTION

The physical laws have reversible mathematical expressions while the physical phenomena are essentially governed by the irreversibility of the second principle of thermodynamics...

A CONCRETE EXAMPLE

Let us open the communication between two neighouring closed vessels full of gas. The brownian motion will egalize the temperatures, the pressures and the compositions while the opposite evolution never appears.

However:

A) The brownian motion and the kinetic theory of gas are conservative and reversible.

B) Henri Poincaré has demonstrated that for bounded and conservative systems almost all initial conditions lead to an infinite number of returns in the vicinity of these initial conditions (the mathematicians specify: "in any vicinity of the initial conditions").

These returns to the vicinity of initial conditions are of course contradictory with the equalization of temperatures, pressures and compositions.

CLASSICAL AND UNSATISFACTORY ANSWERS

A) "There exist perhaps some very small, irreversible and dissipative hidden phenomena that forbid the application of Poincaré results...".

This rejection of a major symmetry of nature is not justified and our present knowledges are sufficient for the resolution of the observed contradiction.

B) "In principle Henri Poincaré is right and for strictly isolated system there is indeed this mysterious correlation between initial and final conditions (after the return time of Poincaré). But our systems are not isolated and even very small perturbations, such as the attraction of planet Pluto, destroy this correlation...".

These "mysterious correlations" are imaginary and it is in a natural fashion that the system returns towards all attainable states from the given initial conditions. The invoked "very small perturbations" have nothing to do there and will not modify the order of magnitude of Poincaré return time, even if it is true that they can modify very much the evolution in a relatively short interval of time (the "Liapounov time") and thus contribute to the disparition of correlations.

THE TRUE ANSWER

There are several other classical answers, all of them unsatisfactory, but the true answer is unexpected: It is because a system is "sensible to initial conditions" and because it depends on billions of parameters, while we measure only a few of them, that we ascertain almost no correlation between successive states at large intervals and that the Poincaré return time is very large, much larger than the age of Universe.

We thus reach the practical irreversibility of our experiments in spite of reversible physical laws.

The following, extremely simplified model with only didactic purposes, will help to understand this answer.

Notice that this model satisfies practically the Boltzmann hypothesis of "molecular chaos" (no correlation between successive variations) but reaches opposite conclusions: the molecular chaos doesn't forbid the Poincaré return.

A SIMPLIFIED MODEL

Let us consider one billion billion of molecules (that is p18 molecules with the notations of the above list of symbols). This number is the number of molecules in 37 mm^3 of air in "normal conditions", which is a very small volume, and in most experiments the effects will be even greater.

These p18 molecules will be put in the two identical vessels A and B and will be numbered from 1 to p18.

The evolution of our simplified model will be the following: at each step an integer number between 1 and p18 will be will be chosen and the corresponding molecule will be transferred from its present vessel to the other one.

We will assume that the rate of these exchanges will be one million billion (that is p15) per second.

We will measure only the number of molecules in the two vessels A and B, *i.e.* the local pressure. We will for instance start with 70% of molecules in A and 30% in B.

With an average pressure of one bar the corresponding initial pressures will be:

$$(1) \qquad P_A(0)=1.4 \text{ bar}; \qquad P_B(0)=0.6 \text{ bar}$$

What will be the evolution?

EVOLUTION OF THE PRESSURES P_A and P_B

This problem has an obvious integral of motion, the total number of molecules is constant and thus at any time:

$$(2) \qquad P_A(t)+P_B(t)=P_A(0)+P_B(0)=2 \text{ bars}$$

An essential question is the mode of choice of the successive 18 digit integer numbers.

A first possibility is a purely random choice (this hypothesis corresponds to the "molecular chaos" of Boltzmann), but since this first possibility leads to several philosophical objections we will also consider deterministic choices such as those given by:

"The kth choice will be given by the decimals of rank (18k – 17) to 18k of a given real number x".

For instance with x = π, that is:

$$(3) \qquad x=3.141\,592\,653\,589\,793\,238\,462\,643\,\ldots$$

the first choice will be 141 592 653 589 793 238.

The purely random choice leads to a simple analysis.

A) Average evolution

Because of (2) it is sufficient to consider the evolution of $P_A(t)$. This evolution is governed by the following:

"At each step we have the probability $P_A/2$ of a variation $\delta P_A=-2n18$ and the probability $1-(P_A/2)$ of a variation $\delta P_A=+2n18$" (the pressures P_A and δP_A are expressed in bars and, according to the above list of symbols, 2n18 means 2×10^{-18}).

Hence the average evolution $P_{AM}(t)$ is given by:

A.1. From the kth choice to the (k+1)th choice:

$$(4) \qquad P_{AM,k+1}=P_{AM,k}+(1-P_{AM,k})\times 2n18$$

that is:

$$(5) \qquad P_{AM,k+1}-1=(P_{AM,k}-1)(1-2n18)$$

and thus:

$$(6) \qquad P_{AM,k}-1=(P_{AM,0}-1)(1-2n18)^k = (1-2n18)^k \times 0.4 \text{ bar}$$

A.2. Since there are p15 choices and exchanges per second:

$$(7) \qquad \begin{cases} P_{AM}(t)=1 \text{ bar}+\exp\left(\dfrac{-t}{500}\right)\times 0.4 \text{ bar} \\ \text{with t expressed in seconds} \end{cases}$$

Thus for t = 500 s we obtain P_{AM} = 1.1471 bar and for t = 1000 s = 16 mn 40 s we obtain P_{AM}=1.0541 bar, the average value P_{AM} of the pressure P_A converges exponentially towards one bar.

B) Evolution of the variance V(t)

This evolution is of course more complex than that of $P_{AM}(t)$ but remains simple.

B.1. From the kth choice to the (k+1)th choice:

$$(8) \qquad V_{(k+1)}=(1-4n18)\,V_k + 4n36.[1-(P_{AM.k}-1)^2]$$

B2. With (6) and with V_0 = 0 we obtain:

$$(9) \qquad V_k=[n\,18.\{1-(1-4n18)^k\} - 0.16\{(1-2n18)^{2k}-(1-4n18)^k\}]\text{bar}^2$$

B3. In terms of the time t (expressed in seconds) the exact expression (9) gives almost:

$$(10) \qquad V(t)=\{1-(1+0.00064\,t).\exp\left(\frac{-t}{250}\right)\}.n\,18\,\text{bar}^2$$

Thus the evolution of the variance V(t) is monotonic, it increases from 0 to n18 bar² and at t=1000s it is already at 97% of its final value.

The main result is that the variance V(t) remains forever very small.

The standard deviation σ(t), the square root of the variance, will also remain forever very small, its maximum is n9 bar that is one billionth of a bar or one deci-millipascal (one pascal = 1 Pa = 1N/m² = n5 bar).

If we measure the pressure with the accuracy of one millipascal, *i.e.*, ten standard deviations, we will notice from time to time a fluctuation with respect to the average evolution. The average frequency of these fluctuations is about one per two years.

If we measure the pressure with the accuracy of five millipascals, *i.e.*, fifty standard deviations, we will have the probability n200 (that is 10^{-200}) to meet a fluctuation before the time t = 4.625p329 seconds that is t = 1.465p322 years ... We will never meet such a "large" fluctuation of five millipascals and the evolution appears as irreversible.

In these conditions the return time of Poincaré is purely theoretical, but it can be computed. The a priori probability of $P_A \geq 1.4$ bar is 10^{-M} with M=3.57350p16, this of course an extremely small but non-zero probability.

If we neglect events with a probability of n200 (which corresponds to the "threshold of certainty of observable Universe") we can write that the first return of Poincaré at $P_A \geq 1.4$ bar will occur after k_R exchanges with:

$$(11) \quad 1/2 \times 10^{(M-200)} \leq k_R \leq 10^M \times 5 \times 2Ln(10^{200})$$

Hence the first return of Poincaré will occur after 10^Q seconds with:

$$(12) \quad 3.57349p16 \leq Q \leq 3.57351p16$$

These results corespond to the random choice of successive exchanges and we have also to consider the cases of deterministic choices as explicited with equation (3). The computations (4) – (12) give then that the values of x that satisfy $0 \leq x \leq 1$ and that doesn't satisfy (12) have a total measure smaller than n200. This set of values is then completely negligible even if many remarkable values of x (such as $x = 0.5$) belong to that very small set.

CONCLUSION

The irreversibility of the second principle of thermodynamics agree fully with our experiments and our measures that are by far neither long nor numerous enough to lead to a contradiction. However some very small temporary fluctuations appear from time to time in very accurate experiments.

Thus the paradox of reversible physical laws associated with irreversible phenomena can be explained without "perfect isolation", "hidden correlations" and/or "small hidden irreversibilities". The main reason of irreversibility is the very large number of parameters of irreversible systems.

The Boltzmann's hypothesis of "molecular chaos" is excellent and allows very accurate computations. The correlations will not increase slowly and insidiously after a very long time and we can almost write that the return of Poincaré occurs by chance which usually requires such a large delay, much larger than the age of Universe, that the corresponding decrease of entropy never appears in our experiments.

The return time of Poincaré is exponentially related to the number of independent parameters of the system of interest and we can thus write:
"If after the usual mathematical simplification (integral of motions, decomposability, etc.) a system:
A) remains with N independent parameters where $N > 10\,000$.
B) is sensible to the initial conditions (chaotic system).
Then its evolution will physically appear as irreversible for measures of accuracy worse than $(50/\sqrt{N})$ even if its laws are mathematically reversible and conservative."

ANALYSE NON LINEAIRE ET DYNAMIQUE DU VOL

Ph. Guicheteau
ONERA
Boite Postale 72
92322 Chatillon CEDEX
France

1. RESUME

Le mouvement d'un avion peut être décrit de manière rigoureuse par une ensemble d'équations différentielles non linéaires, dépendant de paramètres, liant les variables d'état (incidence, dérapage, vitesse ...) et les variables de commandes (braquage des gouvernes) par l'intermédiaire des équations de la mécanique du vol, du modèle aérodynamique et du système de contrôle du vol. La communication évoque les travaux réalisés en France et à l'Etranger en vue d'améliorer la compréhension et de prédire avec précision le comportement de l'avion dans des situations de vol particulières pour lesquelles l'analyse linéarisée habituelle des équations différentielles est insuffisante ou inadaptée.

2. ABSTRACT

In a rigorous manner, aircraft motion can be described by a set of nonlinear differential equations, depending of parameters, associating the state vector (angle of attack, sideslip angle, speed ...) with the control vector (motivators) through flight dynamics equations, aerodynamic aircraft model and flight control system. The communication presents some works which aim at a better understanding and at the precise prediction of aircraft behaviour in particular flight phases for which classical linearized analysis of differential equations is insufficient or not valid.

3. INTRODUCTION

Le mouvement d'un avion peut être décrit de manière rigoureuse par une ensemble d'équations différentielles non linéaires, dépendant de paramètres, liant les variables d'état (incidence, dérapage, vitesse ...) et les variables de commandes (braquage des gouvernes) par l'intermédiaire des équations de la mécanique du vol, du modèle aérodynamique et du système de contrôle du vol.

Depuis le début de l'aviation, de nombreux chercheurs se sont attachés à expliquer certains phénomènes de la dynamique du vol en appliquant, de façon analytique, des méthodes d'analyse non linéaires de la stabilité sur des systèmes d'équations de dimension réduite. Indépendamment de l'intérêt de ces approches pour aider à la compréhension des phénomènes, les hypothèses simplificatrices adoptées nuisent souvent à la qualité du résultat.

Plus récemment, divers auteurs ont mis en oeuvre une méthode numérique globale d'analyse de la stabilité fondée sur le principe de la résolution d'équations algébriques non linéaires dépendant de paramètres au moyen d'un processus de continuation associé à la théorie des bifurcations.

La communication évoque les travaux réalisés en France et à l'Etranger en vue d'améliorer la compréhension et de prédire avec précision le comportement de l'avion dans des situations de vol particulières pour lesquelles l'analyse linéarisée habituelle des équations différentielles est insuffisante ou inadaptée.

Après avoir rappelé brièvement les travaux antérieurs fondés sur des analyse simplifiées de la stabilité, les aspects théoriques de la démarche méthodologique adoptée ici et les éléments de la théorie des bifurcations indispensables à la compréhension des phénomènes rencontrés dans son application à la dynamique du vol sont présentés.

Il s'agira tout d'abord de retrouver les limites usuelles de stabilité de l'avion naturel en montrant, au passage, qu'elles sont des approximations des bifurcations du système complet. A cette occasion, on montrera que la méthodologie est apte à prédire le comportement du système complet au delà de ces limites de stabilité; la vrille et le couplage inertiel en seront des exemples.

Puis, la corrélation avec des essais en vol obtenue dans le cadre d'une opération réalisée à l'ONERA en collaboration avec Dassault Aviation et le CEV viendra confirmer les prédictions énoncées plus haut.

Ensuite, l'influence des commandes de vol et de ses non-linéarités sera abordée par le biais d'une brève synthèse bibliographique et des travaux menés à l'ONERA. Il sera alors montré que la méthodologie globale d'analyse des systèmes est, dans une certaine mesure, à même de fournir des éléments de réponse.

Enfin, en relation avec les problèmes posés par les non-linéarités dans les commandes de vol, quelques remarques sur les méthodes modernes de détermination du domaine d'attraction d'un point d'équilibre stable d'un système différentiel non linéaire viendront conclure cette communication. Sur ce dernier point, des résultats obtenus à l'ONERA et dans la littérature seront présentés.

4. ANALYSE NON LINEAIRE SIMPLIFIEE DU COMPORTEMENT DE L'AVION

La prédiction et l'analyse des pertes de contrôle et des vrilles sont des problèmes anciens qui remontent au début de l'aviation. Cependant, faute de moyens numériques adéquats, les chercheurs ont étudié ces problèmes en les simplifiant suffisamment pour, ensuite, pouvoir utiliser l'arsenal des méthodes analytiques, exactes ou approximatives, d'analyse d'équations non linéaires.

Parmi les précurseurs, il convient de citer les travaux de Phillips [1] sur le couplage inertiel dont la difficulté réside dans la prise en compte des couples gyroscopiques dans les équations du mouvement. En négligeant la pesanteur et en ne considérant que les équations de mouvement, il a pu dégager un critère de stabilité encore utilisé de nos jours. Repris et étendus par Pinsker [2], ces travaux sont décrits en détail dans [3].

A incidence plus élevée, les phénomènes qui apparaissent sont généralement associées à des instabilités de nature aérodynamique. Ils en résulte des pertes de stabilité sur les mouvements longitudinaux et transversaux ou sur les deux en même temps [4]. Parmi ceux ci, l'instabilité de l'oscillation de dérapage a été l'objet de très nombreuses communications dont [5,6] dans lesquelles les non-linéarités aérodynamiques sont introduites dans les équations sont la forme de non-linéarités polynomiales.

A grande incidence, la vrille calme a rapidement été identifiée comme un état d'équilibre du système complet. Ces études avaient montré que la vrille était un mouvement sur une hélice d'axe vertical pendant lequel la portance et la traînée équilibrent respectivement la force centrifuge et le poids. Cependant, malgré l'apparition de moyens de calcul puissants [7], et la similarité entre le mouvement d'Euler-Poisot et la vrille de certains avions de combat [8], les régimes de vrille agitée étaient encore considérés comme dépendre en grande partie du hasard.

En marge de ces travaux consacrés individuellement à l'étude d'un seul phénomène, Schy et Hannah [9,10] ont montré l'existence de solutions multiples au problème de la détermination de l'équilibre du système non linéaire représentant des équations simplifiées de la mécanique du vol; certaines solutions étant stables, les autres instables. Ces derniers travaux sont à la base de la méthode d'analyse globale présentée dans cette communication.

5. METHODE GLOBALE D'ANALYSE NON LINEAIRE DU COMPORTEMENT

Des analyses simplifiées, généralement analytiques, sur un modèle réduit sont souvent suffisantes pour aider à la compréhension de certains phénomènes non linéaires. Cependant, pour être appliquées, ces techniques nécessitent des hypothèses simplificatrices sur le modèle qui ne sont pas toujours faciles à faire a priori ou/et qui ne sont pas souhaitées.

En vue de lever les hypothèses simplificatrices et d'accroître la précision de l'analyse des phénomènes, une nouvelle approche a été développée.

5.1 Méthodologie

La méthodologie globale d'analyse du comportement est fondée sur l'analyse des états asymptotiques des systèmes d'équations différentielles non linéaires autonomes, dépendant de paramètres et exprimés sous la forme:

$$(1) \qquad \frac{dX}{dt}=F(X,U)$$

dans laquelle:

- X : vecteur d'état de dimension n,
- U : vecteur de paramètres de dimension m
- F : n fonctions non linéaires de X et U.

Par opposition aux systèmes différentiels linéaires, un ensemble d'équations différentielles non linéaires peut présenter plusieurs états asymptotiques différents pour une combinaison donnée (U_0) des paramètres.

Dans les cas les plus simples, notamment ceux rencontrés au cours des applications à la dynamique du vol des avions, les états asymptotiques correspondent généralement aux solutions de l'équation algébrique non linéaire:

$$(2) \qquad F(X,U_0)=0$$

lorsque le système est immobile ou aux solutions de l'équation :

$$(3) \qquad X(T)=X(0)+\int_0^T F(X,U_0)dt$$

dans le cas d'orbites périodiques (cycles limites) dont il convient de noter que la période (T) est a priori inconnue puisque le système considéré est supposé autonome.

Lorsque l'analyse linéarisée des petits mouvements autour de ces états asymptotiques a un sens, la caractérisation de leur stabilité est donnée par un calcul de valeurs propres.

La stabilité des états d'équilibres "ponctuels" (2), est assurée si toutes les valeurs propres du système linéarisé sont à partie réelle négative. Pour les orbites périodiques (3), la théorie de Floquet indique que la stabilité est assurée si toutes les valeurs propres de la matrice de transition (sauf une) ont un module inférieur à 1; la dernière valeur propre de la matrice de transition à son module égal à l'unité pour assurer la périodicité du mouvement.

Lorsque l'analyse linéarisée des petits mouvements est insuffisante (limites de stabilité) la théorie des bifurcations permet de conclure sur la stabilité de l'état asymptotique considéré et sur la modification du comportement asymptotique global du système sous l'effet de la variation du paramètre.

5.2 Apport de la théorie des bifurcations

La théorie des bifurcations est constituée par un ensemble de résultats mathématiques qui visent à analyser et à expliquer des changements du comportement asymptotique d'un système différentiel non linéaire quand les paramètres varient de façon quasistatique.

Bien entendu, il n'est pas dans l'esprit de cette communication de se substituer aux ouvrages spécialisés traitant du sujet [11,12,13]. Cependant, il paraît opportun d'en rappeler quelques aspects qui recouvrent quasiment tous les phénomènes rencontrés au cours de son application particulière à la dynamique du vol.

5.2.1 Bifurcations des points d'équilibre

Les phénomènes rencontrés le plus couramment avec des points d'équilibre apparaissent lorsqu'une valeur propre réelle (μ) négative devient positive sous l'effet d'une variation du paramètre. Deux cas peuvent alors se produire.

Si le point d'équilibre est régulier, le théorème des fonctions implicites est applicable et un point de retournement apparait (figure 1). Ce point est structurellement stable vis à vis de perturbations du système différentiel.

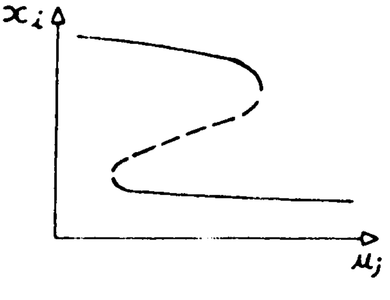


Figure 1 : Point de retournement.
—— équilibres stables, - - - équilibres instables

Si le point d'équilibre est singulier, de nouvelles branches d'équilibres apparaissent. Suivant les particularités du système, le point de bifurcation peut être un point fronce ou un point de bifurcation/retournement (figure 2). Ces points de bifurcation sont structurellement instables vis à vis des perturbations du système et ils se transforment généralement en point de retournement [13]. En dynamique du vol, compte tenu du système d'équations et des modèles d'avions adoptés jusqu'à présent, ces points de bifurcation particuliers apparaissent seulement quand les équations différentielles étudiées sont symétriques.

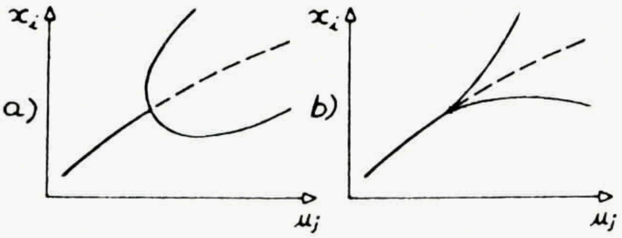


Figure 2 : Points de bifurcation, (a) retournement/bifurcation; (b) bifurcation fourchette.
—— équilibres stables, - - - équilibres instables

Un autre point de bifurcation souvent observé est la bifurcation de Hopf. Elle survient quand une paire de valeurs propres complexes conjuguées ($\mu_{1,2}$)a sa partie réelle qui devient positive lorsque le paramètre du système varie. Au voisinage de ce point (correspondant à une paire de valeurs propres imaginaires pures), des orbites périodiques apparaissent.

Suivant les caractéristiques du système, la bifurcation pourra être supercritique ou subcritique (figure 3). Dans le premier cas, l'orbite périodique est stable et de petite amplitude. Dans le second cas, l'orbite située au voisinage du point de bifurcation est instable et l'orbite stable du système, si elle existe pour cette valeur particulière du paramètre, est de grande amplitude même pour une faible variation du paramètre autour du point de Hopf.

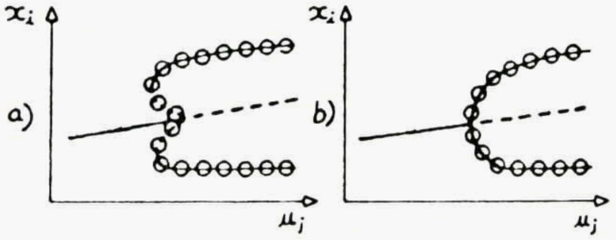


Figure 3 : Bifurcation de Hopf, (a) subcritique; (b) supercritique.
—— équilibres stables, - - - équilibres instables,
-0-0- enveloppe des orbites périodiques

5.2.2 Bifurcations des orbites périodiques
Comme nous venons de l'entrevoir plus haut, le dernier cas très intéressant de bifurcations rencontrées en dynamique du vol correspond à la perte de stabilité des orbites périodiques (figure 4).

Parmi toutes les causes possibles d'instabilité, trois sont usuellement rencontrées.

a. Une valeur propre réelle traverse le point +1. Dans ce cas, la perte de stabilité de l'orbite se traduit par l'apparition d'une orbite de retournement analogue au point de retournement des surfaces d'équilibre.
b. Une valeur propre réelle traverse le point -1. Alors une bifurcation de doublement de période survient. Au voisinage l'orbite périodique initialement stable et de période T apparait une nouvelle orbite généralement stable de période 2T. Généralement cette bifurcation est accompagnée par d'autres semblables (cascade de Feigenbaum) et conduit à l'apparition d'un mouvement chaotique.
c. Deux valeurs propres imaginaires conjuguées quittent le cercle de rayon unité. Au delà de cette bifurcation le mouvement du système décrit une surface toroïdale qui entoure l'orbite périodique devenue instable.

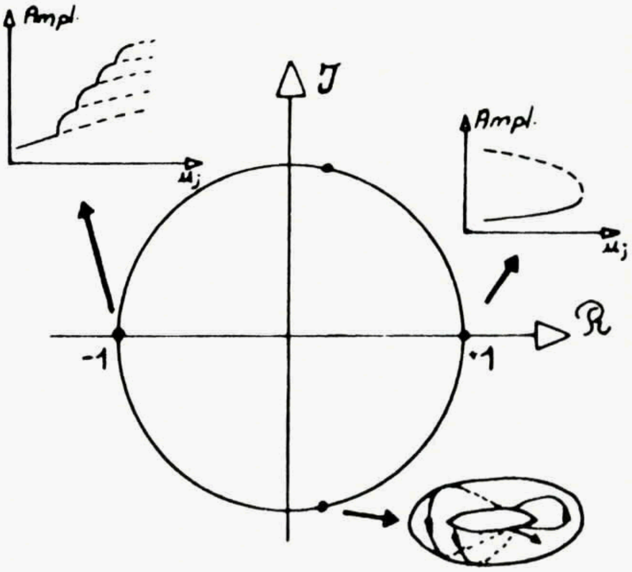


Figure 4 : Bifurcations des orbites périodiques.
—— orbites stables, - - - orbites instables

5.2.3 Caractérisation de la stabilité des limites de stabilité
La stabilité des points de retournement et de bifurcation ne peut pas être fournie par l'analyse linéarisée classique.

L'approche préconisée par la théorie des bifurcations est fondée sur l'hypothèse selon laquelle tous les modes amortis (associés à des valeurs propres à partie réelle négative) du système disparaissent peu de temps après que l'état d'équilibre ait été perturbé. Alors, la méthode [11] consiste à effectuer un développement non linéaire de la projection du mouvement du système dans l'espace engendré par le(s) vecteur(s) propre(s) associé(s) à la (aux) valeur(s) propre(s) à partie réelle nulle.

Pour un point d'équilibre avec une valeur propre nulle, la méthode permet de caractériser sa stabilité et de fournir une indication par le temps mis par une perturbation pour doubler ou être divisé par deux. A noter que contrairement au cas d'un point d'équilibre d'un système linéaire, ce temps est proportionnel à la valeur initiale de la perturbation.

Pour un point de bifurcation de Hopf, il est possible de conclure sur le caractère subcritique ou supercritique de la bifurcation suivant que le point considéré est respectivement instable ou stable.

Tous ces calculs nécessitent une formulation analytique du système. Cependant, moyennant une approximation non linéaire numérique du système différentiel au voisinage du point étudié, la démarche précédente peut être automatisée et fournir encore des résultats intéressants [14].

5.2.4 Remarque sur les mouvements chaotiques
Mises à part quelques expériences numériques, l'analyse fine des mouvements chaotiques n'est pas incorporée dans la méthodologie présentée ici. Une des raisons de ce choix est que, dans la pratique, le temps d'observation de mouvements complexes de l'avion (vrille par exemple) est trop limité pour pouvoir observer le comportement asymptotique et l'analyser complètement.

Cependant, après cette remarque, il ne faudrait pas conclure hâtivement sur l'inutilité de l'analyse de ces comportements. En effet, si l'on aborde l'identification de certains paramètres du système à partir d'essais réels, il est alors indispensable de pouvoir dissocier un comportement ressemblant à un mouvement périodique, pendant le temps d'observation du système, d'un mouvement chaotique. Dans le premier cas le modèle restitué fournira toujours une réponse ayant la même allure alors que, dans la deuxième hypothèse, la dépendance vis à vis des conditions initiales sera peut être à même d'expliquer des résultats d'essais très différents bien que les conditions initiales soient très voisines.

5.3 Procédure numérique développée à l'ONERA

Le calcul de tous les états asymptotiques mentionnés dans le paragraphe précédent est réalisé par un code numérique développé à l'ONERA. Ce code utilise une méthode de continuation initialement proposée par Kubicek [13] et modifiée par, entre autres, les contributions de Keller [15] et Kernevez [16].

Partant d'une approximation initiale d'un état asymptotique, pour une combinaison donnée des paramètres du système, le code détermine la courbe X(U) solution d'un ensemble d'équations algébriques non linéaires qui dépend du cas de calcul réalisé.

a. point d'équilibre:

$$F(X,U)=0$$

b. point de retournement:

$$\begin{aligned} F(X,U)&=0 \\ \mu&=0 \end{aligned}$$

c. point de Hopf:

$$\begin{aligned} F(X,U)&=0 \\ \mu_{1,2}&=\pm 2i\pi/T \end{aligned}$$

d. orbite périodique:

$$X(T)=X(0)+\int_0^T F(X,U)dt$$

Les fonctions non linéaires considérées ici sont supposées continues et dérivables jusqu'à un ordre suffisamment élevé pour ne pas mettre en défaut le principe de la description globale des courbes d'équilibre par continuation. Pratiquement, suivant le type de calcul envisagé et sous réserve que les fonctions non linéaires ne varient pas brusquement en fonction de l'état et des paramètres, on observe qu'il est possible de diminuer les conditions sur la dérivabilité des fonctions sans courir le risque de dégrader notablement le processus de continuation.

5.4 Mise en oeuvre de la méthodologie

L'application de la méthodologie décrite plus haut est usuellement réalisée en cinq étapes successives.

a. Détermination de l'ensemble des états asymptotiques, tels qu'ils ont été définis dans le § 5.1, pour toutes les combinaisons désirées des paramètres et caractérisation de leur stabilité.
b. Représentation graphique des résultats dans des sous espaces particuliers de dimension 2 ou 3 (généralement la dimension du vecteur d'état est très supérieure à 2).
c. Analyse des tracés graphiques et prédiction d'éventuels phénomènes non linéaires sans recourir à de nombreuses simulations du système.
d. Validation des prédictions au moyen de simulations numériques du mouvement.
e. Eventuellement corrélation avec des essais réels en vue de discuter la validité de la représentation du système réel par le système différentiel étudié.

6. ANALYSE NON LINEAIRE GLOBALE DU COMPORTEMENT DE L'AVION NATUREL

L'étude de la dynamique du vol classique repose essentiellement sur l'analyse linéarisée des équations de la mécanique du vol. Sans revenir sur son intérêt didactique, il convient de noter que cette vision simplifiée des équations ne permet pas de montrer la genèse des pertes de stabilité et de mettre en évidence des phénomènes complexes.

En s'appuyant sur des exemples simples et d'autres plus complexes, il est proposé de montrer l'apport de cette méthodologie dans la compréhension du comportement qualitatif de l'avion naturel pour lequel les braquages des gouvernes aérodynamiques font partie des paramètres de l'équation (1).

6.1 Equations du mouvement et classification des non-linéarités

Les équations classiques du mouvement d'un avion rigide dans l'espace sont constituées:

- des équations de moment (en supposant ici que les produits d'inertie D et F sont nuls)

$$\begin{aligned} A\dot{p}-E\dot{r}+(C-B)qr-Epq&=L^A+L^F \\ B\dot{q}+(A-C)rp+E(r^2-p^2)&=M^A+M^F \\ C\dot{r}-E\dot{p}+(B-A)pq+Erq&=N^A+N^F \end{aligned}$$

- des équations de force

$$\begin{aligned} m(\dot{u}+qw-rv)&=X^A+F_X-mg\sin\theta \\ m(\dot{v}+ru-pw)&=Y^A+F_Y+mg\cos\theta\sin\phi \\ m(\dot{w}+pv-qu)&=Z^A+F_Z+mg\cos\theta\cos\phi \end{aligned}$$

- auxquelles sont adjointes deux équations cinématiques

$$\begin{aligned} \dot{\phi}&=p+\tan\theta(q\sin\phi+r\cos\phi) \\ \dot{\theta}&=q\cos\phi-r\sin\phi \end{aligned}$$

dans lesquelles:

- u, v, w : composantes de la vitesse-air dans le trièdre avion,
- p, q, r : composantes du vecteur rotation instantanée dans le trièdre avion,
- θ : assiette longitudinale,
- ϕ : angle de gîte,
- X^A, Y^A, Z^A : composantes de la force aérodynamique dans le trièdre avion,
- F_X, F_Y, F_Z : composantes de la poussée dans le trièdre avion,
- L^A, M^A, N^A : composantes du moment aérodynamique dans le trièdre avion,
- L^F, M^F, N^F : composantes du moment de propulsion dans le trièdre avion.

Le vecteur d'état est de dimension 8. Il est constitué de la façon suivante:

$$X^T=(u,v,w,p,q,r,\theta,\phi)^T$$

Le vecteur de paramètres, au moins de dimension 3, comprend les gouvernes aérodynamiques de l'appareil:

$$U^T=(\delta l,\delta n,\delta m)^T$$

dans lequel:

- δ_l : gouverne de gauchissement,
- δ_m : gouverne de profondeur,
- δ_n : gouverne de direction.

Les non-linéarités de ce système sont nombreuses. Elles peuvent être classées en deux groupes.

Le premier comprend celles qui sont intrinsèques à la formulation des équations du mouvement d'un solide dans l'espace (non-linéarités trigonométriques et inertielles).

Le deuxième groupe est constitué des non-linéarités du modèle aérodynamique de l'avion considéré. Il convient cependant de faire une distinction entre les non-linéarités de courbure de certains coefficients ($C_z(\alpha)$, $C_l(\beta)$, etc..) et les non-linéarités de couplage qui induisent des relations entre les variables longitudinales et les variables latérales ($C_m(\beta)$, $C_l(\alpha)$, etc..).

6.2 Analyse de phénomènes simples

L'objet de cette partie est de montrer que des phénomènes simples de la mécanique du vol peuvent être analysés avec profit au moyen de la théorie des bifurcations.

6.2.1 Mouvement spiral

Dans ce mouvement très lent et à faible incidence, les couplages inertiels sont négligeables et l'aérodynamique peut être considérée comme linéaire. Seules la pesanteur et l'assiette londitudinale de l'appareil influent sur la stabilité du mouvement à dérapage nul.

Gouvernes transversales au neutre, le tracé de la surface d'équilibre montre que la manifestation de l'instabilité spirale coïncide avec l'apparition d'une bifurcation fourchette sur les variables transversales en raison de la symétrie supposée de l'appareil [17,18]. En conséquence, la surface d'équilibre fournit les états d'équilibre stable que l'avion instable spiral ira rejoindre (figure 5).

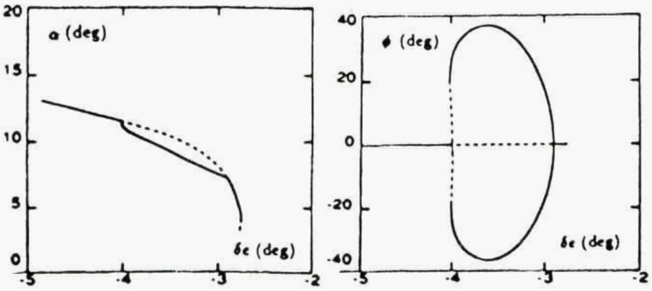


Figure 5 : Bifurcation spirale, (a) , (b)
—— équilibres stables, - - - équilibres instables

Par rapport à la mécanique du vol linéarisée, cette analyse permet de prédire le comportement du système au delà de l'instabilité. Elle offre également le moyen de mettre en évidence les non-linéarités qui créent le point de bifurcation. Ainsi, en conservant les seuls termes non linéaires essentiels, le comportement de l'avion peut se ramener à l'étude d'une équation différentielle non linéaire scalaire dérivée d'un champ de gradient exprimant l'évolution de l'assiette latérale de l'appareil en fonction des gouvernes transversales (δ_l, δ_n):

$$\dot{\phi}=(A+B\cos\phi)\sin\phi$$
$$+(C_{\delta_l}\delta l+C_{\delta_n}\delta n)\cos\phi+D_{\delta_l}\delta l+D_{\delta_n}\delta n$$

En dépit de sa simplicité, cette formulation a le mérite de synthétiser le comportement global de l'appareil au voisinage de l'instabilité spirale et de montrer que même si l'avion est stable lorsque les gouvernes transversales sont au neutre ($\delta_l=\delta_n=0$), la bifurcation spirale existe pour des braquages non nul des gouvernes transversales (figure 6).

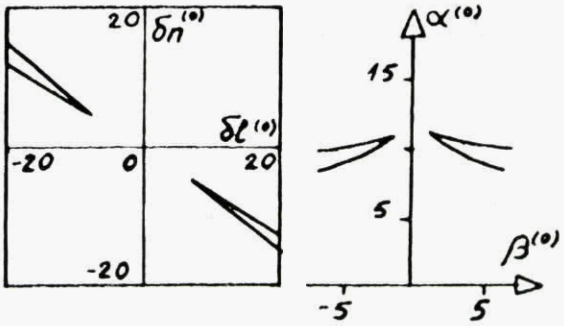


Figure 6 : Surface de bifurcation spirale dans le plan gauchissement-direction.

6.2.2 Couplage inertiel

Ce phénomène est bien connu depuis longtemps en mécanique du vol [1,2]. Cependant, son analyse à la lumière de la théorie des bifurcations [18-23] complète utilement les analyses antérieures.

En supprimant toutes les non-linéarités des équations exceptées celles relatives au couplage inertiel et en considérant que la vitesse de roulis (p) est beaucoup plus grande que les vitesses de tangage (q) et de lacet (r), la prédiction du comportement de l'appareil peut encore se ramener à l'étude d'une équation différentielle polynomiale scalaire liant la vitesse de roulis aux gouvernes aérodynamiques, généralement de degré 5, dérivant d'un champ de gradient:

$$\begin{aligned}\dot{p}=&f_6(\delta m)p^5+(f_{5_{\delta l}}\delta l+f_{5_{\delta n}}\delta n)p^4\\ &+f_4(\delta m)p^3+(f_{4_{\delta l}}\delta l+f_{4_{\delta n}}\delta n)p^2\\ &+f_2(\delta m)p+(f_{1_{\delta l}}\delta l+f_{1_{\delta n}}\delta n)\end{aligned} \tag{4}$$

dont on montre qu'elle se ramène simplement à la catastrophe "papillon" bien connue de la théorie des catastrophes (figure 7).

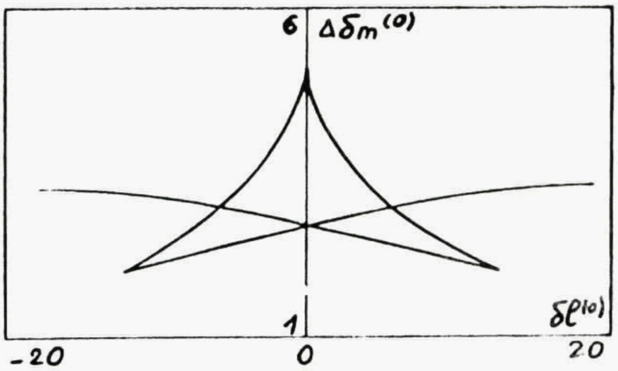


Figure 7 : Surface de bifurcation dans le plan gauchissement profondeur (δ_n = 0).

En outre, de cette formulation on peut extraire un critère de stabilité du mouvement et des lois d'interconnexion entre les gouvernes transversales qui évitent la perte de contrôle.

Ces lois sont issues directement de l'examen de la surface de bifurcation du système pour faire en sorte que les combinaisons de commande ne la franchissent pas (figure 8). Elles peuvent également être fournies en recourant aux propriétés des polynômes qui assurent l'unicité de la solution de l'équation 4 [24,25].

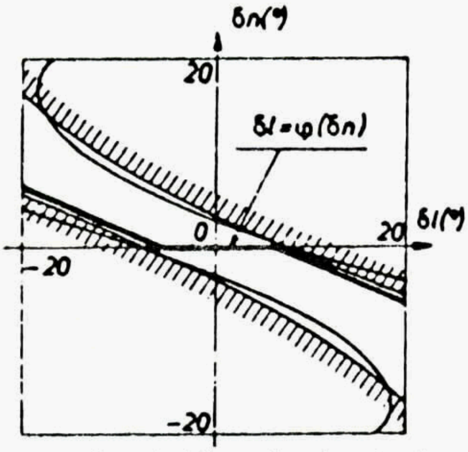


Figure 8 : Surface de bifurcation dans le plan gauchissement-direction et loi d'interconnexion entre les gouvernes transversales : $\delta_l = \varphi(\delta_n)$

6.2.3 Instabilité de l'oscillation de dérapage

Dans ce phénomène très longuement décrit dans la littérature, l'influence des non-linéarités aérodynamiques est prépondérante. L'instabilité correspond généralement à l'apparition d'un point de bifurcation de Hopf [26] à partir duquel la méthodologie présentée précédemment permet de décrire l'enveloppe des orbites périodiques du système [21,27] sans aucune hypothèse simplificatrice contrairement aux méthodes analytiques (figure 9).

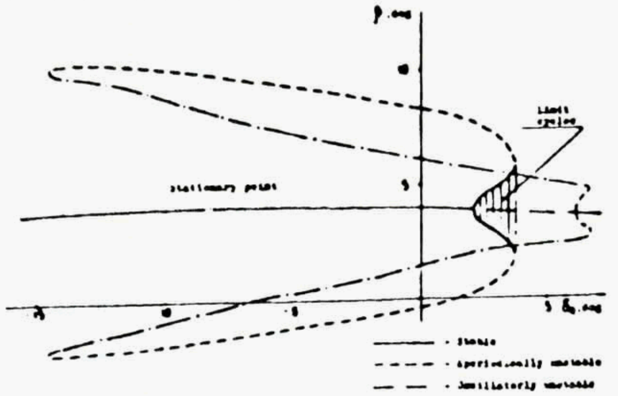


Figure 9 : Instabilité de l'oscillation de dérapage et enveloppe des orbites périodiques.

6.3 Analyse de phénomènes plus complexes

Après avoir montré brièvement que la théorie des bifurcations avait la possibilité d'améliorer la compréhension de certaines limites de stabilité qui sont usuellement obtenues par des calculs analytiques simplifiés, et de dépasser largement les frontières de la mécanique du vol linéarisée, cette partie se propose de présenter les phénomènes les plus complexes observés sur les avions de combat au moyen de cette méthode.

Le calcul par continuation des orbites périodiques et de leurs bifurcations pour différents modèles d'avions de combat révèle que certains comportements très agités peuvent être analysés avec la théorie des bifurcations.

La bifurcation la plus courante est l'orbite de retournement (figure 9).

Lorsque deux valeurs propres imaginaires conjuguées quittent le cercle de rayon unité, il apparait un tore qui entoure l'orbite nouvellement instable et sur lequel se déroule le mouvement du système (figure 10). Dans ce cas particulier étudié dans [21], l'évolution temporelle des variables se comporte comme la superposition des deux mouvements périodiques très différents ($T_1 \approx 4$ secondes et $T_2 \approx 160$ secondes pour l'exemple considéré). Il s'ensuit donc une difficulté potentielle d'interprétation d'un tel phénomène en vol puisque le temps d'observation maximal est souvent inférieur à la plus grande période.

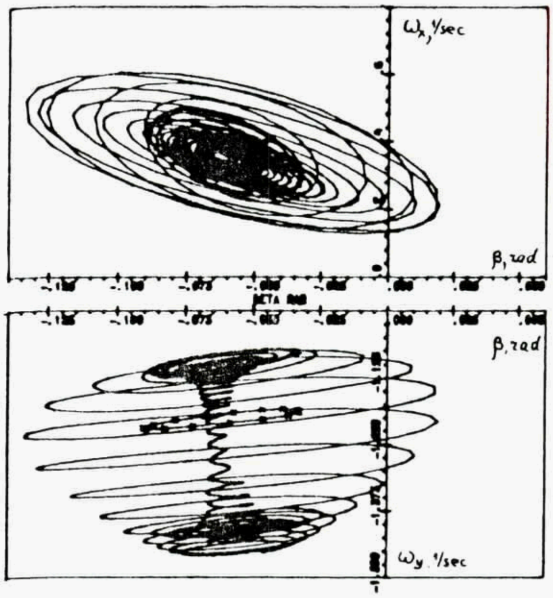


Figure 10 : Trace de l'évolution temporelle de variables d'état particulières sur un tore [21].

-•-•- orbite périodique instable

La seconde illustration est relative à la bifurcation de doublement de période (figure 11) qui peut notablement modifier le comportement en vrille d'un avion pour de très petites variations des commandes.

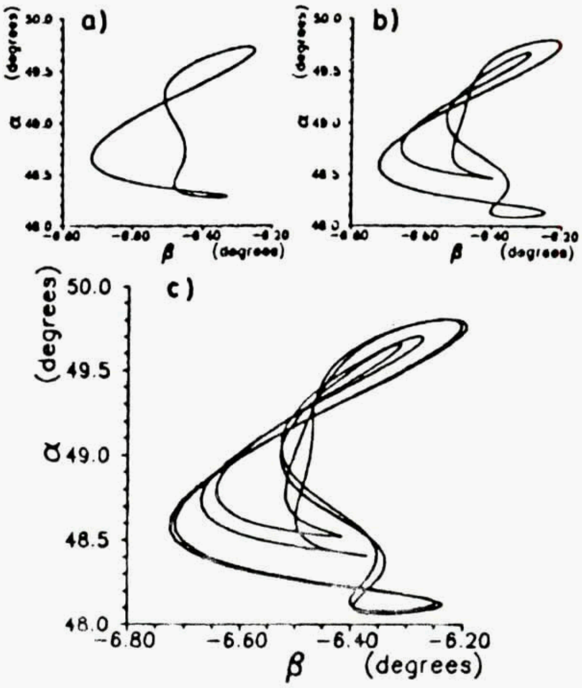


Figure 11 : Doublement de période des orbites périodiques, (a) δ_m = -18.784°, (b) δ_m = -18.971°, (c) δ_m = -19.002°

Enfin, quand tous les états d'équilibre en présence sont instables à l'exception d'un seul faiblement stable, il peut en résulter des mouvements transitoires de longue durée et très complexes pouvant être apparentés à des comportements chaotiques (figure 12).

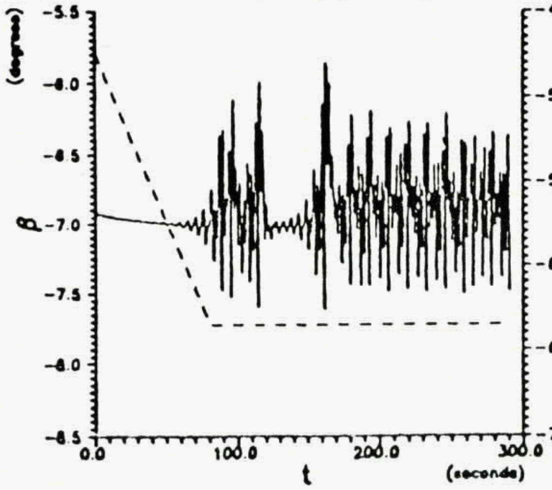


Figure 12 : Oscillations transitoires chaotiques.

6.4 Validation de la méthodologie avec un avion réel [28]

A l'ONERA, la démarche méthodologique présentée dans le paragraphe 5.4 a été parcourue complètement avec un avion réel.

L'application a porté principalement sur le comportement en vrille d'un avion de combat Franco-Allemand, Alpha-Jet, dans sa version biplace d'entraînement en raison de sa disponibilité pour réaliser des essais en vol orientés vers la validation de la méthodologie. Elle a été réalisée en collaboration entre l'ONERA/DES et l'ONERA/IMFL avec le soutien des services officiels et le concours de Dassault Aviation et du Centre d'Essais en vol d'Istres.

Après avoir défini le modèle aérodynamique de l'appareil dans un large domaine, quatre équations différentielles représentant l'évolution de certains coefficients aérodynamiques instationnaires ont été adjointes au système habituel des équations de la mécanique du vol.

Le calcul des comportements asymptotiques du système ainsi obtenu a permis de retrouver tous les phénomènes observés antérieurement sur un modèle d'avion de combat typique [18] et d'approfondir la compréhension du comportement en vol de l'appareil notamment dans les phases de vrille.

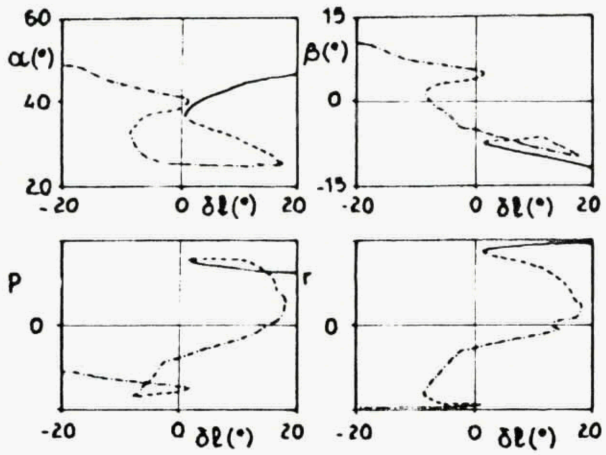


Figure 13 : Courbe d'équilibre pour $\delta_n = 17°$.
—— stable, - - - instable divergent,
-·-·- instable oscillatoire

A titre d'illustration, profondeur en butée à cabrer et direction à gauche, lorsque le gauchissement varie, la courbe des équilibres à incidence positive montre d'une part une vrille piquée à droite stable ($r > 0$) et, d'autre part, une vrille à gauche ($r < 0$) et une tonneau à incidence modérée ($\alpha \approx 25°$) instables oscillatoires (figure 13).

Autour des régimes instables de vrille à gauche des orbites périodiques apparaissent (figure 14). Leur amplitude et leur stabilité varie beaucoup en fonction de la valeur du gauchissement, ce qui conduit aux nombreux comportements très divers démontrés en vol, au Centre d'Essais en Vol d'Istres, par des essais particuliers.

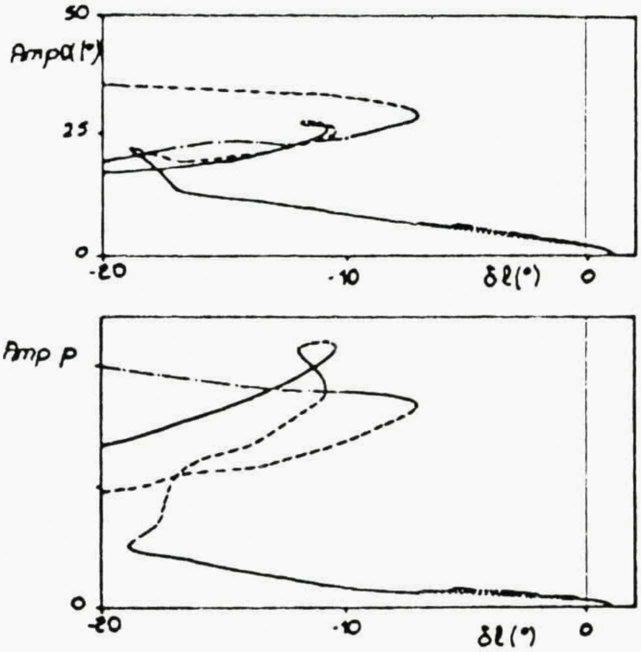


Figure 14 : Enveloppe des orbites périodiques quand δ_l varie pour $\delta_n = 17°$.
—— stable, - - - instable divergent,
-·-·- instable oscillatoire, ·····instable (-1)

Plus précisément, pour une valeur de gauchissement voisine de -10° et suivant l'historique des braquages des gouvernes défini à l'avance après analyse des surfaces d'équilibre, il est possible de mettre en évidence trois allures de vrille.

La première est une vrille calme. Elle correspond à une orbite périodique de petite amplitude (figure 15) qui existe pour quasiment toutes les valeurs négatives du gauchissement, sous réserve de conserver les mêmes braquages de la profondeur et du gauchissement.

La seconde est une vrille agitée (figure 16) avec des variations importantes sur les paramètres du vol. Elle correspond à une orbite dont l'amplitude varie beaucoup en fonction de la valeur du gauchissement.

Quant à la troisième, c'est une vrille très agitée. Elle est associée à une orbite périodique instable oscillatoire et elle correspond à l'évolution du système sur un tore entourant l'orbite instable (figure 17).

Dans cette application, seul le phénomène de doublement de période, prédit pour une valeur de gauchissement voisine de -5°, n'a pu être mis en évidence lors de ces essais en vol compte tenu de la durée limitée d'une phase de vrille ($T \approx 50$ secondes). Cependant, il convient de noter que les pilotes savaient que le comportement de cet appareil est très sensible et peu fidèle pour ces conditions particulières de vol et de braquage des gouvernes.

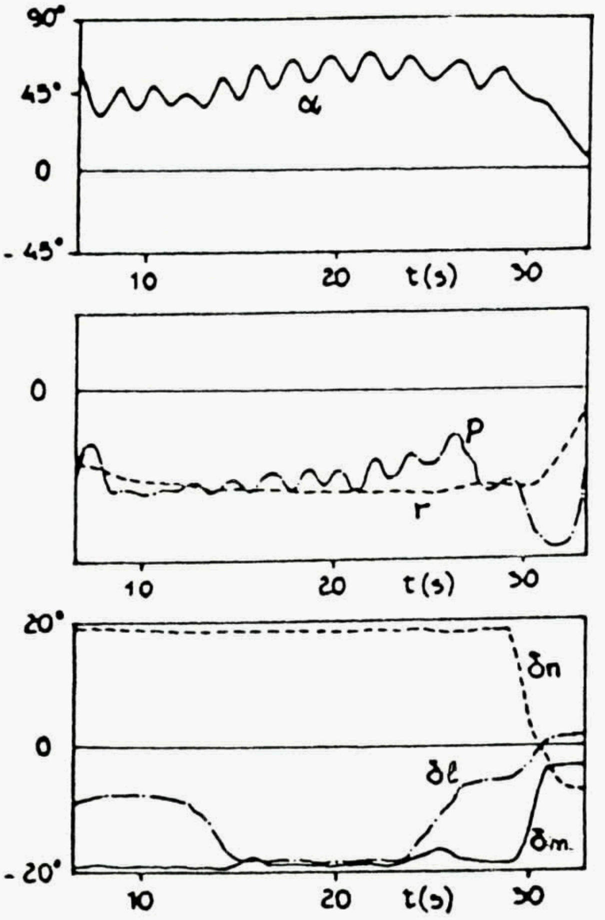


Figure 15 : Vrille calme pour δ_l variant de -20° à 0°

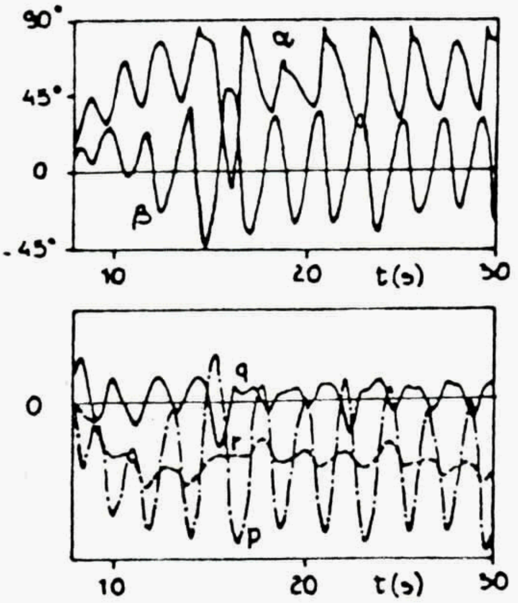


Figure 16 : Vrille agitée.

Outre ces quelques illustrations, cette étude a permis également d'estimer l'influence de nombreux paramètres tels que:

- le déplacement du centre de gravité de l'appareil en dehors de son plan de symétrie sur les vrilles,
- la poussée,
- les couples gyroscopiques dus aux moteurs,
- certains coefficients aérodynamiques,

et de montrer les potentialités de la méthode sur un cas réel.

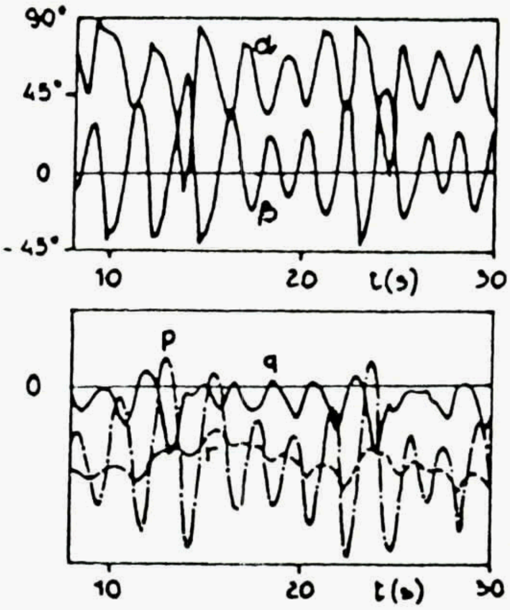


Figure 17 : Mouvement sur un tore.

7. QUELQUES REMARQUES A PROPOS DES COMMANDES DE VOL

Il n'est pas dans le propos de cette communication de traiter de l'élaboration des lois de commande pour les systèmes non linéaires. Cependant, compte tenu de lois de commande linéaires on non linéaires (d'après leur formulation ou par l'introduction d'éléments non linéaires) données, on peut se demander si la méthodologie développée pour l'avion naturel est encore applicable à l'avion piloté. Plus précisément, si les lois de commande entraînent la disparition de certaines bifurcations du système, ne risquent elles pas de créer de nouvelles instabilités?

Même sans employer le terme de bifurcation, ces problèmes de commande sont le souci de nombreux chercheurs et des résultats importants sont d'ors et déjà disponibles dans la littérature.

Dans ce paragraphe, on s'intéresse au système (1) pour lequel on suppose que la commande U n'est plus indépendante de l'état et qu'elle est rebouclée sur l'état selon la relation:

$$U=G(X,P)$$

la fonction G étant continue, dérivable dans les mêmes conditions que F et éventuellement non linéaire. Le vecteur P représente les nouveaux paramètres du système en boucle fermée, caractéristiques de la loi de commande, à faire varier de façon quasistatique

De part la définition de U, il apparait que tout équilibre

$$(\dot{X}=0)$$

est aussi équilibre du système en boucle fermée et réciproquement; la loi de commande ne modifie pas les états d'équilibre du système original. Par contre, la stabilité des solutions est différente et des orbites périodiques peuvent disparaître ou apparaître suivant la formulation de G(X,P).

7.1 Commandes de vol linéaires

Une illustration de l'influence de commande de vol linéaires sur une avion de combat est présentée par Planeaux [29].

Cette loi de la forme:

$$(U=KX)$$

a pour but de modifier l'apparition de la divergence de l'oscillation de dérapage.

Gouvernes transversales au neutre, l'instabilité survient pour un braquage de la *profondeur voisin de* -19°. En assimilant les gains de la boucle de contrôle à des paramètre du système, le processus de continuation permet d'une part de suivre l'apparition de la bifurcation de Hopf en fonction de la profondeur et des gains et, d'autre part de montrer que l'effet d'un gain linéaire n'est pas nécessairement linéaire (figure 18).

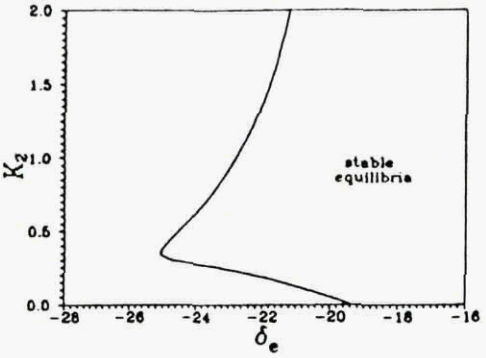


Figure 18 : Evolution de l'instabilité de l'oscillation de dérapage en fonction de la profondeur et d'un gain de la boucle de contrôle.

Au delà de la nouvelle limite de stabilité de l'avion muni de sa loi de commande, il est à noter que l'amplitude des orbites périodiques est également influencée par le gain de la boucle (figure 19).

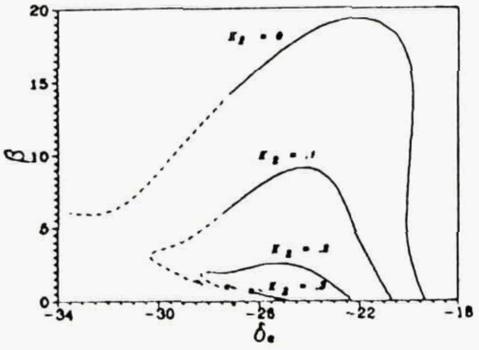


Figure 19 : Effet du gain sur l'amplitude maximale des orbites périodiques (dérapage uniquement) en fonction de la profondeur.

Cet exemple simple montre que la méthode d'analyse globale de stabilité a également la possibilité d'améliorer la compréhension du comportement de systèmes non linéaires muni de lois de commande même complètement linéaires.

7.2 Commandes de vol non linéaires

Dans les commandes de vol, les non-linéarités peuvent apparaître de façon naturelle ou accidentelle. Elles sont "naturelles" lorsque le concepteur les introduit volontairement dans le système en vue de réaliser un objectif donné. Elles sont "accidentelles" quand elles surviennent incidemment du fait que tous les éléments intervenant dans une chaîne de commande ne sont pas naturellement linéaires (butées, saturations, hystérésis).

Par exemple, un gain n'est généralement pas constant quelle que soit l'amplitude du signal d'entrée. On peut alors se demander qu'elle est l'influence de la saturation sur le comportement du système lorsqu'il est soumis à des perturbations de grande amplitude.

Sans oublier de mentionner les travaux, entre autres, de Hirai [30], Holmes [31] et Mitobe [32], l'application de la théorie des bifurcations au cas simple suivant peut laisser augurer d'intéressants développements ultérieurs.

Soit un système différentiel linéaire du second ordre instable en boucle ouverte, d'amortissement ζ_0 et pulsation ω_0. Ce système est stabilisé par un retour en vitesse (figure 20).

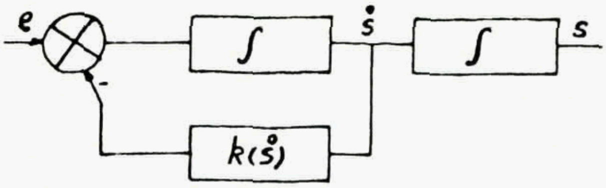


Figure 20 : Système du second ordre instable et stabilisé en vitesse.

L'analyse linéarisée:

$$\ddot{s}+(2\zeta_0\omega_0+k)\dot{s}+\omega_0{}^2s=e$$

montre que le gain k joue sur l'amortissement du système *en boucle fermée.*

Lorsque le gain passe de la valeur zéro à la valeur (k^*) qui réalise l'amortissement souhaité, il existe donc une valeur critique (k_c) pour laquelle l'amortissement est nul; le système en boucle fermée possède alors deux valeurs propres imaginaires pures.

Supposons maintenant que le gain k est non linéaire et, pour faciliter les calculs analytiques, de la forme:

$$k(\dot{s})=\frac{2}{\pi}k(\dot{s})_{\lim}\arctan\frac{k\pi\dot{s}}{2k(\dot{s})_{\lim}}$$

ce qui correspond à une saturation adoucie en vitesse.

Il est aisé de montrer qu'une orbite périodique instable apparait au delà du point de Hopf. Lorsque le gain varie, cette orbite entoure le point de fonctionnement stabilisé du système en boucle fermée et peut éventuellement subsister pour la valeur du gain choisie pour obtenir une "bonne" réponse du système vis à vis des petits mouvements (figure 21). Dans ce cas particulier, de dimension 2, l'orbite instable représente la frontière du domaine de stabilité du système commandé.

Appliquée à un modèle d'avion de combat typique instable en latéral et stabilisé par des retours saturés en vitesses de roulis et de lacet et, également en braquage des gouvernes transversales, la description de l'orbite périodique instable entourant le point de vol (figure 22) stabilisé donne une première information sur la robustesse de la loi vis à vis de perturbations [33]. *Contraire-*

ment à l'exemple précedent, l'orbite instable n'est plus maintenant qu'un élément de la frontière du bassin d'attraction.

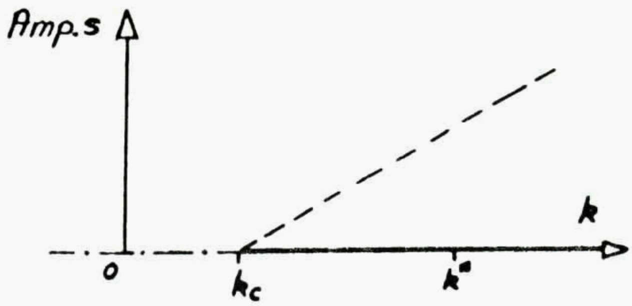


Figure 21 : Amplitude de la solution asymptotique du système en fonction du gain non linéaire.
—— stable - - - instable -·-·- instable oscillatoire

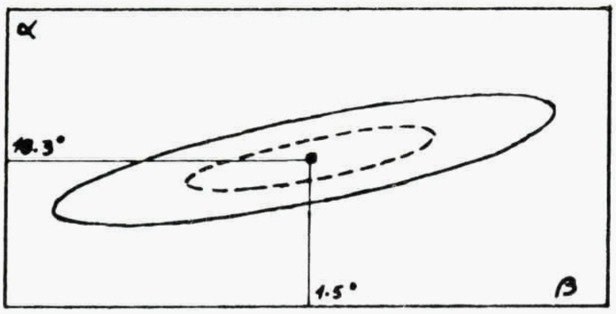


Figure 22 : Etat d'équilibre stabilisé et orbites périodiques pour un avion de combat typique.
• point d'équilibre stabilisé, - - - orbite instable,
—— orbite stable

8. DOMAINE D'ATTRACTION DES ETATS D'EQUILIBRE

Le problème général de la détermination du domaine d'attraction d'un point d'équilibre stable d'un système différentiel autonome est étudié depuis longtemps. Une synthèse des différentes méthodes employées est réalisée dans [34]. Parmi celles ci, il y a lieu de distinguer celles qui sont fondées sur la théorie de Liapounov et les autres.

La théorie de Liapounov est très générale et très puissante. Néanmoins son efficacité dépend beaucoup du choix d'une "bonne" fonction de Liapounov et, étant fondée sur des conditions suffisantes de stabilité, elle ne donne accès qu'à une sous partie, généralement convexe, du domaine d'attraction d'un point d'équilibre.

Pour contourner la difficulté, Génésio et al [34] proposent une méthode fondée sur l'intégration à rebours du système différentiel à partir d'un ensemble de points situés dans le voisinage du point d'équilibre et appartenant au domaine d'attraction. Plus récemment ces méthodes ont été amendées par deux contributions [35,36] dans lesquelles les auteurs cherchent à construire un sous domaine d'attraction constitués par l'ensemble des points du systèmes situés à un temps fini du point d'équilibre stable considéré. Dans [36],le sous domaine est approché par un polyèdre et une application aéronautique est présentée.

Muni d'un moyen de calcul des comportements asymptotiques et en faisant l'hypothèse de l'hyperbolicité des solutions, il est possible de construire directement la frontière de stabilité en réunissant les variétés stables des solutions stationnaires instables situées sur la frontière cherchée [37].

A titre d'exemple, après avoir déterminé les deux points d'équilibre stable, (0., 0., 0.) et (-7.45, -7.45, -7.45), et le point d'équilibre instable (-2.45, -2.45, -2.45) du système suivant:

$$\dot{x}=-x+y$$
$$\dot{y}=0.1x-2y-x^2-0.1x^3$$
$$\dot{z}=-y+z$$

la méthode utilisée à l'ONERA permet de construire directement la frontière de stabilité du point d'équilibre stable (figure 23).

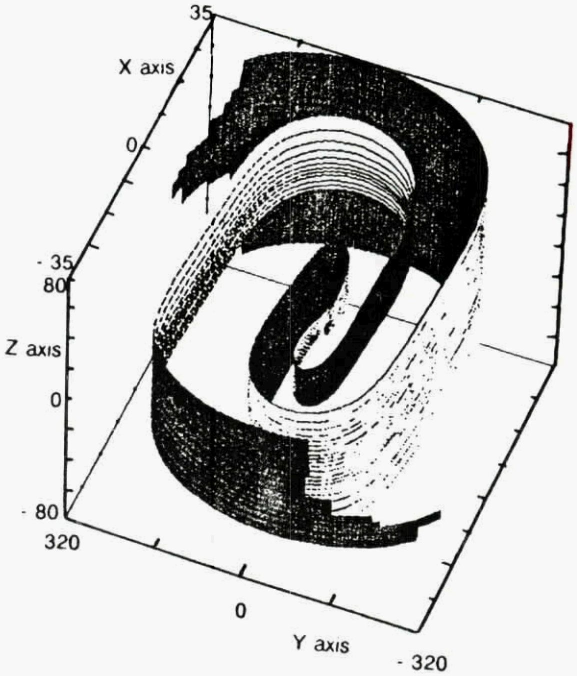


Figure 23 : Vue partielle de la frontière de stabilité.

Lorsque le système est de dimension plus élevée, la représentation graphique de la frontière de stabilité présente des difficultés. Cependant, sa projection dans des sous espaces particuliers fournit des informations utiles [38] (figure 24). A noter que ce résultat est similaire à ceux obtenus par une approximation du domaine d'attraction au moyen de polyèdres et présentés dans [36]

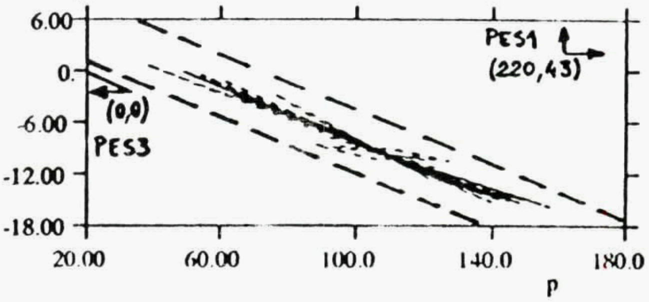


Figure 24 : Projection de la frontière de stabilité entre les deux points d'équilibre stables (PES1 et PES3) dans le plan (p, r)

9. CONCLUSION

Lorsqu'un avion évolue avec une forte dynamique ou/et à grande incidence, l'importance des non-linéarités des équations de la mécanique du vol et du modèle aérodynamique rend la prédiction du comportement de l'appareil délicate. Généralement, ces phases de vol sont étudiées au moyen de nombreuses simulations numériques avant de les aborder en vol.

Face à cette difficulté, de nombreux chercheurs se sont attachés à traiter ces problèmes de façon partielle en appliquant, de façon analytique et sur des systèmes simplifiés de dimension réduite, des méthodes classiques d'étude de la stabilité de systèmes d'équations différentielles non linéaires. Malgré l'intérêt de ce type d'approche pour appréhender un phénomène complexe, les hypothèses simplificatrices nécessaires nuisent souvent à la qualité du résultat.

Parallèlement à l'accroissement des capacités de calcul des ordinateurs, le développement de la théorie des bifurcations a montré la voie vers une méthodologie globale d'analyse du comportement asymptotique des systèmes différentiels non linéaires quelconques.

L'application de cette méthodologie aux avions permet d'expliciter de nombreux phénomènes très divers, simples et complexes, avec la même procédure de calcul et sans avoir besoin de recourir à des hypothèses simplificatrices et à des simulations exhaustives. Quant à la corrélation avec les essais en vol, elle confirme brillamment les prédictions (sous réserve, toutefois, d'utiliser un modèle d'avion de bonne qualité).

Dans le but d'affiner encore l'analyse en se rapprochant les plus possible de systèmes plus complexes, il apparait que cette méthodologie est susceptible d'aider également à la compréhension de l'action de chaînes de pilotage, même simples, sur le comportement global du système. Dans le même esprit, le recherche du domaine d'attraction des états d'équilibre stable permet d'aborder l'analyse des comportements transitoires.

De part ces capacités actuelles et potentielles, cette méthodologie d'étude globale des systèmes différentiels est déjà très utile pour analyser le comportement des avions. Dans l'avenir, elle devrait être appliquée à d'autres systèmes complexes tels, par exemple, des missiles, des hélicoptères ou des sous marins.

10. REMERCIEMENTS

Cette communication rend compte d'études en partie réalisées à l'ONERA dans le cadre d'actions de recherche soutenues par la DRET (Direction des Recherches et Etudes Techniques) et le STPA (Service Technique des Programmes Aéronautiques).

Je voudrais les remercier de leur confiance dans les travaux menés à l'ONERA et, tout particulièrement, d'avoir donné l'opportunité de montrer l'intérêt de ces études à travers une application très poussée sur un avion réel en collaboration avec Dassault Aviation, le centre d'essais en vol d'Istres et l'ONERA/IMFL.

11. REFERENCES

1. Phillips, W.H., "Effect of steady Rolling on Longitudinal and Directional Stability", NASA TND-1627, 1948.

2. Pinsker, W.J.G., " Critical Flight Conditions and Loads Resulting from Inertia Cross-coupling and Aerodynamic Stability Deficiences", ARC-TR-CP-404, 1958.

3. Hacker, T. et Oprisiu, C., "A discussion of the roll-coupling problem" *dans* "Progress in Aerospace sciences", Vol n° 15, Pergamon Press, 1974.

4. Kalviste, Y et Eller, B., "Coupled Static and Dynamic Stability Parameters", AIAA-89-3362, 1989.

5. Ross, J.A. et Beecham, L.J., "An approximate analysis of the nonlinear lateral motion of a slender aircraft (HP115) at low speed", ARC R&M 3674, 1971.

6. Padfield, G.D., "The application of Perturbation Methods to Nonlinear Problems in Flight Mechanics", PhD Thesis, Cransfield Institute of Technology, 1979.

7. Adams, W.M., "SPINEQ: A Program for Determining Aircraft Equilibrium Spin Characteristics Including Stability", NASA TM 78759, 1978.

8. Laburthe, C., "Une nouvelle analyse de la vrille basée sur l'expérience française sur les avions de combat" *dans* "Stall/Spin Problems of Military Aircraft", AGARD CP 199, Papier n° 15A, 1975.

9. Schy, A.A. et Hannah, M.E., "Prediction of Jump Phenomena in Roll-Coupling Manoeuvers of Airplane", J. of Aircraft, Vol. 14, avril 1977.

10. Schy, A.A., Young, J.W. et Johnson, K.G., "Pseudo-state Analysis of Nonlinear Aircraft Manoeuvers", NASA TP 1758,1980.

11. Ioos, G. et Joseph, D.D., "Elementary Stability and Bifurcation Theory", Springer Verlag, New York, 1980.

12. Guckenheimer, J. et Holmes P., "Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields", Springer Verlag, New York, 1983.

13. Kubicek, M. et Marek, M., "Computational Methods in Bifurcation Theory and Dissipative Structure", Springer Verlag, New York, 1983.

14. Guicheteau, P., "Etude du comportement transitoire d'un avion au voisinage d'un point de bifurcation" *dans* "Unsteady Aerodynamics - Fundamentals and Applications to Aircraft Dynamics", AGARD CP 386, Papier S10, 1986.

15. Keller, H.B., "Numerical Solution of Bifurcation and Nonlinear Eigenvalue Problems" *dans* "Applications of Bifurcation Theory", Academic Press, New York, 1977.

16. Kernevez, J.P. et Doedel, E.J., "Software for Continuation Problems in Ordinary Differential Equations", Technical Report, Applied Mathematics, CALTECH, 1985.

17. Jahnke, C.C., "Application of Dynamical Systems Theory to Nonlinear Aircraft Dynamics", Thesis, Engineering and Applied Sciences, CALTECH, 1990.

18. Guicheteau, p., "Application de la théorie des bifurcations à l'étude des pertes de contrôle sur avion de combat" *dans* "Combat Aircraft Manoeuvrability", AGARD CP 319, Papier 17, 1981.

19. Mehra, R.K., Carrol, J.V. et Kessel, W.C., "Global Stability and Control Analysis of Aircraft at High Angle of Attack", Technical Report, ONR CR215-248-1, 1978.

20. Hawkins, C.A., "Application of Bifurcation and Catastrophe Theories to Near Stall Flight Mechanics", AFIT/CR/NR 86-54T, 1985.

21. Zagaynov, G.I. et Goman, M.G., "Bifurcation Analysis of critical Aircraft Flight Regimes, ICAS-80-4.2.1, 1984.

22. Gonzales Blazquez, A.L., "Mathematical Modelleing for Analysis of Nonlinear Aircraft Dynamics",

Computers and Structures, Vol 37, N° 2, 1990.

23. Lowenberg, M.H., "Bifurcation Methods-A Practical Methodology for Implementation by Flight Dynamicists", ICAS-90-5.10R, 1990.

24. Casti, J.L., "Bifurcations, Catastrophes and Optimal Control, IEEE Trans. on Automatic Control, Vol. AC-25, N° 5, 1980.

25. Behtash, S. et Sastry, S., "Stabilization of Nonlinear System with Uncontrollable linearization", IEEE Trans. on AC, Vol. 33, n° 6, 1988.

26. Gao, H., Wang, Z.J. et Zhang, S.G., "A study of wing rock", ICAS-90-5.10.3, 1990.

27. Planeaux, J.B. et Barth, T.J., "High Angle of Attack Dynamic Behaviour of a Model of High Performance Fighter Aircraft", AIAA-88-4368,1988.

28. Guicheteau, P., "Bifurcation Theory in Flight Dynamics - An Application to a Real Combat Aircraft, ICAS-90-5.10.4, 1990.

29. Planeaux, J.B., Beck, J.A. et Baumann, D.D., "Bifurcation Analysis of a Model Fighter Aircraft with Control Augmentation", AIAA-90-2836, 1990.

30. Hirai, K., Iwai, M. et Ushio T, "Catastrophic Jump Phenomene in a Nonlinear Control System", IEEE Trans. on Automatic Control, Vol. AC-26, N° 2, 1981.

31. Holmes, P, "Bifurcation and chaos in a simple feedback control system", IEEE WP5-4:00, 1983.

32. Mitobe, K et Adachi, N., " Hopf bifurcation in an adaptative d.c. servo system", Int. j. Control, Vol 54, N° 4, 1991.

33. Guicheteau, P., "Application de la théorie des bifurcations à des systèmes non linéaires particuliers" *in* "Etude de mécanique du vol relative à la supermanoeuvrabilité", RT ONERA n° 45/5148 SY, 1989.

34. Genesio, R., Tartaglia, m. et Vicino, A., "On the estimation of Asymptotic Stability Regions: State on the art and New Proposal", IEEE Trans. on Automatic Control, Vol. AC-30, n°1, 1985.

35. Skowronski, J.M. et Guttalu, R.S., "Real-time Attractors", Dynamic and Stability of systems, Vol. 5, n° 4, 1990.

36. Piaski, M.L. et Luh, Y.P., "Nonconvex Polytope Approximation of Attracting Basin Boundaries for Nonlinear Systems", AIAA-90-3512-CP, 1990.

37. Chiang, H.D., Hirsh, M.W. et WU, F.F., "Stability Regions of Nonlinear Autonomous Dynamical Systems", IEEE Trans. on Automatic Control, Vol. AC-33, 1988.

38. Guicheteau, P., "Domaine d'attraction d'un point d'équilibre" *in* "Mécanique du vol aux grandes incidences", RT ONERA n° 47/5148 SY, 1990.

DECOUPLING OF AIRCRAFT RESPONSES

David J. Moorhouse
Wright Laboratory
Wright-Patterson AFB OH 45433-6553
United States

Introduction

The theory of aircraft stability is at least as old as powered flight. The original impetus was to design the configuration to have acceptable characteristics. Flight control technology was next applied to tailor, i.e. improve, the classical response characteristics; a yaw damper being a simple example. More recent flight control technology has provided the means to change the character of the aircraft response so that it no longer exhibits the classical behavior of a short period, a phugoid, a Dutch roll, a roll mode and a spiral. With additional control effectors and the use of feedback there is an infinite number of ways to change the basic stability and response characteristics.

Pilot preferences are also well established, however, so that care is needed in implementing any theoretical improvements. The technology exists to provide the pilot with the capability to individually control all six degrees of freedom. Obviously, a requirement for a pilot to integrate six different control effectors would be likely to increase his workload. In that sense decoupling all six axes would be detrimental. Conversely, any unwanted coupling that can be eliminated should reduce the pilot task.

The Wright Laboratory has had a series of flight demonstration programs that have evaluated new technologies in the most realistic tasks. First an F-16 was provided with the additional control effectors to allow independent control of all six degrees of freedom. A variety of control modes was mechanized, so that the pilot could evaluate both coupled and decoupled modes. Following a subjective in-flight assessment, a ground-based piloted simulation experiment was performed to evaluate all modes for both offensive and defensive combat use. Second, an F-15 was modified to facilitate precise landing in adverse conditions. A special short landing mode was implemented to feature decoupling of airspeed and glideslope responses plus the integrated coupling of direct lift and sideforce control. This paper presents results from both these programs, to illustrate the benefits of either decoupling or a new coupling of aircraft responses.

Background

Very shortly after the Wright Brothers flew, aircraft designers standardized on a configuration of aft-mounted elevator (then stabilator) and rudder and wing-mounted ailerons. Together with the propeller or jet-engine thrust, this arrangement gives four effectors to control six degrees of freedom. In matrix notation the equations of motion are written

$$\dot{x} = A'x + Bu$$

The state vector, x, contains some form of the basic six degrees of freedom. The matrix A' contains the stability derivatives which can include both natural airframe characteristics and terms due to stability augmentation. The control vector, u, has the four terms discussed above plus two zero terms for conventional aircraft. Control of the z- and y-axis is then implicit through changes in the angular states. This is obviously typified by using pitch attitude changes to control flight path. This is natural to pilots, but it is possible to postulate conflicting requirements for nose pointing independent of flight path control in combat. We can complete the u vector with the addition of Direct Lift Control (DLC, z-force) and Direct Sideforce Control (DSFC, y-force). Next, we can devise an augmentation scheme (theoretically, at least) to make A and B essentially into diagonal matrices. The pilot could then have control of each axis independent of all the others:

$$\dot{x}_i = a_i x_i + b_i u_i$$

This can be considered perfect decoupling, although it remains to be seen whether a pilot would consider it perfect. In addition, given the same control effectors an infinite number of ways exist for integrating (i.e. coupling) the additional capabilities to produce unconventional responses.

Direct Force Modes for Combat

Each of the flight control modes shown in Figure 1 was evaluated qualitatively in the flight programs of References 1 and 2. The A_N mode changes the flight path without the conventional pitching motion to change angle of attack. The change in flight path was felt to be essential, with the potential benefit of some deception in the lack of pitch cues. The α_1 mode changes pitch attitude without any change in flight path angle. This was assessed as an offensive mode, the only possible defensive benefit would be through deception, but it would be vulnerable unless combined with other maneuvering. Finally the α_2 mode commands vertical velocity without changing pitch attitude. For short-duration control inputs this mode should be similar to the A_N mode. Each of the above three modes decouples the pitch and normal axes. An additional Maneuver Enhancement mode used DLC to "quicken" the conventional load factor response to pitch changes. Overall, none of the longitudinal modes were considered to have any strong defensive potential for a gun encounter.

The A_Y mode controls directional flight path with wings level and zero sideslip angle. This mode has been shown to be advantageous for ground attack, however, there is some possible defensive potential from the ability to turn without first banking the airplane. The β_1 mode changes yaw attitude without changing flight path and is equivalent to the α_1 mode. The β_2 mode commands lateral velocity without changing yaw attitude. This mode was felt to have more defensive potential than any other mode, especially if combined with other maneuvers. This mode could be commanded through the rudder pedal, with conventional responses available through the side-stick. As an example, the pilots particularly liked the ability to translate out of a banked turn.

Following the initial assessment above, it was decided to evaluate the modes quantitatively in a motion-based simulator. The task was to track as aggressively as possible a computer-generated target which maneuvered according to either computer-generated or pre-recorded motions using similar control modes. The pilots were required to give a pilot rating according to the Cooper-Harper scale, i.e., rate <u>task</u> performance not the aircraft configuration or flight control mode. Thus, by comparing appropriate runs a difference in pilot rating could indicate an effect of target motions on the tracking task or an effect of the tracking aircraft's flight control mode. Pilot comments were also used as aids in interpreting the pilot ratings.

The overall objective of the simulation effort was to generate basic data on pilot use of the different modes. Based on the prior assessment plus consideration of the limitations of the visual system, the target was programmed to move in either vertical or lateral translations with various acceleration levels. In this way the experiment was kept reasonably "pure" - tracking the vertical target motions with the longitudinal modes and the lateral motions with the directional modes.

The data can be viewed in one of two lights; by making a mode-by-mode comparison of tracking performance, pilot rating and commentary for a particular target maneuver or set of maneuvers (e.g., all elevation targets) we can evaluate the effect of direct force control capability in the tracker airplane on offensive performance. Conversely by comparing the type of target (i.e., elevation or azimuth) and target authority with tracking performance, ratings and commentary we can determine the impact of target motion on task performance and workload. An increase in rms tracking error, a degradation in ratings, or both indicates the target's defensive effectiveness.

<u>Piloted Simulation Results: Offensive</u>

Figures 2 and 3 show tracking performance for various elevation and azimuth target acceleration capabilities, respectively, for each tracker mode. For the baseline case the pilots were only able the keep the target within the pipper (25 mils radius) for a quarter of the 12 runs against elevation targets (Fig. 2). Both the A_N and ME modes show improved performance over the baseline mode for vertically-accelerating targets. The pilots were able to maintain the target within the pipper for almost all of the runs when using the A_N or ME modes. Pilot commentary states "the mechanization through the sidestick" and "the quickened g response" of the ME mode, and "the excellent dead beat pitch response in fine tracking" were the primary reasons for the improvement in performance. For elevation target capabilities of less than 1g pilots tracked with the A_N mode alone. This was possible since mode authority at the flight condition was + 1g normal acceleration. For larger target authorities pilots had to blend in baseline response through the sidestick controller, though pilot commentary states "the pilot had no problem blending the A_N mode through the trim button and conventional baseline response through the side-stick". Even so, pilots "preferred mechanizing the direct force control through the conventional sidestick controller (i.e., the fully integrated ME mode) over having to use a secondary controller (i.e., the A_N mode)".

Tracking results in Fig. 2 show a significant improvement when using direct lift control modes as compared to the baseline mode. This observation is supported by the pilot ratings (Figs. 4a and 5a) and comments. Almost all of the runs for the A_N mode against various elevation targets are rated Level 1 (Fig. 4a). Level 1 indicates "the pilot has satisfactory task performance and workload". Similar results are found for the ME mode (Fig. 4a). For the baseline mode a majority of the ratings are Level 2, indicating "the pilot can still accomplish the task but he must increase his workload to do so". The A_N mode on the other hand, had a "deadbeat pitch response" resulting in a "stop where you release", removing the problem of having to put in additional control to stop the pitch response from overshooting the target. The ME mode, which automatically blends the direct lift control, has the same favorable characteristics as the A_N mode, resulting in a decreased pilot workload (i.e., lower pilot ratings) evident from the pilot commentary.

Figs. 3, 4b and 5b show the effects of direct sideforce control capability on tracking performance and pilot rating for an azimuth step target with various lateral acceleration capabilities. Again the ME mode shows an overall improvement in tracking performance as compared to the baseline. In order to track the azimuth target the pilot must bank the airplane and then track the target with pitch. Hence, tracking an azimuth target is essentially a pitch tracking task. The ME mode "improves the baseline pitch characteristics", resulting in improved performance and ratings. The ratings and performance for the azimuth target tasks are generally worse than for the elevation targets due to an increase in task complexity over the elevation target tracking task. The A_y mode results indicate the pilot can only maintain the pipper near the target for 1 out of 6 runs. Pilot ratings for the A_y mode task performance lie between 6 and 7 (Figs 4b and 5b). A rating of 6 means "adequate performance requires extensive pilot compensation" with the airplane having "very objectional but tolerable deficiencies". For a pilot rating of 7 the pilot "cannot achieve adequate performance regardless of pilot workload". A "major deficiency" is present requiring improvement. Pilot commentary for the baseline and ME mode against azimuth targets disclose that "the pilot must correctly blend 3 control inputs to follow the target motion". This increases the pilot workload resulting in the degraded ratings.

Pilot commentary indicates the mechanization of the modes was "satisfactory". A_y mode on the rudder pedal was a good way to mechanize that mode since it is "consistent with the way pilots are trained to fly". Pilots use the conventional rudder for directional heading corrections, and the A_y mode also changes heading. Comments on the baseline and the ME mode versus azimuth targets are identical to the comments for the modes versus elevation targets, although ratings are no better than for the baseline aircraft.

In summary, results from the offensive evaluation of the direct force control mode show that the blended direct force mode in the tracking airplane (i.e., ME) improves the tracking performance while lowering the workload. The open-loop direct-lift mode shows similar results, with commentary (but not rms error or pilot rating) indicating a further improvement in fine tracking. The open-loop sideforce mode does not show the same tracking improvement compared to the baseline as does its longitudinal equivalent. This is partly due to an increased task complexity for the azimuth target tracking task. Controller mechanization of the ME mode blended through the sidestick received very favorable pilot comment, as did mechanizing the A_y mode through the rudder pedals. The A_N mode on the button was acceptable, but pilots preferred to command all the direct force modes through the conventional controllers, removing the control harmony problems frequently encountered when using a secondary controller.

Piloted Simulation Results: Defensive

Since the same response capabilities were programmed into the target motions, if the target motions cause a degradation in tracking performance or pilot rating, we interpret this to indicate defensive potential in that target motion. Figs. 4 and 5 are plots of pilot rating versus increasing target acceleration capability for an elevation target and an azimuth target. There is a lot of scatter in the results, as may be expected - on a given run there is an element of luck in whether the pilot acquires and stays with the target. In addition, there is no apparent difference due to the mode of the tracking aircraft, so the results are included together in Fig. 6.

Vertical translation magnitude shows no consistent effect on pilot rating (Fig 4a). Vertical target accelerations from 1/2 to 4g produce pilot ratings between 2 and 5 for all longitudinal modes, with the majority remaining Level 1. Even the worst rating indicates that "acceptable performance can be achieved", although the workload of the pilot tracking that target has to be increased. This is also supported by the mean pilot ratings of both pilots for all elevation targets for a particular mode (Fig. 5a). Typically the elevation target motions were considered as "a simple, single-axis pitch tracking task as long as no inadvertent roll/yaw inputs are made". Pilots learned that minor roll/yaw inputs created enough of a lateral offset from the target plane of motion to dramatically increase the complexity of his task. To correct the lateral offset they sometimes had to blend 2 and 3 control inputs to get back in the target's plane of the motion. Another consideration is that maximum elevation target acceleration exceeds any probable direct lift control authority, even though it is still less than the normal load factor capability of modern fighters. The implication is that pure elevation maneuvers by the target are not going to be effective defensively.

By contrast, Figs. 4b and 6b show that alternating azimuth steps are much more effective defensive maneuvers. Scatter again is evident, with the ratings varying from 3 to 7; however, a third of the results are in the Level 3 region (Fig. 6b). This worst rating indicates, of course, that acceptable performance cannot be achieved regardless of how hard the pilot works. Pilot comments also indicated an "increase in task complexity" relative to tracking the elevation step targets. The difficulties of tracking were expressed as: "the azimuth target forced me to simultaneously blend the three axes of airplane response in pitch, roll and yaw in order to match the target motions". Thus although the experiment was simplified into a single-axis target maneuver, the pilots were faced with a multi-axis tracking task and found the workload often unacceptable. Finally, and rather surprisingly, as little as 1/2g azimuth target acceleration produced pilot ratings of 7 while increases up to 4g produced no further degradation in pilot ratings.

It is interesting to note that direct force control capability in the tracking airplane shows no appreciable improvement in task performance or pilot rating against the azimuth targets although this capability gives improvement in both performance and pilot rating against elevation targets. Hence, we interpret these results to indicate that lateral maneuvering (i.e. y-axis translation) effectively increases defensive capability due to the drastic increase in complexity of the tracking task for the pursuing airplane.

<u>Precision Landing</u>

The preceding section indicates the increases in workload that are caused by the pilot having to coordinate multiple control inputs. In applying control technology to improve handling qualities, it would appear fruitful to address any conditions that require a pilot to coordinate his inputs in more than one axis. A good example of this requirement is the coordinated stick and throttle inputs required to achieve a precision touchdown. It is instructive, first, to consider an effect from theory of constrained stability. For a conventional aircraft the longitudinal motion can be separated into two distinct modes. The short period mode consists of angle of attack and pitch attitude variations and is well damped, while the phugoid consists of lightly damped variations in airspeed and pitch attitude. The common approximation for the phugoid mode gives a system damping of

$$2\zeta\omega_n = -X_u$$

from which the time constant of the oscillatory phugoid amplitude envelope is

$$\tau_p = -2/X_u$$

Now, if the pitch attitude is held constant, the normal longitudinal equations reduced to:

$$(s - X_u)u - X_\alpha \alpha = 0$$
$$-Z_u u + (s - Z_\alpha)\alpha = 0$$

The result is that the classical modes became an approximate "angle of attack short period" and an approximate "airspeed phugoid" mode. The time constant of this aperiodic airspeed mode is

$$\tau_a = -1/X_u$$

i.e. half the time constant of a classical phugoid. The two different responses are compared in Figure 7. It is seen that even with the reduced time constant the aperiodic mode does not give the pilot any appearance of airspeed stability, whereas the oscillatory mode shows a significant initial reduction in the airspeed perturbation. Because of the frequency separation of the two modes, the normal pilot action of tightly controlling pitch attitude to maintain the glideslope is equivalent to holding attitude constant and can be expected to generate the aperiodic airspeed mode. It is suggested that the apparent lack of airspeed stability in the aperiodic mode induces overcontrol, i.e. larger power changes than necessary to correct an airspeed transient.

Thus, when we look at the effect of constraining the pitch axis, the theoretical result is that the speed axis is more stable. The practical effect is that there is less apparent stability to the pilot and there is a natural tendency to produce over control and coupling of the two axes.

This is typified by the requirements for a carrier landing. The requirements are met with carrier-based guidance, and either a manual or coupled approach using a flight control system known as an Approach Power Compensator System (APCS). Reference 3 analyzes and documents development problems with APCS: "Trial and error is the essence of the design procedure now employed in APCS development. While such a procedure is usually to be avoided and certainly not condoned, it has proven necessary and reasonably effective under the circumstance". Each new aircraft is the subject of intense development to meet the operational requirements. The current APCS implementations are angle-of-attack based which implies coupled airspeed and flight path responses. Quoting from Reference 3: "The basic manual control technique, involving operation of both stick and throttle, emphasizes the necessity of using thrust plus attitude as the combined means of maintaining reference airspeed and angle of attack on the glide slope. The (old) notion of using throttle to control altitude and nose attitude or stick input to control airspeed is pointed out as not being wholly valid in all situations along the approach". Navy pilots do confirm that their training is concentrated on learning the required coordinated inputs. Note, however, that the APCS throttle inputs are to an angle-of-attack reference, as are conventional autothrottles. There is an obvious steady-state correspondence, but such mechanizations do nothing to reduce the short term coupling tendencies.

A different approach was directed for the Wright Laboratory's STOL & Maneuver Technology Demonstrator (S/MTD, References 4 and 5), because of a primary requirement to land in less than 1500 ft. In-flight thrust reversing provided a high bandwidth control capability in the x-axis. Airspeed was used as a direct feedback parameter, so that airspeed was held constant independent of stick input. Once the aircraft is "on speed" there is no reason for pilot input into the speed axis; at the same time a speed change commanded by the pilot does not produce a pitch transient. The above approach to decoupling pitch and airspeed control was implemented in a Short Landing (SLAND) mode of the S/MTD control laws. Reference 4 documents results of a piloted simulation effort performed to define the optimum value of pitch axis bandwidth for tracking the glidepath that was used. This mode was designed using multivariable techniques (see Reference 6). Figure 8 shows the decoupling achieved using different orders of compensator.

The conventional aircraft configuration produces a downforce from the stabilator to rotate the aircraft to generate lift to climb. The flight path response is non-minimum phase, i.e. initial motion is opposite to the final motion. For the SLAND mode, therefore, Direct Lift Control was incorporated to produce a minimum-phase flight path response. It can also be shown that this application reduces the conventional lag of flight path response to pitch rate. This also should make glideslope control easier for the pilot. The S/MTD glideslope control, therefore, incorporated the latest control technology to make it the best it theoretically could be for manual control. This makes the evolution of pilot reactions very interesting.

First flight of this mode produced very negative pilot reactions, especially concerning ride qualities. It was characterized as a carnival ride, not like an airplane, and he even felt something that seemed like direct lift! To the engineers it was obvious that the pilot was high gain and exciting the unusual characteristics.

A later flight was performed in reported winds of 18 kts gusting to 24 kts. Even in these severe conditions the pilot reported that it was not as bothersome and that precise touchdowns were possible (with pilot compensation!)

Then back-to-back landings were made to measure touchdown dispersion in the new (SLAND) mode and also in a control mode with conventional flying qualities. With this comparison, done in turbulence conditions, the pilot found that SLAND mode was much easier and offered a significant decrease in pilot workload over conventional characteristics. Even so, the pilot still commented on the "unusual flying qualities".

Two flights later, wind shear encounters gave another boost in pilot confidence. On back-to-back landings, a wind shear was encountered about 200 ft above the runway as the pilot was concentrating on the touchdown point for the precise landing. With the conventional mode, the aircraft dropped approximately 75 ft before the pilot arrested the sink rate by applying power. The pilot also commented that he was continually working to control airspeed and angle of attack throughout the landing. In the SLAND mode, the control system quickly arrested the disturbance caused by the wind shear with the aircraft losing less than 25 ft.

The final flight of the program completely convinced the pilot of the value of the new technology. A short landing was performed at night using only on-board guidance (Reference 7), no runway lights and simulating a breakout from weather at 200 ft above the runway. After this successful demonstration, the pilot said that the integration of the total system was wonderful and certainly much better than the sum of the individual would indicate. He also commented that day, VFR testing was not indicative of the system's true worth - the approach to a totally black airfield was no more difficult that a simple video game.

In summary, pure decoupling of airspeed control from glideslope control was designed to simplify the landing task in adverse conditions. All flight test programs buildup the severity of the tasks, but the initial pilot reactions from flights in benign conditions were very negative. In a demonstration program, it was easy to continue without questions or reviews and reach the final successful approval.

Conclusions

For conventional, fixed-wing aircraft the stability and control characteristics are well established. Flight control technology is able to alter or tailor these characteristics with an infinite number of possibilities. This paper has discussed two different Wright Laboratory applications of control technology. For combat use, direct force control modes were beneficial offensively when they were integrated to simplify pilot control actions. Decoupled responses were not liked. Such modes were beneficial defensively when they could be used to generate out-of-plane maneuvers and force a more complex task on the attacker. For this application, the defensive maneuver could be open loop so that a decoupled mode would be acceptable.

The second application consisted of pure decoupling of airspeed control from glideslope control for precise landing. Simple analyses show that this approach improves the "apparent" stability of the aircraft response and eliminates the requirement for the pilot to coordinate inputs in two axes. Initial pilot reactions were very negative until some experience was acquired. Landings in adverse conditions finally produced enthusiastic pilot comments about the reduced pilot workload that accompanies not having to coordinate two different control inputs.

Finally, this paper has attempted to show how pilot likes and dislikes can guide the development of control strategies. The theoretical possibilities are unlimited for tailoring aircraft stability, in its broadest context. The pilot is still the final judge.

References

1. McAllister, J.D., et al, "Fighter CCV Phase II Report - Detail Design", AFFDL-TR-76-119, October 1976

2. Barfield, A.F., B.W. VanVliet and D.C. Anderson, "AFTI/F-16 Advanced Multimode Control System Design for Task-Tailored Operation", AIAA Paper No. 81-1707, August 1981

3. Craig, S.J., R.F. Ringland and I.L. Ashkenas, "An Analysis of Navy Approach Power Compensator Problems and Requirements", Systems Technology Inc., STI Technical Report No. 197-1, March 1971

4. Moorhouse, D.J., D.B. Leggett and K.A. Feeser, "Flying Qualities Criteria for Precise Landing of a STOL Fighter", AIAA Paper No. 89-3390, August 1989

5. Moorhouse, D.J., J.A. Laughrey and R.W. Thomas, "Aerodynamic and Propulsive Control Development of the STOL and Manoeuvre Technology Demonstrator", AGARD Conference Proceedings AGARD-CP-465, October 1989

6. Moomaw, R., and Lowry, D., "Application of Multivariable Control to the STOL and Maneuver Technology Demonstrator", AIAA Paper No. 87-2403, August 1987

7. Moorhouse, D.J., "Design and Flight Test of On-Board Guidance for Precision Landing", AIAA Paper 91-2641, August 1991

	OFFENSIVE		DEFENSIVE	
	PRIORITY	PILOT COMMENTS	PRIORITY	PILOT COMMENTS
Direct Force: A_N	1 (High)	"Well damped, good authority for tracking, quick response, stop on release, no oscillation".	2 (High)	"change in flight path essential".
ME	2 (High)	"Well damped, command through sidestick, cut down characteristic pitch bobble, gross acquisition is good".	1 (High)	"good path response, ease of command via sidestick".
A_y	3 (High)	"deadbeat response, A_y control via the rudder pedal is satisfactory".	3 (High)	"advantage in turn without bank, no visual cues, unconventional".
Fuselage Pointing: α_1, β_1	4 (Low)	"potentially high priority if mechanized automatically, poor as currently mechanized and controlled".	5 (Low)	"does not change flight path might be good for deception if large enough pitching/yawing motion can be generated, vulnerable unless combined with path changes".
Translation α_2, β_2	5 (Low)	"authority limited, authority mismatch due to greater authority in baseline response, might be good vs. a/c with similar capability".	4 (High)	"if authority is attainable, unconventional response coupled with minimal visual cues, lateral motions have more potential than longitudinal".

Figure 1. Pilot Assessment of Coupled/Decoupled Modes

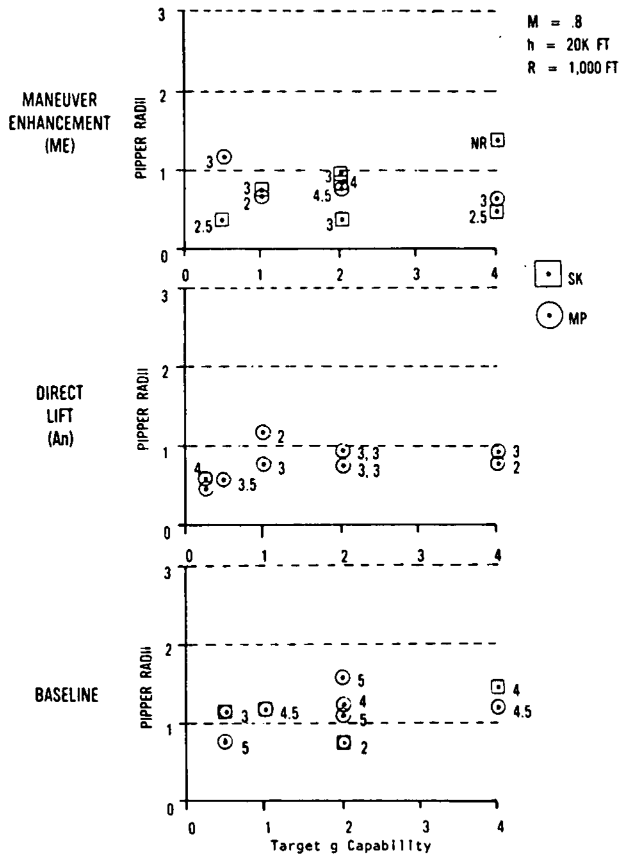


Figure 2. Tracking Error vs Elevation Targets

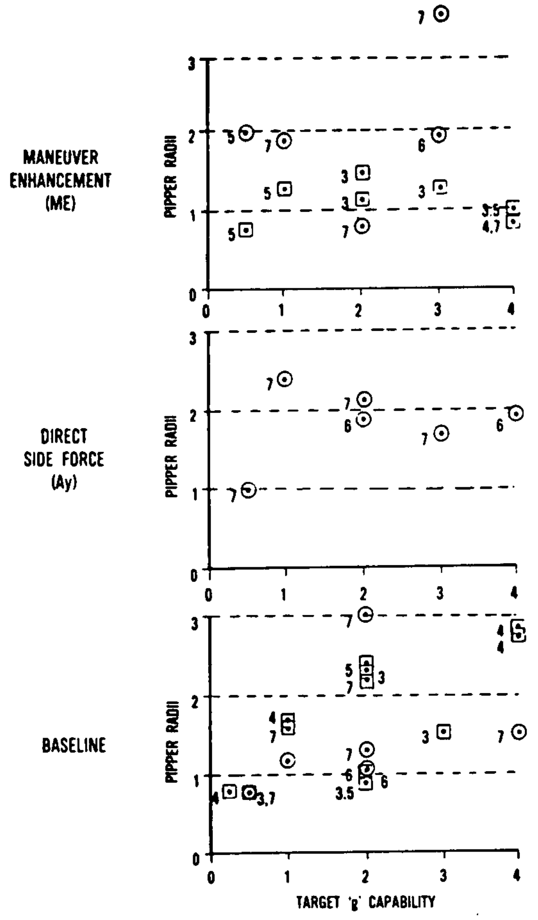


Figure 3. Tracking Error vs Azimuth Targets

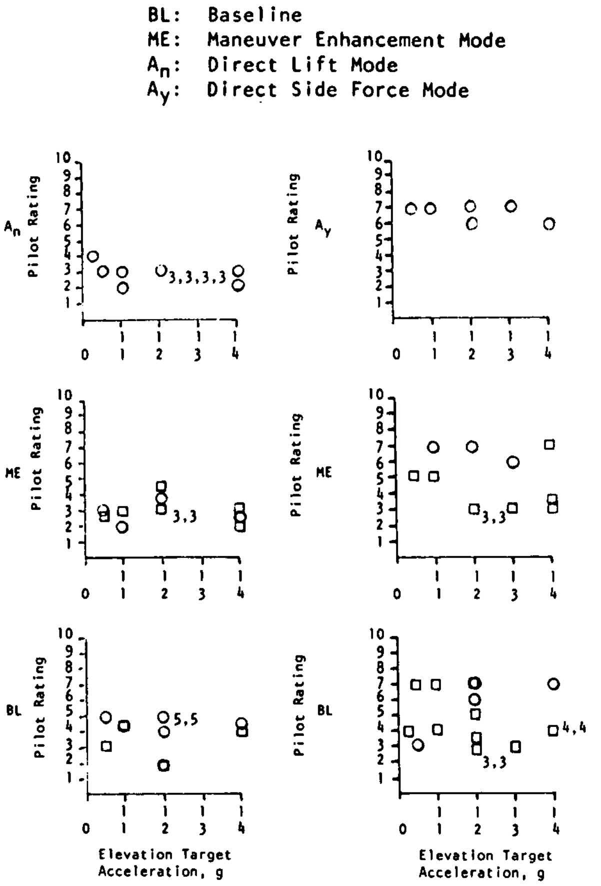
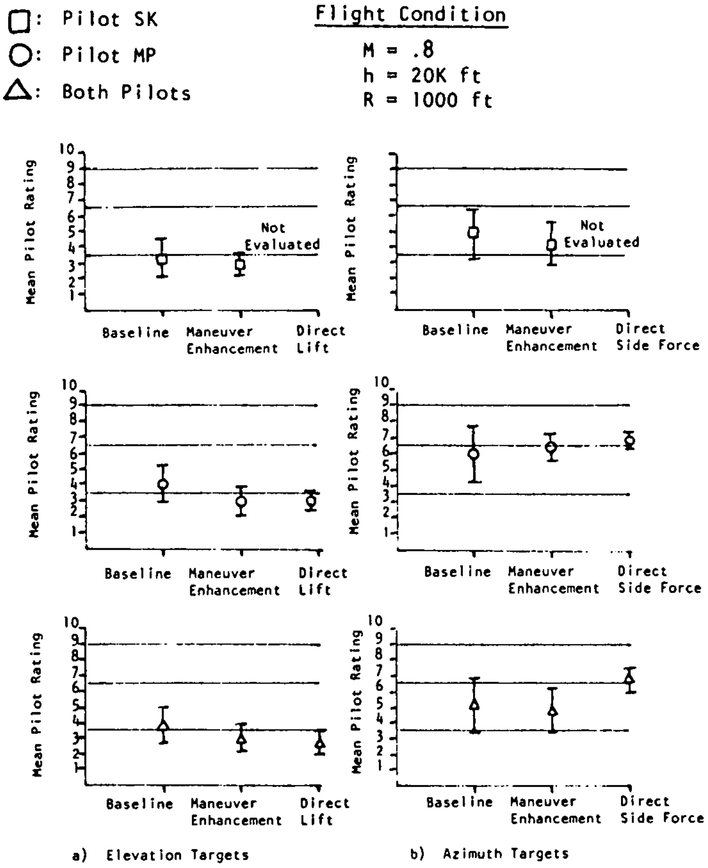


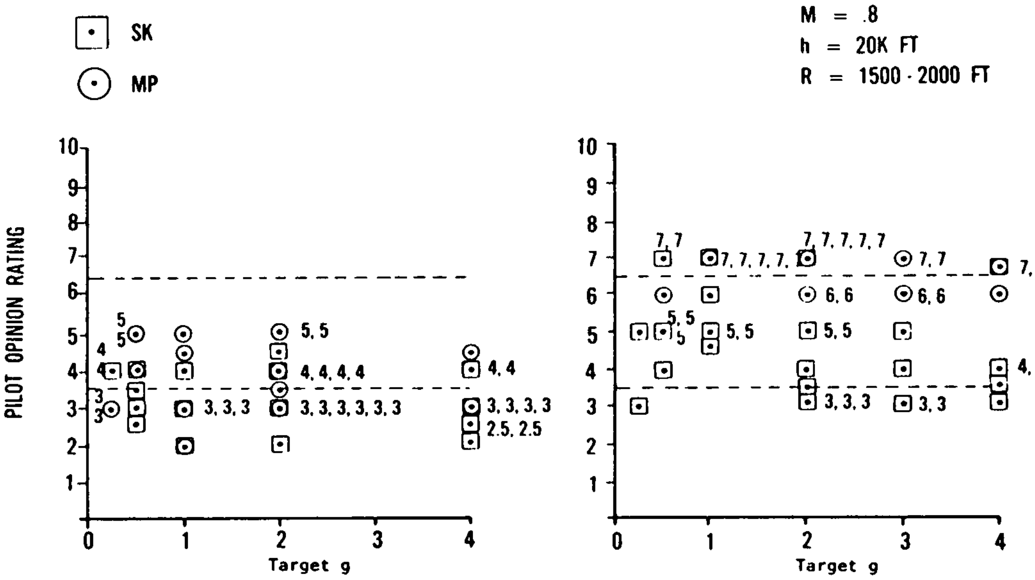
Figure 4. Pilot Rating vs Target Acceleration



a) Elevation Targets

b) Azimuth Targets

Figure 5. Pilot Rating vs Tracker Mode



a) ELEVATION STEP TARGET VS BASELINE, ME, An

b) AZIMUTH STEP TARGET VS BASELINE, ME, Ay

Figure 6. Pilot Rating vs Target Acceleration

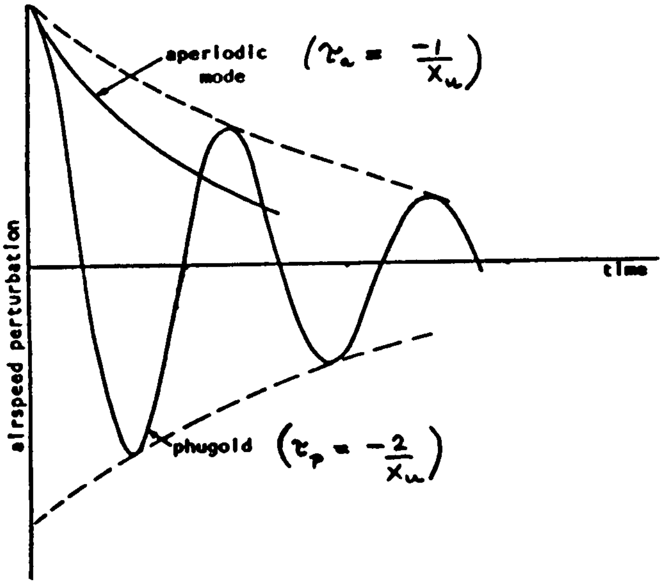
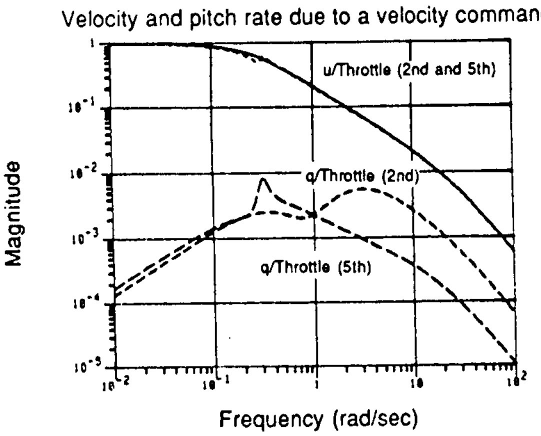


Figure 7. Comparison of Periodic & Aperiodic Responses


Velocity and pitch rate due to a velocity command

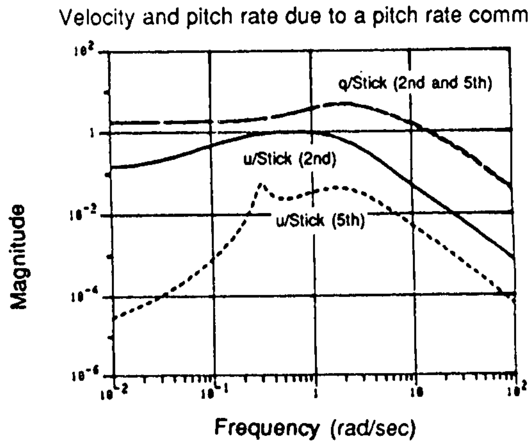

Velocity and pitch rate due to a pitch rate command

Figure 8. Decoupling with Five- and Two-State Compensators

Discussion

SESSION II - QUESTIONS & ANSWERS (PAPERS 8, 10, 11)

PAPER 8: G BERKOOS

Question:

Two fundamental characteristics of active control are measurement error and delay in applying control forces. Does your approach to controlling the flow take these factors into account?

Answer:

Noise comes into the model in a very natural way, namely through the tree boundary term, therefore, as long as noise is not overwhelming we do not think it will be a problem. The delay in applying central forces is accounted for in the model that gives us the control vector in phase space, (neglecting the computation time which will be much smaller than the turbulent time scales).

Question:

In order to implement your control scheme for a turbulent boundary layer, you have to sense where you are in the phase space of your loci dimensional dynamical model. Could you outline how you go about doing this?

Answer:

We plan to have an array of hot film wall mounted sensors. These sensors give the wall shear rate. The shear rate footprint is then decomposed into the footprints of the different eigen functions and thus an estimate for the phase space point is obtained.

PAPER 10: Ph GUICHETEAU

Question:

Can you predict all spin modes of an aircraft? Vertical wind tunnel tests or flight tests can never be sure!

Answer:

The methodology can predict almost all (because of the absence of results about asymptotic states for a set of general nonlinear differential equations) behaviour of the used model. Flight dynamics experience on several aircraft help to look for all the spin modes. Obviously, a vertical wind tunnel tests is not failsafe and this is one of the tasks of aerodynamicists. Also, flight tests are not completely trustworthy, but, if a 'catastrophic state' found with numerical prediction has not been exhibited by the aircraft during the flight test, one must be careful as under special conditions (in operational operations, for example) the real aircraft may exhibit this behaviour.

Question:

Which were the four unsteady aerodynamic coefficients in your set of 12 parameters?

Answer:

The four unsteady coefficients are:

- $C_{z\,unst}$ (lift)
- $C_{l\,unst}$, $C_{m\,unst}$, $C_{n\,unst}$ (moment coefficients)

Question:

How critical to your spin predictions is the accuracy of aerodynamic coefficients?

Answer:

The accuracy needed for 'good' predictions is not the same for all the coefficients. Some of them are very sensitive. We have to study this problem using aerodynamic coefficients as parameters, but the results are not within the scope of this presentation.

Question:

Is asymptotic stability analysis sufficient, or can the aircraft get in trouble before?

Answer:

It is not sufficient to study asymptotic states, but it is a first step. Transient motions are now being studied at ONERA.

Question:

Can you study the behaviour of a stable or unstable point to find the rates? A slow stability may be unattainable, and a slow instability may be flyable.

Answer:

For strong stability or instability, transient behaviour in the vicinity of the equilibrium point is provided by eigen value analysis and associated eigen vectors. For weak stability states or bifurcation points nonlinear approximations are used to study or predict the motion (see [14]).

In practical situations, the designer must choose his own boundary of stability (which is different from mathematical definition) considering handling quality criteria (for aircraft). However, he must be careful when reducing the domain of interest.

Question:

How certain are you about spin type predictions which were not measured in a lab?

Answer:

We are not very confident about spin modes which are not measured in a wind tunnel. To reduce the uncertainty, the aerodynamic model has been measured over a wide range of the state domain:

$- 10^\circ < \alpha < 90^\circ$
$- 40^\circ < \beta < 40^\circ$

and numerous values of angular rates representative of angular rates encountered during spin (see [28]). The inverted spin modes were out of the scope of the study.

PAPER 11: D J MOORHOUSE

Question:

There was a preceding aircraft which could land within 500m: the SAAB Viggen. It had a high-sink undercarriage to land without flare, within 105m of the runway threshold. Did the F-15 landing approach have flare? This makes precision touchdown much more difficult.

Answer:

The STOL and manoeuvre F-15 did not perform a straight-in approach, it did flare and the flare guidance was programmed on the head-up display.

Question:

A difficulty with the Viggen was control on the ground. At what speeds could you use thrust reverser on a wet runway? How was control on the ground exercized?

Answer:

Thrust reversing was scheduled to prevent hot gas ingestion, but was active down to 40 knots. A good paper was published on the Viggen problems and we used some of the lessons learned, such as YAW rate feedback to nosewheel steering. For the STOL F-15, direct sideforce was commanded by the rudder pedals on approach, after touchdown sideforce was commanded by lateral stick into the wind (natural pilot action). This reduced the crosswind effects by 50%.

Question:

We saw a PIO of the F-15 on the ground. More PIOs continue to appear (e.g., YF-22) after having discussed the topic for a decade and claimed to have found the solution. Were the time relays in the digital system above 120ms? If not, what caused the PIO? How was it cured?

Answer:

First, I have strong prejudice that equivalent systems time delay should be less than 100 msecs. There was no PIO in the STOL F-15, the film showed a large pitch oscillation after touchdown that was caused by unsteady ground effect/jet interactions. It was cured by adding damping through pitch rate feedback.

Question:

In this F-15 there were many possibilities for control, e.g., pitch could be controlled using thrust vectoring, canards, tailerons, etc. How did you select the combination of controls for each task? Judgement or some kind of systematic analysis of all possiblities?

Answer:

There were 22 control effectors, but not all were available in all flight regimes. The integration was done more by judgement than a rigorous optimization procedure.

COMMENT ON PAPER 11:

While reporting results as a function of roll rate (P_s) and roll angle (ϕ_s)for fixed sting angle of attack (α_s) completely specifies the variable motions of the wind tunnel experiments, it is suggested that a better understanding of the flow physics would be gained by showing results as functions of the wing angle of attack (α), sideslip angle (β), and associated rates ($\dot{\alpha}, \dot{\beta}$) as well. Static effects are best viewed in terms of wing α and flow asmmetry (β), while dynamic effects are separate functions of rotary motion (P_s) and flow adjustment lags ($\dot{\alpha}, \dot{\beta}$). [The latter effects are analogous to the distinct effects of q and $\dot{\alpha}$ in longitudinal motion.] Cross flow effects are quite different on left and right wing panels -- in fact, for 65° sweep angle, the leeward wing experiences reverse flow for high α_s and ϕ_s; this would be more apparent in the (α, β) description.

BOUNDARY LAYER TRANSITION : PREDICTION AND WIND TUNNEL SIMULATION

D. Arnal
CERT/ONERA
Aerothermodynamics Department
2 avenue E. Belin
31055 TOULOUSE CEDEX - FRANCE

SUMMARY

This paper gives a survey of theoretical and experimental results related to the problem of boundary layer transition ; emphasis is given on applications of practical prediction methods. In the first part of the paper, it is shown that the linear stability theory can provide a good estimate of the transition location if the free stream disturbance level is low enough ; the difficulties to properly simulate free flight conditions in ground facilities is underlined. The second part of the paper is devoted to the problem of boundary layer tripping in the presence of large external disturbances ; in this case the linear theory no longer applies and empirical criteria need to be developed.

1 INTRODUCTION

Since the classical experiments performed by O. REYNOLDS (1883), the instability of laminar flows and the transition to turbulence have maintained a constant interest in fluid mechanics problems. This interest results from the fact that transition controls important hydrodynamic quantities such as drag or heat transfer. The objective of this paper is to give an overview of the transition problems, and of the prediction methods which are currently available for engineering purposes.

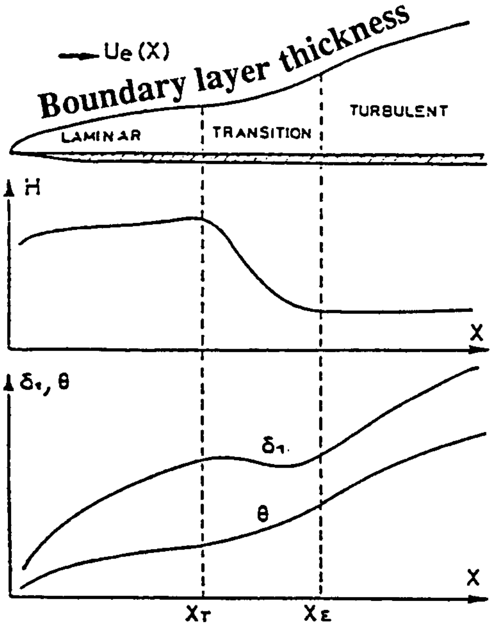


Fig. 1 - Boundary layer development

An overall picture of the boundary layer development is shown on figure 1. From the leading edge to a certain distance x_T, the flow remains laminar. At x_T, turbulent structures appear and transition occurs. From x_T to x_E, there is a noticeable change in the boundary layer properties. The transition process involves a large increase in the momentum thickness θ and a large decrease in the shape factor H. The displacement thickness $\delta_1 = H\theta$ exhibits a more complex evolution. The skin friction coefficient Cf increases from a laminar value to a turbulent one, the latter being in some cases an order of magnitude larger than the former. It is obvious that the location and the extent of transition depend on a large range of parameters, such as external disturbances, vibrations, pressure gradient, roughness ...

When a laminar flow develops along a given body, it is strongly affected by various types of disturbances generated by the model itself (roughness, ...) or existing in the freestream (turbulence, noise, ...). These disturbances are the sources of complex mechanisms which ultimately lead to turbulence. In fact, two kinds of transition processes are usually considered :

a) If the amplitude of the forced disturbances is small (low freestream turbulence level for instance), one can observe at first two-dimensional oscillations developing downstream of a certain critical point. After a linear amplification of these waves, three-dimensional and non linear effects become important, inducing secondary instabilities and then transition. In these cases of "natural" transition, the transition Reynolds numbers are usually very large.

b) If the amplitude of the forced disturbances is large (high freestream turbulence level, large roughness elements), non linear phenomena are immediately observed and transition occurs a short distance downstream of the leading edge of the body. This mechanism is called a "bypass", in this sense that the linear stages of the transition process are ignored (bypassed), (see MORKOVIN, /1/).

Both aspects of the transition mechanisms will be discussed successively. Section 2 is devoted to the problem of "natural" transition. After a short description of the transition processes, applications of the so-called "e^n method" are given. Typical examples of transitions induced by large disturbances (bypass) are given in Section 3.

2. "NATURAL" TRANSITION

2.1. General Description of the Transition Mechanisms

a) Linear stage - As it was pointed out, the instability leading to transition starts with the growth of two-dimensional disturbances, the existence of which was first demonstrated by the now classical experiments of SCHUBAUER and SKRAMSTAD, /2/. In fact, the existence of small, regular oscillations travelling in the laminar boundary layer was postulated many decades ago by Lord RAYLEIGH (1887) and PRANDTL (1921). Some years later, TOLLMIEN and SCHLICHTING worked out a linear theory of boundary layer instability, so that the waves are usually referred to as the "TOLLMIEN-SCHLICHTING waves" (TS waves). Nevertheless, the linear stability theory received little acceptance, essentially because of a lack of experimental results. The measurements of SCHUBAUER-SKRAMSTAD completely revised this opinion by demonstrating the real existence of the TS waves.

A complete account of this linear stability theory is out of the scope of this paper (see MACK, /3/, for complete information). Only some of the basic features will be briefly described.

In two-dimensional, incompressible flow, it is assumed that the TS waves can be expressed by :

$$q = \hat{q}(y)\exp(\sigma x)\exp\left[i\,(\alpha x - \omega t)\right] \qquad (1)$$

q represents a velocity or a pressure fluctuation ; σ, α and ω are the spatial amplification rate, the wavenumber and the circular frequency, respectively. Introducing (1) into the linearized NAVIER-STOKES equations leads to a system of ordinary differential equations, the combination of which gives the well known ORR-SOMMERFELD equation. Due to the homogeneous boundary conditions (the disturbances must vanish at the wall and in the freestream), the problem is an eigenvalue one. When the mean velocity profile U(y) is specified, a non zero solution of the stability equations exists for particular combinations of the four real parameters R, σ, α and ω, where R is the REYNOLDS number.

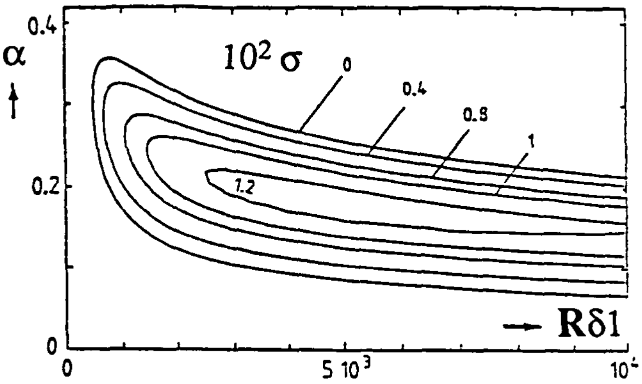


Fig. 2 - Stability diagram (BLASIUS flow)

The ORR-SOMMERFELD equation was solved by many authors. The results of such computations are represented on figure 2 for the BLASIUS flow. All the parameters were made dimensionless with the freestream velocity U_e and the displacement thickness δ_1, so that $R\delta_1 = U_e\delta_1/\nu$. The figure shows some curves of constant amplification rate σ in the (α, $R\delta_1$) plane ; curves of constant frequency ω are not represented for clarity. The curve $\sigma = 0$ is called the neutral curve ; it separates the region of stable ($\sigma < 0$) from that of unstable ($\sigma > 0$) disturbances. There is a particular value of the REYNOLDS number below which all disturbances decay ; it is the critical REYNOLDS number, $R\delta_{1cr}$, which is equal to 520 for the BLASIUS flow.

In three-dimensional and/or compressible flow, the problem becomes more complex, because the most unstable disturbances are oblique waves. For the sake of simplicity, let us consider first the simplest example of three-dimensional flow, i.e. the flow on a swept wing of constant chord and infinite span. φ is the sweep angle, x and z represent the directions normal and parallel to the leading edge, respectively. Relation (1) is now replaced by :

$$q = \hat{q}(y)\exp(\sigma x)\exp\left[i\,(\alpha x + \beta z - \omega t)\right] \qquad (2)$$

The previous expression contains the assumption that there is no amplification in the spanwise direction. An important parameter is the wavenumber direction $\psi = \tan^{-1}(\beta/\alpha)$. At each chordwise position, one has to compute the value of the amplification rate σ for each value of ψ. ψ_M will denote the most unstable direction, i.e. the wavenumber direction associated with the largest value of σ.

In compressible flow (at least for free stream Mach numbers M_e larger than 0.6), the problem is quite similar, even for two-dimensional mean flows ; in this case, the value of ψ_M is no longer equal to 0° as for low speed flows, but it takes a value which depends essentially on the local free stream Mach number. For instance, ψ_M is close to 70° for $M_e = 3$.

One has to keep in mind that the non linear phase and the breakdown process occur over a relatively short distance. For typical flat plate conditions, the streamwise extent of linear amplification covers about 75 to 85 per cent of the distance to the beginning of transition. This explains that calculation methods based on linear theory only (e^n method, paragraph 2.2.) give good results for predicting the transition location.

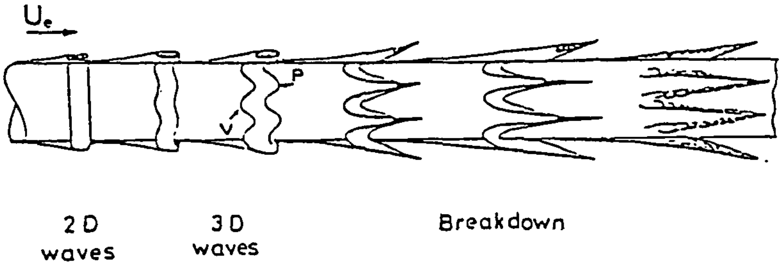


Fig. 3 - Overall picture of the transition process

b) Secondary instability, breakdown, turbulent spots In order to illustrate the downstream evolution of the TS waves, figure 3 presents an example of smoke visualization obtained by KNAPP et al., /4/. A laminar boundary layer develops in natural conditions on an ogive nose cylinder aligned with the freestream. It can be seen that the two-dimensional TS waves take the form of concentrated bands of smoke around the cylinder (left part of the sketch). These "rings" become more distinct as they move down the body, indicating the existence of a strong amplification. When the initially weak disturbances reach a certain amplitude, their evolution begins to deviate from that predicted by the linearized theory : the waves are distorted into a series of "peaks" and "valleys". As the flow proceeds downstream, this pattern becomes more and more pronounced. CRAIK, /5/, used a weakly non linear theory in order to explain the appearance of this peak-valley system ; his model was consistent with some experimental observations, but was inoperative in other cases. More satisfactory results were obtained by HERBERT, /6/, who developed a linear secondary instability theory based on the FLOQUET theory.

Further downstream, non linear mechanisms become dominant. The peak-valley system gives rise to a vortex filament (horseshoe vortex), which breaks down into smaller vortices, which again break down into smaller vortices. The fluctuations finally take a random character and form a so-called "turbulent spot" : it is the transition onset. In the transition region (between x_T and x_E, figure 1), the turbulent spots are swept along with the mean flow ; they grow laterally and axially, overlap and finally cover the entire surface.

c) Receptivity - It has been shown that the development of the TS waves can be correctly predicted by the linear stability theory. The main problem is now to explain the birth of these waves, i.e. to establish the link between their initial amplitude A_0 and the forced disturbances. The concept of receptivity, introduced by MORKOVIN /7/, describes the means by which these forced disturbances (sound, freestream turbulence) enter the laminar boundary layer and impose their signature in the disturbed flow. If they are small, they will tend to excite the TS waves, which constitute the normal modes of the boundary layer. Recent works by GOLDSTEIN /8/ and KERSCHEN /9/ have shown that the receptivity process occurs in regions of the boundary layer where the mean flow exhibits rapid changes in the streamwise direction. This happens near the body leading edge and in any region farther downstream where some local feature forces the boundary layer to adjust on a short streamwise length scale.

Up to now, receptivity studies were restricted to two-dimensional disturbances (sound, two-dimensional convected gusts ...). In many practical situations, however, the forcing disturbances are three-dimensional. In such cases, the receptivity mechanisms are unknown. The only available information come from experiments and illustrate the effect of the freestream disturbances amplitude on the transition REYNOLDS number R_{xT}. Figure 4 shows the evolution of R_{xT} as a function of the freestream turbulence level $Tu = \tilde{u}_e/U_e$, where $\tilde{u}_e$ is the rms value of the freestream disturbances. At first sight, the experimental data seem to collapse into a single curve and it is clear that transition moves rapidly upstream when Tu increases. However, as

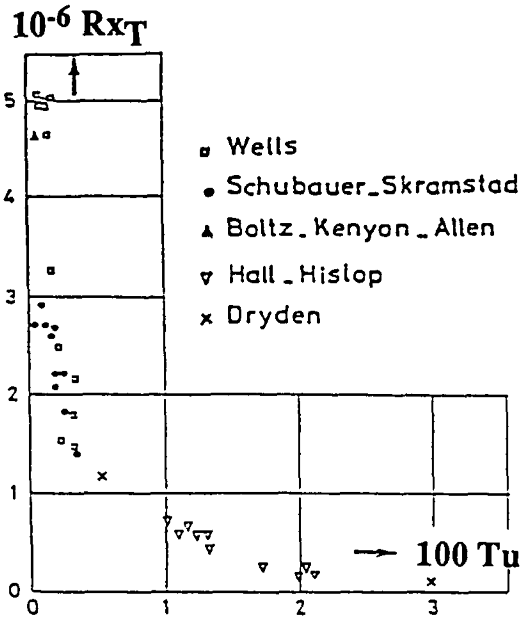


Fig. 4 - Effect of freestream turbulence on transition Reynolds number

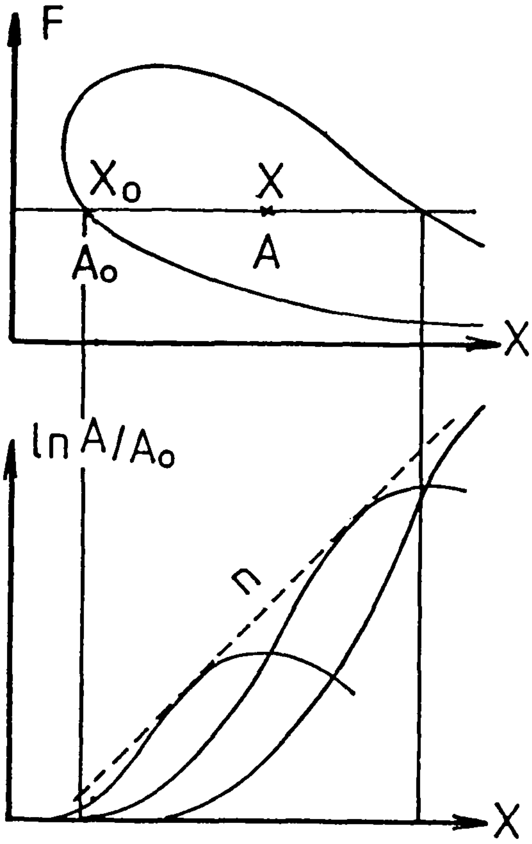


Fig. 5 - Typical stability diagram in physical coordinates - Definition of the total amplification rate and of the envelope curve

Tu becomes small, R_{x_T} reaches a constant value which depends on the experimental set-up : this value is about 2.8 10^6 for SCHUBAUER-SKRAMSTAD, /2/, and 5 10^6 for WELLS, /10/. In fact, sound component controls transition when Tu is very small and the effect of "true" freestream turbulence (vorticity fluctuations) can be only observed at values of Tu greater than 0.1 10^{-2}. On the other side, TS waves are never observed as soon as Tu exceeds 2 or 3 10^{-2} : transition becomes triggered by "bypass" mechanisms.

2.2. Transition Prediction : The e^n Method

a - Two-dimensional, incompressible flows

Let us recall that the general expression of a Tollmien-Schlichting wave is :

$$q = \hat{q}(y) \exp(\sigma x) \exp\left[i\,(\alpha x - \omega t)\right]$$

For a given mean flow, it is possible to compute a stability diagram (figure 5) showing the range of unstable frequencies f as a function of the streamwise distance x. Let us consider now a wave which propagates downstream with a fixed frequency f. This wave passes at first through the stable region ; it is damped up to x_0, then amplified up to x_1, and it is damped again downstream of x_1. At a given station x, the total amplification rate of a spatially growing wave can be defined as :

$$\ln(A/A_0) = \int_{x_0}^{x} \sigma \, dx \qquad (3)$$

A is the wave amplitude and the index 0 refers to the streamwise position where the wave becomes unstable. As an example, figure 5 shows total amplification curves corresponding to various frequencies. The dashed line represents the envelope of these curves, which will be called n :

$$n = \operatorname*{Max}_{f}(\ln(A/A_0)) \text{ at a given } x \qquad (4)$$

The so-called e^n method was developed independently by SMITH-GAMBERONI /11/ and by VAN INGEN /12/. SMITH and GAMBERONI compared the theoretical value of the n factor with transition locations measured on airfoils ; in all cases, transition was found to occur when $n \approx 9$; this means that turbulent spots appear when the most unstable frequency is amplified by a factor e^9. The same result was obtained by VAN INGEN, with a slightly lower value of n (7 to 8).

The e^n method is currently used for the case of natural transitions. The success of this method is certainly due to the fact that many experimental data are obtained in wind tunnels where the disturbance environment is similar, at least for low speed flows ; in particular, the freestream turbulence level is usually rather low, let say Tu ≈ 0.1 %. For larger values of Tu, the n factor at transition onset decreases. MACK /13/ suggested an empirical relationship between Tu and the value of n at the transition location :

$$n = -8.43 - 2.4 \ln Tu \qquad (5)$$

For $Tu < 10^{-3}$, sound disturbances may become the factor controlling transition rather than turbulence and relation (5) may give poor results. On the other side, if $Tu = 2.98\ 10^{-2}$, relation (5) implies that n = 0, i.e. transition occurs at the critical Reynolds number. It is the bypass limit.

b - Three-dimensional and/or compressible flows

Let us recall that it is now necessary to take into account oblique waves, because the streamwise direction is not always the most unstable one. Due to the appearance of this additional parameter, several strategies can be used to compute the n factor, see discussion in /14/ for instance. The numerical results presented in this paper have been obtained with the so-called envelope method. At each streamwise location and for a fixed frequency, the disturbance growth rate is maximized with respect to the wavenumber direction, i.e. the total amplification rates are computed with $\sigma(\psi_M)$. In other words, relation (3) becomes :

$$\ln(A/A_0) = \int_{x_0}^{x} \sigma(\psi_M)\, dx \qquad (6)$$

with $\sigma(\psi_M) = \underset{\psi}{\mathrm{Max}}\ (\sigma)$

The e^n method is then applied as in two-dimensional, incompressible flows.

2.2. Applications of The e^n Method

a - AEDC cone in free flight conditions

This 10-deg sharp cone, 1.1 m long, was mounted on the nose of an F-15 aircraft and flown at Mach numbers from 0.5 to 2, and at altitudes from 1 500 to 15 000 m ; transition was detected with a surface pitot tube which was displaced along a cone ray (FISHER and DOUGHERTY, /15/). As a typical example, figure 6 shows the total amplification rates computed for four frequencies at a free stream Mach number $M_\infty = 1.3$ for free flight conditions corresponding to a unit Reynolds number close to $10\ 10^6$. The measured transition onset is located at x ≈ 0.55 m; this gives a value of the n factor close to 9. Similar computations have been performed for other configurations ; in all cases, the transition locations were correlated with n factors between 9 and 11, so that it can be assumed that the value n = 10 is a more or less "universal" value for free flight conditions. The problem is to know if similar values can be reached in wind tunnels, i.e. if wind tunnel experiments can properly simulate free flight conditions.

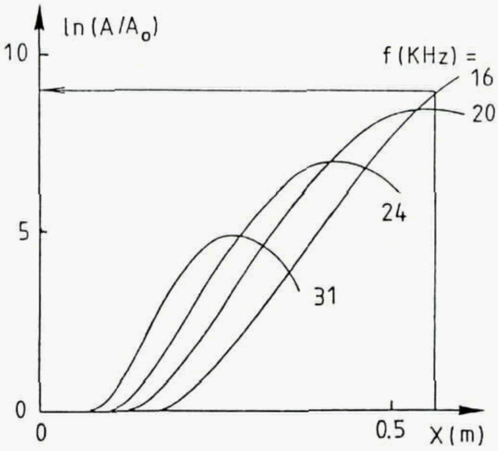


Fig. 6 - AEDC cone experiments in free flight conditions

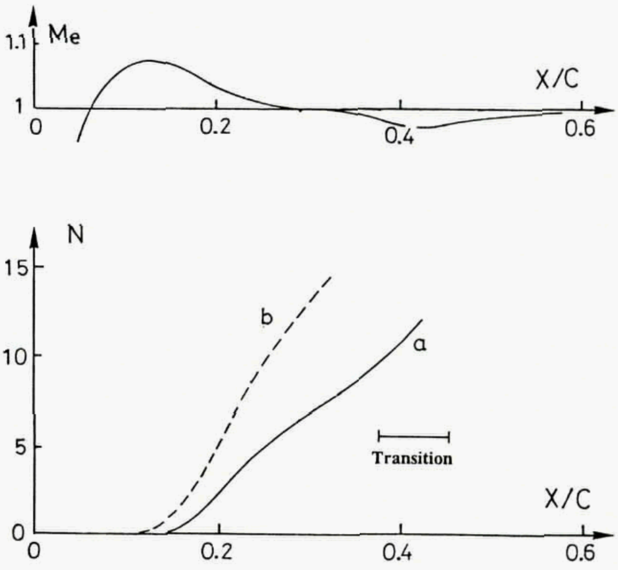


Fig. 7 - CAST10 airfoil in the T2 wind tunnel

b - CAST10 airfoil in the T2 wind tunnel

The e^n method was also used in the case of a two-dimensional airfoil in transonic flow. The experiments were carried out in the pressurized T2 wind tunnel of CERT/ONERA /16/. The model is a CAST10 airfoil with a chord C equal to 0.18 m. Figure 7 shows results obtained in the following conditions : free stream Mach number $M_\infty = 0.73$; angle of attack = 0° ; chord Reynolds number $Rc = 4\ 10^6$. The upper part of the figure shows that the local Mach number is close to 1. Two theoretical curves are presented on the lower part of the figure ; they give the evolution of the n factor as a function of the streamwise distance. Curve (a) was computed by solving the compressible stability equations ; the n factor is about 9 at the beginning of the measured transition region. This means that the flow quality of this wind tunnel is rather good, since the value of n is close to the value corresponding to free flight conditions. Curve (b) was deduced from incompressible stability computations, i.e. the free stream Mach number was set equal to zero. The stabilizing effect of compressibility is rather strong, since it reduces the amplification rates by a factor 2 !

c - ONERA D airfoil with a cambered leading edge

The next example is a three-dimensional, low speed configuration. The model is an ONERA D airfoil equipped with a cambered leading edge, figure 8a. The experiments were carried out in the F1 and in the F2 wind tunnels at Le Fauga-Mauzac Center near Toulouse. In both series of experiments, the wing was mounted on a half fuselage with an angle of sweep φ of 49° and an angle of attack of - 2°. Ten hot films were used to detect the transition location (figure 8b). Detailed results in the F2 wind tunnels can be found in /14/, /17/.

The experimental results obtained in both wind tunnels are reported on figure 9, where the transition location is plotted as a function of the free stream velocity Q_∞. These results are compared with theoretical curves associated with several values of the n factor. It appears that the flow quality in the F2 wind tunnel ($n \approx 11$ at transition onset) is slightly better than in the F1 wind tunnel ($n \approx 8$ to 9 at transition onset), even if the free stream turbulence level Tu is practically the same ($Tu \approx 0.1$ %). A spectral analysis of the free stream velocity fluctuations would be necessary to understand these differences. It must be noticed, however, that the flow quality in both facilities is good enough to give a satisfactory simulation of the free flight environment.

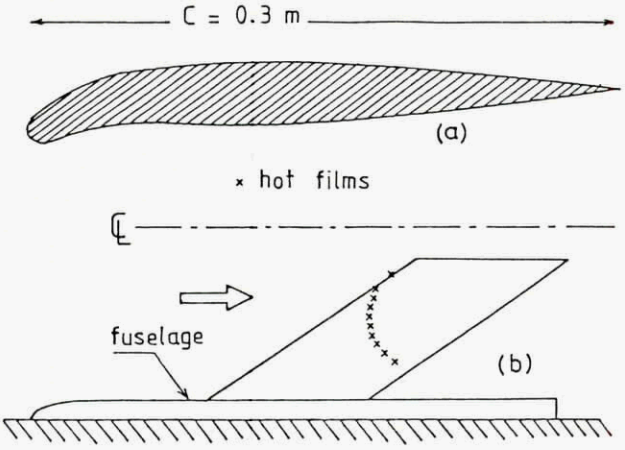


Fig. 8 - ONERA D airfoil with a cambered leading edge - a) Airfoil - b) Experimental arrangement in the F2 wind tunnel

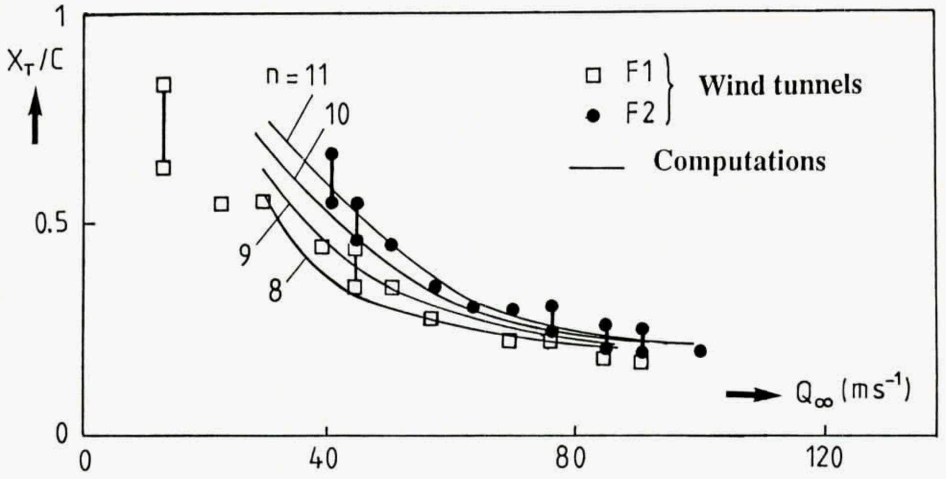


Fig. 9 - Comparison between predicted and measured transition locations

d - Two-dimensional flat plate flow at high Mach numbers

Systematic stability computations have also been performed for flat plate flows on adiabatic walls for supersonic and hypersonic Mach numbers ; the detailed results can be found in /18/. Figure 10 shows an application of the e^n method illustrating the effect of Mach number on the transition Reynolds number. The stability results were used to compute the theoretical streamwise Reynolds numbers $Rx = \frac{U_e x}{\nu_e}$ which correspond to different values of the n factor. The dotted lines represent theoretical results given by Mack /19/ for n = 4.6 and 6, and the crosses correspond to computations performed by Chen and Malik /20/ for $M_e = 3.5$ and n = 2, 4, 6, 8 and 10. If it is assumed that transition occurs for a fixed value of the n factor, each curve represents the evolution of the transition Reynolds number when M_e increases.

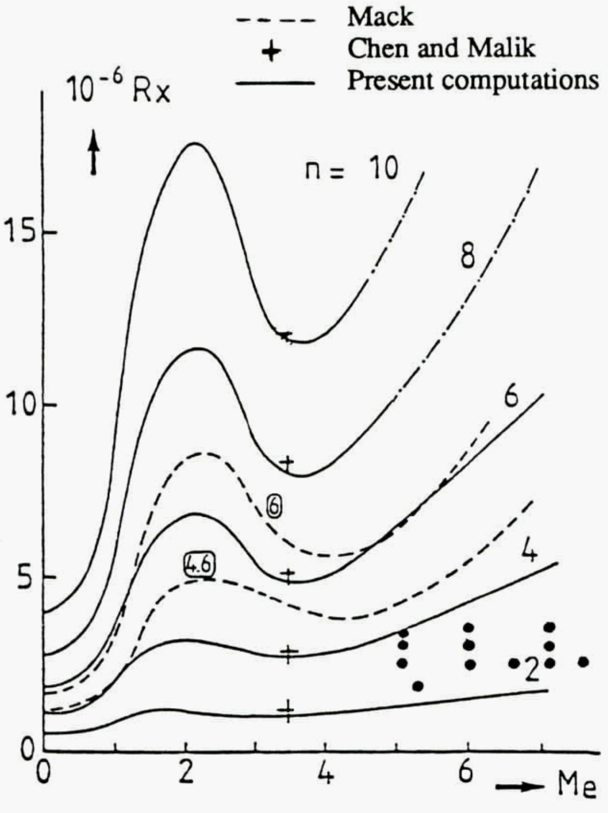


Fig. 10 - Application of the e^n method for flat plate flow on adiabatic wall : Mach number effect

It has been noticed previously that the value of n at transition onset is of the order of 10 for a low disturbance environment, at least for subsonic or transonic flows. The problem is to know if similar values of n are observed at high speeds. The solid symbols in figure 10 represent experimental data obtained at ONERA by Juillen /21/ : they correspond to low values of the n factor, between 2 and 4. Many examples could be given for illustrating the fact that in conventional hypersonic wind tunnels, the transition Reynolds numbers are several times lower than those which are observed in flight conditions. The reason of these discrepancies are now well known : in the conventional wind tunnels, the transition Reynolds numbers are strongly reduced by the noise radiated by the turbulent boundary layers developing along the nozzle walls. As a laminar boundary layer is less noisy than a turbulent one, a possible solution is to delay transition on the nozzle walls. This was done in the "quiet tunnel" developed at NASA Langley with a free stream Mach number $M_\infty = 3.5$ /22/ ; by reducing the noise level by one or two orders of magnitude, n factors close to 10 were obtained on cones and on flat plates.

3 TRANSITIONS INDUCED BY BYPASS MECHANISMS

It has been shown in the previous paragraph that the linear stability theory, associated with the e^n method, can give a fairly satisfactory estimate of the transition location, provided transition is triggered by the breakdown of instability waves. But these waves no longer appear when the amplitude of the forcing disturbances is too large, so that the problem becomes more complex. As there is no general bypass theory, it is necessary to develop different criteria for each type of bypass (the word criterion must be interpreted as a more or less empirical correlation between boundary layer and flow parameters at transition onset). In this paragraph, two examples of transitions induced by a bypass process are described ; the first one is the problem of boundary layer tripping by large roughness elements, the second one deals with leading edge contamination.

3.1. Boundary Layer Tripping By Isolated Roughness Elements

We first consider the problem of boundary layer tripping by large roughness elements embedded in a supersonic laminar boundary layer. The experiments have been performed on a flat plate in the R2Ch wind tunnel at Chalais-Meudon /23/. Figure 11 presents a typical wall visualization of the flow pattern around and downstream of a large, three-dimensional roughness element (sphere). A group of horseshoe vortices initiates ahead of the tripping device and develops around it. The "legs" of the vortices remain parallel to the main flow direction up to a certain distance L from the roughness element. At this point, one can observe the onset of a "turbulent" wedge which spreads slowly downstream. The exact mechanism of this breakdown is not very well known.

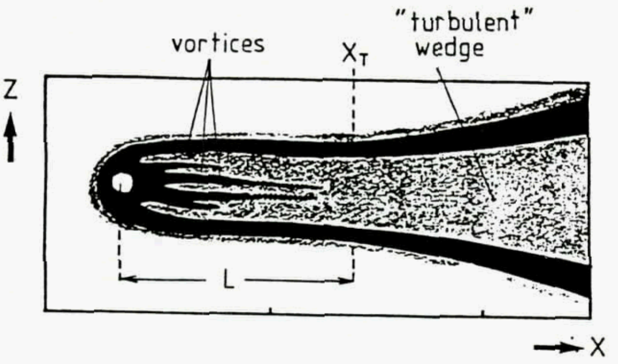


Fig. 11 - Example of wall visualization using thermosensitive paint

When the roughness size k increases (for fixed wind tunnel conditions), the distance L first decreases. This movement can be described by the empirical correlation proposed by Potter and Whitfield /24/. However, as k exceeds a critical value k_{eff} (effective roughness height), L reaches a constant value L_{min} which depends on many parameters such as the Mach number, the wall temperature, the roughness location ... This means that for $k > k_{eff}$, the apex of the turbulent wedge remains fixed at a constant distance L_{min} from the tripping device.

Van Driest and Blumer /25/ performed a series of experiments in order to deduce empirical correlations between Rk_{eff}, Rx_k are the Reynolds numbers based on k_{eff} and the roughness element location x_k, respectively. The measurements were made on cones for M_e between 1.9 and 3.65 ; they have been correlated with the following relationship :

$$Rk_{eff} = 33.4\left[1 + \frac{\gamma - 1}{2} M_e^2 - 0.81\left[\frac{T_{aw} - T_w}{T_e}\right]\right] Rx_k^{1/4} \qquad (7)$$

T_e, T_w and T_{aw} are the static temperature, the wall temperature and the adiabatic wall temperature. Van Driest and Blumer assumed that (7) was also applicable to flat plates by using Mangler's transformation, which simply consists in replacing the coefficient 33.4 by 33.4 (3) $^{1/4}$ = 44. Vignau /26/ demonstrated that the modified correlation largely underestimates the effective roughness height for flat plate flows, because Mangler's transformation is valid for mean flow properties, but it cannot be used for stability and transition problems.

3.2. Leading Edge Contamination

To introduce the notations, figure 12 shows a sketch of a circular cylinder placed at a sweep angle φ with respect to the incoming flow. The free stream velocity Q_∞ has a component $U_\infty = Q_\infty \cos\varphi$ normal to the leading edge and a component $W_\infty = Q_\infty \sin\varphi$ parallel to the leading edge. Z is the spanwise direction and X the direction normal to it. X = 0 corresponds to the atttachment line, which is a particular streamline which separates the flow into one branch following the upper surface and another branch following the lower surface. If it is assumed that the potential flow does not exhibit any spanwise variation, the free stream velocity components U_e and W_e in the X and Z directions are given by :

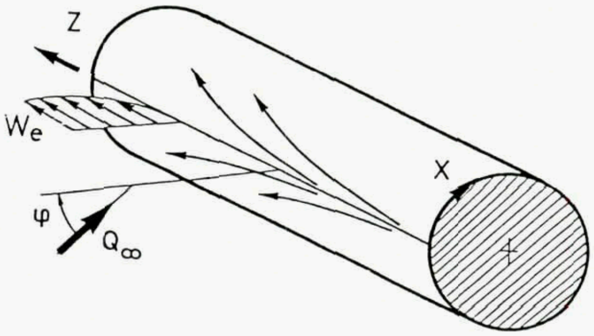


Fig. 12 - Attachment line flow on a swept cylinder

$$U_e = kX \qquad (8a)$$

$$W_e = W_\infty = \text{constant} \qquad (8b)$$

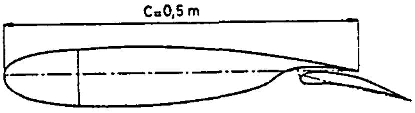
If this cylinder (or a swept wing the leading edge of which can be represented by such a cylinder) is in contact with a solid wall (fuselage, wind tunnel wall ...), it has been observed that the large turbulent structures coming from the wall may develop along the attachment line : it is the so-called leading edge contamination.

A leading edge contamination criterion was proposed by Pfenninger /27/ and then verified by many authors (/28/, /29/ for instance). It is based on the value of a Reynolds number $\overline{R}$ defined as :

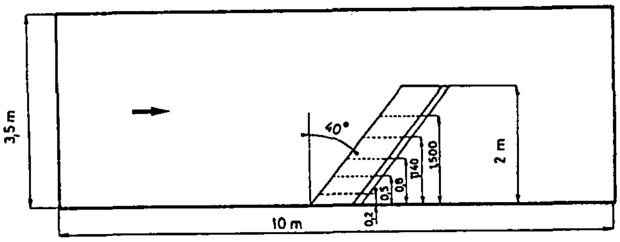
$$\overline{R} = W_e\eta/\nu, \text{ with } \eta = (\nu/k)^{1/2} \qquad (9)$$

For $\overline{R} > 250$, leading edge contamination occurs : the turbulent structures coming from the wall become self-sustaining ; they develop in the spanwise direction as they are convected along the leading and the whole wing can be contaminated. They are damped and disappear for $\overline{R} < 250$; in this case, the attachment line flow remains laminar. The leading edge contamination process is an example of bypass in this sense that the value $\overline{R} = 250$ is much lower than the linear critical Reynolds number of the attachment line boundary layer which is close to 600.

To illustrate the leading edge contamination process, we present typical results obtained by Arnal and Juillen in the F1 wind tunnel at Le Fauga-Mauzac Center /30/. The wind tunnel speed can be varied from 0 to about 100 ms^{-1}. The stagnation temperature is the ambient temperature, but the stagnation pressure P_i can be prescribed between 1 and 3 bars.



Airfoil



Test section (F1 wind tunnel)

Fig. 13 - Experimental setup in the F1 wind tunnel

Figure 13 shows the shape of the airfoil as well as a sketch of the experimental arrangement. The model is a RA16SC1 airfoil. The chord normal to the leading edge is constant and equal to 0.5 m. The wing is directly mounted on the wind tunnel floor ; the thickness of the floor turbulent boundary layer is about 10 cm. With a 40° sweep angle, the tip of the model is 2 m above the floor.

The upper part of figure 14 shows the location of the four hot films (labeled A, B, C and D) which were used to detect leading edge contamination. They were glued on the lower side of the wing, at one or two percent chord from the attachment line. The wind tunnel speed was progressively increased in order to determine accurately the leading edge contamination onset. The results which are described below were obtained for $\varphi = 40°$, $\alpha = 10°$, $P_i = 1$ bar (α is the angle of attack).

The lower part of figure 14 presents typical hot film signals recorded for $Q_\infty = 35\ ms^{-1}$; these signals

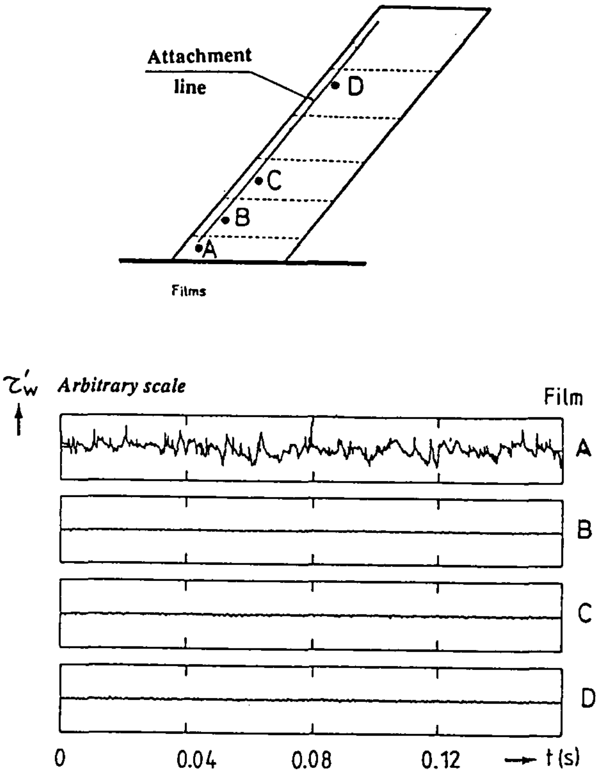


Fig. 14 - Hot films outputs ($Q_\infty = 35\ ms^{-1}$)

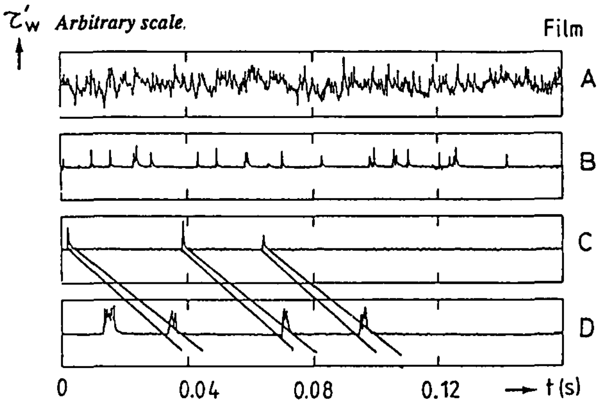


Fig. 15 - Hot films outputs ($Q_\infty = 61\ ms^{-1}$)

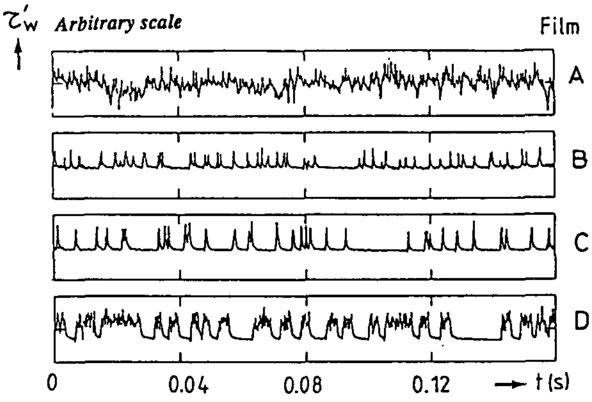


Fig. 16 - Hot films outputs ($Q_\infty = 64\ ms^{-1}$)

represent the wall shear stress fluctuations τ_w. The hot film A exhibits turbulent fluctuations which are generated by the floor turbulent boundary layer in which the sensor is embedded, but the other three signals are of the laminar type. For $Q_\infty = 61\ ms^{-1}$ (figure 15), turbulent structures are observed on film B. The number of these "spots" decreases from film B to film C, then it remains constant. The signals delivered by films C and D are essentially characterized by the spreading of the spots which develop along the leading edge : it is the beginning of leading edge contamination. When the wind tunnel speed increases from 61 to 64 ms^{-1}, a comparison between figures 15 and 16 reveals a rapid increase in the number of turbulent spots. For $Q_\infty = 95\ ms^{-1}$, all signals have a fully turbulent character.

Similar measurements were done for several combinations of φ, α and P_i. The result is that leading edge contamination appears for $\overline{R} = 251 \pm 11$, and that the leading edge is fully turbulent for $\overline{R} = 318 \pm 22$. These values are in good agreement with those given by other investigators.

This rather simple criterion is valid for incompressible flows only. Poll /31/ made an extension to compressible flows by introducing a transformed Reynolds number $\overline{R}^*$ which has the same definition as $\overline{R}$, except that the kinematic viscosity is replaced by ν^* which is computed at a reference temperature T^*. The validity of this criterion was checked by Da Costa /32/ and also by Arnal et al /33/.

It must be kept in mind that leading edge contamination is the first problem to solve for maintaining laminar flow over a wing : if the atttachment line boundary layer is turbulent, turbulence will spread over the whole wing, and the benefits of laminar flow control systems will be lost.

Poll suggested also that leading edge contamination was responsible for transition on the windward face of the Columbia space shuttle during reentry /31/.

4 CONCLUSION

This paper demonstrated that the problems associated with boundary layer transition are numerous but that many of them can be satisfactorily solved for practical applications.

On the theoretical point of view, the linear stability theory constitutes a very efficient tool to understand the fundamental mechanisms leading to a "natural" transition in a low disturbance environment. But the key problem lies in the understanding of the receptivity mechanisms and the exact relation between instability and transition is not very well known ; these issues are fairly well documented for low speed, two-dimensional flows, but very little is known about the receptivity and the non linear mechanisms of compressible and/or three-dimensional flows. We can expect that direct numerical simulations would provide us with interesting information in the next future.

Even if all the transition mechanisms are not completely explained, practical prediction methods need to be developed. In the case of "natural" transition, the e^n method gives surprisingly good results to "predict" transition onset. This technique can also be used to give a measure of the flow quality in ground facilities by comparing the value of the n factor deduced from wind tunnel experiments with that which is obtained in free flight conditions. Good simulations can be achieved in low speed and transonic facilities ($n \approx 10$), but high speed wind tunnels cannot duplicate free flight conditions due to the noise radiated from the nozzle walls.

If the breakdown to turbulence occurs without resorting to linear processes (bypass), the e^n rule no longer applies, but empirical correlations are available for design purposes. Two typical examples have been discussed in this paper (boundary layer tripping by isolated roughness elements and leading edge contamination). Fundamental experiments need to be performed in order to reach a more unified approach of these problems. Let us note that wind tunnel experiments dealing with bypass mechanisms are certainly representative of flight conditions, even for high speed flows. This is due to the fact that the large disturbances which trigger transition overwhelm the effects of the residual fluctuations which are naturally present in the free stream. In other words, the flow quality is of less importance than for "natural" transitions.

REFERENCES

1. Morkovin, M.V., "Bypass transition to turbulence and research desiderata", Symp. "Transition in Turbines", Cleveland, Ohio, 1984

2. Schubauer, G.B., Skramstad, H.K., "Laminar boundary layer oscillations and transition on a flat plate", Report 909 NACA, 1948

3. Mack, L.M., "Boundary layer linear stability theory", AGARD Report N° 709, 1984

4. Knapp, C.F., Roache, P.J., Mueller, T.J., "A combined visual and hot wire anemometer investigation of boundary layer transition", UNDAS-TR-866 CK, 1966

5. Craik, A., "Non linear resonant instability in boundary layers", J. Fluid Mech., Vol. 50, Part 2, 1971

6. Herbert, T., "Secondary instability of shear flows", AGARD Report N° 709, 1984

7. Morkovin, M.V., "Critical evaluation of transition from laminar to turbulent shear layers with emphasis on hypersonically travelling bodies", Report AFFDL-TR-68-149, Wright Patterson Air Force Base, 1968

8. Goldstein, M.E., "The evolution of TS waves near a leading edge", J. Fluid Mech., Vol. 127, pp. 59-81, 1983

9. Kerschen, E.J., "Boundary layer receptivity", AIAA Paper 89-1109, 1989

10. Wells, C.S., "Effects of free stream turbulence on boundary layer transition", AIAA J., Vol. 5, N° 1, pp. 172-174, 1967

11. Smith, A.M.O., Gamberoni, N., "Transition, pressure gradient and stability theory", Douglas Aircraft Co. Rept ES 26388, El Segundo, California, 1956

12. Van Ingen, J.L., "A suggested semi-empirical method for the calculation of the boundary layer transition region", Univ. of Techn., Dept. of Aero. Eng., Rept. UTH-74, Delft, 1956

13. Mack, L.M., "Transition and laminar instability", Jet Propulsion Lab. Publication 77-15, Pasadena, California, 1977

14. Casalis, G., Arnal, D., "Prévision de la transition sur une aile en flèche", La Recherche Aérospatiale n° 1990-1, 1990

15. Fisher, D.F., Dougherty, N.S., Jr., "In-flight transition measurements on a 10° cone at Mach numbers from 0.5 to 2", NASA TP-1971, June 1982

16. Blanchard, A., Mignosi, A., Dor, J.B., Seraudie, A., "Essais de profils CAST7 et CAST10 dans la soufflerie cryogénique à parois adaptables T2. Etude de l'effet Reynolds en transition naturelle et en transition déclenchée", 22ème Colloque d'Aérodynamique Appliquée, Lille, 1985

17. Arnal, D., Juillen, J.C., "Three-dimensional transition studies at ONERA/CERT", AIAA Paper N° 87-1335, 1987

18. Arnal, D., "Stabilité et transition des couches limites laminaires bidimensionnelles en écoulement compressible, sur paroi athermane", La Recherche Aérospatiale n° 1988-4, 1988

19. Mack, L.M., "Linear stability and the problem of supersonic boundary layer transition", AIAA J., Vol. 13, N° 3, pp. 278-289, 1975

20. Chen, F.J., Malik, M.R., "Comparison of boundary layer transition on a cone and flat plate at Mach 3.5", AIAA Paper 88-0411, 1988

21. Juillen, J.C., "Détermination expérimentale de la région de transition sur une plaque plane à M = 5, 6 et 7", ONERA Technical Report n° 10/2334 AN, 1969

22. Beckwith, I.E., Creel Jr., T.R., Chen, F.J., Kendall, J.M., "Freestream noise and transition measurements on a cone in a Mach 3.5 pitot low-disturbance tunnel", NASA TP 2180, 1983

23. Arnal, D., Vignau, F., Juillen, J.C., "Boundary layer tripping in supersonic flow", IUTAM Symp. "Laminar-Turbulent Transition", Toulouse, Ed. Arnal and Michel, Springer Verlag, 1990

24. Potter, J.L., Whitfield, J.D., "Effects of slight nose bluntness and roughness on boundary layer transition in supersonic flows", J. Fluid Mech., Vol. 12, Part 4, pp. 501-535, 1962

25. Van Driest, E.R., Blumer, C.B., "Boundary layer transition at supersonic speeds : roughness effects with heat transfer", AIAA J., Vol. 6, N° 4, pp. 603-607, 1968

26. Vignau, F., "Etude théorique et expérimentale de la transition en écoulement bidimensionnel compressible", Thesis, ENSAE, Toulouse, 1989

27. Pfenninger, W., "Flow phenomena at the leading edge of swept wings", AGARDograph 97, Part 4, 1965

28. Gaster, M., "On the flow along leading edges", The Aeron. Quarterly, Vol. XVIII, Part 2, May 1967

29. Poll, D.I.A., "Some aspects of the flow near a swept attachment line with particular reference to boundary layer transition", Cranfield, College of Aeronautics Report 7805, August 1978

30. Arnal, D., Juillen, J.C., "Etude de la transition et de la contamination de bord d'attaque sur ailes en flèche", AGARD CP N° 438, 1988

31. Poll, D.I.A., "Boundary layer transition on the windward face of space shuttle during reentry", AIAA Paper 85-0899, 1985

32. Da Costa, J.L., "Contribution à l'étude de la transition de bord d'attaque par contamination en écoulement hypersonique", Thesis, Poitiers University, 1990

33. Arnal, D., Vignau, F., Laburthe, F., "Recent supersonic transition studies with emphasis on the swept cylinder case", Boundary Layer Transition and Control, Cambridge, April 1991

Modeling Nonlinear Aerodynamic Loads for Aircraft Stability and Control Analysis

Jerry E. Jenkins
James H. Myatt
Wright Laboratory (WL/FIGC)
Wright-Patterson AFB, OH 45433-6553
U.S.A.

SUMMARY

Results from systematic wind tunnel tests of the dynamic roll behavior of a 65° swept delta wing at moderate (15° to 35°) angles of attack are reviewed. These tests, conducted in both the IAR 2 x 3 m low-speed wind tunnel and the 7 x 10 ft SARL facility at WPAFB, included static, forced oscillation and free-to-roll experiments with flow visualization. Multiple stable trim points (attractors) for body-axis rolling motions and other hard-to-explain dynamic behavior were observed. These data are examined in light of the nonlinear indicial response theory advanced by Tobak and his colleagues. The current analysis shows that force and moment, free-to-roll motion, and flow visualization data all confirm the existence of "critical states" with respect to the static roll angle. When these singularities are encountered in a dynamic situation, large and persistent transients are induced. Conventional means of representing the nonlinear force and moments in the equations of motion are shown to be inadequate in these cases. Alternative approaches based on simplification of the nonlinear indicial model are briefly discussed.

LIST OF SYMBOLS

b	wingspan (ft)
C_l, C_m	nondimensional body-axis rolling moment and pitching moment coefficients
k	reduced frequency, i.e. $\omega b/2U_\infty$
t	time (seconds)
U_∞	freestream velocity (ft/sec)
ϕ	body-axis roll angle (deg)
τ	time at step onset (sec)
ω	circular frequency (rad/sec)

1. INTRODUCTION

Dynamic coupling between aircraft motion and aerodynamic forces and moments acting on the aircraft is at the heart of the stability and control problem. Realistic representations of these interactions are prerequisite for aircraft and flight control system design and evaluation. Maintaining sufficient fidelity in aerodynamic models (for the equations of motion) has become an increasingly difficult problem in the face of flight envelope expansion.

1.1 Mathematical Modeling

A theoretical method for studying the nonlinear aspects of the flight dynamics problem has been under development by Tobak[1] and his colleagues since the 1960s. Their initial approach[2] introduced two important new concepts: (1) a nonlinear indicial response and (2) a generalized superposition integral. As with linear indicial response methods, the idea is to represent aerodynamic responses (force or moment) due to arbitrary motion inputs as a summation of responses to a series of "step" motions. The nonlinear indicial response, as opposed to its linear counterpart, accounts for changes induced by the motion history leading up to step onset. Under a wide variety of circumstances, the summation of indicial responses approaches the generalized superposition integral in the limit.

Subsequently,[3,4] results from the growing body of nonlinear dynamical system theory were used to greatly strengthen the model. The key idea of these extensions has been to accommodate the existence of "critical states," i.e., specific values of the motion variables where discrete changes in *static* aerodynamic behavior occur. These are singular points that require special handling in the superposition integral. *Critical states are important* because potentially large and persistent transient effects can be anticipated when they are encountered in a dynamic situation.

Truong and Tobak[5] have also demonstrated that, for static aerodynamic characteristics that are time-invariant, the nonlinear indicial response, together with the generalized superposition integral, can be derived directly from the Navier-Stokes equations. Thus not only does the theory have a sound mathematical basis, the physics of the aerodynamic contribution are captured as well. There is still much work to be done, especially for cases involving time-dependent equilibrium states. However, the theory is rich in its ability to represent a wide range of physically realizable nonlinearities.

Independently, Hanff[6] proposed the "reaction hypersurface" model. As opposed to the time-domain indicial response model, the hypersurface is expressed in terms of a set of independent variables consisting of the instantaneous values of the motion variables and their time derivatives. It was, at its inception, designed to be experimentally based and primarily intended for simulations of aircraft motion in nonlinear settings where the classical stability derivative approach breaks down.[7] More recent work[8] has been aimed at establishing the theoretical connection between the reaction hypersurface and nonlinear indicial response models.

1.2 Experimental Investigations

Experiments designed to study either of these mathematical models demand a dynamic test capability (including an appropriate data reduction system) that can efficiently collect the necessary nonlinear and time-dependent data. A large-amplitude high-rate roll oscillation system,[9] developed by the Canadian Institute for Aerospace Research (IAR), meets these requirements.

Hanff and S. B. Jenkins[10] used this rig to study the roll dynamics of both a 65° delta wing and a 80°-65° double-delta wing at the IAR. Their experiments produced some extremely interesting results which require further explanation.

The 65° delta wing configuration was found to have multiple stable trim points in roll (depending on roll-axis inclination) as reported by Hanff and Ericsson.[11] Attractor locations found in these tests are shown in Table I. They argue (based on an analysis of the static rolling moment data at 30° incidence) that asymmetric vortex breakdown, induced by differing effective sweep angles on each wing panel, is the root cause. However, the dynamic behavior observed in "free-to-roll" experiments is harder to explain, although Hanff and Huang[12] have shown that the instantaneous loads are largely driven by the dynamics of leading-edge vortex breakdown.

Table I
Roll Attractor Locations

Sting Angle (deg)	Roll Angle (deg)
20	0
25	|1.5|
30	0, |21|
35	|11|
40	0

In free-to-roll tests, the model is given an initial roll displacement, then released by disengaging a remotely actuated clutch. The model is then free to roll about its body axis, restrained only by a small amount of bearing friction in this degree-of-freedom. The resulting roll motion time-history is recorded. For the IAR tests, mass was added to the moving part of the sting to increase its roll-axis moment of inertia. This was done to ensure that the free-to-roll responses were in the same frequency range as the force measurements (about 7 Hertz).

Two free-to-roll time histories for the 65° configuration at 30° incidence, plotted in the phase plane, are shown in Fig. 1 (taken from Ref. 8). Note that the trajectory for the -66° release angle (solid curve) finds the stable equilibrium point at about 0° roll, while the 57° release trims at about 21°. Both trajectories pass quite close to attractors (21° and 0° respectively) with very low rates *but do not trim there.* This behavior was highly repeatable. Furthermore, the trajectories intersect at several points. Similar intersections of phase-plane trajectories (for wing-rock motions) have been observed *only when vortex breakdown occurs over the wing.*[13] Clearly, some phenomenon, not explicitly accounted for in the two-dimensional phase-plane representation, affects the motion. Persistent motion history effects, perhaps related to vortex breakdown dynamics, that require more than a knowledge of the instantaneous roll angle and roll rate are a strong possibility.

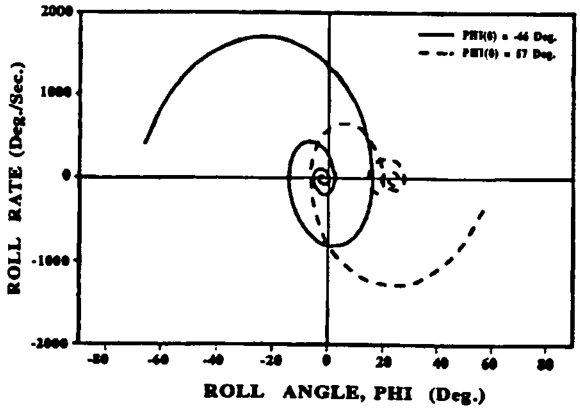


Figure 1. Free-to-Roll Trajectories

Finally, forced oscillation motions about zero-mean roll angle tended to produce distinctively different rolling-moment responses than those measured for motions with non-zero mean roll angles.[10] An analysis of the static and dynamic force data[8] suggested that this behavior could be explained by the existence of critical states. This view was supported by evidence of extremely large and persistent transients following encounters with the suspected critical states. Static and dynamic effects were of the same order. Transients were seen to persist for at least a quarter cycle at $k = 0.08$.

Follow-on wind-tunnel tests, using the 65° delta-wing configuration, were conducted in the SARL facility at Wright-Patterson AFB. These were designed to confirm the IAR results and to further investigate the behavior discussed above.

In this paper, relevant aspects of critical state theory are discussed. An overview of the SARL tests is presented and SARL data confirming the existence of roll-motion critical states are briefly reviewed. Significance of these findings to aerodynamic modeling for application to aircraft simulation and analysis are addressed in the final sections.

2. Critical States - Theoretical Basis

Some key properties of the nonlinear indicial response, pertaining to the present discussion, are summarized below. The interested reader is referred to Refs. 1, 3 and 4 for a complete development of the theory.

1. The nonlinear indicial response (NIR) is represented mathematically as a functional (to incorporate the motion history effect).

2. The NIR is a derivative (called the Fréchet derivative) of the functional representing an aerodynamic response in terms of its motion history. It is the limit, as input step height goes to zero, of the incremental response (due to the step input) divided by step height. Following Tobak's notation,[1] the rolling moment due to an infinitesimal step in roll angle is written

$$C_{l_\phi}(t) = C_{l_\phi}[\phi(\xi); t, \tau]$$

where:

a) square brackets denote a functional,

b) the first argument is the independent function defining the motion history (roll angle in this case), and

c) arguments following the semi-colon give, respectively, the times at which (1) the response is to be evaluated (observed) and (2) the step motion was initiated.

Therefore, the function $\phi(\xi)$ is to be interpreted as the motion history from $t = -\infty$ to step onset, τ, and the motion is to be held constant at $\phi(\tau)$ thereafter.

3. If the Fréchet derivative exists everywhere on a time interval (i.e., for the range of motion variables encountered on that interval) the generalized superposition integral may be used to construct the net aerodynamic response over the interval. Thus,

$$C_l(t) = C_l[\phi(\xi); t, 0] + \int_0^t C_{l_\phi}[\phi(\xi); t, \tau]\frac{d\phi}{d\tau}d\tau$$

Following the notation introduced above, the first term on the RHS is the rolling moment at time t resulting from the roll-angle variation $\phi(\xi)$ which is the motion history *prior* to $\xi = 0$, and is held constant at $\phi(0)$ for all $\xi \geq 0$. The functional in the second term is the NIR, as defined above. In this case, τ is the variable of integration *and* the time at step onset. Thus, the integral sums the effects of all indicial responses over the interval 0 to t.

4. If, on the other hand, there are specific points, τ_c, within the interval where Fréchet differentiability is lost (with a corresponding critical state, ϕ_c), the integration may not be carried beyond the instant at which a critical state is encountered without acknowledging the existence of the singularity.

5. Loss of Fréchet differentiability is handled by allowing the response to change discretely to a new state. Thus the integral must be split to isolate the critical state, i.e.,

$$\begin{aligned} C_l(t) &= C_l[\phi(\xi); t, 0] \\ &+ \int_0^{\tau_c - \varepsilon} C_{l_\phi}[\phi(\xi); t, \tau]\frac{d\phi}{d\tau}d\tau \\ &+ \int_{\tau_c + \varepsilon}^{t} C_{l_\phi}[\phi(\xi); t, \tau]\frac{d\phi}{d\tau}d\tau + \Delta C_l(t; \phi_c), \end{aligned} \tag{1}$$

where

$$\begin{aligned} \Delta C_l(t; \phi_c) &= C_l[\phi(\xi); t, \tau_c + \varepsilon] \\ &- C_l[\phi(\xi); t, \tau_c - \varepsilon] \end{aligned} \tag{2}$$

ΔC_l, as given by Eq. (2), is the transient response associated with ϕ_c. Note that it depends on the motion history from $-\infty$ to just beyond τ_c. However, its effect persists for times $t > \tau_c$.

6. Fréchet differentiability may be lost in several ways.[1] A very important case is when *time-invariant equilibrium flows* lose their analytic dependence on a motion parameter. There are at least two ways this can happen:

(1) The static aerodynamic response can develop a *fold* at a critical value of the parameter (possibly an indirect result of bifurcation). The response slope becomes infinite at the fold, invalidating the Fréchet derivative.

(2) There can be a *change in flow topology* (a change in the number of singular points in either the external flow or in surface skin-friction lines) when the motion parameter reaches a critical value. Tobak *et al.*[1] *anticipate* that not only will the equilibrium response cease to be analytically dependent on the motion parameter at such points, but there will also be, "a significant increase in the time required for the ... response to reach a new equilibrium state."

Flow-field structure changes are evident at the critical state in both cases. However, there is no *net* change in the number of flow-field singular points at a fold (e.g., an asymmetric vortex system could be replaced by one of opposite sense at a bifurcation point).

3. THE SARL EXPERIMENTS

IAR's high-amplitude high-rate roll apparatus was used for the SARL tests. These experiments were conducted by Hanff and his IAR colleagues solely with the 65° delta wing configuration (Fig. 2).

A comprehensive series of tests was conducted involving over 800 runs. Types of data taken along with the range of test conditions are summarized in Table II. Since SARL is an open-return atmospheric tunnel, the IAR Mach number, Reynolds number, and reduced frequencies could not be matched simultaneously. However, test conditions were chosen to match those at the IAR as closely as possible. Since model support systems for the two facilities are quite different, the very good data correlation between tunnels at the low-speed condition eliminated the possibility of significant sting interference effects.

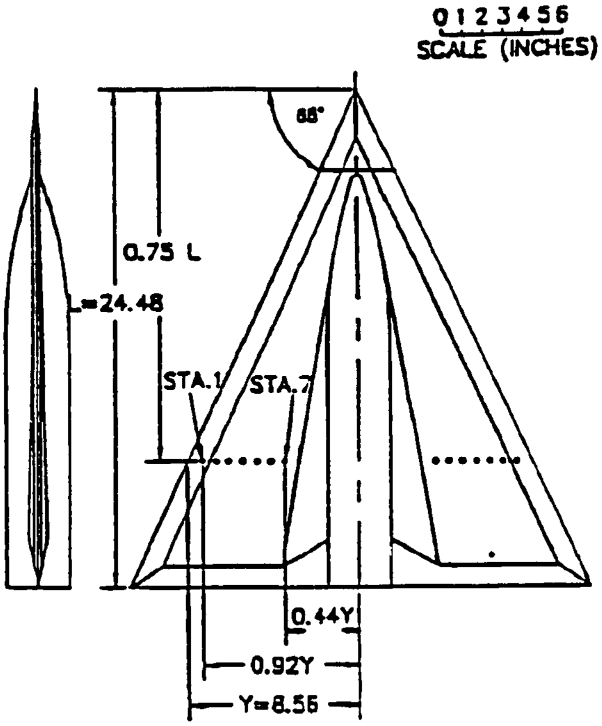


Figure 2. 65° Delta-Wing Model

Dynamic force and moment measurements were taken with the model forced in constant amplitude harmonic motion. Data were taken at 4.4 and 7.7 Hertz ($k = 0.08$ and 0.14) to match IAR conditions. The IAR test results at these frequencies showed little tendency to "track" the static data whenever the motion included small roll angles.[8] Therefore, in the SARL experiments, dynamic tests were also conducted at 1.1 Hertz ($k = 0.02$) to determine how these very large dynamic effects approach quasi-steady behavior at low reduced frequency.

In addition, laser-sheet flow visualization data were taken using a high-speed video camera. Thus, a comprehensive data set that allows a coordinated study of vortex dynamics (including breakdown) and the resulting unsteady aerodynamic forces and moments was created.

Table II
SARL Test Conditions

Test Type	Roll Offset (deg)	Amplitude (deg)	Frequency (Hz)	Total AOA (deg)
Static Force	-70 to 70	NA	NA	15, 30, 35
Dynamic Force	0 to 42	5 to 40	1.1, 4.4, 7.7	15, 30, 35
Free-to-Roll	-65 to 65	NA	"7.7"	30, 35
Flow Vis.	0 to 42	5 to 40	0, 1.1, 4.4, 7.7	30, 35

Free-to-roll experiments were also repeated in the SARL tunnel. These data provide an independent check on the accuracy of the dynamic force measurements (and the mathematical model used to represent them) since the measured forces, together with the known model/test-rig inertia, can be used to "predict" the free-to-roll motion.

4. EVIDENCE - CRITICAL STATE EXISTENCE

The most important results (confirming critical states for the 65° delta wing) based on an analysis of the SARL data are summarized below. A more extensive description of the analysis[14] has been submitted (for consideration) to the 31st AIAA Aerospace Sciences Meeting to be held in January 1993. All data discussed in this section were taken at the same condition (0.3 Mach number and 30° roll-axis inclination).

Based on the theory presented in Section 2, critical states should exhibit the following properties:

(1) Static flow visualization studies should show a change in flow structure at the critical state.

(2) Static data should exhibit non-analytic behavior across critical states; i.e., there should be discontinuities in the force/moment curves and/or their derivatives with respect to the motion variable. Discontinuities are located at the critical states.

(3) Transient effects should be observed following dynamic critical-state encounters. The transient, ΔC_l in Eqs. (1) and (2), will in general depend on motion history.

4.1 Static Flow Visualization Results

Vortex breakdown locations for the *left wing* as a function of roll angle, (from Hanff and Huang)[12] are shown in Fig. 3.

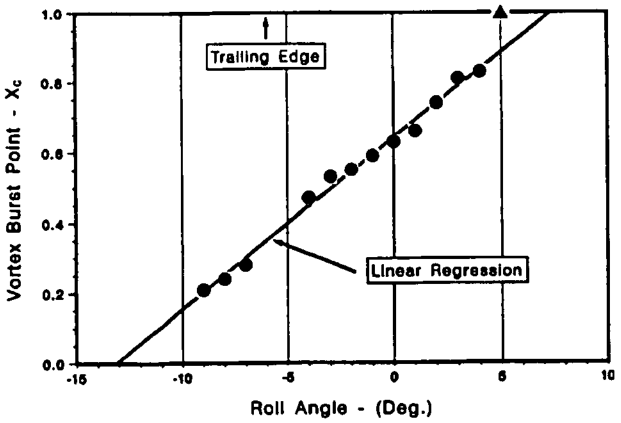


Figure 3. Vortex Breakdown Location (Left Wing)

Two important events are noted:

(1) The vortex breakdown point at $\phi = 5°$ (triangular symbol) departs significantly from the linear behavior shown by the rest (circles). Thus, breakdown location is seen to be a strong (perhaps discontinuous) function of static roll angle in the 4 to 5 degree range. For ϕ greater than 5 degrees the breakdown point is well aft of the trailing edge.

(2) Vortex breakdown reaches the wing vertex at about $\phi = -13$ degrees, as suggested by an extrapolation based on the linear regression (see Fig. 3).

When the leading-edge vortex structure on *both* wings is considered, the conditions, $|\phi| \sim 5°$ and 13°, are strong critical state possibilities.

The first pair, $|\phi| \sim 5°$, must be considered because a "jump" in vortex breakdown position on the lee wing would cause a discontinuous force/moment response. Note that the corresponding windward wing vortex-breakdown movement is both well behaved and small, as shown in Fig. 3 (the change from $\phi = -4°$ to $-5°$).

When breakdown reaches the wing vertex, the axial-flow stagnation point in the vortex-core is lost[15]. Therefore the second pair is almost certainly a critical state (flow topology change). The precise roll angle where this occurs is unknown (13° is based on a linear extrapolation of Hanff's data).

4.2 Static Force Data

Rolling moment coefficient vs. roll angle is presented in Fig. 4a. Also shown are critical state locations corresponding to both conditions. For clarity only those for $\phi > 0$ are given. Note the steep slopes between $|\phi| = 4°$ and $|\phi| = 5°.0$

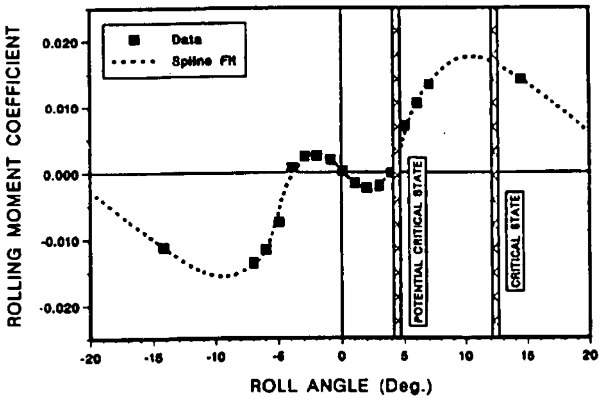


Figure 4a. Static Rolling Moment

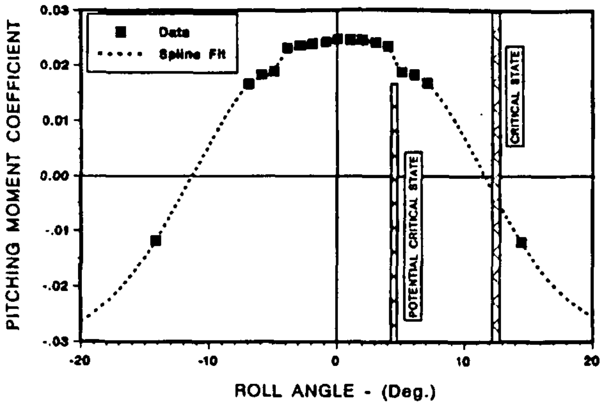


Figure 4b. Static Pitching Moment

Moreover, Fig. 4b (pitching moment vs. roll angle) reinforces the notion that there is a discontinuity between these points. Thus, the static behavior of both moment coefficients supports the notion that |ϕ| ≈ 5° represents a critical-state pair. Similar observations concerning the critical states at about 13° are not possible because of the gap in static data for roll angles between 7° and 14°. Further static testing is required to determine the behavior in this region.

4.3 Dynamic Force Data

Previous analyses[8] of IAR dynamic data for this configuration suggested that there was at least one critical state at these conditions. "Significant" dynamic effects were observed when the rolling motion included "small" roll angles.

Dynamic force data taken at both facilities (IAR and SARL) are "steady-state" responses to harmonic motion, i.e., starting transients have dissipated and the data for each cycle are repeatable. If the aerodynamic responses are slow compared to the period of the motion, the measurements are an aggregate of effects initiated during earlier cycles. Under these conditions, significant phase and amplitude variations with frequency are observed. Furthermore, at "high" frequencies, the effects of events occurring at discrete points during each cycle (e.g. critical state encounters) can be masked by the lingering responses from previous cycles. This was the case with the IAR dynamic data.

SARL dynamic rolling-moment data taken for a 12° amplitude body-axis rolling motion, centered about a mean roll angle of 14°, is presented in Fig. 5a. The abscissa is the argument of the cosine function (ωt) which defines the motion. Thus, precisely one cycle of motion is presented regardless of frequency. Dynamic data taken at three frequencies (1.1, 4.4 and 7.7 Hertz.) are shown. In addition, the roll-angle time history and static data rolling moment data (plotted as a function of the instantaneous roll angle) are presented in this figure for reference.

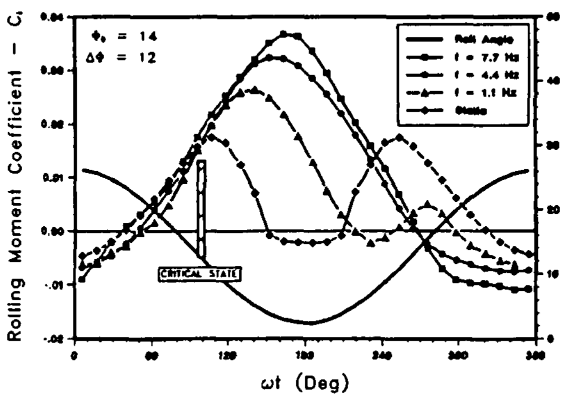


Figure 5a. Dynamic Rolling Moment

Note the dramatic differences in wave form between responses at the two highest frequencies and the 1.1 Hertz data (where distinct transients originating at critical states become more apparent). The critical state encounter at an ωt of about 95°, ϕ = 13, is readily discernable. Later in the cycle, where transient effects overlap, a positive identification of other critical states is more difficult.

Significantly, responses for all three frequencies follow the static data closely for ωt in the range 0 to about 95° (26° ≥ ϕ ≥ 13°), then depart from the static curve. Note that the "dynamic overshoot" (from the static response at the ϕ = 13° critical state) increases with frequency. Also, the slopes of the dynamic responses just beyond the critical state are equal for all three frequencies, Since the independent variable in this plot is ωt, the initial part of the transient (ΔC_l) is proportional to ω. ΔC_l clearly depends on the motion history leading up to the critical state.

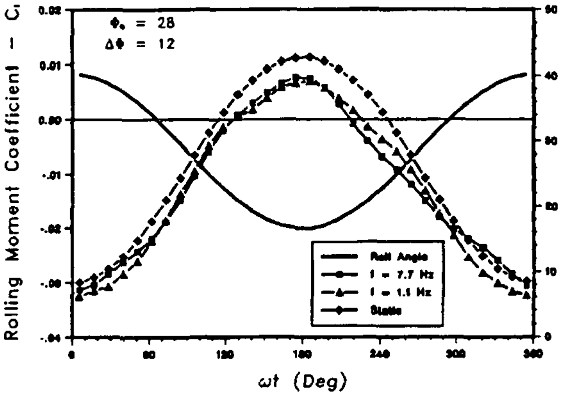


Figure 5b. Dynamic Rolling Moment

In addition, all three responses approach the static values at the end of the cycle, the deviation increasing with frequency. Thus, the rolling-moment responses are essentially quasi-steady *when transients due to critical state encounters have had time to die out.* This notion is supported by Fig. 5b which shows static and dynamic rolling-moment data for 12° oscillations about a mean roll angle of 28°. Thus, the roll-angle range covered by this motion, 16° to 40°, excludes all suspected critical states. Again, the rolling-moment behavior is quasi-steady. Differences between the static and dynamic data are well within interpolation errors caused by the sparsely spaced static data in this region.

Even in the nonlinear case, with critical states present, the total response under dynamic conditions may be separated into a static component (evaluated at the instantaneous roll angle) and a dynamic component[14]. This is a very useful device because changes in dynamic behavior (relative to the equilibrium state) are highlighted. Since transients at critical states represent the dynamic transition between dramatic changes in the static behavior, their presence is more easily detected in the dynamic component. Hereafter, the term "dynamic," when applied to a force/moment coefficient, is used in this more restrictive sense. "Total," when used in the same context, implies the sum of the static and dynamic components.

The results discussed above (Figs. 5a and 5b) for two motions also apply over a wide range of test conditions. Typical results are shown in Figs. 6a and 6b. Both of these are contour plots of the *dynamic* rolling moment coefficient presented in the phase plane. Data at 4.4 Hertz (k = 0.08) is presented for the range of test amplitudes at a given roll-angle offset. A series of tests with fixed offset and frequency generates a family of ellipses, centered about the offset angle. As the rolling motion proceeds, the ellipses are traversed in the clockwise direction. Offsets of 0° and 14° are shown in Figs. 6a and 6b respectively.

In both cases, the contour lines turn rapidly, becoming essentially parallel to the phi-dot axis at $|\phi| = 5°$. Moving clockwise in the bottom half of the ellipse through $\phi = -5°$, the contour lines again turn rapidly between -10° and -15° to run nearly parallel to the phi axis. Note that this "turning point" is less distinct than the first, perhaps because the static data has been faired in this region. The pattern repeats in the upper half of the figure. Significantly, the loci of turning points (critical states) are independent of roll rate (dependent only on roll angle). This is, of course, because critical states by definition represent discontinuous changes in the *equilibrium* state.

Following a critical state encounter, the *dynamic* response becomes weakly dependent on roll angle (contour lines nearly parallel to phi axis). The contours in this region are likely the result of expressing the implicit time variation in the phase plane (rather than an actual dependence on either roll angle or roll rate).

Finally, note that as the offset roll angle is increased (allowing more time to elapse between encounters with the critical state at 10° to 15°, first with positive phi-dot, then with negative roll rate) the region of negligible dynamic rolling moment in the lower-right quadrant expands. (There is time for the response to become quasi-steady).

5. SIMULATION AND ANALYSIS

In this section, the implications of critical state encounters to flight simulation and/or analysis of aircraft stability and control are discussed. Although the "locally linear model" has come into question in recent years,[7] the consequences of employing this technique when critical states are present are addressed. This is followed by a brief examination of plausible alternatives.

5.1 Locally Linear Representation

Often, even in flight simulations that are billed as "fully nonlinear," a locally linear model for the aerodynamic forces and moments is used. In this model, the static forces/moments are represented by a nonlinear function of the instantaneous values of angle of attack and sideslip. Dynamic effects are calculated by using locally linear "damping" derivatives (linearized about the instantaneous vehicle state). Linearization of the damping derivative is effected by using small amplitude

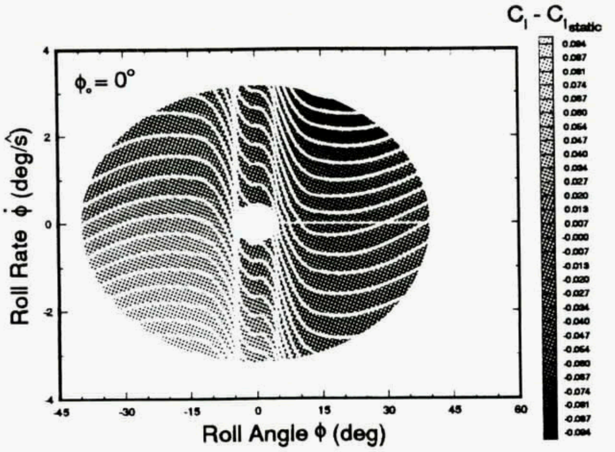


Figure 6a. Dynamic Component of Rolling Moment

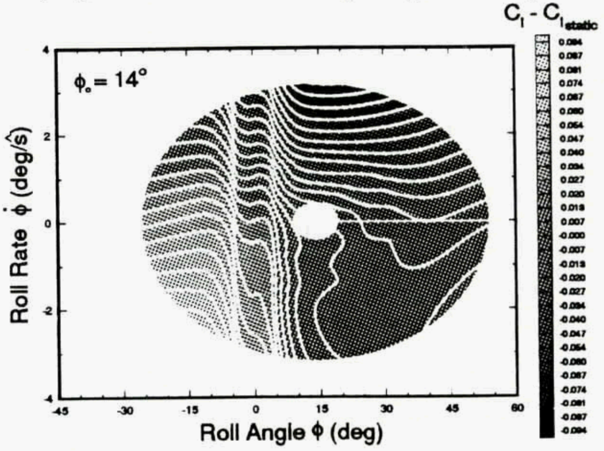


Figure 6b. Dynamic Component of Rolling Moment

test data and retaining only the measured aerodynamic loads at the forcing frequency (and 90° out of phase with the motion). These tests are repeated for a variety of offset angles. Thus, dynamic nonlinearities are accounted for by allowing the derivatives to be functions of angle of attack and sideslip. This procedure becomes questionable when there are significant motion-history-dependent transient effects such as those discussed above.

A dramatic example[7] of the errors that can result from the application of the locally linear model under inappropriate circumstances is presented in Fig. 7. Comparisons between the measured free-to-roll time history and predicted motions (based on the dynamic force and moment data) are shown. With the locally linear model, there are gross errors in the frequency and damping of the free-to-roll motion. The calculated response even finds the wrong trim condition. On the other hand, the motion predicted using the "reaction hypersurface" model[6,7] correlates well with the actual free-to-roll response. Both aerodynamic models used for this comparison were based on data taken at the same reduced frequency (k = 0.14).

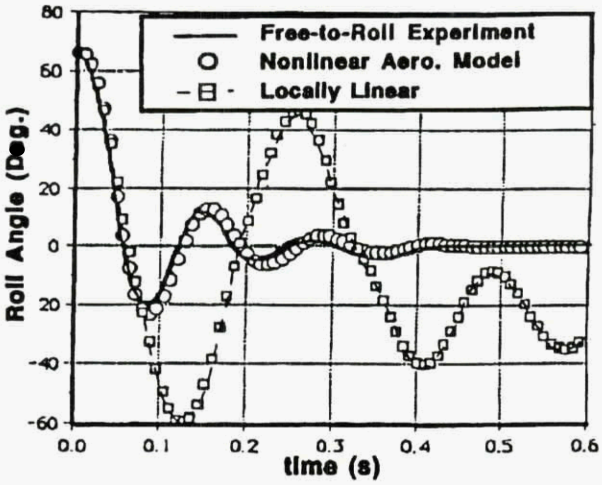


Figure 7. Free-to-Roll Time History

The reason for such poor results (with locally linear damping derivatives) is illustrated in Figs. 8a and 8b. Here, the locally linear model was used to "predict" the measured rolling moment (with all harmonics) over the same motion as used to determine the damping derivative. Nonlinear static rolling moment data (from SARL) was used together with roll damping derivatives obtained from 5° amplitude tests. The damping derivative was calculated by retaining only the out-of-phase rolling moment at the forcing frequency, although up to 20 harmonics were recorded.

Figure 8a shows the comparison for an offset roll angle of 3° and a frequency of 7.7 Hertz (k = 0.14). The result is totally unacceptable. Over much of the cycle, the model predicts a positive rolling moment, while the actual response is the opposite. Note that there is little evidence of a rate effect in the actual response. Rather, it appears to be a response nearly 180° out of phase with the motion. Yet the (negative) in-phase component does not agree with the static rolling moment at all, a fact previously observed in Ref. 8. (The calculated damping derivative is small and the locally linear model does not deviate much from the static data).

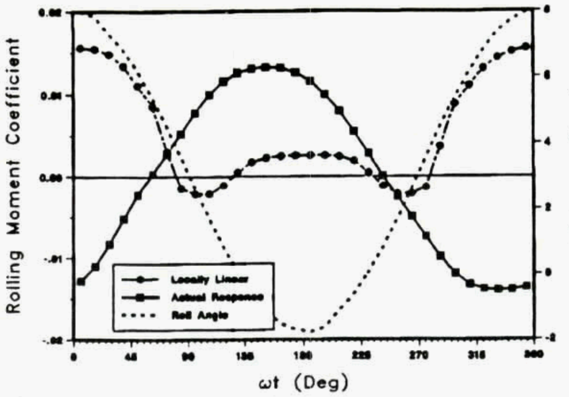


Figure 8a. Actual Load and Locally Linear Model (k = 0.14)

Agreement at 1.1 Hertz (Fig. 8b) is better but the result is still poor. However, the improvement is almost entirely due to the fact that the actual response follows the static behavior more closely. The difference between static and total response is still quite surprising, given the low reduced frequency (0.02). At this condition the *total* response crosses the static curve at only two points, suggesting perhaps that discernable critical state transient effects persist for at least a half cycle.

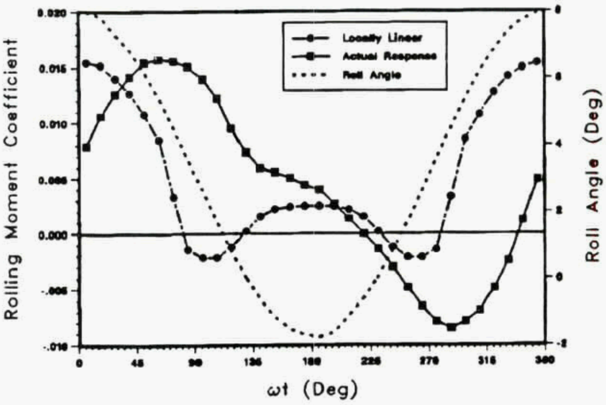


Figure 8b. Actual Load and Locally Linear Model (k = 0.02)

This particular motion, 3° offset and 5° amplitude, was chosen to include the critical state at $\phi = 5°$. Thus, the attempt to use conventional methods to represent the *transient* behavior provoked by the critical state encounter was ill-advised (except for illustrative purposes) from the beginning. Errors incurred with the locally linear model were *not* due to a poor representation for the damping moment, rather they were caused by the assumption that the in-phase component reflects the static aerodynamic contribution. This is simply *not* the case when significant transient effects are present.

5.2 Alternative Approaches

Clearly, the success of the nonlinear indicial response theory, in providing a rational framework for understanding the data reviewed above, establishes this approach as an extremely valuable "diagnostic" tool. For applications to simulation or analysis, simplifications to the model have been examined by several authors.[3,8] In Ref. 8, for example, the connection between the hypersurface model and the NIR was established. (The hypersurface representation can be derived from the NIR if the motion is analytic in the strict mathematical sense). Even so, based on the analysis of the present data, certain additional simplifications may be possible.

The 65° delta wing (at the flight conditions studied to date) has shown little important dynamic effects *except following critical state encounters.* Therefore, in this case, the integral terms in Eq. (1) may be accurately modeled by the stability derivative (locally linear) model. Of course, the resulting simplification is substantial because the need to include nonlinear superposition integrals in the equations of motion would be avoided. However, critical states must still be acknowledged as shown above and correct handling of the transient terms, $\Delta C_l(t{:}\phi_c)$ is still a problem. This is less difficult for simulation than analysis. The principal difficulty in the former case is identification of the transient term (including history effects) from test data.

Hanff's hypersurface model offers an approach; however, the complexities of tracking elapsed times[8] for multiple critical state encounters have not been worked out. The number of hypersurface dimensions (derivatives of the motion variable) required to both provide sufficient accuracy in the implicit time function and to account for history effects on $\Delta C_l(t{:}\phi_c)$ is unknown.

For analysis applications, Etkin[16] has shown (for the linear case) that "aerodynamic transfer functions" can be used when transient effects are present. However, for rates typical of "rigid-body" aircraft dynamics, significant transients seldom appear in a linear fashion. Nevertheless, it may be possible to extend this approach to the nonlinear case by the use of higher-order transfer functions.[17] Much more work is needed in this area.

6. CONCLUSIONS

An analysis of static and dynamic roll-motion data for a 65° delta wing has shown the existence of critical states, defined in terms of the configuration's static behavior. This finding was entirely consistent with the nonlinear theory advanced by Tobak and his colleagues. The particular characteristics observed for the delta-wing model were that:

(1) The critical states correspond to static roll angles where the leading-edge vortex breakdown position is either at the wing vertex or at the trailing edge.

(2) Significant dynamic effects (of the same order as the static rolling moments) were observed. These effects are in fact critical state transients. The transients persist for surprisingly long times, becoming identifiable only at extremely low reduced frequencies.

(3) Outside of critical state encounters, the dynamic contribution to the aerodynamic loads are small. In this case, considerable simplification of mathematical models describing the behavior is possible.

(4) Serious errors result from applying the locally linear model to cases involving critical state encounters. These errors are caused by the large contribution of critical state transients to the in-phase component of the measured loads.

ACKNOWLEDGEMENTS

This work was conducted under a Joint Research Program of the U.S. Air Force Office of Scientific Research, Wright Laboratory, the Institute for Aerospace Research and the Canadian Department of National Defence. The support of these organizations is gratefully acknowledged.

REFERENCES

1. Tobak, M., Chapman, G. T. and Unal, A., "Modeling Aerodynamic Discontinuities and Onset of Chaos in Flight Dynamical Systems," *Annales des Telecommunications*, tome 42, no. 5-6, 1987, pp. 300-314. Also NASA TM-89420, 1986.

2. Tobak, M., Chapman, G. T. and Schiff, L. B., "Mathematical Modeling of the Aerodynamic Characteristics in Flight Mechanics," Proc. Berkeley-Ames Conference on Nonlinear Problems in Control and Fluid Dynamics, *Math. Sci. Press*, 1985. Also NASA TM-85880, 1984.

3. Tobak, M. and Chapman, G. T., "Nonlinear Problems in Flight Dynamics Involving Aerodynamic Bifurcations," *AGARD Symposium on Unsteady Aerodynamics - Fundamentals and Applications to Aircraft Dynamics*, AGARD CP 386, 1985, pp. 25-1 to 25-15.

4. Tobak, M. and Unal, A., "Bifurcations in Unsteady Aerodynamics," NASA TM-88316, 1986.

5. Truong, K. V. and Tobak, M., "Indicial Response Approach Derived from Navier-Stokes Equations: Part 1 - Time-Invariant Equilibrium State," NASA TM-102856, 1990.

6. Hanff, E. S., "Dynamic Nonlinear Airloads - Representation and Measurement," *AGARD Symposium on Unsteady Aerodynamics - Fundamentals and Applications to Aircraft Dynamics*, AGARD CP 386, 1985, pp. 27-1 to 27-14.

7. Hanff, E. S., "Large Amplitude Oscillations," *Special Course on Aircraft Dynamics at High Angles of Attack: Experiments and Modelling*, AGARD Report No. 776, 1991, pp. 4-1 to 4-19.

8. Jenkins, J. E. and Hanff, E. S., "Non-linear Airloads Hypersurface Representation - A Time Domain Perspective," *Manoeuvring Aerodynamics*, AGARD CP 497, 1991, pp. 7-1 to 7-9.

9. Hanff, E. S., Kapoor, K. and Anstey, C. R., "Large-Amplitude High-Rate Roll Oscillation System for the Measurement of Non-linear Airloads," AIAA Paper 90-1426, 1990.

10. Hanff, E. S. and Jenkins, S. B., "Large-Amplitude High-Rate Roll Experiments on a Delta and Double-Delta Wing," AIAA Paper 90-0224, 1990.

11. Hanff, E. S. and Ericsson, L. E., "Multiple Roll Attractors of a Delta Wing at High Incidence," *AGARD Symposium on Vortex Flow Aerodynamics*, AGARD CP 494, October 1990.

12. Hanff, E. S. and Huang, X. Z., "Roll-Induced Cross-Loads on a Delta Wing at High Incidence," AIAA Paper 91-3223, September 1991.

13. Arena Jr., A. S., Nelson, R. C. and Schiff, L. B. "An Experimental Study of the Nonlinear Dynamic Phenomenon Known as Wing Rock," *AIAA Atmospheric Flight Mechanics Conference*, AIAA-90-2812-CP, August 1990, pp. 173 to 183.

14. Jenkins, J. E., Myatt, J. H., and Hanff, E. S., "Body-Axis Rolling Motion Critical States for a 65-Degree Delta Wing," (to be published).

15. Payne, F. M. and Nelson, R. C., "An Experimental Investigation of Vortex Breakdown on a Delta Wing," *Vortex Flow Aerodynamics - Vol. I*, NASA CP 2416, October 1985, pp. 135 to 161.

16. Etkin, B., *Dynamics of Atmospheric Flight*, Wiley, New York, 1972, pp. 165-168.

17. Jenkins, J. E., "Simplification of Nonlinear Indicial Response Models: Assessment for the Two-Dimensional Airfoil Case," *Journal of Aircraft*, Vol. 25, No. 2, 1991, pp. 131-138.

STABILITY MODEL OF THE ATMOSPHERE

by

Andreas Knüppel and Daniel Martens
Institut für Flugführung
Technical University Braunschweig
Germany

Andreas H. Siemer
Institut für Meteorologie und
Klimatologie
University of Hanover
Herrenhauser Strasse 2
W-3000 Hanover
Germany

Abstract

The atmosphere of the earth is a very complex system covering a wide range of interacting scales. The concept of stability is applied to subsystems thereof where the quantities involved in a stability analysis depend on the particular question to be answered. But all stability considerations share a common basic structure. After giving a short account to the nature of stability investigations some examples of stability related atmospheric phenomena are presented which are relevant for flight operation. Finally the impact of atmospheric instabilities on aircraft performance with special regard to flight safety is illustrated.

List of Symbols

c_p	$J/(kg\,K)$	spec. heat of dry air at const. pressure
c_L	—	lift coefficient
f	Hz	frequency
f_w	Hz	gust frequency
g	m/s^2	gravitational acceleration
l_μ	m	mean aerodynamic chord
m	kg	mass
n	—	normalized wavenumber
p	Pa	pressure
p_0	Pa	reference pressure = $10^5\ Pa$
s	$1/s$	Laplace variable
t	s	time
u_i	m/s	velocity component
u_*	m/s	friction velocity
w'	m/s	vert. velocity fluctuation
x_i	m	Cartesian coordinate
z	m	altitude
A_α	kg/s	lift factor
E_{kin}	J/kg	kinetic energy
H	m	height
K	kg/s	gain
L	N	lift
L_w	m	gust wave length
L_*	m	Monin-Obukhov-Length
R	$J/(kg\,K)$	gas constant for dry air
S	m^2	wing area
S_{u_i}	m^3/s^3	spectral density of u_i as a function of wavenumber
T	K	temperature
T	s	time constant
T_k	s	time constant of Küssner
V	m/s	air speed or wind speed
V_k	m/s	kinetic velocity
V_w	m/s	wind speed
W	N	aircraft weight
$\mathcal{L}$	J/kg	generalized Lagrangian function
α	deg	angle of attack
γ	deg	flight path angle
κ	—	von-Karman-constant
ρ	kg/m^3	density
θ	K	potential temperature
θ'	K	pot. temperature fluctuation
Δt	s	pilot's time lag
Λ	—	aspect ratio
Φ	m^2/s^2	geopotential
Φ_ϵ	—	normalized dissipation rate
Ψ	J/kg	generalized potential
Ψ	—	Küssner function

1 Introduction

The concept of stability is a very elementary one in nearly all branches of everyday life (including such difficult fields as politics). This is especially true for the natural sciences and engineering applications, presumably because many stability related problems in these fields can be tackled by mathematical methods. Much of our notion of stability derives from the occupation with problems of classical mechanics and – on a more sophisticated level – those of thermodynamics. Elementary examples of stability problems can be found in nearly every introductory textbook on physics or engineering. These elementary problems are characterized by the fact that they can be represented by a finite number of points in the phase space. To the contrary, hydrodynamic processes are characterized by the simultaneous existence of phenomena with different space and time scales, the characteristic measures of which cover a continuous subset of the space-time continuum, and beyond that which are interacting with each other.

Therefore the first step in a stability investigation is to confine the consideration to a distinct phenomenon with a definite space and time scale – which clearly must be an abstraction of physical reality – and to apply the assumption that the interaction with other scales can be neglected. The governing equations are thus reduced by neglecting terms of minor influence while keeping them consistent with the physical reality observed in a given scale. We thus make a priori assumptions about the solutions based on the perception of physical reality and feed them back into the governing equations in order to reduce the manifold of solutions. This invokes the assumption that small causes have small effects, which is clearly not true for nonlinear systems such as the atmosphere if time intervals extending beyond a critical value are considered[1] [1, 2, 3]. But nevertheless it is possible to get much insight into the short range future of distinct atmospheric disturbances by investigating the instantaneous forces or energy exchanges, respectively. In the following we shall confine ourselves to local disturbances in the atmosphere and study the immediate interaction with their environment.

2 Stability concepts

The concept of stability has a wide range of application in atmospheric science extending over nearly all scales of atmospheric motion. Basic instability mechanisms such as static stability, inflection point instability of shear flows, baroclinic instability, inertial instability etc. are treated in almost every textbook on dynamic meteorology (see e.g. [4]). Besides there is a number of classical monographs on hydrodynamic stability in general, as for instance the excellent books of Chandrasekhar [7] and Drazin & Reid [8]. But much has still to be understood so that problems of hydrodynamic stability – especially in geophysical fluid dynamics – are still a matter of intense research, which received increasing interest in recent years (see e.g. [5]).

Without going into the formal theory of stability (see e.g. [9]) we present an approach here which is rather heuristic, but in our view nevertheless appealing. It elucidates the underlying concept of stability, which is shared by all stability problems, and suggests that stability considerations can be viewed as merely a special aspect of a Hamiltonian formulation of atmospheric dynamics as the fascinating overall concept.

A mandatory prerequisite for all stability investigations is the definition of a basic state with regard to a given question, and the definition of what will be considered as a perturbation thereof. Once having defined these we shall use the term "stable" for any situation where the energy of a perturbation is "consumed" by the basic state, and therefore the perturbation is damped out, and we shall use the term "unstable" for any situation where a perturbation – once initialized – extracts energy from the basic state. This is indicated in fig. 1 by the bubbles on the right hand side, where the bubble we start with is growing or shrinking depending on the direction of the flow of energy.

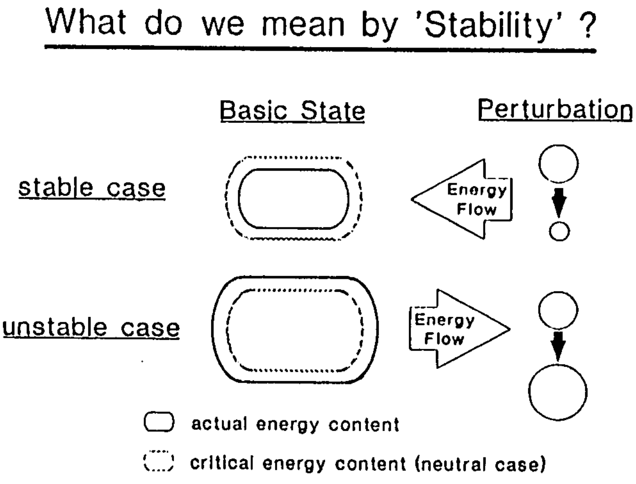


Figure 1: Scheme of energy tranfer between a perturbation and its environment depending on stability.

But when is a given basic state stable or unstable? Obviously the preceeding definitions do not suffice to answer this question and an additional term must be introduced. Experience shows that in all considerations of stability related processes a unique value of the energy content of the local basic state (with respect to an arbitrary reference level) can be defined. If the energy content of the basic state exceeds this value then perturbations are promoted by an energy flux toward the perturbations, if it falls short of this value then perturbations are supressed. We shall call this value the "critical energy content" of the basic state which is marked in fig. 1 by the dotted symbol. It coincides with the so called "neutral" case of stability when a perturbation is neither growing nor ceasing and is just left what it is.

This conceptual model immediately suggests a procedure for stability investigations: we start with an arbitrary equilibrium state and make some kind of virtual perturbation. We suspect that the decisive quantity we have to look at is the difference

$$\Psi - E_{kin} \qquad (1)$$

between the potential energy Ψ and the kinetic energy E_{kin}, where we use the term "potential

---

[1] It is interesting to mention that speculations about intrinsic features of the atmosphere that limit its predictability and that are related to the uncertainty in initial conditions have already been made a long time before the study of chaotic phenomena (see e.g. [6]).

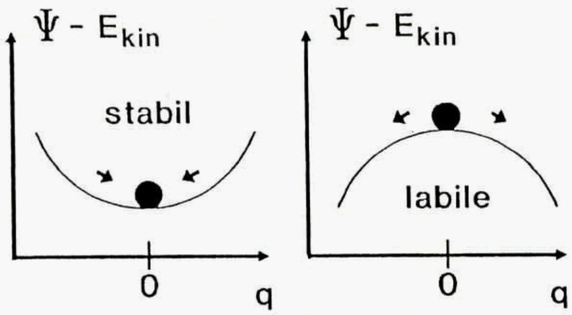
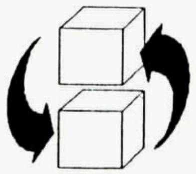


Figure 2: Local equilibrium states of a system. Assumptions: a) the "forces" on a perturbation derive from a generalized potential, b) a displaced fluid element always adopts the velocity from its immediate environment.

energy" in a generalized sense as will become apparent below.[2] Therefore we consider the new value of the difference between the potential energy and the kinetic energy of the perturbed element (or in other words, the new value of the excess potential energy over the kinetic energy). Fig. 2 presents a mnemonic scheme: a one-dimensional system with coordinate q investigated at an arbitrary point $q \equiv 0$; we consider the "landscape" of $\Psi - E_{kin}$ over q; the "height" of the black ball represents the local state of the system, and the expected reaction of the ball under an imaginary gravity force to a virtual displacement parallels the expected reaction of the system state to virtual perturbations (either away or back to the equilibrium position it started from). If the value of the excess potential energy is increased by the virtual perturbation then the basic state should be stable, meaning that the energy content of the system is below the critical value; if it is decreased we should have instability, and the basic state would try to get rid of its excess energy by putting it into the perturbation; if it remains constant we should just have the neutral case. We shall see below that this criterion can indeed be derived from a Hamiltonian formulation of the stability problem.

In completing our picture we assume that the overall basic state is not seriously affected by a small perturbation and can therefore be considered to remain unaffected for the moment of disturbance.

If we talk about stability in case of a liquid or gas we refer to a small element which is displaced from its equilibrium position, and we shall examine the subsequent flow of energy between this small element and its environment. But wait – we have to take into account the conservation of mass while making the virtual perturbation. Therefore it will be simply assumed that two small fluid elements of equal mass, which are adjacent to each other, are interchanged (see marginal fig.).



Then, if we want the system to be stable at the point around which the interchangement has taken place we must assure that stability is achieved by both fluid elements. So we have to focus our attention to that element which is most "likely" to become unstable. For a one-dimensional problem the foregoing statement is equivalent to the instruction to make a displacement of a fluid element in both directions and to study stability in both cases.

How can we cast this concept into mathematical terms? The procedure just described reminds us of D'Alembert's principle of virtual work. Since this principle can be derived from the Lagrangian equations, which in turn are an implication of Hamilton's principle of least action [10], it may be argued that our stability considerations can be derived in a similar manner from a general "Hamilton-like" principle.

And indeed, if we displace an arbitrary fluid element by a small amount starting from an equilibrium position then the forces acting on the element will result from the change of its position in the potential field Ψ and from the momentum transfer between the particle and its environment so that

$$\frac{du_i}{dt} = u_k \frac{\partial u_i}{\partial x_k} - \frac{\partial \Psi}{\partial x_i} \quad , \tag{2}$$

where we have applied the tensor notation for the Cartesian coordinates x_i (with $i = 1, \ldots, 3$) and the respective velocity components u_i employing Einstein's summation rule.[3] If for the present we disregard the force term deriving from a potential it can be shown that the remaining expression $du_i/dt = u_k \partial u_i / \partial x_k$ can be rewritten as (see appendix)

$$\frac{d}{dt}\frac{\partial E_{kin}}{\partial u_i} - \frac{\partial E_{kin}}{\partial x_i} = 0 \quad , \tag{3}$$

where E_{kin} is the kinetic energy of the fluid element. Reintroducing the potential force term on the right hand side and assuming that the potential Ψ is independent of velocity this leads to

$$\frac{d}{dt}\frac{\partial}{\partial u_i}\left[E_{kin} - \Psi\right] - \frac{\partial}{\partial x_i}\left[E_{kin} - \Psi\right] = 0 \quad . \tag{4}$$

After we have defined the general Lagrangian function to be

$$\mathcal{L} := E_{kin} - \Psi \tag{5}$$

---

[2] If we talk about kinetic energy in this context we always refer to the kinetic energy of the perturbed element which is aquainted from or lost to the basic state. We do not mean something like the kinetic energy of the displacement itself, which doesn't actually exist, since no variation of time is performed.

[3] The summation rule requires that a summation must be carried out over equal indices in a product, e.g. $u_k \partial u_i / \partial x_k \equiv \sum_{k=1}^{3} u_k \partial u_i / \partial x_k$ and $\partial u_j / \partial x_j \equiv \vec{\nabla} \cdot \vec{u}$.

we are simply left with

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial u_i} - \frac{\partial \mathcal{L}}{\partial x_i} = 0 \quad . \tag{6}$$

Notice that $\mathcal{L}$ is equal to expression (1) with its sign reversed so that a well of (1), which indicates stability, corresponds to a ridge of $\mathcal{L}$. These are the well-known Euler-Lagrange differential equations for $\mathcal{L}$, which are a direct implication of Hamilton's principle, the variational formulation of which is

$$\delta \int_{\Delta t} \mathcal{L}\, dt = 0 \quad . \tag{7}$$

This states that the mean kinetic energy of a system produced by external forcing is always kept at a minimum. Stull ([11], p. 173) notes remembering LeChatelier's principle that "for ... many instabilities it is interesting to note that the fluid reacts in a manner to undo the cause of instability" — and we add here: this is accomplished by the fluid in the most effective way.[4]

It also shows that the quantity given by expression (1) provides indeed a stability criterion as we have suspected above. We conclude that the atmosphere is stable at a given point if $\mathcal{L}$ is at a maximum. Hence we arrive at the equilibrium condition

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0 \tag{8}$$

since $\mathcal{L}$ is at an extremum and we have the general stability criterion

$$\frac{\partial^2 \mathcal{L}}{\partial x_i^2} \cdot |\delta x_i| \;\Big\}\!\!\left[\begin{array}{lcl} < 0 & : & \text{stable} \\ = 0 & : & \text{neutral} \\ > 0 & : & \text{unstable} \end{array} \right. \quad . \tag{9}$$

Although classical Hamiltonian theory only applies to finite-dimensional phase spaces spanned by the generalized coordinates and momenta the concepts can be extended to continua [13, 14]. But we shall not go into details here.

Since the forces occuring in the atmosphere are not purely mechanical, but also derive from thermodynamic processes and are a consequence of the earth's rotation, we must postulate the existence of a generalized potential from which the forces can be derived (perhaps remembering that the Lorentz-force in electrodynamics can also be derived from a generalized potential.) This would naturally lead to a generalized Lagrangian function, which would also include thermodynamic and rotation dependent terms. Such a generalized Lagrangian has already been used by Ertel [15] (see also [16]).

Let us now illustrate this view in three examples: the first may serve as an introductory one, the second concerns static, and the third dynamic stability in the atmosphere. In selecting the last two examples we have also taken into account that they ought to be relevant to flight operation.

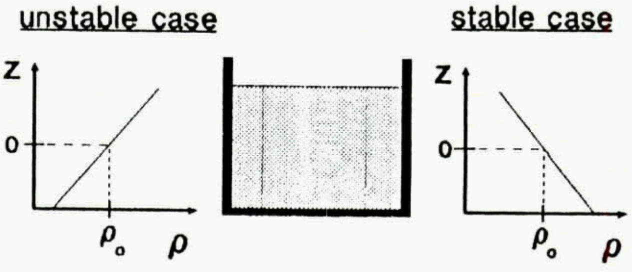


Figure 3: Incompressible fluid with stable and unstable density stratification.

Consider an incompressible fluid with a linear density profile

$$\rho(z) = \rho_0 + \gamma z \tag{10}$$

and assume horizontal homogeneity (see fig. 3) so that the problem can be reduced to one dimension (along the vertical axis). In this case the potential Ψ results from gravity and buoyancy forces and the Lagrangian is simply

$$\mathcal{L}(z) = -\Psi = \frac{\gamma g}{2\rho_0} z^2 \quad . \tag{11}$$

Stricly speaking this equation describes the potential experienced by single fluid elements starting at $z = 0$ (where $\rho = \rho_0$) whose density differs from the surrounding liquid by γz according to eq. (10). But this means no loss of generality since we are free to choose $z = 0$ arbitrarily at our point of investigation. According to eq. (9) the system is just neutral at $z = 0$ if we have

$$\frac{\partial^2 \mathcal{L}}{\partial z^2} = 0 \quad .$$

This results in

$$\frac{\gamma g}{\rho_0} = 0 \quad . \tag{12}$$

Since g and ρ_0 are $\neq 0$ we conclude that $\gamma = 0$ corresponds to the neutral state, whereas a stably stratified fluid corresponds to negative values of γ and vice versa, which surely is clear by our everyday physical experience. The result can also be interpreted as follows: the parameter γ appearing in eq. (11) is a shape parameter of the potential function which has a potential well if $\gamma < 0$, is flat for $\gamma = 0$, and has a potential hill for $\gamma > 0$. We can generalize this result by saying that overall stability is guaranteed if the density gradient is positive at every point within a fluid.

As a second example let us consider a somewhat more complicated case: a static dry atmosphere which is characterized by horizontal homoge-

---

[4] A variational formulation has already been used by Raethjen [12] to explain the dynamics of squall lines.

neity and therefore reduces to a one-dimensional system where no heat transfer occurs and hence reversible thermodynamics is assumed. The so called "adiabatic temperature lapse rate" represents neutral stability of this atmosphere. The derivation of an equation for this lapse rate – applying the first law of thermodynamics, the equation of state, and the hydrostatic equation – is standard since the early days of geophysical thermodynamics and can be found in almost any introductory course on meteorology (see e.g. [17, 18]). But we can also apply the concept stated above. In this case the generalized Lagrangian is

$$\mathcal{L} = -(\Phi(z) - c_p T(z)) \quad , \qquad (13)$$

where $\Phi(z) \approx gz$ is the geopotential height and $c_p T(z)$ is the enthalpy.[5] The equilibrium condition $\partial\mathcal{L}/\partial z = 0$ leads to

$$\frac{\partial T}{\partial z} = -\frac{g}{c_p} \qquad (14)$$

for the temperature lapse rate and since $\partial^2 T/\partial z^2 = 0$ we conclude according to the stability criterion (9) that this corresponds to the neutral state. Likewise we can show that $\partial T/\partial z > -g/c_p$ results in stable and $\partial T/\partial z < -g/c_p$ in unstable conditions so that we finally get the stability criterion for a dry atmosphere

$$\frac{\partial T}{\partial z} \left\{ \begin{array}{lcl} > -g/c_p & : & \text{stable} \\ = -g/c_p & : & \text{neutral} \\ < -g/c_p & : & \text{unstable} \end{array} \right. \quad , \qquad (15)$$

where $g/c_p \approx 1\ K/100\,m$ (see fig. 4).

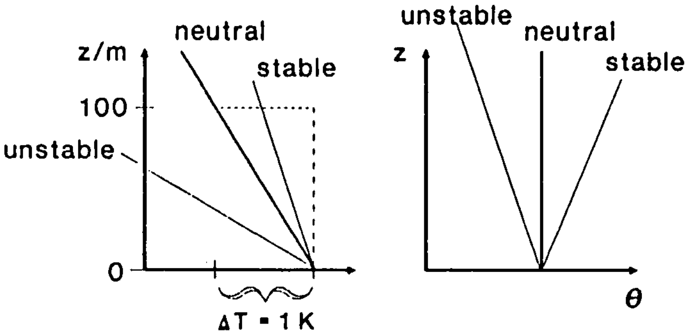


Figure 4: Actual temperature profile and potential temperature profile for different stabilities.

Often this condition may be expressed more conveniently in the so called p-System with pressure as a vertical coordinate. By considering the temperature changes within an air parcel under adiabatic pressure changes we can define the so called "potential temperature" by

$$\theta := T \cdot \left(\frac{p_0}{p}\right)^{R/c_p} \quad , \qquad (16)$$

which is merely a special version of Poisson's equation relating the temperature of an ideal gas in an adiabatic process to a reference pressure of $p_0 = 10^5 Pa$. From this formula we can derive the condition that an atmosphere is in neutral equilibrium if the potential temperature is constant for all p, and that it is stable for θ increasing with p, and unstable otherwise [18]. Since p and z are monotonous in the atmosphere the same statement applies by replacing p by z in the preceding statement. So if we employ the potential temperature instead of the actual temperature for atmospheric problems we get (see fig. 4)

$$\frac{\partial \theta}{\partial z} \left\{ \begin{array}{lcl} > 0 & : & \text{stable} \\ = 0 & : & \text{neutral} \\ < 0 & : & \text{unstable} \end{array} \right. \quad . \qquad (17)$$

(Notice that if we replace θ by the actual temperature T this criterion holds for incompressible fluids! So the atmosphere is by this view in a way "reduced" to an incompressible fluid.)

Whereas in the preceding examples the Lagrangian comprises only potential energy terms, we shall now turn to a case which involves kinetic energy, too. We consider a stratified shear flow with a constant gradient of density γ_ρ and horizontal velocity γ_v according to

$$\rho(z) = \rho_0 + \gamma_\rho z \qquad (18)$$

$$v(z) = v_0 + \gamma_v z \qquad (19)$$

and shall study the virtual displacement of a fluid element in the vertical. The potential of an element displaced from $z = 0$ is given by

$$\Psi = -\frac{\gamma_\rho g}{2\rho_0} z^2$$

just as in the first example (compare eq. (11)!), and the kinetic energy that the fluid element gains from its "new" environment if it is displaced to a larger value of z is

$$E_{kin} = \frac{\gamma_v^2}{2} z^2$$

so that the Lagrangian is

$$\mathcal{L}(z) = \frac{\gamma_v^2}{2} z^2 + \frac{\gamma_\rho g}{2\rho_0} z^2 \quad . \qquad (20)$$

Fig. 5 shows an example of a stably stratified fluid where the horizontal velocity increases with height and a fluid element is raised to a higher level thereby increasing its potential energy as well as its kinetic energy; the decisive question is, whether the

---

[5] In the early days of atmospheric thermodynamics there has been a debate whether the "heat content" $c_p T$ or the internal energy $c_v T$ is to be used in the derivation of the adiabatic lapse rate. It is interesting to mention that the advocates of both approaches arrived at the same result because the error of using $c_v T$ was just compensated by a second error in the derivation (cf. e.g. [19]).

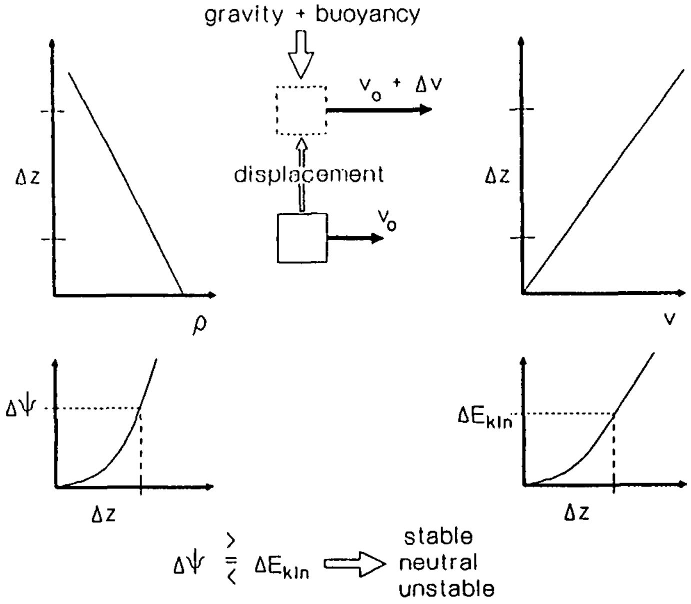


Figure 5: Displacement of a particle in a stably stratified shear flow.

potential energy increase exceeds the kinetic energy increase, or vice versa.

If we again apply our stability criterion (9) we get

$$\frac{(g/\rho_0)\gamma_\rho}{\gamma_v^2} \left\{ \begin{array}{lcl} < -1 & : & \text{stable} \\ = -1 & : & \text{neutral} \\ > -1 & : & \text{unstable} \end{array} \right. \tag{21}$$

In the atmosphere instead of density the potential temperature is used to describe the buoyancy. If we neglect pressure changes to the first order then we have from the logarithmic version of the equation of state of a gas and from (16)

$$\frac{d\rho}{\rho} \approx -\frac{dT}{T} \approx -\frac{d\theta}{\theta}$$

or equivalently

$$\frac{\gamma_\rho}{\rho_0} \approx \frac{1}{\theta_0}\frac{\partial\theta}{\partial z}$$

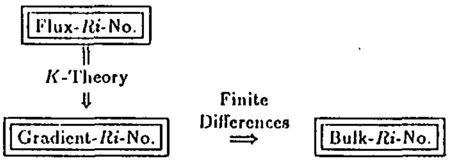
and we get instead of (21)

$$Ri := \frac{g}{\theta_0}\frac{\partial\theta}{\partial z} \Bigg/ \left(\frac{\partial v}{\partial z}\right)^2 \left\{ \begin{array}{lcl} > 1 & : & \text{stable} \\ = 1 & : & \text{neutral} \\ < 1 & : & \text{unstable} \end{array} \right. \tag{22}$$

The fraction in expression (22) is called the Richardson Number Ri (after L.F. Richardson). In applications instead of the Richardson number as defined in eq. (22) the so-called bulk Richardson number Ri_b is used which we get if we replace the differentials in (22) by finite differences. The critical value of Ri_b determined experimentally is about 0.25 so that we arrive at the stability condition

$$Ri_b = \frac{g}{\theta_0}\frac{\Delta\theta \cdot \Delta z}{(\Delta v^2)} \left\{ \begin{array}{lcl} > 0.25 & : & \text{stable} \\ = 0.25 & : & \text{neutral} \\ < 0.25 & : & \text{unstable} \end{array} \right. \tag{23}$$

The "traditional" approach for deriving the Richardson number starts from the prognostic equation for the turbulent kinetic energy where several simplifying assumptions are introduced until only the buoyancy term and the production term describing the generation of turbulent kinetic energy from shear are left (see fig. 6). The momentum flux occuring in the shear term and the heat flux in the buoyancy term are then expressed by the velocity and the potential temperature gradients, respectively, employing K-theory.



Flux Richardson Number Ri_f

- Start with balance equation of turbulent energy:

  Local change of turb. energy = − divergence of fluxes
  + production
  − consumption
  − viscous dissipation

- Making simplifying assumptions:
  - homogeneity ⇒ divergence of fluxes = 0
  - reversibility ⇒ viscous dissipation = 0
  - production by shear $-\overline{u_i'u_j'} \cdot \partial\overline{u_i}/\partial x_j$
  - production and consumption by buoyancy $g \cdot \overline{u_3'\theta'}/\overline{\theta}$

  leaves (considering only the main horizontal direction of flow)

$$\frac{\partial e}{\partial t} = -\overline{u'w'} \cdot \partial\overline{u}/\partial z + g \cdot \overline{w'\theta'}/\overline{\theta}$$

- Define

$$\boxed{Ri_f := \frac{g}{\overline{\theta}} \cdot \frac{\overline{w'\theta'}}{\overline{u'w'} \cdot \partial\overline{u}/\partial z}}$$

  then

  $\partial e/\partial t = 0$ (stationarity) ⇒ $Ri_f = 1$
  $\partial e/\partial t < 0$ (stability) ⇒ $Ri_f > 1$
  $\partial e/\partial t = 0$ (instability) ⇒ $Ri_f < 1$

Figure 6: Sketch of derivation of the Ri-number from the turbulent kinetic energy budget equation.

If the Ri-number exceeds its critical value then turbulent fluctuations are suppressed, if it is below the critical value then they give rise to fully developed turbulence.

3 Stability related phenomena with a view to flight operation

In the following we shall show that the overall mean state of the atmosphere is statically stable as well as dynamically stable with respect to the *Ri*-criterion. But we shall also see that there are strong deviations from the mean state. We shall start with a simple picture of atmospheric reality and study its implications concerning static stability. In order to come to the result first we state that we shall arrive at an atmosphere which is unstable, and subsequently we shall ask ourselves why this is not in agreement with the mean (stable) conditions found in the earth's atmosphere. Finally we shall throw some light upon the deviations from stable conditions and illustrate the dynamic consequences.

Well now, let us start with the promised oversimplified picture: it is widely known that the atmosphere is transparent for solar radiation to a large extent (and in fact, the amount of short wave radiation absorbed at the ground is approximately three times as large as the amount absorbed in the atmosphere [20]). To keep our picture simple we assume a dry clear atmosphere which is entirely transparent to solar radiation, i.e. absorption occurs only at the ground. Furthermore we shall ignore the (almost) spherical surface of the earth and anticipate a plane earth with constant insolation.[6] To prevent the earth from permanent heating the heat must be taken away from the surface. We assume that there is some "heat transporting mechanism" in the atmosphere, which may include short range terrestrial radiation. This mechanism transports energy from the surface to the upper region of the atmosphere, from where it is finally reradiated to space by long wave emission. But regardless of what the particular nature of the transport is, it is necessary that the potential temperature decreases with height; however, as we have seen in the preceding section, this means that the atmosphere is unstable.

As already indicated above, this result is not in agreement with the mean state observed in the atmosphere, which is well reflected by the U.S. Standard Atmosphere [22], which is shown in fig. 7 together with its potential temperature profile for the troposphere and lower stratosphere.

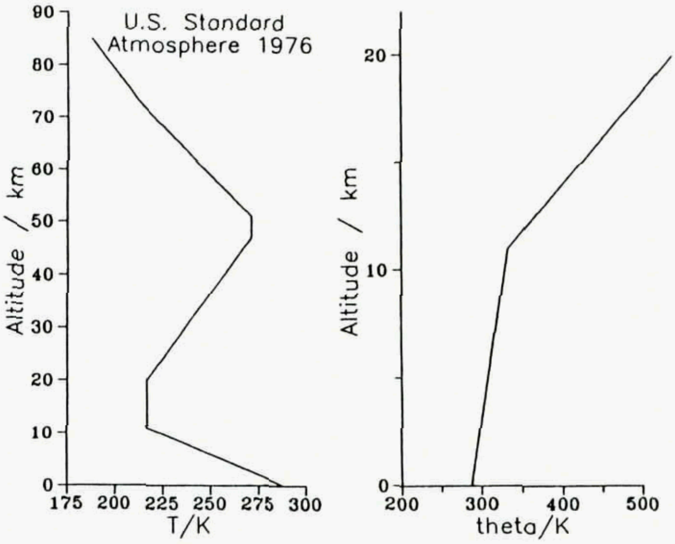


Figure 7: U.S. Standard Atmosphere [22] (left) and derived profile of potential temperature for the lowest 20 km (right).

Now, what is wrong with our model? Aside from the fact that we have neglected the differential heating of the earth as a consequence of its almost spherical shape and also the earth's rotation, the decisive oversimplification was that we have neglected the water and its phase changes within the atmosphere. Lifting moist air with vapour pressure below saturation starts cooling according to the dry adiabatic lapse rate until it reaches the condensation level. Above this level a continuous release of latent heat by condensation of water vapour takes place, thereby heating the air. A detailed analysis shows that the potential temperature increases above the condensation level by about $6\,K/km$ in the lower troposphere and by $4-3\,K/km$ in the middle troposphere. This is in close agreement with the potential temperature gradient in the standard atmosphere, which is about $3.5\,K/km$.

But we have to be careful: the model of the static and stably stratified atmosphere is an idealization which reflects the mean overall conditions in the earth's envelope; but to maintain this mean state the actual atmosphere cannot always be statically stable everywhere, even then if the earth would be a resting disk.

Fig. 8 shows a meridional cross section of the january northern hemisphere along 80 $deg\ W$ which extends from the equator on the left to the pole on the right. The bold solid lines indicate equal zonal wind speed, the thin solid lines are isotherms, and the dashed are those of equal potential temperature (values on the right margin). We can see that the highest values of vertical wind shear of the mean zonal wind occur at about $30-40\,deg\ N$. To make a crude estimation of *Ri*-number stability of the mean wind we take $40\,ms^{-1}/10\,km$ as a representative value for

---

[6] To be at least to some extend realistic the insolation should be taken as one quarter of the solar constant, since this is the mean energy flux density onto the earth's surface in case of an entirely transparent atmosphere (see e.g. [21], entry: "equivalent blackbody temperature"). To be complete we mention that the incident radiation is not equal to the radiation absorbed because the reflected part must be taken into account.

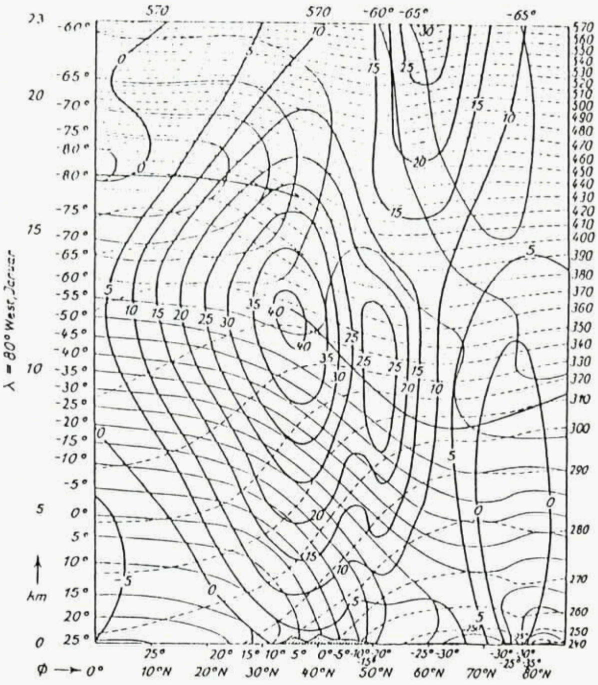


Figure 8: Meridional cross section through the northern hemisphere along 80 *deg W* after Landsberg and Ratner [23]. Bold solid lines = zonal wind speed; thin solid lines = isotherms; dashed lines = potential temperature.

the wind shear, and for the potential temperature a value of $+3.5\,K/km$ as indicated above. This leads to a Ri-number of about 15 so that the mean conditions must surely be considered to be stable.

The conditions on earth do not promote an equally distributed and steady input of latent heat into the atmosphere. Rather, we find temporary and local phenomena, which means that the energy exchanges, which are necessary to maintain the overall mean state, are highly concentrated; so the area related power of local disturbances may exceed the overall mean value by orders of magnitude. This is also true for the kinetic energy involved in these energy exchange processes. Whereas the global mean production rate of kinetic energy is of the order of a few W/m^2, in severe local storms we can find values one order of magnitude larger in severe midlatitude storms [24] and presumeably up to $100\,W/m^2$ and more in extreme cases.

A variety of phenomena which constitute deviations from the mean state is found in the planetary boundary layer because this is the region of the atmosphere where the main energy conversions occur. In the following we shall deal with some stability related phenomena within the planetary boundary layer which are relevant to flight.

To proceed to some flight mechanical considerations we shall relate the Ri-number to the turbulent structure of the planetary boundary layer and show how this number can be used to asses the energy transfer between an aircraft and its atmospheric environment.

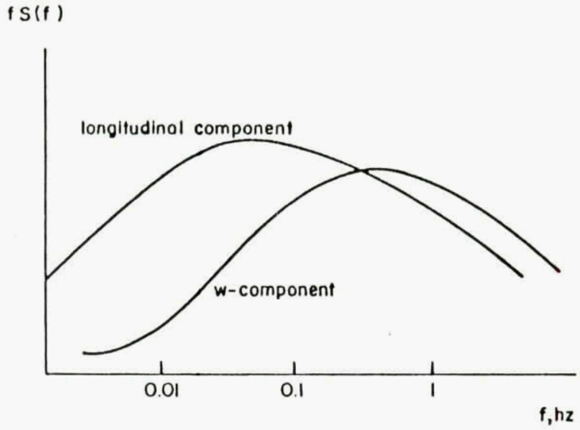


Figure 9: Sketch of turbulent energy spectra (from Panofsky & Dutton [25].

It is well known that the turbulent wind velocity fluctuations within the planetary boundary layer comprise eddies covering a wide range of sizes. The power spectra of the horizontal and vertical velocity components, respectively, show the distribution of energy over a certain size range. A given spectral diagram (with proper scaling of the axes) allows us to estimate the energy contained within a band of eddy sizes. Usually the spectral density is related to and plotted against the reciprocal of the eddy size, which is called the wave number, although we are not dealing with a periodic phenomenon. Assuming that the turbulent elements are advected to a fixed point with the mean wind V and that the turbulent structure doesn't change against the wind the wave-number can be converted into a frequency, which is measured by an observer at a fixed location. This assumption is known as "Taylor's Frozen Turbulence Hypothesis" [25]. A typical plot of a spectral curve is shown in fig. 9, where the abscissa is scaled logarithmically and the spectral density has been multiplied by the frequency so that equal areas under the curves correspond to equal energies.

Measurements within the planetary boundary layer have shown that the particular shape of a spectral curve depends on

— the distance from the ground

— the stability of the boundary layer air .

To find out universal relations that yield the shape of the spectral distribution of turbulent energy as a

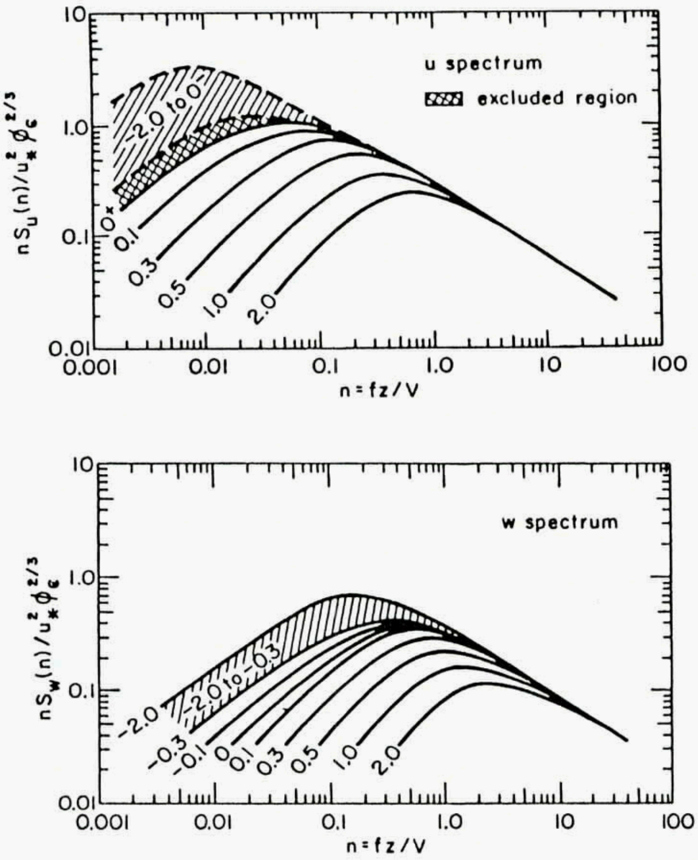


Figure 10: Normalized spectra of horizontal and vertical wind velocity fluctuations fur the surface layer after Kaimal [27] (reproduced by Panofsky & Dutton [25]).

function of a single dimensionless parameter it suggests itself to try a definition a length scale as a measure of stability from all quantities which are relevant for stability . This length scale should then be used to non-dimensionalize the distance from the ground, thereby leading to a dimensionless height as a shape parameter of the spectral curves.

There are different ways of scaling the boundary layer. An example which applies for the lower part of the boundary layer is the so-called Monin-Obukhov-scaling [26] where a characteristic length-scale is defined by

$$L_* = -\frac{u_*^3}{\kappa \dfrac{g}{\theta_0} \dfrac{w'\theta'}{\rho c_p}} \quad , \qquad (24)$$

where u_* is the friction velocity, κ the von-Karman-constant, and $w'\theta'/(\rho c_p)$ the kinematic heat flux. Fig. 10 shows a number of spectral curves for the horizontal and the vertical component of the turbulent fluctuations in the surface layer, respectively.

In the convective boundary layer the influence of large eddies goes down to the surface so that also the inversion height z_i has to be taken into account, but we shall not go into details here.

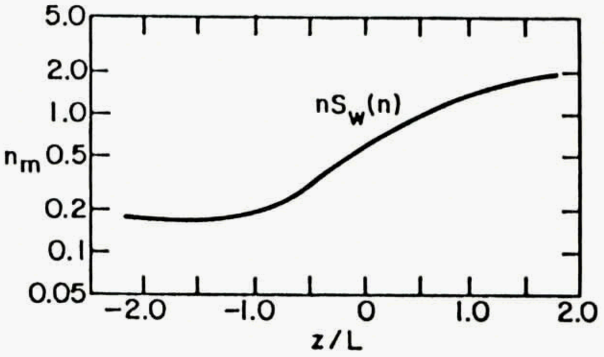


Figure 11: Normalized peak frequency for w Kaimal [27] (reproduced by Panofsky & Dutton [25]).

It is apparent that the location of the curves' maxima strongly depends on the dimensionless height. Fig. 11 shows as an example the loci of the maxima of the w-spectra as a function of z/L_*. Yet for flight mechanical applications the locus of the maximum for a given height and given stability conditions is of greatest interest since it gives together with the air speed of an aircraft the main forcing frequency of the disturbances acting upon the aircraft, from which by the aid of the transfer functions for aeroadmittance and mechanical admittance the aircraft response can be derived.

Unfortunately the determination of the stability-length is difficult because it requires high-resolution instrumentation. But since the quantities from which the stability length is derived describe two phenomena, i.e.

- vertical heat flux $\left\{ \begin{array}{c} \text{against} \\ \text{promoted by} \end{array} \right\}$ buoyancy
- stress ,

we may argue that there is a unique relation between the stability length and the Ri-number. If it would be possible to apply the bulk-Ri-number Ri_b as a stability criterion we simply need to determine the vertical gradients of the horizontal wind and of temperature. And indeed, Businger [28] has given such relationships based upon empirical data.

So to get the wavenumber (or equivalently the wavelength) around which the maximum energy is contained in a given height and with given stability conditions we may follow the subsequent procedure:

- $(\Delta V/\Delta z, \Delta\theta) \quad \Rightarrow \quad Ri_b$
- $Ri_b \quad \Rightarrow \quad L_*$
- $z/L_* \quad \rightarrow$ choose the proper spectral curve and determine its maximum.

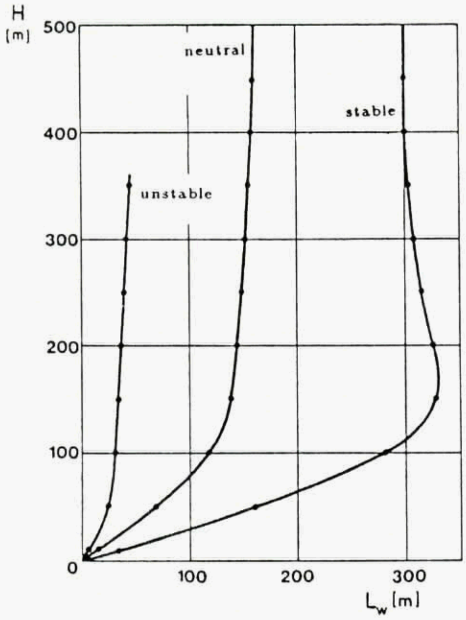


Figure 12: Maximum amplitude gust wave length for different stabilities and different heights.

The result of this procedure can be summarized in a diagram showing the wavelength of the most intensive gusts as a function of height for different values of Ri_b, which occurs as a shape parameter of the curves. Fig. 12 shows such a diagram prepared by Schänzer.

Another stability related phenomenon, which is often found within the lowest few hundred meters of the atmosphere – especially in clear nights – and which is also relevant to flight safety, is the so-called low level jet. It is characterized by the periodic occurence of supergeostrophic windspeeds immediately above the planetary boundary layer. These windmaxima are of special importance to flight safety because they are often accompanied by high values of wind shear.

To give a short explanation of the phenomenon we start with the simple case when no friction is present in the atmosphere and we consider horizontal forces only. In this case the only forces acting on an air parcel are the Coriolis force due to the rotation of the earth and the pressure force. If a parcel starts moving along an isobar, which for a first aprroximation we assume to be straight, then it will continue to do so ad infinitum if its initial speed equals an equilibrium value, which is called the "geostrophic wind speed". But if the initial velocity is not directed along the isobar or if its speed is different from the uniquely determined geostrophic wind speed then the parcel will start oscillations.

If we now allow for friction as a third force, which occurs primarily within the planetary boundary layer, then the wind must have a component directed from high to low pressure, i.e. across the isobars. The reason is that the pressure field is the only energy source to overcome the friction since that the Coriolis force can do no work, for it is always directed perpendicular to the velocity. But in order to receive energy from the pressure field the motion must have a component along the pressure gradient. This component is usually called the "ageostrophic component" of the wind.

We shall now restrict ourselves to the nocturnal low level jet. Hence we consider the diurnal evolution of the planetary boundary layer and start e.g. with the convective boundary layer on a nice summer day. In this case the boundary layer may extend to heights up to $2\,km$ above the ground with turbulent friction within the whole range that differs significantly from the frictional forces above. After sunset the convection ceases, the boundary layer shrinks (possibly to heights below $100\,m$), and following the emission of terrestrial radiation into space the surface cools and stable stratification develops near the ground. As a consequence the friction within the range above the nighttime boundary layer becomes very small compared to the daytime values; and at the same time the layer above the nighttime boundary layer but below the upper boundary of the daytime boundary layer is decoupled from the ground.

Since during daytime there has been an ageostrophic component within this layer and during the night the only forces are Coriolis' and pressure the motion in the range immediately above the nighttime boundary layer is not in geostrophic balance. This leads to inertial oscillations which are characterized by a cyclic deviation of the actual wind from the geostrophic wind. Within such a cycle under- and super-geostrophic wind speeds occur, whereby especially the super-geostrophic windspeeds are accompanied by sharp gradients of horizontal velocity.

A low level jet can extend over hundred of kilometers making it like an horizontally wobbling pancake. Fig. 13 shows two extreme cases of boundary layer wind maxima which were measured by Shelkovnikov [29], and whose impact on aircraft during landing approaches has been studied by Swolinsky [30].

Nocturnal low-level-jets have also been studied experimentally by Roth [31], Kottmeier [32] and Kraus [33]. An elementary analytical approach to nocturnal low level jets has been given by Blackadar [34]. Numerical simulations were performed by Thorpe [35] and Malcher & Kraus [36]

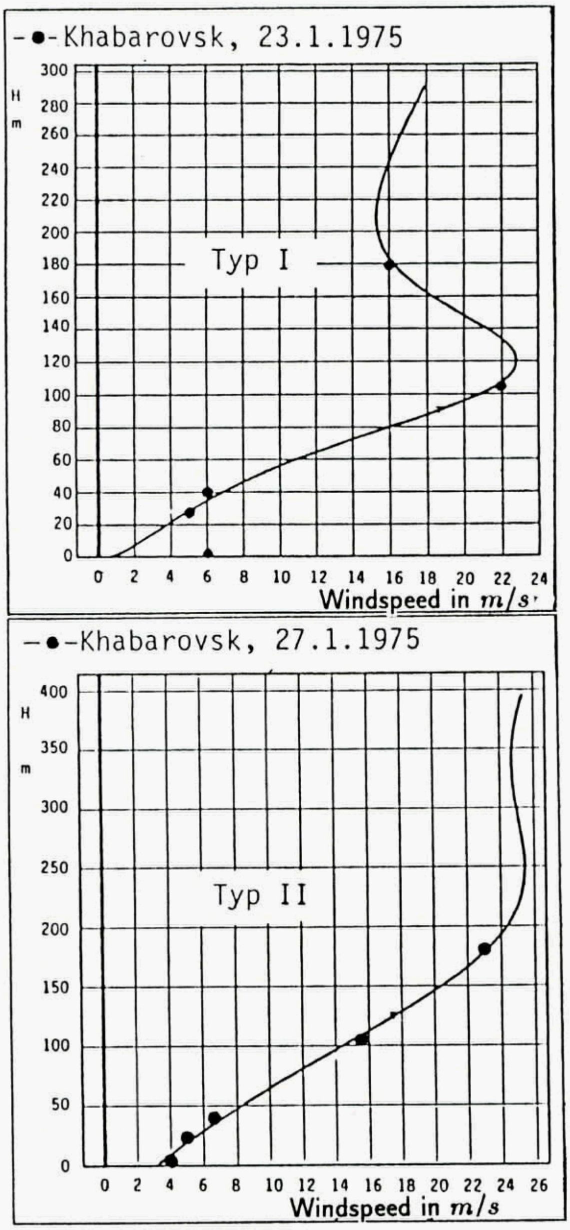


Figure 13: Worst case LLJ after [29] (diagram from Swolinsky [30]).

4 Stability Phenomena, and its influence on flight operations

We now want to know how a certain type of turbulence affects the aircraft motion. A flight at a given height in a given atmospheric environment may be subject to the investigations. The model, we have just seen in fig. 13 leads us to a special characteristic gust-wavelength L_w of the wind-field passed by the aircraft with a given airspeed V. According to Taylor's hypothesis [25] a windfield can be assumed to be time-fixed, if the windspeed is low compared to the aircraft speed. Hence the airspeed has an important influence on the frequency of the gusts, seen from the aircraft. The wavelength L_w can be transverted to a frequency f_w by using the airspeed V

$$f_w = V/L_w \quad . \tag{25}$$

Thus, a higher airspeed induces a higher gust-frequency in the same windfield. Most important for flight operations and safety are the accelerations and flightpath-changes of the aircraft in the windfield. Therefore transfer-functions from windspeed to the aircraft-response are very helpful for further investigations. As an example we want to set up the transfer-function from the vertical windspeed to the vertical acceleration of the aircraft. Let us first review the phases of an aircraft flying into a step-shaped upwind-gust: in a stationary horizontal flight, the aerodynamic lift equals the weight of the aircraft (fig. 14):

$$W = L = \frac{\rho}{2} V^2 \cdot S \cdot c_{L\alpha} \cdot \alpha_{St} \quad . \tag{26}$$

When the aircraft is penetrated by the vertical upwind-gust, an additional angle of attack $\Delta\alpha$ arises producing additional lift ΔL and an upward acceleration $\ddot{H}$ of the aircraft-mass m:

$$\Delta L = m \cdot \ddot{H} = \frac{\rho}{2} V^2 \cdot S \cdot c_{L\alpha} \cdot \Delta\alpha \tag{27}$$

with the lift coefficient assumed to be linear ($c_{L\alpha} = const$). The pitching motion of the aircraft is neglected, to keep this model simple.

The flightpath-vector rises whereas the vertical windspeed is chosen to be constant and the additional angle of attack $\Delta\alpha$ decreases.

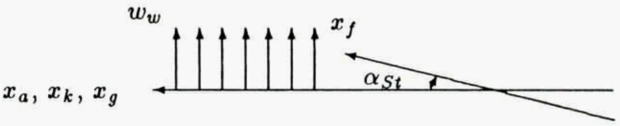


Figure 14: The aircraft flies in steady conditions, horizontal flight at an airspeed V without wind at a certain angle of attack α_{St} to keep the lift L equal to its weight W.

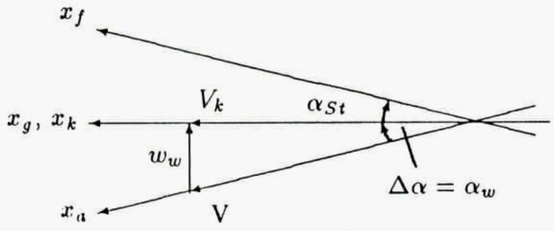


Figure 15: At the time the gust of upwind w_w is reached, the flight path speed V_k initially remains constant, whereas the airspeed is the result of V_k and w_w with an additional angle of attack $\Delta\alpha$, initially equal to the wind-induced angle α_w.

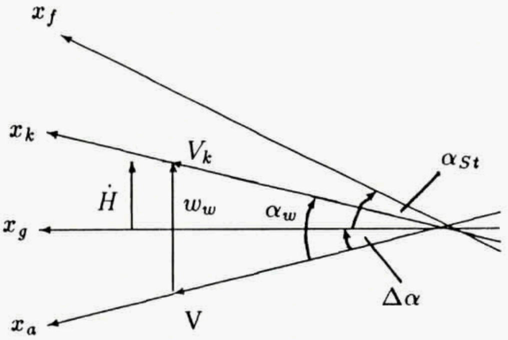


Figure 16: The additional lift accelerates the aircraft upwards, the longitudinal lift component, which arises when the lift-vector is turned into the new direction of airstream, is assumed low and omitted for this view.

Now we can start to develop the transfer-function from equation (27). Assuming small angles, from fig. 14,15,16 can be derived:

$$\Delta\alpha = \frac{w_w - \dot{H}}{V} \quad . \tag{28}$$

With eq. (27):

$$\Delta L = \frac{\rho}{2} V \cdot S \cdot c_{L\alpha} \cdot \left(w_w - \dot{H} \right) = A_\alpha \cdot \left(w_w - \dot{H} \right) \tag{29}$$

Laplace-transformed:

$$\Delta \hat{L} = A_\alpha \cdot \left(\hat{w}_w - s\hat{H} \right) \tag{30}$$

Equation (27) Laplace-transformed:

$$s^2 \hat{H} = \Delta \hat{L} / m \quad . \tag{31}$$

Equations (30) and (31) together yield:

$$\Delta \hat{L} = A_\alpha \left(w_w - \frac{\Delta \hat{L}}{sm} \right) \tag{32}$$

resulting in the transfer-function from vertical wind to additional lift:

$$\frac{\Delta \hat{L}}{\hat{w}_w}(s) = \frac{ms}{1 + \dfrac{m}{A_\alpha} s} \tag{33}$$

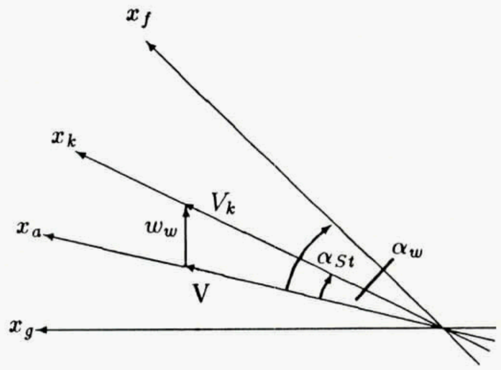


Figure 17: The aircraft has returned to a steady, now climbing flight. $\Delta\alpha = 0$ and $\Delta L = 0$.

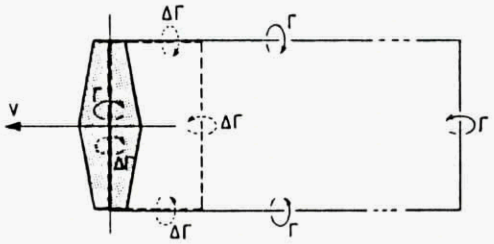


Figure 18: Free and bounded vortex of a wing-circulation [37].

Which is of the form of a DT_1-link:

$$F(s) = \frac{KTs}{1 + Ts} \tag{34}$$

with

$$K = A_\alpha \qquad \text{and} \qquad T = m/A_\alpha$$

Looking back to the step-shaped-gust, it has to be considered, that the lift-force on the wing does not build up immediately, when the angle of attack changes. An additional circulation $\Delta\Gamma$ appears when α changes. A wing-bounded and a free vortex counteract, so initially no additional lift is produced. As the free vortex is left behind, the bounded one becomes effective and lift rises in the same manner (fig.18). An approximation of the time-behavior of unsteady lift can be developed from the approach of the Küssner-function $\Psi(t)$ (see Schänzer [37]).

$$\Delta L(t) = \Delta L_\infty \cdot \Psi(t) \quad . \tag{35}$$

A simple form of this model will be sufficient to approximate gust-loads (see Schänzer [37]).

$$\Psi(t) = 1 - e^{-t/T_k} \tag{36}$$

with

$$T_k = f \frac{l_\mu}{V} \quad .$$

The factor f can be drawn out of optimization-calculations and is depending on aspect-ratio and

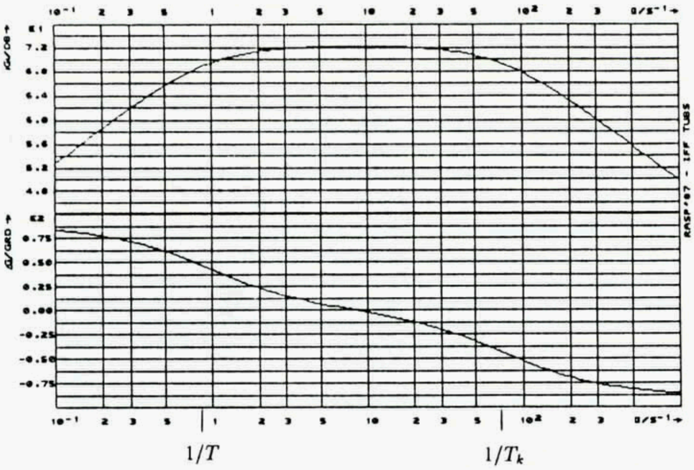


Figure 19: Bode-Diagramm of transfer-function w_w to ΔL.

Mach-number [37]. A Laplace-transformation of eq. (36) together with eq.(35) leads to the transfer-function from stationary to unsteady lift:

$$\frac{\Delta L(s)}{\Delta L_\infty} = \frac{1}{1+T_k s} \tag{37}$$

Combination of eqs. (34) and (36) leads to the transfer-function:

$$F(s) = \frac{\Delta L(s)}{w_w(s)} = \frac{KTs}{1+Ts} \cdot \frac{1}{1+T_k s} \tag{38}$$

with

$$K = A_\alpha \qquad \text{and} \qquad T = \frac{m}{A_\alpha}$$

Fig 19 shows a Bode-Diagramm of this transfer-function for the following set of aircraft data:

$m = 4350\,kg$	$c_{L\alpha} = 4.0$
$V = 60\,m/s$	$l_\mu = 2.0\,m$
$S = 28\,m^2$	$\rho = 1.225\,kg/m^3$

The diagram shows clearly, that the additional lift has a maximum approximately between 1 *Hz* and 50 *Hz* of gust circle-frequency. In this range the wing structure is stressed mostly. At lower frequencies, the aircraft is able to follow the gust waves, at higher frequencies lift cannot build up due to the low-pass behavior of the Küssner-function. With this kind of transfer-functions any wind-model achieved by meteorological modelling can be applied to an aircraft-model for analysing structural loads or flight path deviations. The same effects could be shown with longitudinal gusts. In this case the change of the airspeed ΔV would cause the change in lift ΔL. With this transfer-function the curves of fig. 13 can be transverted into a diagram of additional lift, respectively acceleration, depending on the height and atmospheric stability. Fig. 20 shows how the ratio of unsteady lift variance σ_A^2 to stationary lift variance $\sigma_{A\,St}^2$ increases with height and atmospheric stability, i.e. the *Ri*-number.

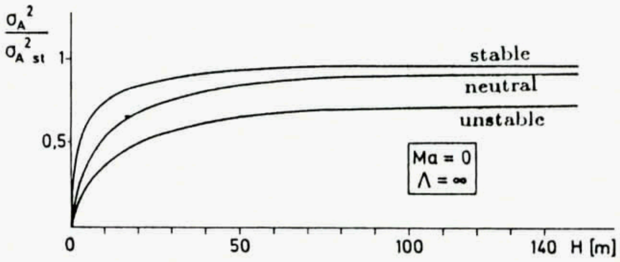


Figure 20: Unsteady lift variance as a function of height and atmospheric stability [37].

5 Energy budget of aircraft and safety criteria

Facing a flying aircraft from the energetic point of view, it can be described as a system of energy stores. Single stores can exchange energy with each other and across the border of the system "aircraft" with its environment, usually the atmosphere. Depending on the case, the exchange can take place in one or both directions, but is combined with a transformation of the energy-form with a certain loss according to the effectivity of the transformation process. The energy flow $\dot{E}$ between the stores or the atmosphere is the Power P.

Let us now take a look at the different energy forms, stores and describing values.

A powered aircraft has stored an amount of chemical energy in its fuel tanks which can be transformed to mechanical energy by the engines. The level of this store can be expressed in the remaining fuel quantity. The energy flow from the engines has to be divided into one stream, which is needed to counteract the drag. The remaining stream, if there is one left, can be taken to feed the other stores. The influence of the chemical energy store on flight safety is evident, everyone knows what happens if this store runs dry.

The level of the potential energy store can be expressed by the height. It becomes marginal for flight safety if an impact on the ground is possible. Energy exchange with the atmosphere takes place in vertical winds. An upward movement of the surrounding air delivers potential energy to the aircraft, downwind takes energy from it.

Kinetic energy in the common physical sense is defined by using the flight path speed. Aerodynamic forces are depending on the airspeed, thus the definition of the aero-kinetic energy by using the airspeed is of much more interest regarding to flight safety. So the airspeed is a value for the level of kinetic energy. A low margin would be the aircraft approaching stall speed, an upper limit is the design speed. Using the aero-kinetic energy, an exchange with the atmosphere would be a flight into a longitudinal gust, e.g. in a wind-shear. Due to the inertia of mass the aircraft initially keeps its flight path speed. The airspeed changes with the wind, resulting in an aerokinetic energy gain or loss until the flight path speed reaches the new value. Most important for flight safety is the quick exchange of kinetic to potential energy or vice versa, which is possible simply by using the elevator.

A certain level of energy can be stored by elastical deformation of the aircraft-cell, especially the wings. The upper margin could be a certain value of the aircraft weight multiplied by the g-load, at which structural damages are to be expected. If the aircraft flies into a gust, a change of lift causes the wings to bend until the aircraft mass is accelerated. In this case energy is buffered until it is passed to the potential and kinetic energy stores. As this energy store is not relevant for flight path applications it will be neglected in our further considerations.

The rotation energy is of minor interest, because its absolute value is low, compared to the kinetic and potential energy, in addition it is not directly related to safety. Of course a spinning aircraft is probably in danger, but more important to its safety are the loss of potential and excess of kinetic and elastic energy. Rotation energy could become important, if wake-turbulence is taken into account.

Concluding the energy considerations, we have two main kinds of aircraft energies which are directly influenced by wind:

- Potential energy: $E_{pot} = mgH$
  which is changed especially by
  vertical wind: $\dot{E}_{pot} = mg\left(V\sin(\alpha_w - \gamma) + w_w\right)$
- Aerokinetic energy: $E_{kin} = mV^2/2$
  which is changed especially by
  horizontal wind:
  $\dot{E}_{kin} = m \cdot \left(V + u_w\cos(\alpha_w - \gamma)\right)^2/2$

Now the question is, how these different energy-stores are affected by a given wind-phenomenon and what kind of wind brings which store to a margin.

Wind shear situations arise mostly during unstable or neutral atmospheric conditions. Typical wind shear situations are those appearing during thunderstorm downbursts, low level jets or due to the earth's surface boundary layer (see figs. 21 and 22).

In the atmosphere the flow is usually three-dimensional. The spatial variation of the windspeed vector $\vec{V}_w$ with its components u_w, v_w, w_w yields a nine element gradient tensor:

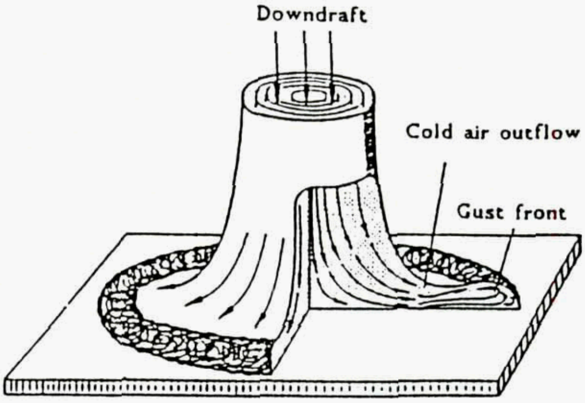


Figure 21: Sketch of a downburst [38].

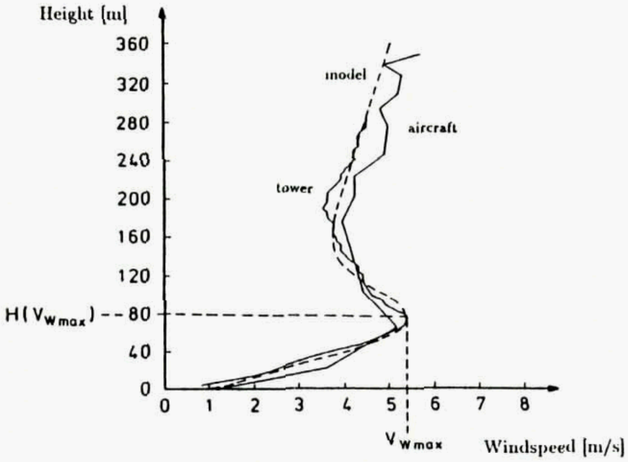


Figure 22: Windspeed profile in a low level jet.

$$\text{grad}\,\vec{V}_w = \begin{pmatrix} \dfrac{\partial u_w}{\partial x} & \dfrac{\partial u_w}{\partial y} & \dfrac{\partial u_w}{\partial z} \\ \dfrac{\partial v_w}{\partial x} & \dfrac{\partial v_w}{\partial y} & \dfrac{\partial v_w}{\partial z} \\ \dfrac{\partial w_w}{\partial x} & \dfrac{\partial w_w}{\partial y} & \dfrac{\partial w_w}{\partial z} \end{pmatrix} \tag{39}$$

Only three of these elements seem to be relevant in wind shear accidents:

- wind variation with height $\partial u_w/\partial z = u_{wz}$
- wind variation along the flight path $\partial u_w/\partial x = u_{wx}$
- spanwise variation of the vertical wind $\partial w_w/\partial y = w_{wy}$

These wind gradients change the aerodynamic forces acting on the aircraft. The first two elements influence the longitudinal motion whereas the latter induces a rolling motion in the lateral-directional mode. The ICAO classifies the wind gradient u_{wz} as

– dangerous if	$u_{wz} \geq 0.200\,1/s$
– difficult if	$u_{wz} \geq 0.130\,1/s$
– significant if	$u_{wz} \geq 0.066\,1/s$

Considering only the aircraft's longitudinal motion the time dependent horizontal wind speed may be expressed as follows:

$$u_w(t) = u_{wx} \cdot \Delta x_k - u_{wz} \cdot \Delta H \quad . \qquad (40)$$

Wind variation with time leads to the following equation [38]:

$$\dot{u}_w(t) = u_{wx} V_k - u_{wz} \dot{H} \quad . \qquad (41)$$

The variation of the horizontal wind with time depends on the aircraft's kinetic velocity V_K and the vertical speed $\dot{H}$. Thus, if an aircraft enters a space-variable windfield it will be transformed into a time variable windfield, which will have an effect on the aerodynamic flow around the aircraft. Considering the unsteady aerodynamic loads the aircraft's reactions to a simple wind shear model are shown on fig. 23. The airspeed changes at the time the windspeed decreases from $-12\,m/s$ to $0\,m/s$ inducing the phugoid motion. The deviations from the flight path are significant ($\Delta H > 100\,m$) and could lead to an accident during a landing approach. The total energy rate with its kinetic and potential components are plotted during this procedure as well.

In order to avoid wind shear accidents, deviations from the flight path and airspeed must be held to a minimum. Mathematically this means:

$$\Delta H = 0$$

and

$$\Delta V = 0$$

The kinetic velocity V_K is the superposition of airspeed V and windspeed V_w. This is true for the time derivatives as well:

$$\dot{V}_k = \dot{V} + \dot{V}_w \qquad (42)$$

Taking into account the requirement that the airspeed V has to be kept precisely it results that the kinetic velocity has to be changed with wind speed variations:

$$\dot{V}_k = \dot{V}_w \quad . \qquad (43)$$

In other words this means that the total energy (kinetic + potential) has to be kept constant throughout a wind shear field. This requirement can be fulfilled by providing or taking thrust energy. The linearized longitudinal equations of motion yield the following dependency of thrust from wind [30]:

$$\frac{\Delta T}{W} \approx \frac{\dot{u}_{wg}}{g} + \frac{\Delta u_{wg}}{V}\gamma + \frac{\Delta w_{wg}}{V} \quad . \qquad (44)$$

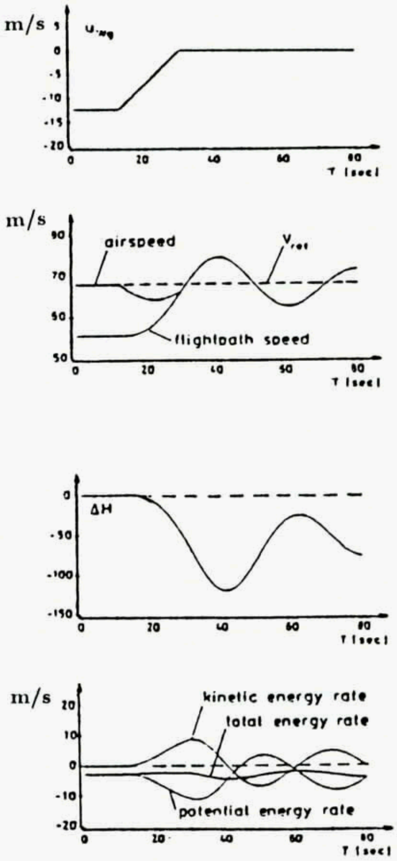


Figure 23: Reaction of an aircraft to a simple horizontal wind shear model.

Where $\dot{u}_{wg} = f(u_{wx}, u_{wz})$ as shown in eq. (41). Fig. 24 shows the flight path during a landing approach through a low level jet with and without thrust compensation. The danger of a crash due to the loss of total energy is obvious. In fig. 25 the reaction time Δt of a number of pilots to compensate thrust is plotted against the maximum wind speed dependent height of a low level jet. It shows clearly the tendency of the pilots not to react correctly with increasing height.

6 Conclusions

It may be concluded that it is a matter of power density and wavelength of the wind-phenomenon resulting from the atmospheric conditions in relation with aircraft type specifications, such as mass and wing loading, if the wind influence results more in a power and energy problem or in a structural one. The first case is relevant for flight safety from the flight dynamics point of view and was discussed by analysing the aircraft's flight path response to a simple wind shear model. The resulting flight safety requirement is that the total energy of the system "aircraft" has to be kept constant, which can be fulfilled by providing or taking thrust energy according

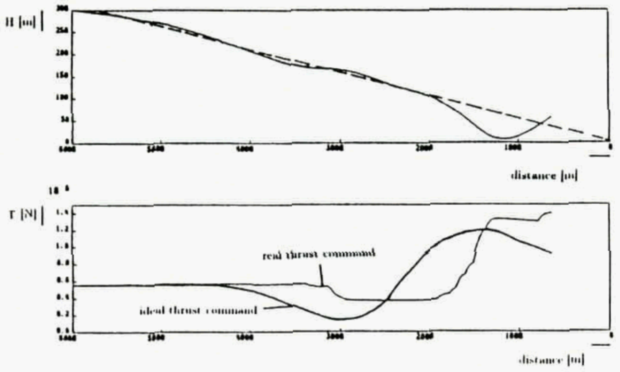


Figure 24: Flight path of a landing approach during wind shear conditions (low level jet) with and without thrust compensation.

to the energy flow resulting from windspeed variations.

Appendix

Derivation of Equation (4)

We refer to the footnote concerning the summation rule for tensors on page 3 and start with the equation of motion of a fluid element (2) in the form

$$\frac{du_i}{dt} - u_k\frac{\partial u_i}{\partial x_k} = -\frac{\partial \Psi}{\partial x_i} \quad . \tag{45}$$

Employing the notation

$$E_{kin} := \frac{u_j u_j}{2}$$

the first term on the left hand side can be written

$$\frac{du_i}{dt} = \frac{d}{dt}\left[\frac{\partial}{\partial u_i}\left(\frac{1}{2}u_i u_i\right)\right] = \frac{d}{dt}\frac{\partial E_{kin}}{\partial u_i} \quad , \tag{46}$$

whereas the second term on the left hand side upon multiplication with u_i yields

$$u_i u_k \frac{\partial u_i}{\partial x_k} = u_k \frac{\partial}{\partial x_k}\left(\frac{1}{2}u_i u_i\right) \quad . \tag{47}$$

After changing indices on the right hand side of (47) to

$$u_i \frac{\partial}{\partial x_i}\left(\frac{1}{2}u_k u_k\right) = u_i \frac{\partial E_{kin}}{\partial x_i}$$

and dividing the equation by u_i we substitute eqs. (46) and (47) into eq. (45) to give

$$\frac{d}{dt}\frac{\partial}{\partial u_i}\left(E_{kin}\right) - \frac{\partial E_{kin}}{\partial x_i} = -\frac{\partial \Psi}{\partial x_i} \quad . \tag{48}$$

If we assume that the potential is independent of velocity we can add Ψ inside the brackets on the left side of eq. (48) and get the equivalent relation

$$\frac{d}{dt}\frac{\partial}{\partial u_i}\left[E_{kin} - \Psi\right] - \frac{\partial}{\partial x_i}\left[E_{kin} - \Psi\right] = 0 \quad , \tag{49}$$

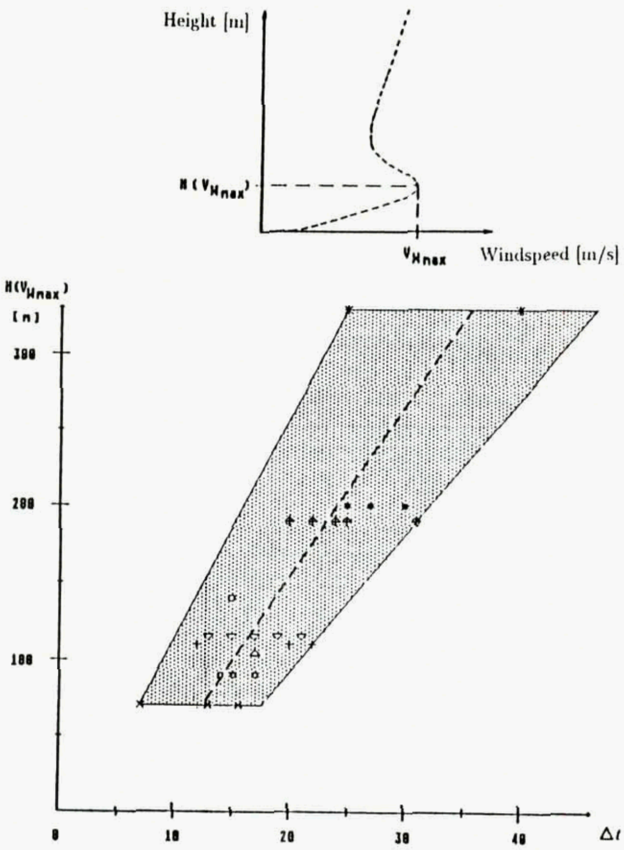


Figure 25: Pilot's reaction delay during landing approaches through low level jets.

and with the Lagrangian function

$$\mathcal{L} := E_{kin} - \Psi$$

we finally have

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial u_i} - \frac{\partial \mathcal{L}}{\partial x_i} = 0 \quad . \tag{50}$$

References

[1] Lorenz,E.N., "The Predictability of a Flow Which Possess Many Scales of Motion", Tellus, **21**, 1969, pp. 289-307.

[2] Lorenz,E.N., "Predictability of Geophysical Flows", In: Italian Physical Society, "Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics", North-Holland, Amsterdam, 1985, pp. 241-265.

[3] Lorenz,E.N., "Deterministic and Stochastic Aspects of Atmospheric Fluid Dynamics", In: Nicolis,C. and Nicolis,G., "Irreversible Phenomena and Dynamical Systems Analysis in Geosciences", D.Reidel Publ.Co., Dordrecht, 1987, pp. 159-179.

[4] Holton,J.R., "An Introduction to Dynamic Meteorology", Academic Press, N.Y., 1979, 391 p.

[5] Tatsumi,T., "Turbulence and Chaotic Phenomena in Fluids", North-Holland, Amsterdam, 1984, 556 p.

[6] Wigand,A., "Zum Problem der Wettervorhersage", Meteorol.Zeitschrift, **50**, 1933, pp. 26-28.

[7] Chandrasekhar,S., "Hydrodynamic and Hydromagnetic stability", Clarendon Press, Oxford, 1961. (Reprint-Paperback available from Dover Publ., N.Y., 704 p.)

[8] Drazin,P.G. and Reid,W.H., "Hydrodynamic Stability", Cambridge Univ. Press, Cambridge, 1981, 525 p.

[9] Haken,H., "Synergetics", Springer-Verlag, Berlin, 1978, 355 p.

[10] Goldstein,H., "Classical Mechanics", Addison-Wesley, Reading (Mass.), 1980.

[11] Stull,R.B., "An Introduction to Boundary Layer Meteorology", Kluwer Acad. Publ., Dordrecht, 1988, 666 p.

[12] Raethjen,P., "Zur Thermo-Hydrodynamik der Böen", Meteorol.Zeitschrift, **48**, 1931, pp. 11-22.

[13] Arnol'd,V.I., "The Hamiltonian Nature of the Euler Equations in the Dynamics of a Rigid Body and a Perfect Fluid", Uspekhi Matematicheskikh Nauk, **24**, 1969, pp. 225-226.

[14] Shepherd,T.G., "Symmetries, Conservation Laws, and Hamiltonian Structure in Geophysical Fluid Dynamics", Adv. in Geophys., **32**, 1990, pp. 287-338.

[15] Ertel,H., "Das Variationsprinzip der atmosphärischen Dynamik", Sitzungsberichte der Preußischen Akademie der Wissenschaften, 1933, pp. 461.

[16] Bretherton,F.P., "A Note on Hamilton's Principle for Perfect Fluids", J.Fluid.Mech., **44**, 1970, pp. 19-31.

[17] Ertel,H., "Thermodynamische Begründung der atmosphärischen Stabilitätskriterien", Meteorol.Zeitschrift, **50**, 1933, pp. 176-177.

[18] Houghton,J.T., "The Physics of Atmospheres", Cambridge Univ. Press, Cambridge, 1986, 271 p.

[19] Bjerknes,V., "Theoretisch-meteorologische Mitteilungen", Meteorol.Zeitschrift, **34**, 1917, pp. 166-176.

[20] Houghton,H.G., "On the Annual Heat Balance of the Northern Hemisphere", J.Met., **11**, 1954, pp. 1-9.

[21] Wallace,J.M. and Hobbs,P.V., "Atmospheric Science", Academic Press, N.Y., 1977, 467 p.

[22] "U.S. Standard Atmosphere 1976", National Oceanic and Atmospheric Administration, Washington D.C., 1976, 227 p.

[23] Landsberg,H. and Ratner,B., "Neue klimatologische Meridionalschnitte der Atmosphäre", In: Linke,F. and Baur,F., "Meteorologisches Taschenbuch", Akad. Verlagsgesellschaft Geest & Portig, Leipzig, 1962, pp. 334-336.

[24] Eddy,A., "Kinetic Energy Production in a Mid-latitude Storm", J.Appl.Met., **4**, 1965, pp. 569-575.

[25] Panofsky,H.A. and Dutton,J.A., "Atmospheric Turbulence", John Wiley & Sons, N.Y., 1984, 397 p.

[26] Monin,A.S. and Obukhov,A.M., "Basic Laws of Turbulent Mixing in the Ground Layer of the Atmosphere", Trans.Geophys.Inst.Akad.Nauk USSR, **151**, 1954, pp. 163-187.

[27] Kaimal,J.C., Wyngaard,D.A., Izumi,Y. and Cote,O.R., "Spectral Characteristics of Surface Layer Turbulence", Quat.J.Roy.Met.Soc., **98**, 1972, pp. 563-589.

[28] Businger,J.A.; Wyngaard,J.C.; Izumi,Y. and Bradley,E.F., "Flux-Profile Relationships in the Atmospheric Surface Layer", J.Atm.Sci., **28**, 1971, pp. 181-189.

[29] Shelkovnikov,M.S., "Strujnye techeniya na malykh vysotakh" ("Jets at low altitudes"), Meteorologiya i Gidrologiya, 1983, no. **11**, pp.44-46.

[30] Swolinsky,M., "Beiträge zur Modellierung von Scherwind für Gefährdungsuntersuchungen", Master-Thesis, Tech.Univ. Braunschweig, 1986, 161 p.

[31] Roth,R., Kottmeier,C. and Lege,D., "Die lokale Feinstruktur des Grenzschichtstrahlstroms", Meteorol. Rundschau, **32**, 1979, pp. 65-72.

[32] Kottmeier,C.,"Die Vertikalstruktur nächtlicher Grenzschichtstrahlströme", Ph.D. Dissertation, Berichte des Institutes für Meteorologie und Klimatologie der Universität Hannover, **21**, 1982, 129 p.

[33] Kraus,H.; Malcher,J. and Schaller,E.,"Nocturnal low-level-jet during PUKK", Boundary Layer Meteorology, **38**, 1985, pp. 1-22.

[34] Blackadar,A.K., "Boundary Layer Wind Maxima and Their Significance of the Growth of Nocturnal Inversions", Bull.Am.Met.Soc. **38**, 1957, pp.283-290.

[35] Thorpe,A.J. and Guymer,T.H., "The Nocturnal Jet", Quat.J.Roy.Met.Soc., **103**, 1977, pp. 633-653.

[36] Malcher,J. and Kraus,H., "Low-Level Jet Phenomena Described by an Integrated Dynamic PBL Model", Bound.Lay.Met., **27**, 1983, pp. 327-343.

[37] Schänzer,G., "Abschätzung von stochastischen Böenlasten unter Berücksichtigung instationärer Luftkräfte", Zeitschrift für Flugwissenschaft und Weltraumforschung, **9**, 1985, pp. 167-177.

[38] Schänzer,G., "Influence of Wind Shear, Downdraft, and Turbulence on Flight Safety", AGARD-CP 470, Gol (Norway), 1989, pp. 7-1 – 7-19.

Utilizing Quantitative Feedback Theory Design Technique For Flight Control System

Constantine H. Houpis
Air Force Institute Of Technology
Wright-Patterson AFB, Ohio, 45433, USA

ABSTRACT

Quantitative feedback theory (QFT) has achieved the status as a very powerful design technique for the achievement of assigned performance tolerances over specified ranges of plant uncertainties without and with control effector failures. This paper presents a brief overview of QFT and its applications to advanced MIMO flight control systems. Desired performance over varied flight conditions may be achieved with fixed compensators (controllers), despite failures of effectors. QFT is the only design technique that has been able to make considerable progress in improving the design of an overall flight control system taking into account the *man-in-the-loop specifications.*

INTRODUCTION

The paper is divided into four parts, the first part presents a qualitative overview of the theorectical concepts of QFT, the second part presents a number of examples to which QFT was utilized to design the control system, the third part presents the concept of a dual reconfigurable control system scenario, and the fourth part presents an overview of QFT computer-aided-design (CAD) programs. A short reference list is presented as a supplement to this list.

PART 1 -- OVERVIEW OF QFT

1.1 Introduction[1-4,10]

Quantitative feedback theory (QFT) is a unified theory that emphasizes the use of fedback for achieving the desired system performance tolerances despite plant uncertainty and plant disturbances. QFT quantitatively formulates these two factors in the form of (a) sets $\mathcal{T}_R = \{T_R\}$ of acceptable tracking or command input-output relations and $\mathcal{T}_D = \{T_D\}$ of acceptable disturbance input-output relations, and (b) a set $\wp = \{P\}$ of possible plants. The object is to guarantee that the control ratio $T_R = Y/R$ is a member of $\mathcal{T}_R$ and $T_D = Y/D$ is a member of $\mathcal{T}_D$, for all P in $\wp$. QFT is applicable for both nonlinear and linear, time-varying and time-invariant, continuous and sampled-data, uncertain MISO and MIMO plants, and for both output and internal variable feedback.

The representation of a MIMO plant with m inputs and ℓ outputs is shown in Fig. 1. The QFT synthesis technique for highly uncertain LTI MIMO plants has the following features[8]:

1. The MIMO synthesis problem is converted into a number of single-loop feedback problems in which parameter uncertainty, external disturbances, and performance tolerances are derived from the original MIMO problem. The solutions to these single-loop problems are guaranteed to work for the MIMO plant. It is not *necessary to consider the complete system* characteristic equation.

2. The design is tuned to the extent of the uncertainty and the performance tolerances.

This frequency-domain design technique is applicable to the following classes: (a) MISO linear time-invariant (LTI) systems; (b) MISO nonlinear systems; (c) MIMO LTI systems for which the performance specifications on each individual closed-loop system transfer function and on all the closed-loop disturbance response functions must be specified; (d) MIMO nonlinear systems; (e) disturbed systems; and (f) sampled-data systems as well as continuous systems for all classes. The MIMO classes are converted into equivalent sets of MISO systems to which the QFT design technique is applied. The objective is to solve the MISO problem, i. e., to find compensation functions which guarantee that the performance tolerances for each MISO problem are satisfied for all P in $\wp$. The amount of feedback designed into the system is then tuned to the desired performance sets $\mathcal{T}_R$ and $\mathcal{T}_D$ and the given plant uncertainty set $\wp$. Also, time-varying and nonlinear uncertain plant sets can be converted into equivalent MISO LTI plant problems to which the MISO frequency-domain technique can be readily applied and where the fundamental tradeoffs are highly visible.[6]

1.2 MIMO Uncertain Plants[1-4,10]

The state-space representation for a LTI MIMO system is:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \qquad (1)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \qquad (2)$$

where **A**, **B**, and **C** are constant matrices. The plant transfer-function matrix **P**(s) is evaluated as

$$\mathbf{P}(s) = \mathbf{C}[s\mathbf{I} - \mathbf{A}]^{-1}\mathbf{B} \qquad (3)$$

This plant matrix **P**(s) = $[p_{ij}(s)]$ is a member of the set $\wp = \{\mathbf{P}(s)\}$ of possible plant matrices which are functions of the uncertainty in the plant parameters. In practice, only a finite set of **P** matrices is formed, representing the extreme boundaries of the plant uncertainty under varying conditions.

Figure 2 represents an m×m MIMO closed-loop system in which **F**, **G**, and **P** are each m×m matrices. There are m^2 closed-loop system transfer functions (transmissions) $t_{ij}(s)$ contained within its system transmission matrix, i.e., $T(s) = \{t_{ij}(s)\}$, relating the outputs $y_i(s)$ to the inputs $r_j(s)$, e.g., $y_i(s) = t_{ij}(s)r_j(s)$. In a quantitative problem statement there are tolerance bounds on each $t_{ij}(s)$, giving a set of m^2 acceptable regions $\tau_{ij}(s)$ which are to be specified in the design, thus $t_{ij}(s) \in \tau_{ij}(s)$ and $\mathcal{T}(s) = \{\tau_{ij}(s)\}$.

From Fig. 2 the following equations can be written

$$\mathbf{y} = \mathbf{P}\mathbf{x} \quad \mathbf{x} = \mathbf{G}\mathbf{u} \quad \mathbf{u} = \mathbf{v} - \mathbf{y} \quad \mathbf{v} = \mathbf{F}\mathbf{r}$$

In these equations **G**(s) and **F**(s) are the matrices of compensator and prefilter transfer functions, respectively, and are often simplified to diagonal matrices, i.e., **G**(s) = diag{$g_i(s)$} and **F**(s) = diag{$f_i(s)$}. The combination of these equations yields

$$\boldsymbol{Y} = [\boldsymbol{I} + \boldsymbol{PG}]^{-1}\boldsymbol{PGFr} \qquad (4)$$

where the system control ratio relating **r** to **y** is

$$\boldsymbol{T} = [\boldsymbol{I} + \boldsymbol{PG}]^{-1}\boldsymbol{PGF} \qquad (5)$$

The QFT objective is to design a system which behaves as desired for the entire range of plant uncertainty. This requires finding three $f_i(s)$ and three $g_i(s)$, for diagonal matrices, such that each $t_{ij}(s)$ stays within its acceptable region $\tau_{ij}(s)$, no matter how $p_{ij}(s)$ may vary. The MIMO system, in order to simplify the design process, is converted into an equivalent set of MISO systems, as shown in the next section.

1.3 The MISO Equivalent Method[8,10]

This section presents an overview of the representation of an m×m MIMO system by m^2 MISO equivalent systems, each with two inputs and one output. One input is designated as a "desired" input and the other as a "disturbance" input. The inverse of the plant matrix is represented by

$$\boldsymbol{P}^{-1} = \begin{bmatrix} p_{11}^* & p_{22}^* & \cdots & p_{1m}^* \\ p_{21}^* & p_{22}^* & \cdots & p_{2m}^* \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ p_{m1}^* & p_{m2}^* & \cdots & p_{mm}^* \end{bmatrix} \qquad (6)$$

where **Λ** is the diagonal part and **B** is the balance of $\mathbf{P}^{-1}$. A **Q** = {q_{ij}} matrix is obtained from (6) by letting $q_{ij} = 1/p_{ij}^*$ thus

$$\boldsymbol{Q} = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1m} \\ q_{21} & q_{22} & \cdots & q_{2m} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ q_{m1} & q_{m2} & \cdots & q_{mm} \end{bmatrix} \qquad (7)$$

The m^2 effective plant transfer functions are formed as

$$q_{ij} = 1/p_{ij}^* = [\det\boldsymbol{P}/adj\boldsymbol{P}_{ij}] \qquad (8)$$

where is a requirement that detP be minimum phase (m.p.). Utilizing (6), (5) is manipulated to

$$\boldsymbol{T} = [\boldsymbol{\Lambda} + \boldsymbol{G}]^{-1}[\boldsymbol{GF} - \boldsymbol{BT}] \qquad (9)$$

This is is used to define the desired fixed point mapping where each of the m^2 matrix elements on the right side of (9) can be interpreted as a MISO problem. Proof of the fact that the design of each MISO system yields a satisfactory MIMO design is based on the Schauder fixed point theorem.[9] This theorem is described by defining a mapping **Y(T)** by

$$\boldsymbol{Y}(\boldsymbol{T}) \triangleq [\boldsymbol{\Lambda} + \boldsymbol{G}]^{-1}[\boldsymbol{GF} - \boldsymbol{BT}] \qquad (10)$$

where each member of **T** is from the acceptable set $\mathcal{T}$. If this mapping has a fixed point, i.e., $\mathbf{T} \in \mathcal{T}$ such that **Y(T)** = **T**, then this **T** is a solution of (9). Figure 3 shows the four effective MISO loops in the boxed area) resulting from a 2×2 system and the nine effective MISO loops resulting from a 3×3 system.[9] The control ratios for the desired tracking of the inputs r_i by the corresponding outputs y_i for each feedback loop of (10), of the MISO equivalent systems, have the form

$$\begin{aligned} y_{ii} &= w_{ij}(v_{ij} + d_{ij}) \\ &= y_{ri} + y_{dij} \end{aligned} \qquad (11)$$

where $w_{ii} = q_{ii}/(1 + g_i q_{ii})$ and $v_{ij} = g_i f_{ij}$. The interaction between the loops has the form

$$d_{ij} = -\Sigma[t_{kj}/q_{ik}] \qquad (12)$$

and appears as a "disturbance" input in each of the feedback loops.

If the plant matrix **P** is not a square matrix then **P** is replaced in the above equations by the "effective plant matrix" $\mathbf{P}_e$ where $\mathbf{P}_e = \mathbf{PW}$ and **W** is an ℓ×m weighting or a squaring down matrix.

1.4 The MISO QFT Design[4]

1.4.1 Performance Specifications -- The overview of the QFT design technique is best presented in terms of the m.p. LTI MISO system of Fig. 4 since an m×m MIMO control system can be represented by m^2 equivalent MISO control systems. The control ratios for tracking (D = 0) and for disturbance rejection (R = 0) are, respectively,

$$\begin{aligned} T_R(s) &= \frac{F(s)G(s)P(s)}{1 + G(s)P(s)} \\ &= \frac{FL}{1 + L} \end{aligned} \qquad (13)$$

$$T_D = \frac{P(s)}{1 + G(s)P(s)} = \frac{P}{1 + L} \qquad (14)$$

or

$$T_D = \alpha_p = a\ constant \tag{15}$$

The tracking thumbprint specifications, based upon satisfying some or all of the step forcing function figures of merit for underdamped (M_p, t_p, t_s, t_r, K_m) and overdamped (t_s, t_r, K_m) responses, respectively, for a simple-second system, are depicted in Fig 5a. The Bode plots corresponding to the time responses $y(t)_U$ and $y(t)_L$ in Fig. 5b represent the upper bound B_U and lower bound B_L, respectively, of the thumbprint specifications; i.e., an acceptable response y(t) must lie between these bounds. Note that for the m.p. plants, only the tolerance on $|T_R(j\omega_i)|$ need be satisfied for a satisfactory design. For nonminimum-phase (n.m.p.) plants, tolerances on $\angle T_R(j\omega_i)$ must also be specified and satisfied in the design process.[5,6,11] It is desirable to synthesize the tracking control ratios corresponding to the upper and lower bounds T_{RU} and T_{RL}, respectively, so that $\delta_R(j\omega_i)$ increases as ω_i increases above the 0 dB crossing frequency of T_{RU}. This characteristic of δ_R simplifies the process of synthesizing a loop transmission $L_o(s) = G(s)P_o(s)$, where P_o is the nominal plant transfer function, that requires the determination of the tracking bounds $B_R(j\omega_i)$ which are obtained based upon $\delta_R(j\omega_i)$. The simplest disturbance control ratio model is $T_D(s) = Y(s)/D(s) = \alpha_p$ a constant (the maximum magnitude of the output based upon a unit step disturbance input).

1.4.2 Plant Templates of $P_j(s)$, $\mathcal{T}P(j\omega_i)$ -- With L = GP, (13) yields

$$Lm\ T_R = Lm\ F + Lm\left[\frac{L}{1 + L}\right] \tag{16}$$

The change in T_R due to the uncertainty in P is

$$\begin{aligned}\Delta(Lm\ T_R) &= Lm\ T_R - Lm\ F \\ &= Lm\left[\frac{L}{1 + L}\right]\end{aligned} \tag{17}$$

By the proper design of $L = L_o$ and F, this change in T_R is restricted so that the actual value of Lm T_R always lies between B_U and B_L of Fig. 5. The first step in synthesizing an L_o is to make templates which characterize the variation of the plant uncertainty, as described by j = 1,2, ..., J plant transfer functions, for various values of ω_i over a specified frequency range. For the simple plant

$$P(s) = \frac{Ka}{s(s + a)} \tag{18}$$

where $K \epsilon \{1,10\}$ and $a \epsilon \{1,10\}$, is used to illustrate how the templates are obtained for a plant with variable parameters. The region of plant uncertainty is depicted in Fig. 6. The boundary of the plant template can be obtained by mapping the boundary of the plant parameter uncertainty region as shown on the Nichols chart (NC) in Fig. 7. A curve is drawn through the points A, B, C, and D and the shaded area is labeled $\mathcal{T}P(j1)$, which can be represented by plastic a template. Templates for other values of ω_i are obtained in a similar manner.

1.4.3 U-Contour -- The specifications on system performance in the frequency domain (see Fig. 5) identify a minimum damping ratio ζ for the dominant roots of the closed-loop system which becomes a bound on the value of $M_p \approx M_m$. On the NC this bound on $M_m = M_L$ (see Fig. 5) establishes a region which must not be penetrated by the template of $L(j\omega)$ for all ω. The boundary of this region is referred to as the universal high-frequency boundary (UHFB), the U-contour, because this becomes the dominating constraint on $L(j\omega)$. Therefore, the top portion, efa, of the M_L contour becomes part of the U-contour. For a large problem class, as $\omega \to \infty$, the limiting value of the plant transfer function pproaches

$$\lim_{\omega \to \infty} [P(j\omega)] = \frac{K}{\omega^{\lambda}} \tag{19}$$

where λ represents the excess of poles over zeros of P(s). The plant template, for this problem class, approaches a vertical line of length equal to

$$\begin{aligned}\Delta &\triangleq \underset{(\omega \to \infty)}{Lim}[Lm\ P_{max} - Lm\ P_{min}] \\ &= Lm\ K_{max} - Lm\ K_{min} = V\ \ dB\end{aligned} \tag{20}$$

If the nominal plant is chosen at $K = K_{min}$, then the constraint M_L gives a boundary which approaches the U-contour abcdefa of Fig. 8.

1.4.4 Bounds $B_o(j\omega)$ on $L_o(j\omega)$ -- The determination of the tracking $B_R(j\omega_i)$ and the disturbance $B_D(j\omega_i)$ bounds are required in order to yield the optimal bounds $B_o(j\omega_i)$ on $L_o(j\omega_i)$. The solution for $B_R(j\omega_i)$ requires that the actual $\Delta T_R(j\omega_i) \leq \delta_R(j\omega_i)$ dB in Fig. 5. Thus it is necessary to determine the resulting constraint, or bound $B_R(j\omega_i)$, on $L(j\omega_i)$. The procedure is to pick a nominal plant $P_o(s)$ and to derive the bounds, by use of templates or a CAD package, on the resulting nominal transfer function $L_o(s) = G(s)P_o(s)$. The disturbance bounds can be determined by the method described in Reference 5. For the case shown in Fig. 9 $B_o(j\omega_i)$ is composed of those portions of each respective bound $B_R(j\omega_i)$ and $B_D(j\omega_i)$ that have the largest dB values. The synthesized $L_o(j\omega_i)$ must lie on or just above the bound $B_oj\omega_i)$ of Fig. 9.

1.4.5 Synthesizing (or Loop Shaping) $L_o(s)$ and F(s) -- The shaping of $L_o(j\omega)$ is shown by the dashed curve in Fig. 9. A point such as Lm $L_o(j2)$ must be on or above $B_o(j2)$. Further, in order to satisfy the specifications, $L_o(j\omega)$ cannot violate the U-contour. In this example a reasonable $L_o(j\omega)$ closely follows the U-contour up to ω = 40 rad/sec and must stay below it above ω = 40 as shown in Fig 9. It also must be a Type 1 function (one pole at the origin). Synthesizing a rational function $L_o(s)$ which satisfies the above specification involves building up the function where for k = 0, $G_o = 1\angle 0°$, and $K = \Pi^{w}_{k=0}K_k$. $L_o(j\omega)$ is built up term-by-term or by a CAD

$$\boldsymbol{L}_o(j\omega) = \boldsymbol{L}_{ok}(j\omega) = \boldsymbol{P}_o(j\omega)\prod_{k=0}^{w}[K_kG_k(j\omega)] \tag{21}$$

loop shaping routine,[12] in order to stay just outside the U-contour in the NC of Fig. 9 [see (21) where k = 0,1,2,...,w]. The design of a proper $L_o(s)$ guarantees only that the variation in $|T_R(j\omega_i)|$ is less than or equal to that allowed, i.e., $\delta_R(j\omega_i)$. The purpose of the prefilter F(s) is to position Lm [T(jω)] within the frequency domain specifications, i.e., that it always lies between B_U and B_L (see Fig. 5) for all J plants. The method for determining F(s) is given Ref. 1, 5, and 8.

1.4.6 Simulation -- By use of a QFT CAD package (see Part 4) a verification of the "goodness of the design" can be readily determined for all J plants.

PART 2 -- QFT DESIGN EXAMPLES

2.1 Introduction

A number of QFT design examples are qualitatively presented in this part to illustrate the power of the QFT design technique. The reader is referred to the references in order to obtain the quantitative aspects of the MIMO control system design procedure.

2.2 YF-16CCV Analog Flight Control System Reconfiguration Design[11]

A linearized 2×2 YF-16CCV open-loop unstable model is used with 4 individually control effectors: 2 elevators and 2 flaperons (they are normally tied together, constituting two inputs). For this QFT design, the command inputs and the controlled outputs are pitch rate and roll rate. The 4 step-response tolerances are shown in Fig. 10, to be achieved over the Mach, altitude flight conditions (F.C.): (1) 0.2, 30 ft; (2) 0.7, 30 K; (3) 0.9, 20 K; (4) 1.6, 30 K; and over as many effector failures as is feasible. These time-domain responses are translated into "equivalent" ω-domain tolerances (a_{ii} and b_{ii}, lower and upper bounds, respectively, on the $|t_{ii}(j\omega)|$ tracking responses and b_{ij} disturbance upper bounds) on the $|t_{ij}(j\omega)|$, shown in Fig. 11. The design proceeds in a manner similar to that of Sec. 1.4. The final results are shown in Fig. 12 in which each figure gives the responses at a single F.C. for the following failures: (1) none, (2) an elevator, (3) a flaperon, (4) an elevator and a flaperon on the same side, (5) as in (4) but on opposite sides, and (6) both flaperons. The specifications of Fig. 10 are satisfied for these 6 cases and the 4 F.C. In Fig. 12a, a 50°/sec roll rate is commanded, and in Fig. 12b, a 10°/sec pitch-rate is commanded. Note that failure of both elevators (case 7) is omitted. The reason is apparent from Fig. 13 which shows loop 1 effective plant set $\mathcal{P}(1)$ with 4×6 = 24 elements, whose size is a measure of the uncertainty range. Case 7, with its 4 F.C., more than doubled this uncertainty range and its inclusion would require about 20 dB larger $|\ell_1(j\omega)|$, which was considered intolerable under the bandwidth constraints. This is an example of the transparency of QFT, revealing the trade-off between the benefits and cost of feedback <u>during the course of the design</u>, and not at the end after the design simulation, as is done by other MIMO control system design techniques.

2.3 F-15STOL Analog Flight Control System Reconfiguaration Design[13]

The results of an QFT design, for a linearized 2×2 F-15STOL open-loop unstable model with forward velocity and angle of attack as outputs, is presented in this section. The canard, stabilator, elevator, and the top and bottom reverser vanes are the 5 effectors. The uncertainty range consists of (a) 3 F.C. (100, 120, 180 knots all at sea level) and (b) no fail and 5 single failure cases, inasmuch double failure provided insufficient authority for control. The time-domain specs and their corresponding frequency-domain specs are available in Ref. 13. The 5×3 **W** matrix is chosen and the design proceeds in a manner similar to that of Sec. 1.4. The final design simulation results are shown in Fig. 14 (6×3 = 18 plots per figure) with the response tolerances being satisfied.

2.4 AFTI/F-16 Digital Flight Control System Reconfiguration Design[14]

A linearized 3×3 AFTI/F-16 unstable model is used for this design having 3 outputs (pitch, roll, and yaw rates) with 6 independent effectors (2 flaperons, 2 horizontal tails, vertical canard pair, and rudder). The uncertainty consists of (a) 4 F.C., (b) no effector failure, and (c) 5 single, 2 double, and 1 triple effector failures. A digital design is inherently n.m.p. so is limited in its feedback benefits[2,15,20]. In the <u>preliminary</u> design effort, it is possible to determine the resulting tolerance range of uncertainty and if some gain scheduling may be advantageous (rather than by repeated trial designs with simulations). This design effort is an example of the transparency QFT offers the designer so he can make intelligent compromises in the course of his design. The pitch channel w'-domain simulations yield robust results. The roll and yaw channel results were robust except for rudder failure.

2.5 Unmanned Research Vehicle Digital Flight Control System Design

2.5.1 Reconfiguration Design[16] -- This example is of an 3×7 unmanned research vehicle (URV) utilizing a sampling frequency of 60 rps, and second-order actuators and first-order sensor models. The effectors are rudder, left and right elevators, ailerons, and flaps. The uncertainty consists of 36 failure cases. For the most part, elevators and flaps can fail up to 100%. But for one case, 2 elevators and 1 aileron fail 95% each and in another case of an elevator failing 100%, the two flaps fail 90% only. The 5% and and 10% residuals are necessary so that **P** would remain m.p. The rudder failure can be up to 90% for **P** to remain m.p., but when limiters are included, only 70% rudder failure is tolerable. Otherwise the failure cases include all combinations of single, of 2 surfaces and of 3 surfaces failing 100% simultaneously. The 3 responses (roll, yaw, and c*) due to roll rate command are shown in Fig. 15. Each figure is for all 36 cases. Fig. 16 shows the responses due to yaw rate command for all failure cases. The pulse command is not long enough in duration for steady-state to be achieved but the 36 responses are almost identical and satisfy the specifications. Fig. 17 shows the responses due to c* command for the 36 failure cases. The specifications are satisfied over the set $\wp$. These simulations are all for the digitally compensated system including practical effector limiters.

2.5.2 URV Robust Digital Flight Control System[17] -- A 50 Hertz robust digital flight control system and an autopilot are designed for the Lambda URV model. The control system is built using a small perturbation 3×3 LTI plant model **P** developed from flight test data. The actuators, also modelled from aircraft (A/C) test data, are second order in roll and in pitch and first order in yaw. 19 separate plants are used to represent the flight envelope of the A/C resulting from variations in speed, altitude, center of gravity location, and weight. The pitch channel is decoupled from the lateral-directional channel resulting in a SISO system for the pitch channel and a 2×2 MIMO system for the lateral-directional channel. Pitch rate, roll rate, and yaw rate controllers and prefilters are designed to satisfy a 45° phase margin throughout the flight envelope, and meet figures of merit including rise time and overshoot requirements. The digital simulations, which include limiters, show that the resulting robust controllers met all specifications throughout the flight envelope without gain scheduling. This design is to be implemented and flight tested.

2.6 Man-in-the-Loop Flight Control System Design Technique

Figure 18 is a block diagram that represents the man-in-the-loop flight control design problem. The inner loop represents the flight control system where **P** is the YF-16 A/C (plant) matrix, **G** is the compensator matrix, and **F** is the prefilter matrix. The inner loop is represented by the matrix $\mathbf{P}_e$ and where $\mathbf{F}_p$ is the pilot compensator that is to be designed.

2.6.1 Flight Controller Design with Nonlinear Aerodynamics, Large Parameter Uncertainty, and Pilot Compensation[18,21] -- This initial QFT design effort, for the man-in-the-loop the control system, is restricted to a SISO design. Nonlinear QFT is used to design the flight control system P_e for the nonlinear YF-16 A/C model with C*, a blend of the normal acceleration at the pilot station and pitch rate, as the controlled output. The first step in the design procedure is to obtain LTI plant models which are rigorously equivalent to the nonlinear plant with respect to the defined set of desired outputs. These equivalent plants are generated from time histories of the input and output from simulator data. This data is used in a program[21] to generate LTI plants that are equivalent to the nonlinear plant in that both models give the same output for the given input. This procedure is repeated for the set of inputs and outputs that are to exist in the operating region of the system. The result is a set of LTI equivalent plants which are used in the QFT design process. It has been proven that the solution to the equivalent plant problem is guaranteed to solve the original nonlinear problem. The resulting closed-loop stability augmentation system (SAS), P_e, becomes part of the outer loop containing the pilot. The Neal-Smith pilot model for a compensatory tracking task is used to develope a technique which allows the designer to synthesize compensation in the outer loop, which includes a free compensator $F_p(s)$ ("pilot compensator)." The latter is chosen to minimize pilot workload, increase system bandwidth, and improve handling qualities ratings as per the Neal-Smith criteria, for the tracking task. The available pilot compensation abilities are then available for further increasing of system bandwidth to improve overall capabilities. This approach can be used at the early stages of flight control design, thus saving time and money over the current practice. Simulations in the time and frequency domains demonstrate that the desired performance is attained.

2.6.2 Multiple-Input Multiple-Output Flight Control System Design for the YF-16 Using Nonlinear QFT and Pilot Compensation[19] -- Nonlinear QFT and pilot compensation techniques are used to design a 2×2 multi-axis flight control system for the YF-16 A/C over a large range of plant uncertainty. The controlled variables are C* and roll rate. The only independent commanded controls are symmetric horizontal tail (elevator) deflection and aileron deflection. The design objective is to obtain responses for step commands in both C* and roll rate at 2 F.C.: 0.9 mach at 20K and 0.6 mach at 30K. The design is based on numerical input-output time histories generated with a FORTRAN implemented six degree of freedom nonlinear simulation of the YF-16 (see Sec. 2.6.1). The original compensation for rudder and leading edge flaps is left in place. Based on these time histories a set of equivalent LTI plant models are generated to represent the nonlinear plant[19]. Standard QFT techniques are then used in the design synthesis, based on these equivalent LTI plant models, to generate the inner loop diagonal compensator matrix **G** and the diagonal prefilter matrix **F**. After the inner loop design is completed, pilot compensation is developed to reduce the pilot's workload. This outer loop design is based on a set of equivalent LTI plant models. This is accomplished by modelling the pilot with parameters that result in good handling qualities ratings, and developing the necessary compensation to force the desired system responses. The designed compensator is implemented in the original non-linear simulation and gives the desired responses over the specified range of uncertainty and beyond.

PART 3 -- DUAL RECONFIGURATION CONTROL SYSTEM SCENARIO

3.1 Introduction

Elevators, aeropspace vehicles, nuclear reactor plants, and heating systems are a few examples that involve control systems which affect the safety and comfort of humans. One aspect of these control systems that has been and is of much interest is the problem of maintainning a stable system under a "control surface (or effector)" failure, i.e., the design of reconfigurable control systems.

3.2 Dual Reconfiguarable Flight Control System

Figure 19 describes the nature of the dual reconfigurability scenario. Control system #1 utilizes QFT designed controllers that are designed to maintain the aircraft stable at the instant of one or more effector failure. At the onset of the QFT design it is possible to determine the degree of effector failure and the number of simultaneous effector failures that can exist and yet be able to achieve a stable system. Thus control system #1 maintains the aircraft in the air and provides time for the "fault, detection and isolation (FDI)" control system #2 to "reconfigure" its controllers to achieve the best possible flying qualities under effector failure(s). A $\Delta t = 30$ msec corresponds to approximately a phase margin frequency of 30 rad/per. Decreasing this FDI Δt interval results in an increase in the phase margin frequency which for flight control systems is undesirable.

PART 4 -- QFT CAD PROGRAMS

4.1 Introduction

The first useable MISO QFT CAD package was developed, in 1986 for the analog design and in 1991, for the discrete design at the Air Force Institute of Technology (AFIT). This CAD package has been a catalyst in assisting the newcomer to QFT to understand the fundamentals of this powerful design technique. A MIMO QFT CAD package has been developed at AFIT this spring. These CAD packages will accelarate the utilization of the QFT design technique for the appropriate control problems.

4.2.1 MISO QFT CAD -- The AFIT package is called "ICECAP/QFT" which is designed for the VAX. Those desiring a copy of this package can contact: Professor Gary B. Lamont, AFIT/ENG, Wright-Patterson AFB, OH 45433. Currently Professor Lamont is developing an extended PC version. These packages have been designed as an "educational tool."

4.2.2 MISO QFT PC CAD -- Dr. Oded Yaniv, Tel-Aviv University, Israel, has a MISO QFT PC CAD package for both analog and discrete system design.

4.2.3 MIMO QFT CAD -- An AFIT graduate student, Mr. Richard Sating, has developed a MIMO QFT PC CAD package for both and analog and discrete designs. The discrete design is done in the w-prime domain. This PC CAD has been designed to be used on a SPARC station utilizing the MATHEMATICA and MATRIXx packages.

Professor F. Bailey, University of Minnesota, Minneapolis, Minnesota, has also developed QFT CAD packages. The QFT CAD packages mentioned in this section will be demonstrated at the First International Quantitative Feedback Symposium to be held On 3-4 Aug '92 at Hope Hotel, Wright-Patterson AFB, Ohio. This symposium is being sponsored jointly by Wright Laboratories and AFIT. Non U.S. citizen must have prior clearance, through their respective foreign office, to attend this symposium.

REFERENCES

1. Horowitz, I. M.: "Advanced Control Theory and Applications," unpublished notes, The Weizmann Inst. of Science, Rehovot, Israel, 1982.

2. Horowitz, I. M. and M. Sidi: "Synthesis of Feedback Systems with Large Plant Ignorance for Prescribed Time Domain; Tolerances," Int. J. Control, vol. 16, pp 287-309, 1972.

3. Horowitz, I.M.: "Synthesis of Feedback Systems, Academic Press, N.Y., 1963.

4. D'Azzo, J. J., and C. H. Houpis: "Linear Control System Analysis and Design," McGraw-Hill, NY, 3rd Ed., 1988.

5. Horowitz, I. M.: "Optimum Loop Transfer Function in Single-Loop Minimum Phase Feedback Systems," Int. J. Control, vol 22, pp. 97-113, 1973.

6. Horowitz, I. M.: "Synthesis of Feedback Systems with Non-Linear Time Uncertain Plants to Satisfy Quantitative Performance Specifications," IEEE Proc., vol. 64, pp. 123-130, 1976.

7. Betzold, R. W.: "Multiple-Input Multiple-Output Flight Control Design with Highly Uncertain Parameters,

Application to the C-135 Aircraft," M.S. Thesis, AFIT/GE/EE(83D-11), School of Engineering, Air Force Inst. of Tech., 1983.

8. Horowitz, I. M., and C. Loecher: "Design of a 3×3 Multivariable Feedback System with Large Plant Uncertainty," Int. J. Control, vol. 33, pp. 677-699, 1981.

9. Horowitz, I. M.: "Quantitative Synthesis of Uncertain Multiple-Input Multiple-Output Feedback Systems," Int. J. Control, vol. 30, pp. 81-106, 1979.

10. Houpis, C. H.: "Quantitative Feedback Theory (QFT): Technique for Designing Multivariable Control Systems," AFWAL-TR-86-3107, AF Wright Aeronautical Laboratories, Wright-Patterson AFB, OH, 1987. (Available from Defense; Technical Information Center, Cameron Station, Alexandria, VA 22314, document number AD-A176883.)

11. Arnold, P. B., I. M. Horowitz, and C. H. Houpis: "YF-16CCV Flight Control System Reconfiguration Design Using Quantitative Feedback Theory," Proceedings of the National Aerospace and Electronics Conference (NAECON), vol. 1. pp. 578-585, 1985.

12. Thompson, D. F. and O. D. I. Nwokah: "Optimal Loop Synthesis in Quantitative Feedback Theory," Proceed. of the American Control Conference, San Diego, CA, pp. 626-631, 1990.

13. Migyanko, B. S., "Design of Integrated Flight/Propulsion Control Laws of a STOL Aircraft During Approach and Landing Using Quantitative Feedback Theory," M.S. Thesis, AFIT/GE/ENG/86D-33, School of Engineering, Air Force Inst. of Tech., 1986.

14. Schneider, D. L.: "QFT Digital Flight Control Design as Applied to the AFTI/F-16," M.S. Thesis, AFIT/GE/EE; (88D-4), School of Engineering, Air Force Inst. of Tech., 1988.

15. Horowitz, I. M. and Y. K. Liao: "Quantitative Feedback Design for Sampled-data Systems," Int. J. Control, vol. 38, May, 1986.

16. Hamilton, S., I. M. Horowitz, and C. H. Houpis: "QFT Digital Controller for an Unmanned Vehicle," Proceedings of the American Control Conference, Pittsburgh, PA, 1989.

17. Lacey, D. J. Jr.: " A Robust Digital Flight Control System for an Unmanned Research Vehicle Using Discrete Quantitative Feedback Theory," M.S. Thesis, AFIT/GE/EE(91D-33), School of Engineering, Air Force Inst. of Tech., 1991.

18. Kobylarz, T. J.: "Flight Controller Design with Nonlinear Aerodynamice, Large Parameter Uncertainty and Pilot Compensation," M.S. Thesis, AFIT/GE/EE(88D-19), School of Engineering, Air Force Inst. of Tech., 1988.

19. Miller, R. B.: "Multi-Input Multi-Output Flight Control System Design for the YF-16 Using Nonlinear QFT and Pilot Compensation," M.S. Thesis, AFIT/GE/EE(90-D-42), School of Engineering, Air Force Inst. of Tech., 1990.

20. Houpis, C. H. and G. Lamont: "Digital Control Systems: Theory, Hardware, Software," McGraw-Hill, NY, 2nd Ed., 1992.

21. Kobylarz, T, I. M. Horowitz, C. H. Houpis, and F. Barfield: "Flight Controller Design with Nonlinear Aerodynamics, Large Parameter Uncertainty, and Pilot Compensation," AIAA Guidance and Control Conference, Portland, Oregon, USA, Aug 1990. Also to be published in the Int. J. of Robust and Nonlinear Control.

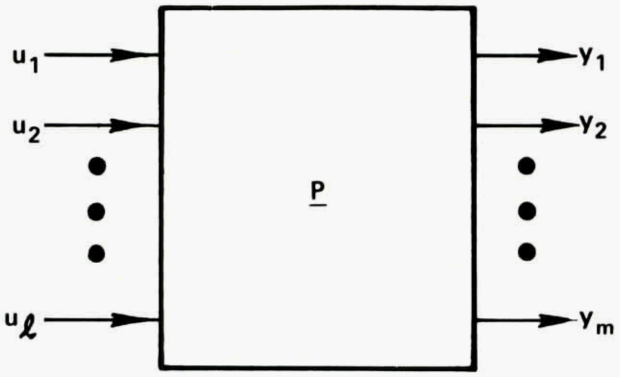


Fig. 1 A MIMO plant

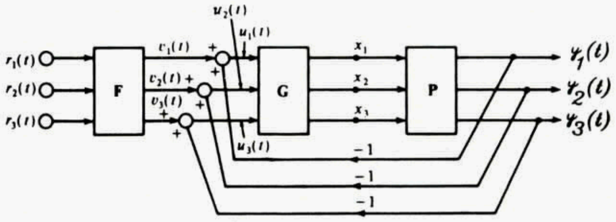


Fig. 2 MIMO control structure (3×3)

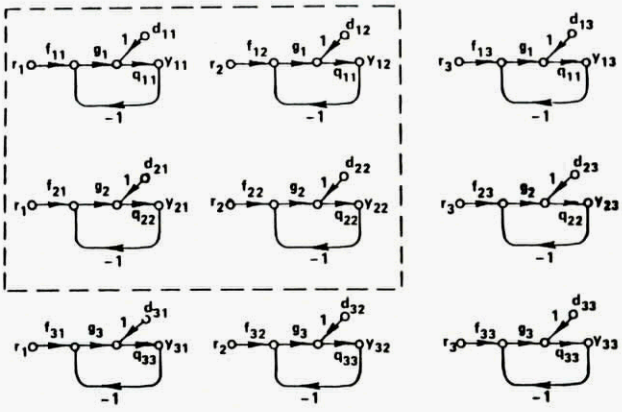


Fig. 3 Effective MIMO 2×2 (boxed in loops) and 3×3 (all 9 loops)

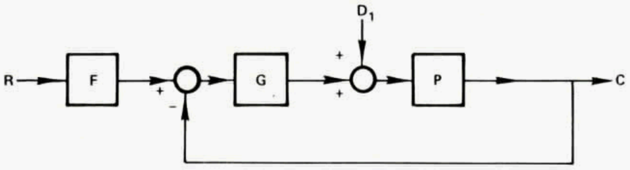


Fig. 4 A MISO control system

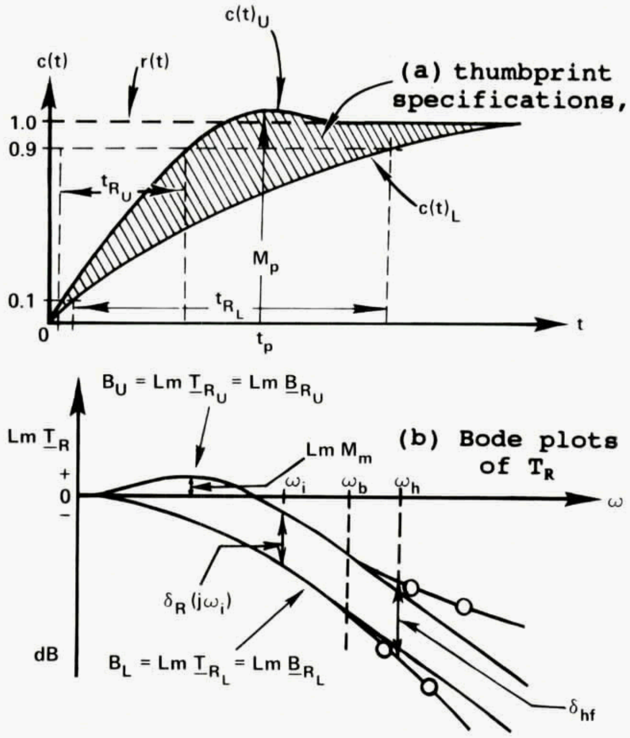


Fig. 5 Desired response characteristics

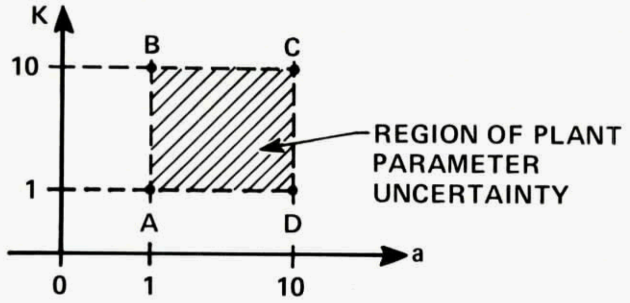


Fig. 6 Plant uncertainty region

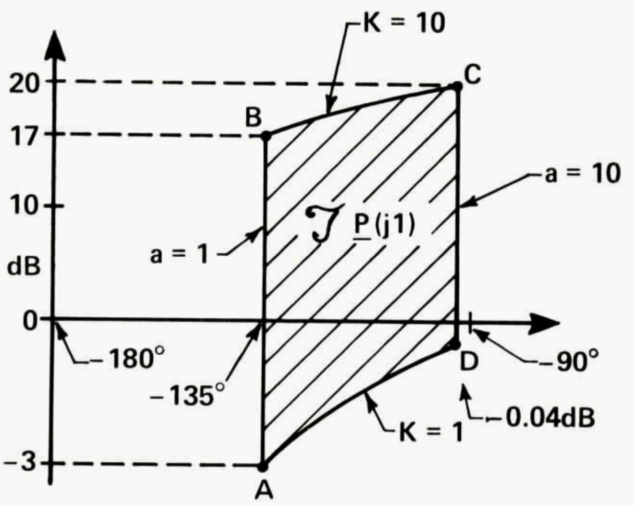


Fig. 7 Plant template

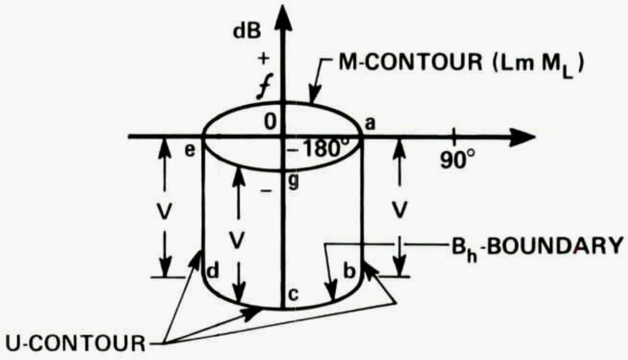


Fig. 8 U-contour construction

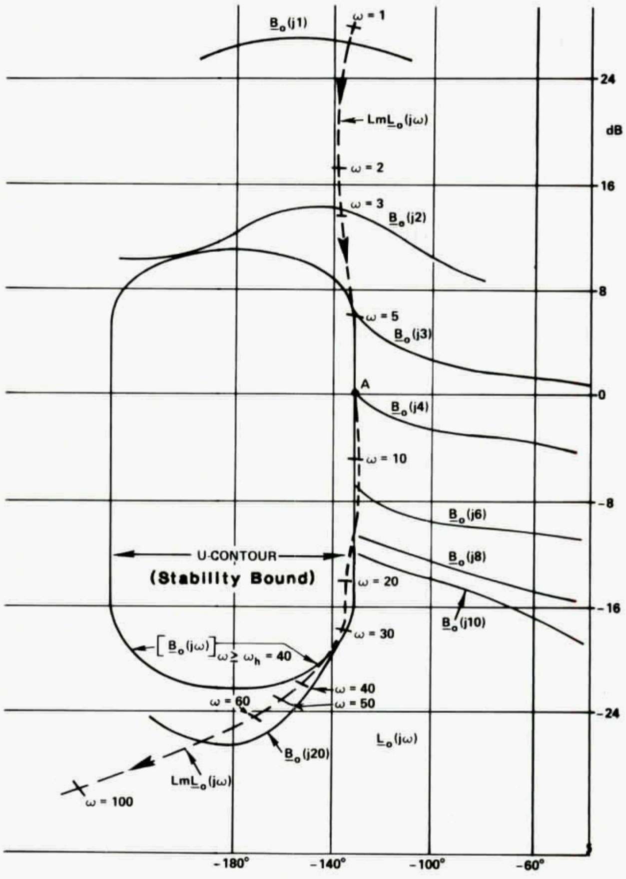


Fig. 9 Bounds $\mathbf{B}_o(j\omega_i)$ and loop shaping

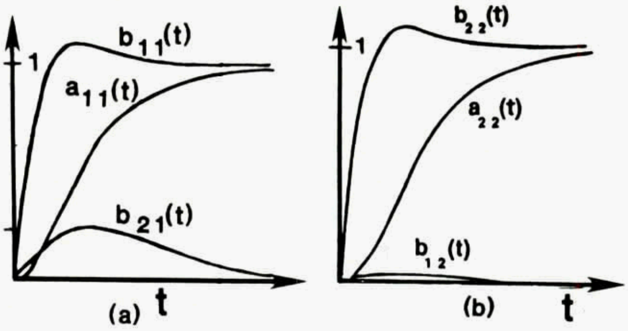


Fig. 10 Time-domain bounds on step responses for:
(a) $y_{1,1}(t)$ and (b) $y_{2,2}(t)$

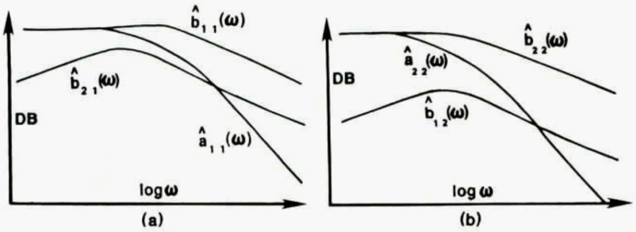


Fig. 11 Frequency-domain equivalents of Fig. 10

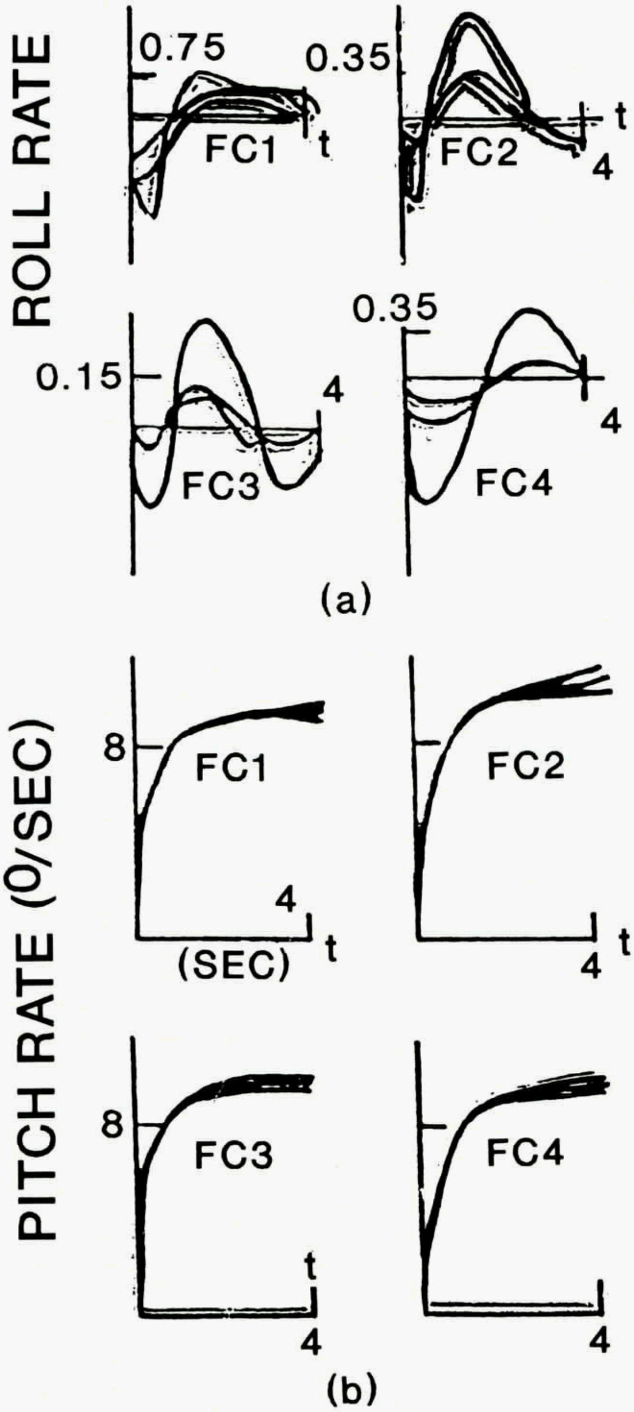


Fig. 12 Time domain simulation results:
(a) 50°/sec roll rate commanded,
(b) 10°/sec pitch rate commanded

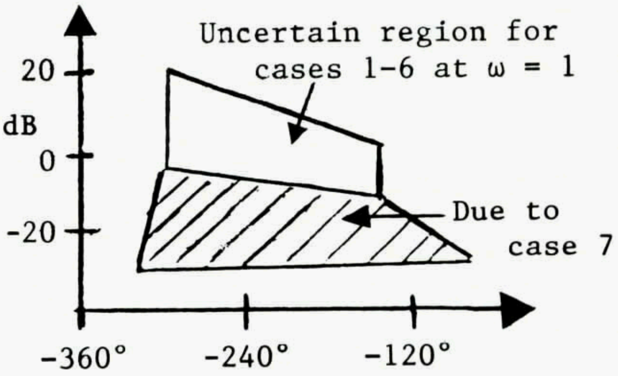


Figure 13 Transparency of QFT: Large uncertainty increase due to case 7 (both elevators fail) requires 20 dB increase in $L_o(j1)$

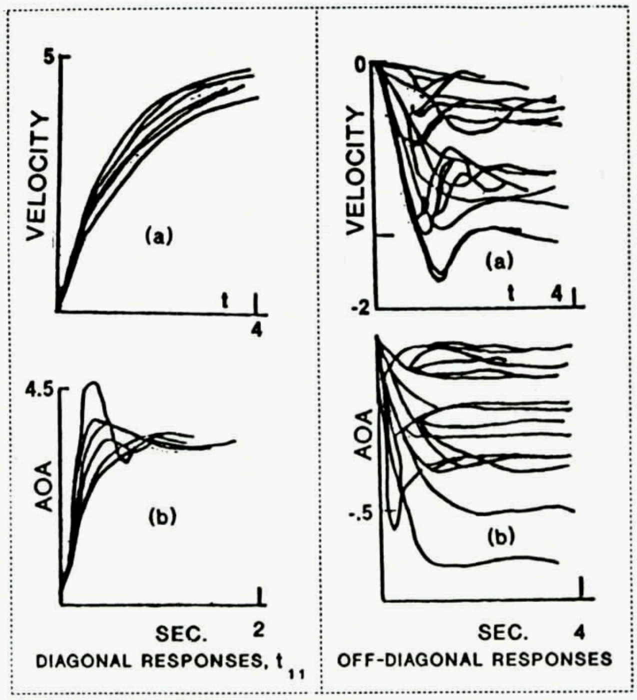


Fig. 14 F-15 STOL simulation results for: (a) velocity, (b) AOA commands

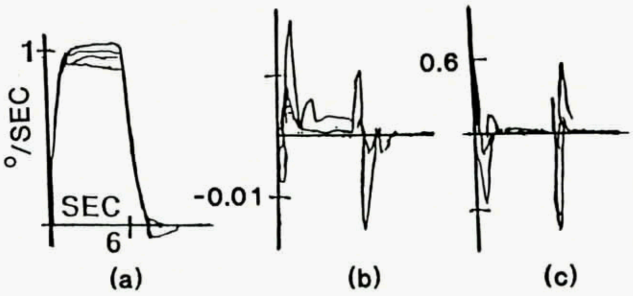


Fig. 15 Responses to roll-rate pulse command:

(a) roll rate, (b) yaw rate, (c) c*

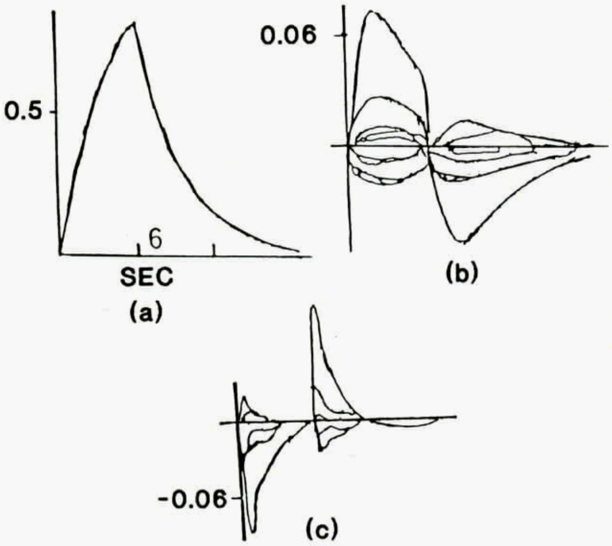


Fig. 16 Responses to yaw-rate command:
(a) yaw rate, (b) roll rate, (c) c*

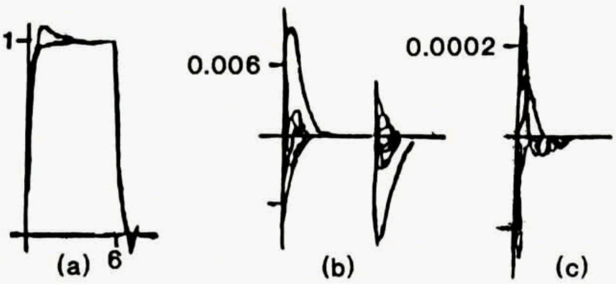


Fig. 17 Responses to c* command:
(a) c*, (b) roll rate, (c) yaw rate

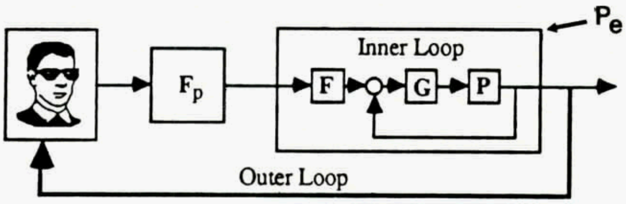


Fig. 18 Man-in-the-loop
flight control system

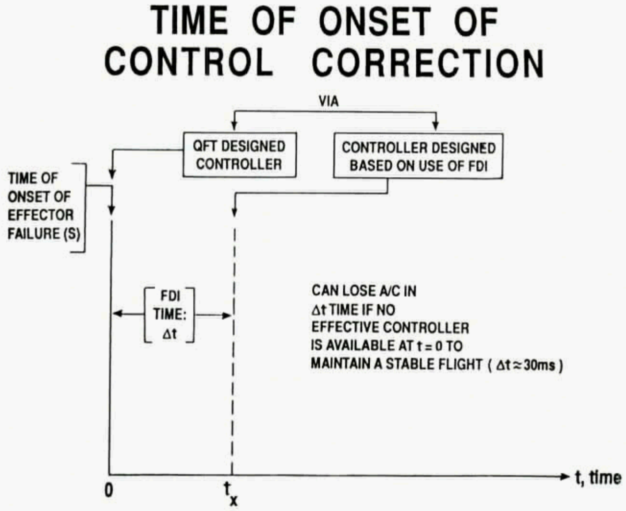


Fig. 19 Dual reconfigurable
flight control system scenario

ADAPTIVE RECONFIGURABLE FLIGHT CONTROLS FOR HIGH ANGLE OF ATTACK AIRCRAFT AGILITY

by

T. Sadeghi[1], **M. Tascillo, A. Simons** and **K. Lai**[2]
Electrical Engineering Department
State University of New York
Binghamton, New York 13902-6000
United States

Abstract

Two control technologies have been developed and integrated towards the definition of an Integrated Flight Control for the year 2000 (IFC2000) aircraft. The control technologies considered for IFC2000 are flight control reconfiguration and post stall maneuvering. The reconfiguration technique allows an aircraft to utilize the inherent redundancy among its control effectors for maintaining aircraft's controllability after loss (or degradation in effectiveness) of one or more of its control effectors. The reconfiguration technique used in this paper consisted of redistribution of control signals after identification of a control effector failure (first stage adaptation). This technique is based on utilizing a pseudo inverse algorithm and minimizing a performance index to redistribute pilot's commands to the remaining control effectors. Two control laws were developed for controlling the aircraft in the post stall region where the aircraft is flying at High Angle Of Attack (HAOA) while allowing reconfiguration in the event of surface damage or actuator failure. A self tuning adaptive control law was developed for parameter estimation and control gain tuning (second stage adaptation). The control law utilizes the Bierman's algorithm for estimating aircraft parameters, and a linear quadratic regulator for tuning the gains. A neural net control law was developed to account for nonlinearity, parameter uncertainties, and disturbances in the flight control system. The reconfiguration, adaptive and neural net control laws have been partially integrated, the results are reported in this paper. Post stall maneuvering is configuration sensitive requiring a high performance aircraft with relaxed static stability, thrust vectoring, and/or additional surfaces (such as canard). A generic high performance aircraft model was modified to incorporate thrust vectoring for generating pitch and yaw moments. A control structure was developed to fly the aircraft with high angle of attack at low speed. The control structure can track α, ϕ, and β commands from pilot's longitudinal and lateral sticks and rudder pedal, respectively. The control laws were designed to give steady state tracking for fuselage pointing. Thrust vectoring was used to produce pitch and yaw moments at HAOA. The control signal distribution function of the control laws was modified to facilitate aircraft transition from and to post stall region.

Introduction

Figure 1 represents a pictorial overview of the proposed Integrated Flight Control 2000 (IFC2000). The upper part of the diagram delineates the fundamental blocks of a flight control law, and the lower part shows the promising techniques which can be developed and integrated together to achieve the objectives of IFC2000. The end result is expected to be a highly nonlinear and adaptive control system which can provide instant controls for precision tracking of the pilot commands.

A nonlinear model of a High Performance Aircraft (HPA) with twin engines [1] was used in this development. Thrust Vectoring (TV) capability was added to HPA to make it usable for HAOA control development. Thrust vectoring moments for HAOA maneuvering were developed analytically, and were added to the HPA model. The Bare airframe nonlinear model was linearized at various flight conditions to generate approximated A, B, and C matrices of linearized aircraft models. A linear control law for HPA was developed from its nonlinear control structure.

Thrust vectoring was assumed to be made available through 3 actuated vanes which were added to each engine at the nozzle locations. Two sets of TV vanes were added around the circumferences of the nozzles in a way that their deflections into the jet streams were capable of generating pitch and yaw moments. Through simple trigonometric relationships and addition of appropriate columns to the control derivative matrix (B), TV vane deflections were translated into variations in pitch and yaw moments. Efficiency factors of 0.5 to 0.75 were used for all TV vane deflections and were incorporated in the appropriate columns of the B matrix.

Mixer technologies for blending control signals after control surface battle damage have been evolving since the early 1980's [2]. A Control Signal Distributor (CSD) function was developed and added to the HPA models. In the case of battle damage causing loss of an entire effector, the failure scenario was emulated by multiplying an appropriate column of the B matrix with a factor $\gamma = .5$. The second order effects on the stability matrix (A) due to battle damage were ignored. In the event of a partially missing effector due to battle damage or an actuator soft failure causing degradation in performance, the effect was represented by multiplying an appropriate column of the B matrix by the effectiveness factor γ $(1 > \gamma > .5)$. The damaged aircraft parameters were then estimated using the Bierman's algorithm. It was assumed that the flight control redundancy management could identify failure of the effectors. In the event of a partially missing surface or an actuator degradation, the parameter estimator of the self

---

[1] Associate Research Professor, Electrical Engineering Department, State University of New York in Binghamton

[2] Research Assistants, EE Department at SUNY, Binghamton, New York

tuning adaptive control law was used to identify the failed surface effectiveness. The CSD was then utilized to redistribute the pilot's commands to the remaining effectors.

A Self Tuning Adaptive Control (STAC) [3] law was developed with the capability of tuning the control gains once the control signal redistribution had taken place through CSD. The results obtained from integration of CSD and STAC had demonstrated the effectiveness of CSD in making gross correction in the control laws, and the role of STAC in estimating aircraft parameters and fine tuning the control gains. The speed of control optimization convergence in STAC was improved significantly when CSD was integrated with the STAC.

A Neural Network Control (NNC) architecture [4] was developed in concurrence with the CSD and STAC. The NNC is capable of producing a highly nonlinear control law for precise control of the aircraft at HAOA. This is done by training the NNC's hidden units with a large set of inputs. NNC offers a means to incorporate the needed nonlinear control and transition logic into the IFC2000 structure along with the CSD and the STAC. The NNC also prescribes the underlying architecture for the flight control computers.

With addition of thrust vectoring capabilities to HPA, the aircraft equations of motion [5] were used to derive analytic expressions for precise nonlinear control of the aircraft at high angle of attack. This control structure has a Proportional plus Integral (PI) feedback and feedthrough structure. The PI control laws are capable of rejecting disturbance and noise as well as holding the aircraft at trim conditions in absence of pilot commands. The feedforward block of this control structure consisted of a static gain matrix to guarantee steady state tracking of pilot's commands. Later on, the feedforward gain matrix will be replaced with either desired linear dynamic models or precise nonlinear dynamic relationships for desired motion.

The IFC2000 must be structurally capable of reconfiguring the control laws after failure, tuning the control gains in real-time, and flying the aircraft at high angle of attack.

Thrust Vectoring Capability

Thrust vectoring capabilities were derived analytically and were integrated into the HPA nonlinear model. The nonlinear model was linearized at high angle of attack flight conditions, and the linearized models were verified for effectiveness of the thrust vectors at those angles [11,12,13]. The HPA's linearized model was obtained at the speed of Mn = .9 and the altitude of H = 10Kft while trimming the aircraft at $\alpha = 70^o$. The following modifications were applied to convert the models to have Thrust Vectoring (TV) capabilities. Thrust vectoring was assumed to be available through 3 TV vanes which were added to the aircraft at each engine nozzle. The TV vanes were assumed to be equally distant around the circumference of the nozzles, and were appropriately sized to have sufficient clearances when they were deflected simultaneously. The TV vanes for one engine are shown in Figure 2. The bottom two vanes (when deflected) can generate yaw moment. The top vane deflection can generate pitch moment. From Figure 2, the Pitch Vane (P_v) deflection of δ_p degree generates a deflected thrust force (F_p) with its component, C_{fy}, projected onto the Z-body axis. C_{fy} thrust force, through its moment arm (R), generates a moment around the Y-body axis with respect to the aircraft's center of rotation. Therefore,

$$C_M(V_p) = \eta_p\, R\, C_{fy} = \eta_p\, R\, F_p\, Sin\delta_p \qquad (1)$$

where δ_p is the V_p vane deflection in degrees.
Similarly, for the Yaw Vane (Y_v) deflections (Y_{vr} and Y_{vl}), the following relationship holds

$$C_N(V_y) = \eta_y\, R\, C_{fz} = \eta_y\, R\, F_y\, Sin\delta_y\, Cos(45) \qquad (2)$$

Two efficiency factors were assumed for translating vane deflections into moments: η_p=0.5 and η_Y =0.35. These efficiency factors were used to create thrust moments from what was available in terms of aerodynamic moments. Equations (1) and (2), along with the converted aero data, were used to modify the nonlinear model of HPA. The modified HPA was linearized analytically to obtain analytic expression for appropriate entries of the B columns. The modified HPA was then linearized numerically to obtain numerical values for the thrust vectoring coefficients. The values of the TV parameters were approximated by assuming that the control effectiveness of the TV vanes are 75% of their counterpart primary control surfaces in normal conditions. These approximations can be replaced with real data when available. Small angle approximation was assumed for thrust vector deflections; that is,

$$Sin\ (\delta_p) \approx \delta_p \qquad (3)$$

$$Sin\ (\delta_y) \approx \delta_y \qquad (4)$$

It was assumed that the yaw vanes move collectively in one direction (coupled) to generate yaw moment in that direction. The pitch and yaw vanes can be deflected simultaneously without colliding with each other.

Aircraft Models

The HPA linear model is represented by

$$\dot{X}(t) = AX(t) + BU(t); \qquad X(t)\subset\mathbb{R}^n; \quad U(t)\subset\mathbb{R}^m \qquad (5)$$

$$Y(t) = CX(t); \qquad Y(t)\subset\mathbb{R}^l; \quad n\geq m;\ n\geq l;\ m\geq l \qquad (6)$$

where $x^T = [p, q, r, V, \alpha, \beta, \theta, \phi, \psi]$,
$u^T = [\delta_{ss}, \delta_{ds}, \delta_A, \delta_R, \delta_{vp}, \delta_{vy}]$,
and $y^T = [\alpha, \beta, \theta, \phi, \psi]$.
Engine thrusts were eliminated from the input vector of the

linear model since it was assumed no thrust modulation was required for the design objectives. For control signal distribution, the control elements were separated by the individual surfaces. HPA has two stabilators which are moved symmetrically to generate pitch moments and differentially to generate roll (i.e., four columns in B matrix with a rank of two). Also, HPA has two ailerons, a single vertical rudder, a pitch thrust vectoring vane, and two yaw thrust vectoring vanes. Therefore, the dimension of the modified B matrix is B: 9X10, but its rank is equal to 6. The control effectors have inherent redundancy in generating the moments; that is, the ailerons and differential stabilators can generate roll moments; the symmetric stabilators and pitch vanes can generate pitch moment; and the rudder and yaw vanes can generate yaw moment. Therefore, this aircraft has a minimum of dual redundancy among its effectors in all three axes. The nonlinear model was modified by moving the engine's center of gravity 30% ahead of its mean aerodynamic cord to fly the aircraft into the high angle of attack regime without violating the α limit.

Control Signal Distributor (CSD)

CSD's primary function is to integrate the reconfiguration strategy with the post stall maneuvering. For reconfiguration, once an effector damage or an actuator failure has been identified by the flight control redundancy management, the CSD redistributes control signals to the remaining healthy effectors. CSD is represented by the matrix M in the control block diagram of Figure 3. CSD's in-flight alteration of matrix M represents the first stage adaptation process to accommodate failures or to transition to and from PSR. For post stall maneuvering and as a function of angle of attack, the CSD redistributes control signals generated from the flight control laws to the surfaces and the thrust vectors in order to give more control authorities to the healthy effectors. Through CSD, an aircraft can smoothly transition to and from high angle of attack region by changing the thrust vectoring effectiveness. CSD is a full time function in a flight control system. It is represented by the matrix M in the linear control law

$$U(t) = -MKX(t) + MLNU_c(t) \qquad (7)$$

where, K is the feedback gain matrix, M is the feedthrough CSD gain matrix, L is the pilot's command-to-surface distribution block, and N is the feedforward tracking matrix. Under normal operating conditions, the CSD matrix is set to identity; $M = I_{6X6}$. The closed loop model of the aircraft is represented by

$$\dot{X}(t) = (A - BMK)X(t) + BMLNU_c(t) \qquad (8)$$

In simulation, an aircraft control effector impairment is represented by multiplying an appropriate column of the B matrix by a factor γ $(1 \geq \gamma \geq .5)$. This factor alters a surface effectiveness after an impairment. The impaired model of the aircraft is represented by the modified B-hat matrix

$$\dot{X}(t) = (A - \hat{B}MK)X(t) + \hat{B}MLNU_c(t) \qquad (9)$$

Note, the rank of the control matrix has not changed under this scenario. That is,

$$Rank(B_{9x6}) = Rank(\hat{B}_{9x6}); \qquad for \;\; 1 \geq \gamma \geq .5 \qquad (10)$$

The redistribution matrix is changed by either a right or a left pseudo inverse of impaired B matrix in order to mask out the effects. The right pseudo inverse solution to CSD is represented by

$$M = [\hat{B}^T Q \hat{B}]^{-1} \hat{B}^T Q B \qquad (11)$$

where Q is a diagonal matrix to be found by minimizing an error index in the Euclidian norm. The reconfigured closed loop system is, therefore, equal to

$$\dot{X}(t) = (A - \hat{B}[\hat{B}^T Q \hat{B}]^{-1} \hat{B}^T Q B K)X(t) + \hat{B}MLNU_c(t) \qquad (12)$$

For a perfect solution, the following equality must hold:

$$\hat{B}[\hat{B}^T Q \hat{B}]^{-1} \hat{B}^T Q B K = B K \qquad (13)$$

This relationship holds only if {BK} belongs to the right range space of {B-hat^{T}Q}; that is,

$$\hat{B}^T Q [\hat{B}[\hat{B}^T Q \hat{B}]^{-1} \hat{B}^T Q B K] = \hat{B}^T Q[B K] \qquad (14)$$

The preceding equation holds, if and only if

$$(I - \hat{B}[\hat{B}^T Q \hat{B}]^{-1} \hat{B}^T Q) B K = 0 \qquad (15)$$

Therefore, for arbitrary and non trivial matrices Q, B and K, the following relationship must hold

$$(I - \hat{B}[\hat{B}^T Q \hat{B}]^{-1} \hat{B}^T Q) = 0 \qquad (16)$$

Defining an error equation in a Euclidean norm,

$$\varepsilon = | I - \hat{B}[\hat{B}^T Q \hat{B}]^{-1} \hat{B}^T Q | \qquad (17)$$

the optimized redistribution is obtained by minimizing the error function in a least squares sense; that is,

$$\frac{Min}{Q}[\varepsilon] = \frac{Min}{Q} | I - \hat{B}[\hat{B}^T Q \hat{B}]^{-1} \hat{B}^T Q | \qquad (18)$$

This optimization must take place once after each impairment and for all flight conditions. A fast parallel algorithm [6] was used for computing the Moore-Penrose pseudo inverse. This algorithm was based on Karmarker's algorithm [7] which was also used to minimize the preceding performance index.

The weighting matrix Q in CSD can be also used to represent the control couplings among the surfaces (e.g., aileron to rudder interconnection). Furthermore, pre-computed Q matrices can be either scheduled as a function of angle of attack or adaptively computed in real time as the aircraft transition from and to post stall region. The entries of the Q matrix represent the emphasis placed on effectiveness of an effector with respect to other effectors.

A left pseudo inverse solution may be obtained only if the number of columns in the B matrix is greater or equal to the number of rows. A solution for matrix M is obtained by

$$M = Q\,\hat{B}^T\,[\hat{B}Q\hat{B}^T]^{-1}\,B \tag{19}$$

The closed loop system in this case has the form

$$\dot{X}(t) = (A - \hat{B}Q\hat{B}^T[\hat{B}Q\hat{B}^T]^{-1}BK)X(t) + \hat{B}MLNU_c(t) \tag{20}$$

or

$$\dot{X}(t) = (A - BK)X(t) + BLU_c(t) \tag{21}$$

which is the original closed loop dynamics before impairment.

In general, employing the left pseudo inverse requires a reduction in the model's order to satisfy the condition on matrix B's dimension. The least squares solution must be upgraded to a larger dimension model by minimizing the effects of the matrix M (i.e., the spillover) on the states that were eliminated through model reduction.

Self Tuning Adaptive Control (STAC)

A Self Tuning Adaptive Control (STAC) was developed for controlling the aircraft at high angle of attack and fine tuning the controller's gains every time the CSD matrix was changed. The STAC law's block diagram shown in Figure 4 relies on its ability to identify the stability and control parameters of the aircraft and tuning the feedback control gains to compensate for any changes that may have occurred in aircraft states during a flight. STAC laws operate continuously to automatically compensate for the aircraft parameter variations due to: changes in flight conditions, changes due to failure of its redundant components, and changes in transitioning to and from PSR.

For reconfiguration after a failure of a redundant effector, the STAC relies on the flight control's redundancy management and Fault Detection/Isolation (FDI) to detect and isolate failures. A positive identification allows adaptation and reconfiguration procedures to take corrective action in controlling the aircraft. For example, should the actuator belonging to a stabilator fail in flight (while in normal flight condition or in post stall region), sensors dedicated to that actuator would indicate that the actuator being out of tolerance. Upon comparing the impaired performance of the aircraft to that of an unimpaired model on-board the flight control computers, a discrepancy would be detected and the underlying failure isolated. The failure or battle damage may result in: 1) loss of an entire surface, 2) partially missing surface, 3) actuator/surface floating, 4) actuator failing in fail-safe position, or 5) actuator failing hardover. The failure/battle damage scenarios 1, 3, and 4 require zeroing out the corresponding column in $\underline{B}$ matrix. Scenario 2 requires an on line estimation of surface effectiveness. Scenario 5 may require either deflection of another surface (out of a pair) to restore aircraft symmetry, or may require re-computing the M matrix in CSD to account for a hardover failure.

Upon an occurrence of one of the preceding failure scenarios, the first stage adaptation is achieved through feedthrough CSD matrix for redistributing the control signals to the remaining healthy surfaces. As shown in Figure 4, the STAC and CSD are integrated together to realize an optimum performance. The Bierman Estimator block in Figure 4 is the heart of the second stage adaptation, and instrumental in making the first stage effective in the case of a partially missing surface. The STAC is capable of estimating effectiveness of surface after a battle damage (i.e., the factor γ). The adaptive law block in Figure 4 is a linear quadratic regulator which computes the necessary adjustments to the feedback gains after matrix M has been modified by the first stage adaptation. The Bierman Estimator updates the entries of A and B matrices in the System's Truth Model (STM). The CSD modifies the error signals used to command the actuators and to deflect the effectors. This function utilizes the redundancy among the effectors to restore the aircraft's stability and tracking characteristics.

The feedforward block can easily be replaced with linearized (or nonlinear) dynamic models of the aircraft for implementing a model reference adaptive control structure to guarantee a fully integrated reconfiguration and high angle of attack control law. The linear quadratic regulator can be replaced with the Optimal, Proportional plus Integral control techniques [10] which were specifically developed for self repairing flight control of an aircraft with dual redundant effectors.

The STAC is structured around the Bierman's algorithm which is known for its numerical stability. The equations governing the Bierman's algorithm are delineated in [3]. The STAC law corresponding to Figure 4 has the following structure:

$$U(t) = -K_a M_a X(t) + L_a M_a N_a U_c(t) \tag{22}$$

where subscript "a" denotes on line adaptation of the corresponding matrices, B_a represents the B matrix effectively in place after the first stage adaptation through CSD, and K_a is the adaptive discrete Kalman gain matrix computed by solving the discrete Riccati equation:

$$P_a = A_a^T[P_a - P_a B_a (B_a^T P_a B_a + R)^{-1} P_a] A_a + S \tag{23}$$

and

$$K_a = (\,B_a^T P_a B_a + R\,)^{-1} B_a^T P_a A_a \tag{24}$$

Referring to Figure 4, the first stage adaptation is represented by CSD, the Bierman's Estimator by *BE*, and the second adaptation law by *AL*. The integrated CSD/STAC law implementation is represented by the following steps:

1- Estimate Impairment and Aircraft Parameter Variations:

$$(A_a , B_a , C_a) = BE(U, X, Y - Y_m) \tag{25}$$

2- Update Truth Model Parameters:

$$(A_m , B_m , C_m) = (A_a , B_a , C_a) \tag{26}$$

3- Minimize Control Signal Distributions:

$$\frac{Min}{Q_a} = |I - \hat{B}_a [\hat{B}_a^T Q_a \hat{B}_a]^{-1} \hat{B}_a^T Q_a B_a| \tag{27}$$

4- Update CSD Matrix, M_a:

$$M_a = [\hat{B}_a^T Q_a \hat{B}_a]^{-1} \hat{B}_a^T Q_a B_a \tag{28}$$

5- Compute Pilot-to-Surface Command Distributions:

$$L = [l_{ij}] \text{ as a function of } \alpha \tag{29}$$

6- Compute Discrete Riccati Solution:

$$P_a = A_a^T [P_a - P_a B_a (B_a^T P_a B_a + R)^{-1} P_a] A_a + S \tag{30}$$

7- Update Feedback Gain:

$$K_a = (B_a^T P_a B_a + R)^{-1} B_a^T P_a A_a \tag{31}$$

8- Update Feedforward Tracking:

$$N_a = [C_a (I - A_a + \hat{B}_a M_a K_a)^{-1} \hat{B}_a M_a L_a]^{-1} \tag{32}$$

In the event of a partially missing surface where the remaining portion of a surface is still functional or in the case of a soft actuator failure which results in performance degradation, it may be desirable to utilize that effector in the reconfigured control law. The effectiveness of the effector is estimated by the Bierman's on-line recursive algorithm which continuously updates the current estimate of the surface parameters with each control loop iteration. The Bierman U-D factorization algorithm [3] is numerically robust with fast recursive steps. This algorithm has an advantage over standard Recursive Least Squares (RLS) algorithms in that instead of recursively computing a square matrix, it calculates a factor of that square matrix.

The estimator structure is based on the following equations [3]:

The input and output measurements are represented by

$$\Omega^T(t) = [U(t-1), \ldots , U(t-n_b-1), \ldots , Y(t-1), \ldots Y(t-n_a)] \tag{33}$$

The parameters to be estimated (A, B, and C Matrix entries) are represented by

$$\Theta^T = [a_{11}, \ldots , a_{nn}, b_{11}, \ldots , b_{nm}, c_{11}, \ldots , c_{nr}] \tag{34}$$

The measurement at discrete time t and the estimates of the measurement are

$$Y(t) = \Omega^T(t)\ \Theta(t) + e(t) \tag{35}$$

$$\hat{Y}(t) = \hat{\Omega}(t)^T\ \hat{\Theta}(t) + \hat{e}(t) \tag{36}$$

The error in estimate is represented by

$$\varepsilon(t+1) = Y(t+1) - X^T\ \hat{\Theta}(t) \tag{37}$$

The parameter estimate is obtained by minimizing a quadratic performance in error terms with respect to the estimated parameters

$$\frac{Min}{\hat{\Theta}}\ J = \sum_{i=1}^{N} \lambda^{t-i} \hat{\varepsilon}^2(i) \tag{38}$$

The new estimates are obtained by solving the following equations recursively

$$P(t+1) = \lambda^{-1} P(t)\ [I_m - \frac{\Omega(t+1)\,\Omega(t+1)^T P(t)}{\lambda + \Omega(t+1)^T P(t) \Omega(t+1)}] \tag{39}$$

$$\hat{\Theta}(t+1) = \hat{\Theta}(t) + P(t+1)\,\Omega(t+1)\,\varepsilon(t+1) \tag{40}$$

Bierman's recursive algorithm for finding P(t) [3] generates a solution to equation (39) in the form $P(t) = S(t)^T S(t)$, where S(t) is an upper triangular matrix. S(t) corresponds to the square root of P(t). A factorization algorithm updates S(t) with each iteration. As this operation is essentially based on the square root of P(t), the precision of the calculation is effectively doubled over that of operating upon P(t).

The magnitude of the state error between the aircraft measurement vector and the model is monitored by the fault detection/isolation function which activates the Bierman's Estimator (*BE*) and the Adaptive Law (*AL*) blocks when a large value of state error indicates that a damage/failure has occurred. When the parameter estimation scheme is invoked, a jitter signal consisting of 30 Hz square waves with an amplitude equal to 1% of the input signal is added to the input to assure convergence of the estimation algorithm. The estimate of B is updated with every iteration of the estimator. In many implementations it would be required that matrix M in CSD be updated only when the system identification algorithm has fully converged and the state error has been reduced to an acceptable level. When the state error is reduced below a threshold, the parameter estimator stops and the jitter signal is removed from the control inputs.

Neural Network Control (NNC)

With recent development in massively parallel and high speed computer technologies, application of Neural Network Architectures (NNA) in developing a precise and highly nonlinear flight control system capable of performing complex and integrated control functions has become a reality. Once it is properly trained, a Neural Network Control (NNC) can fly a high performance aircraft with self repairing and high angle of angle of attack features within its expanded flight envelope.

The NNA principles are founded on Rosenblatt's [8] theorem that if a training set of linearly independent matrix inputs were introduced to a NNA, then the NNA in a finite number of iterations would learn the function that is generating the inputs. A basic NNA shown in Figure 5 represents a function $g_i(x) = x * w^T_i$, where x is a vector of inputs and w_i is a vector of weighting functions. The NNC architecture for the HPA flight control law is based on ANYANET [4] with 108 hidden units (2 hidden units per feedback entry for a feedback gain matrix, K_{6x9}) as shown in Figure 6. The hidden units are those that are not directly accessible as an input or output. They represent either trained or exact nonlinear functions of the control. The weights of the hidden units in NNA are adjusted by utilizing the aircraft actuator commands, its state information, and the derivative of the states. The adjustments are done by comparing the desired feedback entries to the outputs of NNA. The weight adjustments proceed by minimizing the error between the measured and the estimated parameters of NNC.

The NNC in Figure 6 was developed to accurately represent the feedback gain matrix, K, for the aircraft through incremental addition of hidden units which, beyond the original inputs, are trained by other hidden units devoted to the same output vector. If a particular output does not need a highly nonlinear adjustment, its hidden units will have relatively small values. The NNC does not make any preferential assumptions about the inputs or the outputs, but allows the net to decide if there are any preferences. It continues to implement Fahlman's [9] incremental hidden units to eliminate unnecessary connections for on-line updates, but accommodates outputs which exhibit varying degrees of nonlinear behavior. The ultimate goal here is to convolute the NNC with the STAC in developing a neural net adaptive flight control system capable of controlling a high performance aircraft at high angle of attack while providing the aircraft with self repairing capabilities.

Successful utilization of two NNAs controlling linear models of the HPA and a F-16 [10] are represented in the simulation section. The NNC laws for the linear F-16 model required 12 inputs, 42 hidden units to represent entries of A, and B matrices, and 6 outputs representing the entries of the state feedback gain matrix, K. The same NNA architecture was trained with the HPA linear data with 9 states, 6 inputs and 6 outputs. The performance of the NNC for both models were simulated, and the results are discussed in [16].

Nonlinear Control Structure (NLC)

Flying an aircraft in the Post Stall Region (PSR) with a High Angle Of Attack (HAOA) requires a precise nonlinear control that accounts for large angle variations as well as the dynamic couplings exist between the aircraft states. The Post Stall Maneuvering (PSM) for the conventional fighter aircraft is usually limited to 25 to 30 degrees. Expanding the flight envelope for higher angle of attack at higher speeds than Mn=.3 is not of interest. This is because handling aircraft at higher α and Mn is beyond the control of a human pilot. At low speeds (Mn ≈ 0.1 - 0.3) and an altitude range of H ≈ 10K ft - 40K ft, the flight envelope may be expanded up to a theoretical limit of $\alpha \approx 90^o$. In practice it is reasonable to develop control for PSM of about $\alpha \approx 70^o$. Many assumptions used in aircraft modeling do not hold for PSM. This inspires motivation for developing a nonlinear control system for highly nonlinear and dynamically coupled aircraft motion.

A fundamental difference between a HAOA control and conventional controls is the small angle assumption which does not hold for HAOA aircraft; that is,

$$\alpha = Tan^{-1} \frac{W}{U} \neq \frac{W}{U} \qquad (41)$$

$$\beta = Sin^{-1} \frac{v}{V} \neq \frac{v}{V} \qquad (42)$$

Another important difference is that a conventional aircraft with conventional surfaces cannot fly in PSR with an angle of attack greater than 30°. This means modifications must be made to an existing aircraft configuration in order to make it fly beyond its stall margins. PSM requires aircraft configurations that are either statically unstable such as the experimental X29 and X31, or modified with additions of surfaces and thrust vectors such as the F15 and F18 research test beds. Another fundamental change to an existing aircraft requires moving the aircraft's center of gravity 20 to 30 percent ahead of its Mean Aerodynamic Cord (MAC). This modification is done easily in an aircraft's dynamic model. In the real aircraft, weight must be added to the noseboom. This change would allow commanding an aircraft to higher angles of attack without violating the aircraft's stall margin or its angle of attack limiter.

Another fundamental modification to the aircraft configuration is due to loss of rudder's effectiveness and reduced capability to pitch down the aircraft from a HAOA orientation with conventional surfaces. An aircraft's controllability is greatly improved when additional control surfaces such as canard or thrust vectoring vanes are added to its control effector suite. Such effectors can provide additional yaw and pitch moments at HAOA.

The HPA nonlinear and linear models were modified to have thrust vectoring vanes for this purpose. At HAOA, an aircraft is encountered with increased drag and reduced effectiveness of its conventional surfaces for generating yaw, roll or pitch moments. A three dimensional visualization of an aircraft orientation, as shown in Figure 7, considering the aircraft flying at low speed and high angle of attack is indicative of the moment losses from deflection of its conventional surfaces. Furthermore, ineffectiveness of the rudder can cause unwanted and uncontrollable yaw and sideslip angles if the aircraft were to fly with conventional controls into this region. The control laws and structures that are developed using linearized models under small angle assumptions are deemed to fail at HAOA. It is, therefore, logical to address HAOA maneuvering by nonlinear controls derived directly from nonlinear model of the aircraft. Such a control structure can cope with nonlinearity and coupling of equations of motion at HAOA. Linear models and control laws should then be modified to reflect the nonlinear control structure before any analysis or synthesis of the control laws are performed.

As shown in Figure 7, it is desired to fly a modified HPA to a high angle of attack and be able to roll it around its velocity vector. It is realized from this objective that a control system based on acceleration command (g-command) is not feasible for this application. The NLC must be developed to command α, ϕ, and β as these angles are critical in acquiring precise control of the aircraft and meeting the PSM objectives. This can be realized by carefully by visualizing the aircraft's motion in Figure 7 under the conditions specified.

Due to increased drag and reduced control effectiveness of the conventional surfaces, a number of issues may be encountered when developing NLC structures and laws for HAOA maneuvers. These issues are outlined in Table 1 in terms of cause, effects, and remedies.

To handle flight regimes other than HAOA, it is desired to develop a NLC structure capable of transitioning the flight control, for example, from a α-command to a g-command control system when exiting PSM and entering a high speed flight condition, or from α-command to c^*-command and β-dot-command for landing, takeoff, and cruise. These flight control transitions can be constructed in the Control Signal Distribution (CSD) function which goes into effect automatically without requiring any pilot intervention.
In consideration of the issues discussed in Table 1, the nonlinear control (NLC) structure must incorporate the following provisions:

1- NLC structure must have a C^*/β-dot structure modified to have capability of controlling the aircraft at HAOA and low speed.

2- Modifications to the C^*/β-dot structure must be faded in and out through transition logic in the CSD as the aircraft goes into the PSM region and comes out of it.

3- NLC at HAOA must have a fuselage pointing characteristics in order to assume tracking of commanded angles with maximum decoupling of aircraft states.

4- NLC at HAOA must command angles: Longitudinal stick commanding α, lateral stick commanding ϕ, and rudder pedal commanding β-dot with minimizing ψ departure from zero.

5- NLC signals must be distributed to the effective control effectors at HAOA.
6- Stall prevention logic and departure prevention logic must be superimposed on the distributed control signals before reaching the actuators.

7- NLC Laws at HAOA must be of Proportional + Integral (PI) type to assure holding of a commanded angle in the

Table 1: Issues in Developing NLC Laws for HAOA Maneuvers

Cause		**Effects**	**Remedy**
Unanticipated roll	(ϕ)	Oscillatory (dutch roll), & roll X-spin & steep spin	Increase loop gain in roll Adjust Dutch roll damping
Unanticipated yaw	(ψ)	Slow yaw rate spin, Z-Spin	Adjust ARI gains Increase loop gain in yaw
Unanticipated SideSlip	(β)	Flat spin	Use yaw TV to compensate increase yaw TV authority Coordinate yaw TV with aileron
Unanticipated AOA	(α)	Inverted Spin	Use PI control and α-Command increase integral gain
Stall		Uncontrollable	Add stall prevention logic Add g-command and roll feedback Add Pitch TV for pitch down
Limited Sensor Measurement		Opens Control Loops	Modify air data computer, angle of attack and sideslip vanes to
Limited Schedule		Divide by zero Loss of control	Wind tunnel and flight tests Expand flight envelope database
Wing Rock		Oscillatory Spin	Increase dutch roll damping
Nose Slice		Flat Spin	Increase Yaw authority
All angle departure		Rock, roll and tumble	Redesign NLC law and structure Add departure prevention logic

absence of feedback error signal, and must reject unwanted noise and disturbance which could be promoting stall or departure. The integral gains must be adjusted to assure aircraft being capable of pitching into the wind (pitch down) and not contributing to spin tendency of the aircraft.

8- NLC at HAOA must include feedback from α, β, ϕ, and ψ, and their derivatives where possible.

9- NLC at HAOA must be digitized and analyzed for the effects of digital time delay on the phase margins of the aircraft's control loops.

10- The candidate gains for NLC at HAOA must be optimized to assure best performance in this sensitive and highly critical region.

STAC Simulation

The CSD/STAC structure of Figure 4 was simulated in discrete time with a control sampling rate of 100 Hz. The discrete state space system of linear F-16 model was used to represent the airframe, actuators and sensors. Actuator dynamics of 20/(s+20) were augmented into the continuous system before converting into discrete form. The actuator limits incorporated during simulation were:
Maximum Range: ± 20°; Maximum Rate: $24^{o/sec}$.
Each simulation run consisted of 6 seconds of flight time responding to the following step control inputs:
δ_{E1} = -.1; δ_{E2} = -.1; δ_{F1} = 1.0; δ_{F2} = 1.0.

A subset of the simulation results from [14, 15] is shown here. The feedforward gain matrix, N, was computed to have the θ and α angles track the elevator and the flap commands in steady state. All angles in the simulation were measured in degrees and plotted as a function of time in seconds as shown in Figure 8. In all simulations, the commands were made of a 1 second ramp to the desired value with a 1 Hz sinusoidal function superimposed to represent pilot's stick motion.

Figure 8a shows the unimpaired response of the aircraft to the commanded α, and Figure 8b represents the aircraft's response to the commanded θ. In both cases, steady state decoupling of θ from α was realized. Figure 8a is a representative of commanding the aircraft to fly at 70° angle of attack while keeping the pitch angle at zero. Figure 8c represents θ and α responses to the commanded values of 30° and 70° for when a flaperon was totally inoperative after 3.5 seconds. Upon correct FDI, only the CSD was reconfigured to redistribute the control signals to the remaining healthy surfaces. The simulation shows successful recovery of aircraft after reconfigured CSD went into effect at HAOA. Figure 8d represents the simulation results for when a flaperon was partially damaged. In this case, STAC was used to estimate the entries of the B matrix and to update all of the appropriate gain blocks. The impaired and adapted responses of the aircraft are shown in Figure 8d. The results show successful adaptation of STAC without using the CSD function. Here the Bierman's algorithm was excited by the jitter signal, and took about 1 second to estimate the entries of the impaired B matrix. The simulation of combined STAC and CSD (not shown here) improved the aircraft's performance significantly in presence of parameter variations in the A and B matrices. Preceding failure scenarios were repeated for a stabilator failure. The simulation results of Figures 8e and 8f show excellent performance recovery. Figure 8g shows the Bierman's algorithm convergence properties in estimating the impaired B matrix.

NNC Simulation

The Neural Net Architecture (NNA) for controlling a linear F-16 model required 42 hidden units (2 per parameter). Training of the Neural Net Control (NNC) took about 90 hours on a IBM 386 Personal Computer, utilizing 10 training set examples [16]. These examples represented 1% variations in the A and B matrix entries. The objective of training NNC was to develop a robust flight control law over an assumed flight envelope represented by the 20% parameter variations. Because of small training sets and the time required to train NNC, the simulation was stopped after 1% parameter variation training. The step responses of the NNC system and the corresponding gain matrix are shown in Figures 9.

The HPA model used in the simulation [17] has 9 states for coupled longitudinal and lateral dynamics and 6 control inputs. An optimization technique was used to minimize the errors between the outputs of NNC and the outputs of aircraft which, in this case, were the states. The optimization technique in minimizing the errors was a mechanism to train the NNC. Various limits were incorporated in the optimization technique. They included: maximum number of iterations, maximum allowable error, step size per variable per iteration, and limits on the reflect, shrink, and expand parameters. The network was trained by 300 input training sets with calculated NNC weights truncated by 10^9 and 10^{-12} factors.

Figure 10a represents a top level NNC's hidden unit architecture for one output. The HPA's NNC required 54 outputs, each with subsequent Hidden Units 1 and 2 as shown in Figures 10b and 10c.

The linearized HPA model has a non minimum phase transfer function for the angle of attack dynamics. The right half plane zeros associated with the angle of attack dynamics further complicated the NNA training and the tuning process of the hidden units. Figures 11a to 11f represent HPA's transient responses to various step inputs with the trained neural net control. Figure 11a shows the unimpaired aircraft response to a 70° α-command, 30° ϕ-command, and 0° β-command. Figure 11b represents aircraft responses to the same commands for when one stabilator was lost and only the NNC was employed. Note, this failure scenario corresponds to loss of both symmetric and differential stabilators. Although the aircraft remained stable, significant degradation was noticed in the transient responses. Figure 11c shows the aircraft responses after loss of a stabilator, but this time both CSD and NNC were used to control the aircraft. These transient responses show excellent recovery from this failure scenario with minimum failure transients. Figure 11d depicts the aircraft's responses to 35^o α-command, 15^o ϕ-command, and 0^o β-command for when 50% of a stabilator was lost and only NNC was employed. The NNA was trained to assess and use the remaining portion of the stabilator. The training of NNA to recover from this failure scenario showed about 15% degradation in the aircraft's α command response. When 50% loss of aileron was simulated with NNC only, the aircraft response to α command was unsatisfactory as shown in Figure 11e. As shown in Figure 11f, the combined NNC and CSD performed effectively at high angle of attack to provide controllability to the aircraft after loss of an aileron.

Similar results were obtained for loss of rudder and thrust vectors. The simulation showed that the NNC required longer time and more input sets to train. The non-minimum phase characteristics of aircraft's dynamics made it more difficult to train NNC. High gain control was realized from current training as seen in all simulation results.

Conclusion

The control laws presented in this paper are capable of providing aircraft controllability at high angle of attack and reconfiguration after an actuator failure or a damaged effector.

The control structure contained the following constituents. A Control Signal Distributor (CSD) capable of redistributing control signals after an effector failure. This capability can be extended to allow an aircraft to transition to and from post stall region. CSD was shown to be very effective for both self repairing and high angle of attack flight controls. A Self Tuning Adaptive Control (STAC) law was introduced with numerically robust estimator capable of estimating aircraft parameter variations as well as estimating effectiveness of a partially missing surface. It was shown that integration of CSD with STAC helped significantly in an aircraft's recovery from battle damage, and improved the convergence speed of the estimator significantly. Although the Neural Net Control (NNC) did not perform as well as the STAC, its development and improvements could make the approach a viable candidate for IFC2000. Integration of the NNC with the CSD, however, was very successful and comparable to the STAC with CSD. Many modeling assumptions and approximations used in present flight controls are no longer valid for developing high angle of attack control laws. At HAOA, an aircraft's characteristics are highly nonlinear and dynamically coupled. Therefore, precise nonlinear control laws are required to cope with the nonlinear phenomenon. The requirements, cause, and effects of developing such a control law were discussed.

Integration of these promising techniques and better understanding of aircraft behavior at high angle of attack will pave the path to the future in developing an Integrated Flight Control for the year 2000 aircraft.

References

1. Brumbaugh, R.W.,"An Aircraft Model for the AIAA Controls Design Challenge", AIAA Guidance, Navigation, and Control Conference, 8/91
2. Rattan,K.S.,"Evaluation of Control Mixer for Reconfiguration of Flight Control System", NAECON 85
3. Wellstead, P.E. and M.B. Zarrop. Self-Tuning Systems, New York: Wiley. 1991
4. Simons, A.,"Robotics Trajectory Control Employing ANYANET Neural Architecture", Master's Thesis, Renssaeler Polytechnic Institute, Troy, New York, 8/92
5. McRuer, et. al.,"Aircraft Dynamics and Automatic Control", Princeton, 1973
6. Benson, M.W., et. al.,"Fast Parallel Algorithm for Moore-Ponrose Pseudo inverse Computation",Los Alamos Nat. Lab., @nd Conf. on Hypercube Multiprocessors, 9/86
7. Karmarker, N.,"Some Comments on the Significant of the New Polynomial-Time Algorithm", AT&T laboratory Report.
8. Rosenblatt, F., Principles of Neurodynamics, Spartan Books, New York, 1962.
9. Fahlman, S.E.,"The Cascade-Correlation Learning Architecture", Carnegie Mellon University, Pittsburgh, Pennsylvania, 2/90
10. Sadeghi.,T., et. al., "Optimal, Proportional plus Integral Control Technique for Self Repairing Flight Controls", IEEE Transaction on Automatic Control, 1983.
11. Lai. K.,"AIAA High Performance Aircraft Model: Modifications and Integration with Matrix-X Environment", Project Report, No. EE5061, 12/91.
12. Lai. K.,"AIAA High Performance Aircraft Model: Incorporation of Thrust Vectoring", Project Report, No. EE6021, SUNY, Binghamton, New York, 6/92.
13. Lai. K.,"Integrated Post Stall Maneuvering and Self Repairing Flight Control", Master's Thesis, SUNY, Binghamton, New York, 8/92.
14. Tascillo, M.,"Self Tuning Adaptive Control of F-16 Model", Project Report, No. EE5062, SUNY, Binghamton, New York, 2/92.
15. Tascillo, M.,"Integrated Self Tuning Adaptive Control and Control Signal Distributor for a High Performance Aircraft", Project Report No. EE6021, SUNY, Binghamton, New York, 5/92.
16. Simons, A.,"A Neural Net Control Architecture for F-16 Model", Project Report No. EE5063, SUNY, Binghamton, New York, 3/92.
17. Simons, A.,"Neural Net Control Development for a High Performance Aircraft", Project Report No. EE6023, SUNY, Binghamton, New York, 6/92.

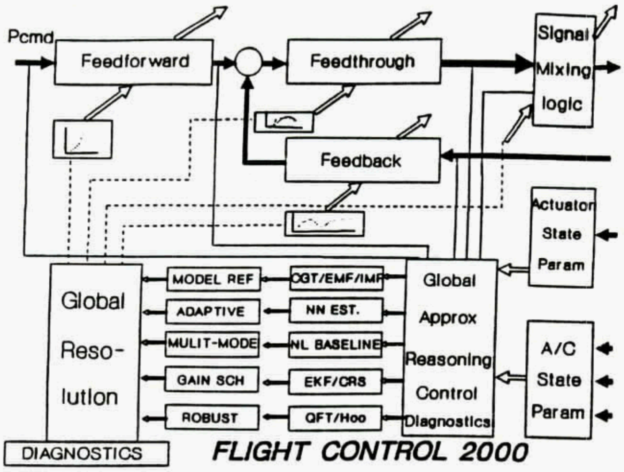


Figure 1: IFC2000 Concept for High Performance Aircraft

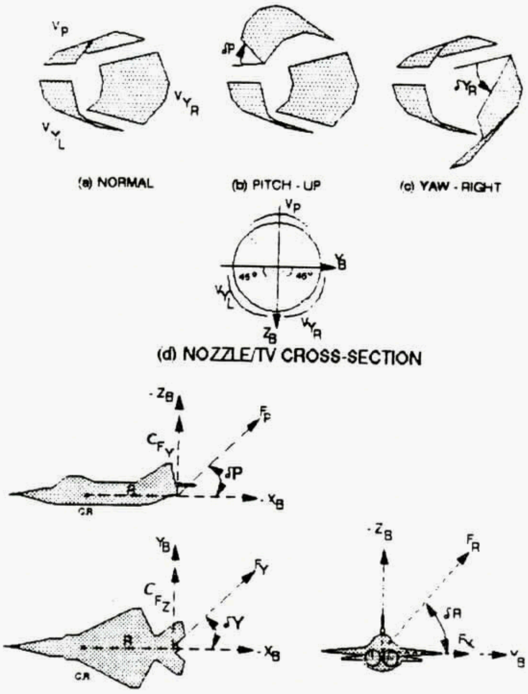


Figure 2: Addition of Thrust Vectoring to the HPA

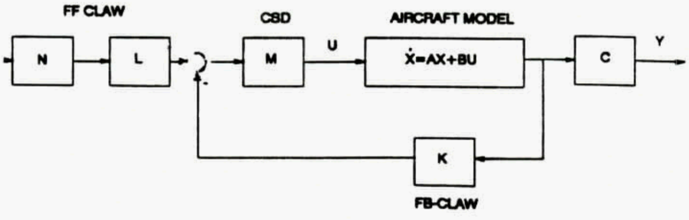


Figure 3: Control Law Block Diagram

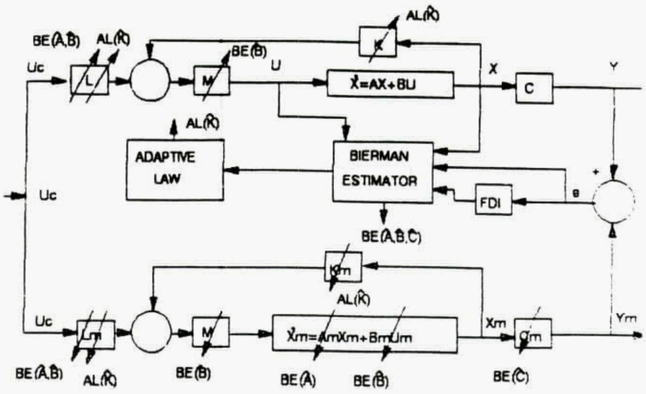


Figure 4: Self Tuning Adaptive Control Law

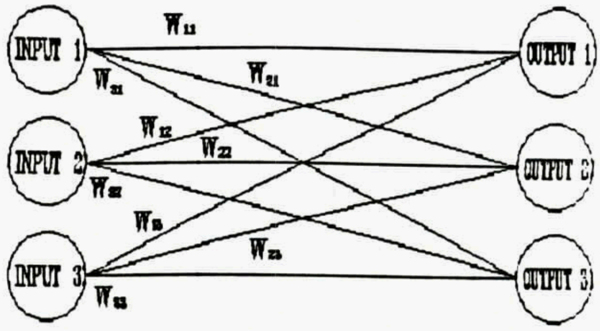
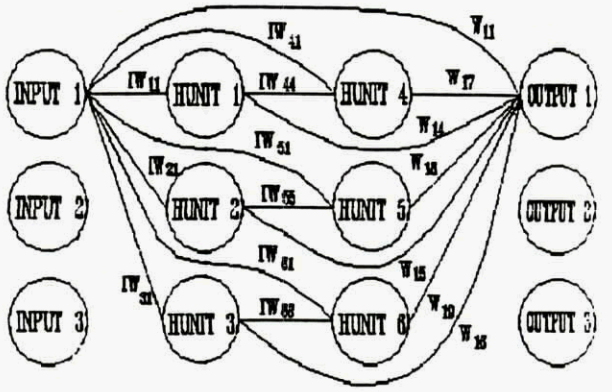


Figure 5: Basic Neural Net with No Hidden Units



(2ND & 3RD INPUTS & OUTPUTS NEGLECTED FOR CLARIFICATION)

Figure 6: ANYANET with 2 Hidden Units per Output

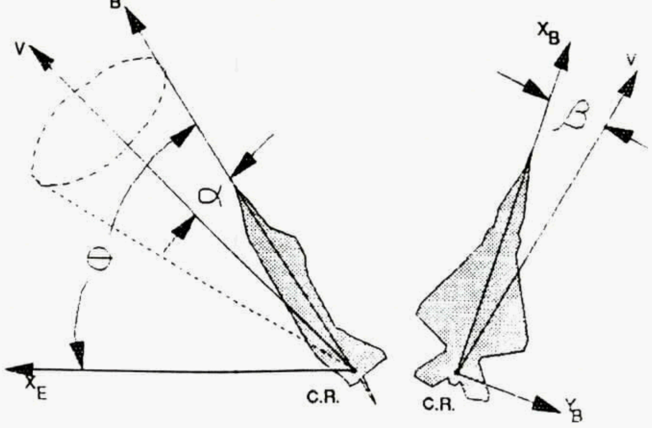


Figure 7: An Aircraft at High Angle of Attack

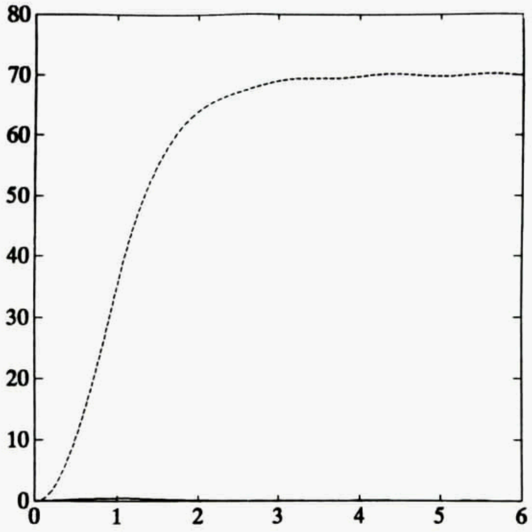


Figure 8a: Unimpaired F-16 Decoupled Response
$\alpha=70^\circ$, $\Theta=0^\circ$; 1 Sec Ramp, 1 Hz Stick

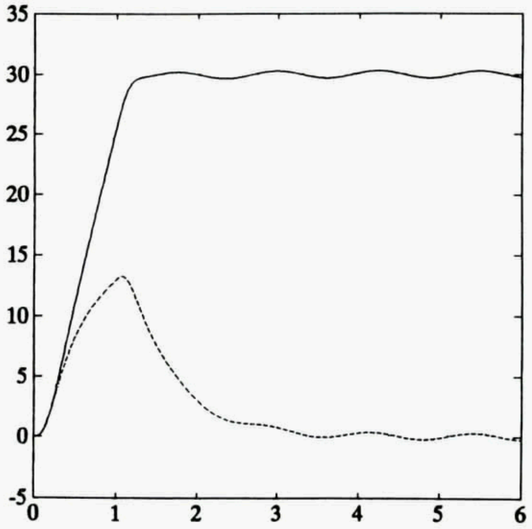


Figure 8b: Unimpaired F-16 Decoupled Response
$\alpha=0^\circ$, $\Theta=30^\circ$; 1 Sec Ramp, 1 Hz Stick

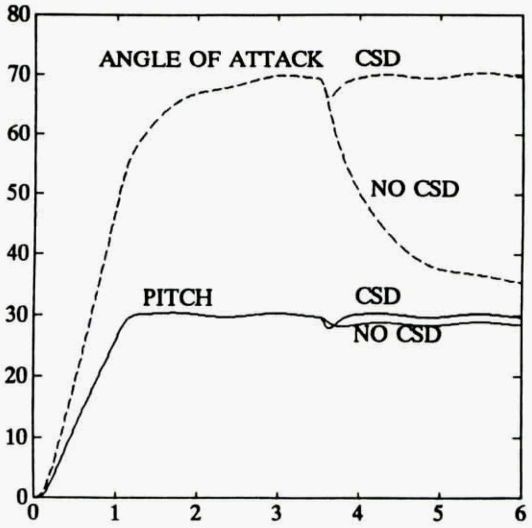


Figure 8c: Impaired F-16, Loss of One Flaperon
With and Without CSD; $\alpha=70^\circ$, $\Theta=30^\circ$

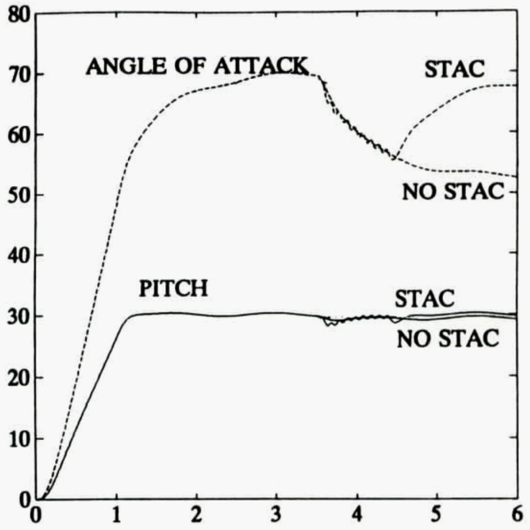


Figure 8d: Impaired F-16, 50% Loss of One Flap
With and Without STAC

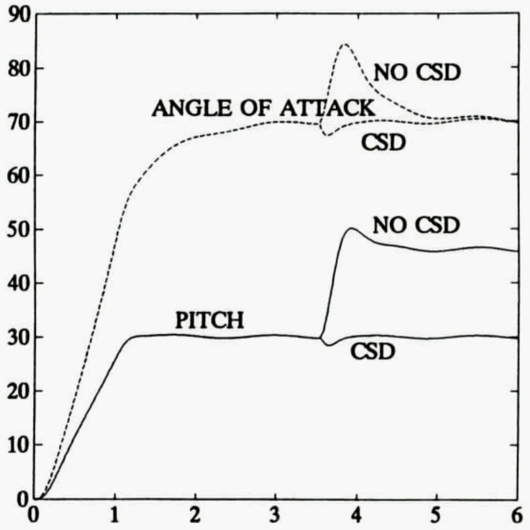


Figure: 8e: Impaired F-16, Loss of One Stabilator
With and Without CSD

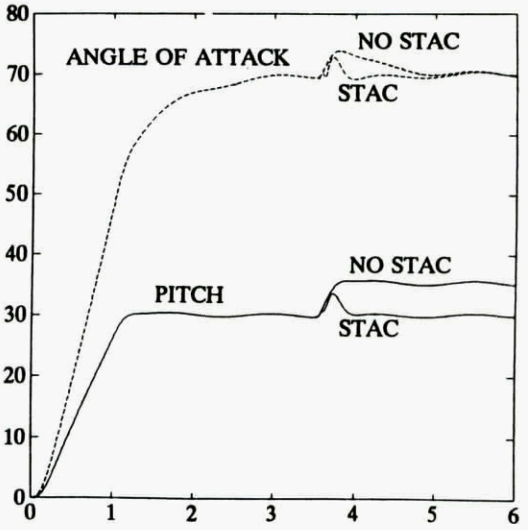
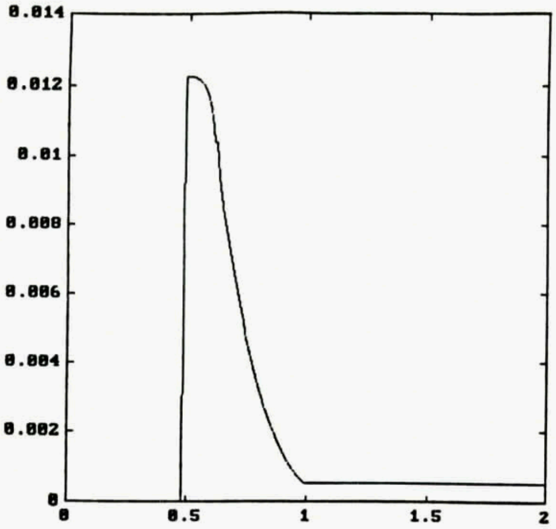


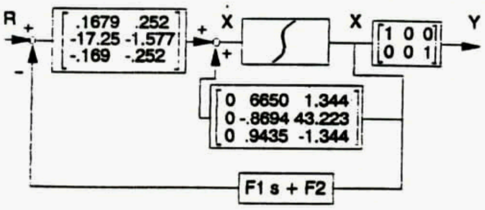
Figure 8f: Impaired F-16, 50% Loss of One Stab.
With and Without STAC



8g) Parameter Estimation Convergence in Euclidean Norm for Variation in B Matrix

F16 CONTROLLER:

Find F1 and F2 to assure 10% overshoot and one second settling time by estimating parameters of A and B matrices.



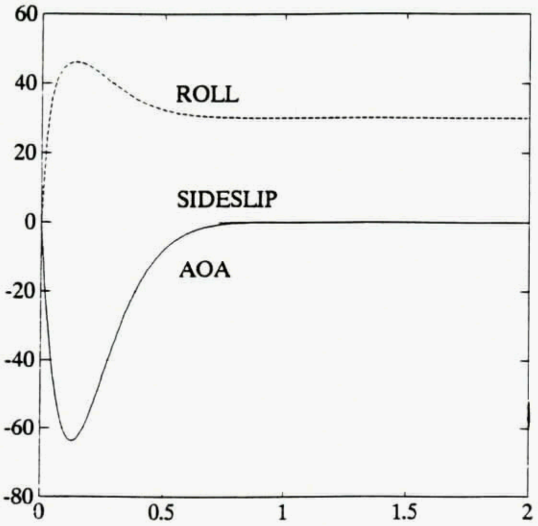
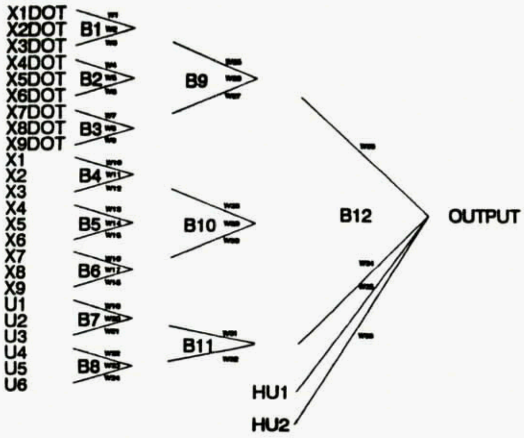
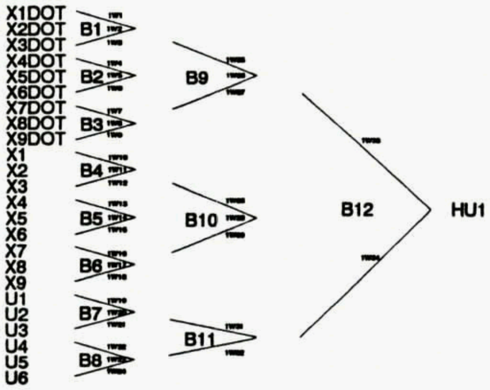


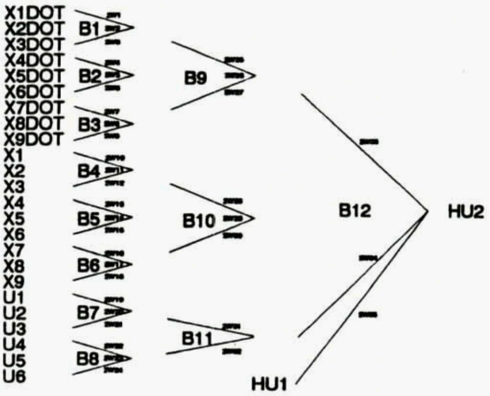
Figure 9: Linear F-16 Model and NNC Control



10a) NNC Requiring 54 Outputs for the HPA



10b) Hidden Unit 1 Training Architecture



10c) Hidden Unit 2 Training Architecture

Figure 10: Neural Net Control for the HPA

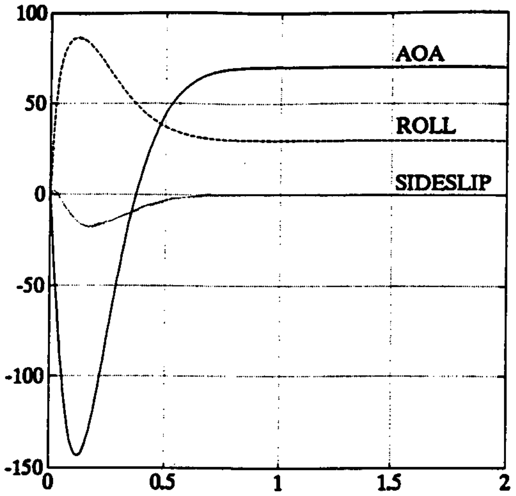


Figure 11a: Unimpaired HPA Decoupled Response
α=70°, ϕ=30°, β=0°, With NNC Only

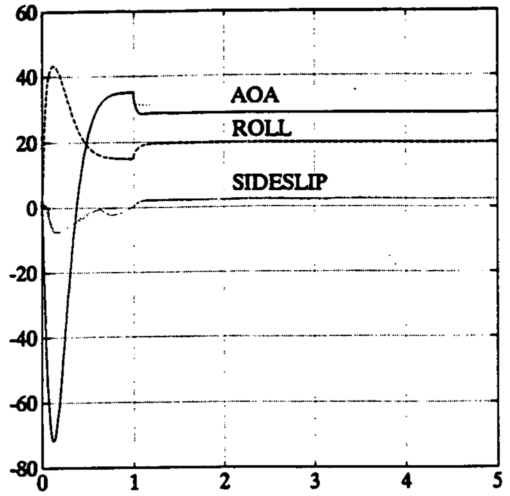


Figure 11d: Impaired HPA, 50% Loss of a Stabilator
α=35°, ϕ=15°, β=0°, With NNC Only

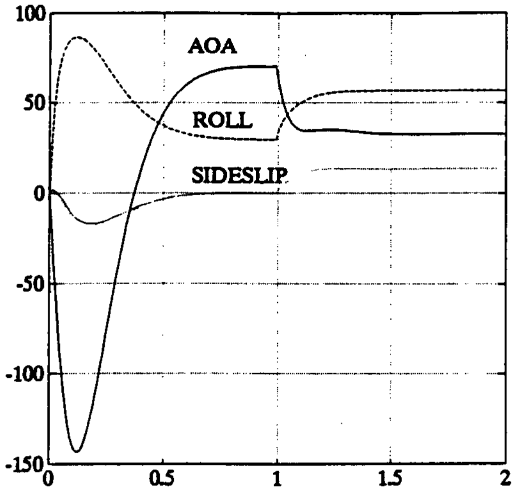


Figure 11b: Impaired HPA, Loss of a Stabilator
α=70°, ϕ=30°, β=0°, With NNC Only

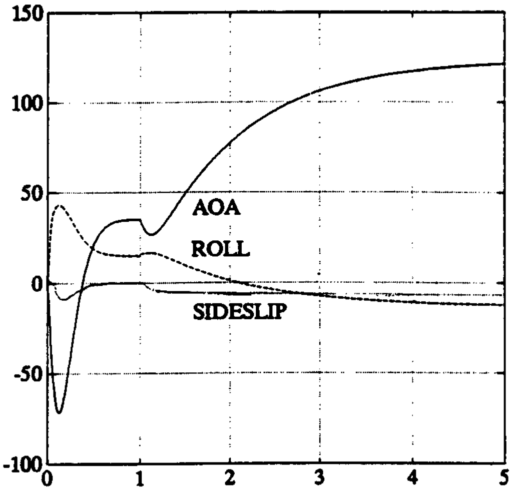


Figure 11e: Impaired HPA, 50% Loss of an Aileron
α=70°, ϕ=30°, β=0°, With NNC Only

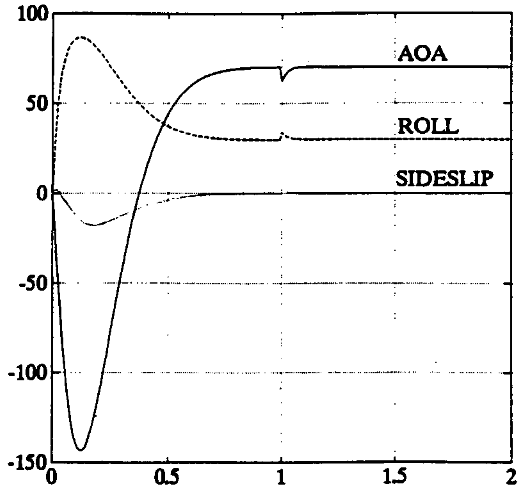


Figure 11c: Impaired HPA, Loss of a Stabilator
α=70°, ϕ=30°, β=0°, With NNC+CSD

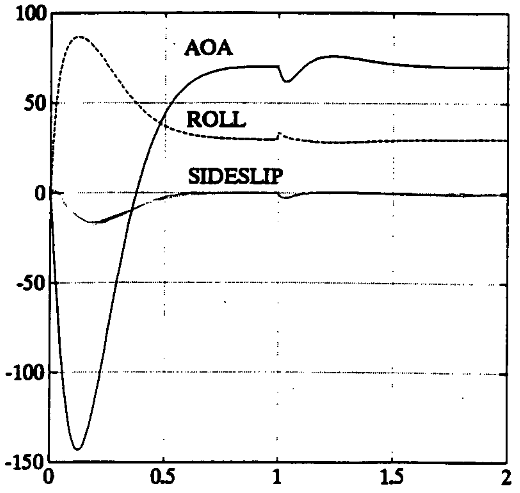


Figure 11f: Impaired HPA, Loss of an Aileron
α=70°, ϕ=30°, β=0°, With NNC+CSD

TOWARD INTELLIGENT FLIGHT CONTROL

by

Robert F. Stengel
Princeton University
Department of Mechanical and Aerospace Engineering
Princeton, NJ 08544
United States

ABSTRACT

Flight control systems can benefit by being designed to emulate functions of natural intelligence. Intelligent control functions fall in three categories: declarative, procedural, and reflexive. Declarative actions involve decision-making, providing models for system monitoring, goal planning, and system/scenario identification. Procedural actions concern skilled behavior and have parallels in guidance, navigation, and adaptation. Reflexive actions are more-or-less spontaneous and are similar to inner-loop control and estimation. Intelligent flight control systems will contain a hierarchy of expert systems, procedural algorithms, and computational neural networks, each expanding on prior functions to improve mission capability, to increase the reliability and safety of flight, and to ease pilot workload.

INTRODUCTION

Human pilots traditionally have provided the intelligence to fly manned aircraft in numerous ways, from applying manual dexterity through informed planning and coordination of missions. As aircraft characteristics have changed, and more importantly as the technology has allowed, an increasing share of the aircraft's intelligent operation has relied on proper functioning of electro-mechanical sensors, computers, and actuators. It has become possible to apply machine intelligence to flight control.

It can be argued that any degree of feedback from sensed motions to control actions instills intelligent behavior because control actions are shaped by knowledge of the system's response, though it was not always so. In contemplating the effects of atmospheric turbulence, one of the Wright brothers wrote, "The problem of overcoming these disturbances by automatic means has engaged the attention of many ingenious minds, but to my brother and myself, it has seemed preferable to depend entirely on *intelligent control*" [1, 2]. The Wright brothers' piloting actions depended on proper interpretation of visual and inertial cues, demonstrating biological intelligent control. Later, panel displays of compass heading, pressure altitude, airspeed, aircraft attitude, and bearing to a radio station enhanced the intelligent behavior of human pilots. Stability augmentation systems that fed pitch rate to elevator or yaw rate to rudder were among the first intelligent systems that did not rely on the human pilot, while automatic bombing and landing systems carried machine intelligence to the point of "hands-off" flying for small portions of the aircraft's mission.

In a contemporary context, intelligent flight control has come to represent even more ambitious plans to

- make aircraft less dependent on proper human actions for mission completion,
- enhance the mission capability of aircraft,
- improve performance by learning from experience,
- increase the reliability and safety of flight, and
- lower the cost and weight of aircraft systems.

This is not to say that earlier systems were not designed and implemented intelligently -- in fact, it required a high degree of human intelligence to squeeze the desired performance out of measurement-and-control devices whose reliability and capabilities were limited. Nevertheless, the march of technology has made it possible to design computer-based control systems that have fewer inherent limitations, that can adapt to changing physical and environmental characteristics, that are more reliable than earlier systems, and that can perform decision-making in addition to feedback control.

The goal of this paper is to present concepts for intelligent flight control in the contemporary context, that is, through the aid of what were once called "artificial" devices for sensing, computation, and control. Emphasis is placed on alternatives for analysis and design of control logic rather than on the equipment that makes it possible (i.e., on software rather than hardware). As in any complex subject, there are many ways to partition and describe intelligent control. The approach adopted here is to distinguish between control functions according to a cognitive/biological hierarchy that is bounded on one end by *declarative functions*, which typically involve decision-making, and on the other by *reflexive functions*, which are more-or-less spontaneous reactions to external or internal stimuli.

In a classical flight control context, declarative functions are performed by the control system's *outer loops*, and reflexive functions are performed by its *inner loops*. We may also define an intermediate level of *procedural functions*, which -- like reflexive functions -- have well-defined input-output characteristics

presented at the NATO-AGARD *Workshop on Stability in Aerospace Systems*, Toulouse, France, June 23-26, 1992.

but of a more complicated structure. Traditional design principles suggest that the outer-loop functions should be dedicated to low-bandwidth, large-amplitude control commands, while the inner-loop functions should have high bandwidths and relatively lower-amplitude actions. There is a logical progression from the sweeping, flexible alternatives associated with satisfying mission goals to more local concerns for stability and regulation about a desired path or equilibrium condition.

Intelligent control appeals to nature for useful design paradigms, for ideas that have been proven to work. In some respects, the field brings closure to parallel yet opposing directions in cognitive science and control theory. On the one hand, psychologists have looked to mathematics, engineering, and computer science for control-theoretic and statistical models that unify and explain observations of human behavior. On the other, control engineers want to design systems that deal with nonlinearity and time variability, that are robust in the face of uncertainty, and that exhibit some degree of autonomy -- that is, to give their control systems human traits. Both groups can benefit from knowledge gained in each other's fields.

FOUNDATIONS FOR INTELLIGENT FLIGHT CONTROL

Intelligent flight control design draws on two apparently unrelated bodies of knowledge. The first is rooted in classical analyses of aircraft stability, control, and flying qualities. The second derives from human psychology and physiology. The goal is to find new control structures that are consistent with the reasons for flying aircraft, that bring flight control systems to a higher level of overall capability.

Aircraft Flying Qualities and Flight Control

An aircraft requires guidance, navigation, and control so that it can perform its mission. As suggested by Fig. 1, a human pilot can interact with the aircraft at several levels, and his or her function may be supplanted by electro-mechanical equipment. The pilot performs three distinct functions: sensing, regulation, and decision-making. These three tasks exercise different human characteristics: the ability to see and feel, the ability to identify and correct errors between desired and actual states, and the ability to decide what needs to be done next. The first of these depends on the body's sensors and the neural networks that connect them to the brain. The second relies on motor functions enabled by the neuro-muscular system to execute learned associations between stimuli and desirable actions. The third requires more formal, introspective thought about the reasons for taking action, drawing on the brain's deep memory to recall important procedures or data. Sensing and regulation are high-bandwidth tasks with little time for deep thinking. Decision-making is a low-bandwidth task that requires concentration. Each of these tasks exacts a workload toll on the pilot.

Pilot workload has become a critical issue as the complexity of systems has grown, and furnishing ideal flying qualities throughout the flight envelope has become an imperative. It is particularly desirable to reduce the need to perform high-bandwidth, automatic functions, giving the pilot time to cope with unanticipated or unlikely events. In the future, teleoperated or autonomous systems could find increasing use for missions that expose human pilots to danger. Intelligent control may make it possible to remove the pilot from the fighter/attack/reconnaissance cockpit altogether, allowing a single operator to supervise the flight of one or more unmanned air vehicles, each capable of performing a complex mission.

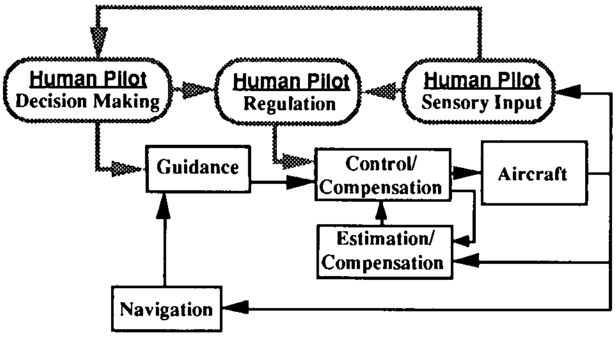


Figure 1. Guidance, Navigation, and Control Structure, Distinguishing Between Human-Pilot and Computer-Based Functions.

Research on the *flying (or handling) qualities of aircraft* has identified ways to make the pilot's job easier and more effective, and it provides models on which automatic systems might be based. The first flying qualities specification simply stated, "(the aircraft) must be steered in all directions without difficulty and all time (be) under perfect control and equilibrium" [3, 4]. Further evolution of flying qualities criteria based on dynamic modeling and control theory has resulted in the widely used U. S. military specification [5] and the succeeding military standard, described in [6].

Our immediate objective is to find control systems that are both automatic (or semi-automatic) and intelligent; we look to experimental studies of the pilot to learn what functions should be implemented. One fruitful result of flying qualities research has been the development of *control-theoretic models of piloting behavior*. Most of these models have dealt with reflexive, compensatory tracking tasks using simple time-lag and transfer function models [7, 8] or linear-quadratic-Gaussian (LQG) optimal-control models [9, 10]. Some of the transfer-function approaches go into considerable detail about neuro-muscular system dynamics [10, 12]. These models often show good correlation with experimental results, not only in compensatory tracking but in more procedural tasks: the progression of piloting actions from single- to multi-input strategies as the complexity of the task increases is predicted in [11], while test-pilot opinion ratings are predicted by a "Paper Pilot" in [13]. These results imply that *computer-based control laws can perform procedural and reflexive tasks within the fit error of mathematical human-pilot models.* Models of the human pilot's declarative actions have yet to receive the same level of attention; however [14-16] introduce the types of decisions that must be made in aerospace scenarios, as well as likely formats for pilot-vehicle interface.

Figure 1 also portrays a hierarchical structure for stability-augmentation, command-augmentation, autopilot, and flight-management-system functions that can be broken into reflexive and declarative parts. Stability augmentation is reflexive control provided by the inner-most loop, typically implemented as a linear feedback control law that provides stability and improves transient response through an *Estimation/Compensation* block. Forward-loop control provides the shaping of inputs for satisfactory command response through a *Control/Compensation* block, again employing linear models. The combination of control and estimation can be used to change the flying qualities perceived by the pilot, or it can provide a decoupled system for simplified guidance commands [17-20]. A basic autopilot merely translates the human pilot's commands to guidance commands for constant heading angle, bank angle, or airspeed, while the *Guidance* block can be expanded to include declarative flight management functions, using inputs from *Navigation* sensors and algorithms.

Intelligent functions have been added to flight control systems in the past. Many stability and command augmentation systems have employed gain scheduling and switching for improved performance in differing flight regimes and mission phases. Navigation and flight management functions usually involve significant nonlinearity related to the geometry of measurements and the earth, as well as to pressure and temperature gradients in the atmosphere. Control theory, heuristics, and reduced-order optimization have been used to achieve near-optimal trajectory management in many flight phases (e.g., [21-23]).

The Guidance, Navigation, and Control (GNC) Systems for Project Apollo's Command/Service and Lunar Modules provide an early example of intelligent aerospace control [24-26]. These two GNC Systems adapted mathematical structures and parameters to changing flight conditions, employed nonlinear control laws, and made low-level decisions. They used a computer that was orders of magnitude slower than today's flight computers and that had two programming choices: machine (assembly) code or interpretive code. The state-of-the-art of aircraft flight control systems has progressed to comparable levels and beyond, as represented by systems installed in modern transport and fighter aircraft (e.g., [27, 28]).

Intelligent flight control[1] can be justified only if it materially improves the functions of aircraft, if it saves the time and/or money required to complete a mission, or if it improves the safety and reliability of the system. A number of interesting philosophical problems can be posed. Must machine-intelligence be better, in some sense, than the human intelligence it replaces in order for it to be adopted? We are willing to accept the likelihood that humans will make mistakes; if a machine has the same likelihood of making a mistake, should it be used? Lacking firm knowledge of a situation, humans sometimes gamble; should intelligent machines be allowed to gamble? When is it acceptable for machine intelligence to be wrong (e.g., during learning)? Must the machine solution be "optimal," or is "feasible" good enough? Which decisions can the machine make without human supervision, and which require human intervention? In a related vein, how much information should be displayed to the human operator? Should intelligent flight control ever be fully autonomous? If the control system adapts, how quickly must it adapt? Must learning occur on-line, or can it be delayed until a mission is completed? All of these questions must be answered in every potential application of intelligent control.

Cognitive and Biological Paradigms for Intelligence

Intelligence is the "ability involved in calculating, reasoning, perceiving relationships and analogies, learning quickly, storing and retrieving information .... classifying, generalizing, and adjusting to new situations" [29]. This definition does not deal with the mechanisms by which intelligence is realized, and it makes the tacit assumption that intelligence is a human trait. Intelligence relates not only to intellectuality and cognition but to personality and the environment [30].

The debate over whether-or-not computers ever will "think" may never be resolved, though this need not restrict our working models for computer-based intelligent control. One argument against the proposition is that computers deal with syntax (form), while minds deal with semantics (meaning), and syntax alone cannot produce semantics [31]. This does not, of course, limit the ability of a computer to <u>mimic</u> natural intelligence in a limited domain. Another contention is that thinking is "non-algorithmic," that the brain evokes consciousness through a process of natural selection and inheritance [32]. Consciousness is required for common sense, judgment of truth, understanding, and artistic appraisal, concepts that are not formal and cannot readily be programmed for a computer (i.e., they are not syntactic).

Conversely, functions that are automatic or "mindless" (i.e., that are <u>un</u>conscious) could be programmed, implying that computers have more in common with "unintelligent" functions.

---

[1] As used here "intelligent flight control" subsumes "intelligent guidance, navigation, and control."

Gödel's Theorem[2] is offered in [33] as an example of an accepted proposition that may be considered non-algorithmic; the statement and proof of the theorem must themselves be non-algorithmic and, therefore, not computable. However, while the human curiosity, intuition, and creativity that led to Gödel's Theorem may not be replicable in a computer, the statement and proof are expressed in a formal way, so they might be considered algorithmic after all.

The notion that syntax alone cannot produce semantics is attacked as being an axiom that is perhaps true but not knowable in any practical sense [34]; therefore, the possibility that a computer can "think" is not ruled out. A further defense is offered in [35], which suggests that human inference may be based, in part, on inconsistent axioms. This could lead to rule-based decisions that are not logically consistent, that are affected by heuristic biases or sensitivities, that may reflect deeper wisdom, or that may be wrong or contradictory. For example, knowledge and belief may be indistinguishable in conscious thought; however, one implies truth and the other bias or uncertainty. One might also postulate the use of meta-rule bases that govern apparently non-algorithmic behavior. The process of searching a data base, though bound by explicit symbolic or numerical algorithms, may well produce results not immediately identifiable as algorithmic.

More to our point, it is likely that a computer capable of passing a flying-qualities/pilot-workload/control-theoretic equivalent of the Turing test[3] [36] could be built even though that computer might not understand what it is doing[4]. For intelligent flight control, the principal objective is improved control performance, that is, for improved input-output behavior. The computer can achieve the operative equivalent of consciousness on its own terms and in a limited domain, even if it does not possess the veiled emotions of the computer *HAL* in the movie *2001* or the apparent concern for human problems of the pseudo-psychoanalytic computer program *Eliza.*

Discussions of human consciousness naturally fall into using the terminology of computer science. From an information-processing perspective, it is convenient -- as well as consistent with empirical data -- to identify four types of thought: conscious, preconscious, subconscious, and unconscious [37]. *Conscious thought* is the thought that occupies our attention, that requires focus, awareness, reflection, and perhaps some rehearsal. Conscious thought performs declarative processing of the individual's knowledge or beliefs. It makes language, emotion, artistry, and philosophy possible. *Unconscious thought* "describes those products of the perceptual system that go unattended or unrehearsed, and those memories that are lost from primary memory through display or displacement" [37]. Within the unconscious, we may further identify two important components. *Subconscious thought* is procedural knowledge that is below our level of awareness but central to the implementation of intelligent behavior. It facilitates communication with the outside world and with other parts of the body, providing the principal home for the learned skills of art, athletics, control of objects, and craft. We are aware of perceptions if they are brought to consciousness, but they also may take a subliminal (subconscious) path to memory. *Preconscious thought* is pre-attentive declarative processing that helps choose the objects of our conscious thought, operating on larger chunks of information or at a more symbolic level. It forms a channel to long-term and implicit memory, and it may play a role in judgment and intuition.

Whether we adopt a single-processor model of consciousness such as Adaptive Control of Thought (ACT* as in [38]) or a connectionist model like Parallel Distributed Processing (PDP from [39]), we are led to believe that the central nervous system supports a hierarchy of intelligent and automatic functions with *declarative actions* at the top, *procedural actions* in the middle, and *reflexive actions* at the bottom. We may assume that declarative thinking occurs in the brain's cerebral cortex, which accesses the interior limbic system for memory [40]. Together, they provide the processing unit for conscious thought. Regions of the cerebral cortex are associated with different intellectual and physical functions; the distinction between conscious and preconscious function may depend on the activation level and duration in regions of the cerebral cortex.

The working memory of conscious thought has access to the spinal cord through other brain parts that are capable of taking procedural action (e.g., the brain stem for autonomic functions, the occipital lobes for vision, and the cerebellum for movement). Procedural action can be associated with subconscious thought, which supports voluntary automatic processes like movement and sensing. These voluntary signals are sent over the somatic nervous system, transmitting to muscles through the motor neural system and from receptors through the sensory neural system.

The spinal cord itself "closes the control loop" for reflexive actions long before signals could be processed by the brain. Nevertheless, these signals are available to the brain for procedural and declarative processing. We are all aware of performing some task (e.g., skating or riding a bicycle) without effort, only to waver when we focus on what we are doing. Involuntary regulation of the body's organs is provided by the autonomic nervous system, which is subject to unconscious processing by the

---

[2] As summarized in [32]: Any algorithm used to establish a mathematical truth cannot prove the propositions on which it is based. Or another [33]: Logical systems have to be fixed up "by calling the undecidable statements axioms and thereby declaring them to be true," causing new undecidable statements to crop up.

[3] Turing suggested that a computer could be considered "intelligent" if it could "converse" with a human in a manner that is indistinguishable from a human conversing with a human.

[4] Searle describes such a computer as a "Chinese Room" that translates Chinese characters correctly by following rules while not understanding the language in [31].

brain stem. "Bio-feedback" can be learned, allowing a modest degree of higher-level control over some autonomic functions.

Declarative, procedural, and reflexive functions can be built into a model of intelligent control behavior (Fig. 2). The *Conscious Thought* module governs the system by performing declarative functions, receiving information and transmitting commands through the *Subconscious Thought* module, which is itself capable of performing procedural actions. Conscious Thought is primed by *Preconscious Thought* [41], which can perform symbolic declarative functions and is alerted to pending tasks by Subconscious Thought. These three modules overlie a bed of deeper *Unconscious Thought* that contains long-term memory. They are capable of intellectual learning, and while their physical manifestation may be like the PDP model, they exhibit characteristics that are most readily expressed by the ACT* model[5].

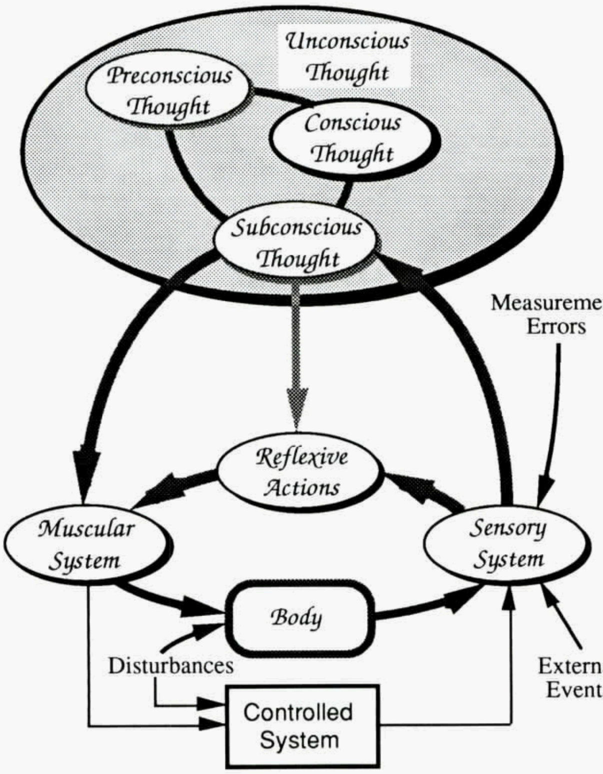


Figure 2. A Model of Cognitive/Biological Control Behavior.

The Subconscious Thought module receives information from the *Sensory System* and conveys commands to the *Muscular System* through peripheral networks. Voluntary *Reflexive Actions* provide low-level regulation in parallel with the high-level functions, responding to critical stimuli and coordinating control actions. High- and low-level commands may act in concert, or one may block the other. Voluntary Reflexive Actions can be trained by high-level directives from Subconscious Thought, while the learning capabilities of involuntary Reflexive Action are less clear. Control actions produce *Body* motion and can affect an external *Controlled System*, as in piloting an aircraft. In learned control functions, Body motion helps internalize the mental model of Controlled System behavior. The Body and the Controlled System are both directly or indirectly subjected to *Disturbances*; for example, turbulence would affect the aircraft directly and the pilot indirectly. The Sensory System observes *External Events* as well as Body and Controlled System motions, and it is subject to *Measurement Errors*.

There are many parallels and analogies to be drawn in comparing the functions of human and computer-based intelligence -- in fact, too many to detail here. However, it may be useful to ponder a few, especially those related to knowledge acquisition, natural behavior, aging, and control. Perhaps the most important observation is that *learning requires error or incompleteness*. There is nothing new to be gained by observing a process that is operating perfectly. In a control context, any operation should be started using the best available knowledge of the process and the most complete control resources. Consequently, learning is not always necessary or even desirable in a flight control system. *Biological adaptation is a slow process*, and proper changes in behavior can be made only if there is prior knowledge of alternatives. If adaptation occurs too quickly, there is the danger that misperceptions or disturbance effects will be misinterpreted as parametric effects. *Rest is an essential feature of intelligent biological systems*. It has been conjectured that *REM (rapid-eye-movement) Sleep* is a time of learning, consolidating, and pruning knowledge[6] [42]. Systems can learn even when they are not functioning by reviewing past performance, perhaps in a repetitive or episodic way.

The cells of biological systems undergo a continuing birth-life-death process, with new cells replacing old; nature provides a means of transmitting genetic codes from cell to cell. Nevertheless, *the central nervous system is incapable of functional regeneration*. Once a portion of the system has been damaged, it cannot be replaced, although redundant neural circuitry can work around some injuries. *Short-term memory often recedes into long-term memory*, where it generally takes longer to be retrieved. With time, *items in memory that are less important are forgotten, possibly replaced by more important information*; hence, information has a half-life that depends upon its significance to our lives (and perhaps to its "refresh rate"). *Humans de-*

---

[5] ... although the actual processing mechanism is not clear. In a recent seminar at Princeton (March 9, 1992), Herbert Simon noted that if you open the cabinet containing a sequential-processing computer, the innards look very much like those of a parallel processor.

[6] "In REM Sleep, the brain is barraged by signals from the brain stem. Impulses fired to the visual cortex produce images that may contain materials from the day's experiences, unsolved problems, and unfinished business." [42]

velop the capability to form chords of actions that are orchestrated or coordinated to achieve a single goal. Response to an automotive emergency may include applying the brakes, disengaging the clutch, steering to avoid an obstacle, and bracing for impact all at once. *We develop "knee-jerk" reactions that combine declarative, procedural, and reflexive functions*, like clapping after the last movement of a symphony.

Nature also provides structural paradigms for control that are worth emulating in machines. First, there is a *richness of sensory information* that is hard to fathom, with millions of sensors providing information to the system. This results in high signal-to-noise ratio in some cases, and it allows symbolic/image processing in others. Those signals requiring high-bandwidth, high-resolution channel capacity (vision, sound, and balance) have *short, dedicated, parallel runs* from the sensors to the brain. This enhances the security of the channels and protects the signals from noise contamination. *Dissimilar but related sensory inputs facilitate interpretation of data.* A single motion can be sensed by the eyes, by the inner ear, and by the "seat-of-the-pants" (i.e., by sensing forces on the body itself), corroborating each other and suggesting appropriate actions. When these signals are made to disagree in moving-cockpit simulation of flight, a pilot may experience a sense of confusion and disorientation.

There are *hierarchical and redundant structures* throughout the body. The nervous system is a prime example, bringing inputs from myriad sensors (both similar and dissimilar) to the brain, and performing low-level reasoning as an adjunct. Many sensing organs occur in pairs (e.g., eyes, ears, inner ears), and their internal structures are highly parallel. *Pairing allows graceful degradation in the event that an organ is lost.* Stereo vision vanishes with the loss of an eye, but the remaining eye can provide both foveal and peripheral vision, as well as a degree of depth perception through object size and stadiametric processing. Our control effectors (arms, hands, legs, feet) also occur in pairs, and there is an element of "Fail-Op/Fail-Op/Fail-Safe" design [43] in the number of fingers provided for manual dexterity.

Structure for Intelligent Flight Control

The preceding section leads to a control system structure that overlays the cognitive/biological model of Fig. 2 on the flight control block diagram of Fig. 1 and adds new functions. The suggested structure (Fig. 3) has *super-blocks* identifying declarative, procedural, and reflexive functions; these contain the classical GNC functions plus new functions related to decision-making, prediction, and learning. The black arrows denote information flow for the primary GNC functions, while the gray arrows illustrate the data flow that supports subsidiary adjustment of goals, rules, and laws.

Within the super-blocks, higher-level functions are identified as conscious, preconscious, and subconscious attributes, not with disregard for the philosophical objections raised earlier but as a working analog for establishing a computational hierarchy. The new functions relate to setting or revising goals for the aircraft's mission, monitoring and adjusting the aircraft's systems and subsystems, identifying changing characteristics of the aircraft and its environment, and applying this knowledge to modify the structures and parameters of GNC functions.

The suggested structure has implications for both hardware and software. Declarative functions are most readily identified with single-processor computers programmed in LISP or Prolog, as decision-making is associated with list processing and the statement of logical relationships. Procedural functions can be conceptualized as vector or "pipelined" processes programmed in FORTRAN, Pascal, or C, languages that have been developed for numerical computation with subroutines, arrays, differential equations, and recursions. Reflexive functions seem best modeled as highly parallel processes implemented by neural networks, which apply dense mappings to large masses of data almost instantaneously. Nevertheless, parallel processes can be implemented using sequential processors, and sequential algorithms can be "parallelized" for execution on parallel processors. The choice of hardware and software depends as much on the current state-of-the-art as on the closeness of computational requirements and GNC functions.

In the remainder of the paper, declarative, procedural, and reflexive control functions are discussed from an aerospace perspective. In practice, the boundaries between mission tasks may not be well defined, and there is overlap in the kinds of algorithms that might be applied within each group. A number of practical issues related to human factors, system management, certifiability, maintenance, and logistics are critical to the successful implementation of intelligent flight control, but they are not treated here.

DECLARATIVE SYSTEMS

Goal planning, system monitoring, and control-mode switching are declarative functions that require reasoning. Alternatives must be evaluated, and decisions must be made through a process of *deduction*, that is, by inferring answers from general or domain-specific principles. The inverse process of learning principles from examples is *induction*, and not all declarative systems have this capability. Most declarative systems have fixed structure and parameters, with knowledge induced off-line and before application; declarative systems that learn on-line must possess a higher level of reasoning ability, perhaps through an internal declarative module that specializes in training.

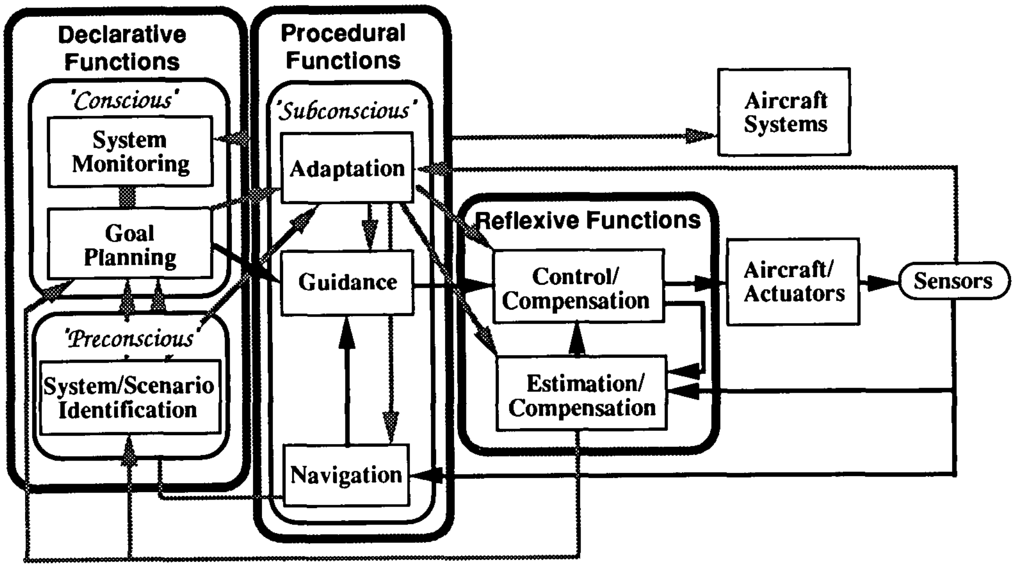


Figure 3. Intelligent Flight Control System Structure.

Expert Systems

Expert Systems are computer programs that use physical or heuristic relationships and facts for interpretation, diagnosis, monitoring, prediction, planning, and design. In principal, an expert system replicates the decision-making process of one or more experts who understand the causal or structural nature of the problem [44]. While human experts may employ "nonmonotonic reasoning" and "common sense" to deduce facts that apparently defy simple logic, computational expert systems typically are formal and consistent, basing their conclusions on analogous cases or well-defined rules[7].

A *rule-based expert system* consists of *data*, *rules*, and an *inference engine* [46]. It generates actions predicated on its data base, which contains measurements as well as stored data or operator inputs. An expert system performs deduction using *knowledge* and *beliefs* expressed as parameters and rules. *Parameters* have values that either are external to the expert system or are set by rules. An "IF-THEN" *rule* evaluates a *premise* by testing values of one or more parameters related by logical "ANDs" or "ORs," as appropriate, and it specifies an *action* that set values of one or more parameters.

The rule base contains all the cause-and-effect relationships of the expert system, and the inference engine performs its function by searching the rule base. Given a set of premises (evidence of the current state), the logical outcome of these premises is found by a data-driven search (*forward chaining*) through the rules. Given a desired or unknown parameter value, the premises needed to support the fixed or free value are identified by a goal-directed search (*backward chaining*) through the rules. Querying (or firing) a rule when searching in either direction may invoke procedures that produce parameter values through *side effects*.

Both search directions are used in rule-based control systems [47]. Backward chaining drives the entire process by demanding that a parameter such as *CONTROL CYCLE COMPLETED* have a value of *true*. The inference engine works back through the rules to identify other parameters that allow this and, where necessary, triggers side effects (procedural or reflexive functions) to set those parameters to the needed values. Backward chaining also is invoked to learn the value of *ABNORMAL BEHAVIOR DETECTED*, be it *true* or *false*. Conversely, forward chaining indicates what actions can be taken as a consequence of the current state. If *SENSOR MEASUREMENTS REASONABLE* is *true*, and *ALARM DETECTED* is *false*, then failure identification and reconfiguration side effects can be skipped on the current cycle.

Rules and parameters can be represented as *objects* or *frames* using ordered lists that identify names and attributes. Specific rules and parameters are represented by lists in which values are given to the names and attributes. The attribute lists contain not only values and logic but additional information for the inference engine. This information can be used to compile *parameter-rule-association lists* that speed execution [48]. Frames provide useful parameter structures for related productions, such as analyzing the origin of one or more failures in a complex, connected system [49]. Frames possess an *inheritance property*; thus a particular object lays claim to the properties of the object type.

---

[7] Expert systems can have *tree* or *graph* structures. In the former, there is a single *root* node, and all final (*leaf*) nodes are connected to their own single *branch*. In the latter, one or more branches lead to individual nodes. Reasoning is *consistent* if an individual node is not assigned differing values by different branches [45].

Crew/Team Paradigms for Declarative Flight Control

Logical task-classification is a key factor in the development of rule-based systems. To this point, we have focused on the intelligence of an individual as a paradigm for control system design, but it is useful to consider the hypothetical actions of a multi-person aircraft crew as well. In the process, we develop an expert system of expert systems, a hierarchical structure that reasons and communicates like a team of cooperating, well-trained people might. This notion is suggested in [50] and is carried to considerable detail in [51-53]. The Pilot's Associate Program initially focused on a four-task structure and evolved in the direction of the multiple crew-member paradigm [54-56].

AUTOCREW is an ensemble of nine cooperating rule-based systems, each figuratively emulating a member of a World War II bomber crew: executive (pilot), co-pilot, navigator, flight engineer, communicator, spoofer (countermeasures), observer, attacker, and defender (Fig. 4) [53]. The executive coordinates mission-specific tasks and has knowledge of the mission plan. The aircraft's human pilot can monitor AUTOCREW functions, follow its suggestions, enter queries, and assume full control if confidence is lost in the automated solution. The overall goal is to reduce the pilot's need to regulate the system directly without removing discretionary options. AUTOCREW contains over 500 parameters and over 400 rules.

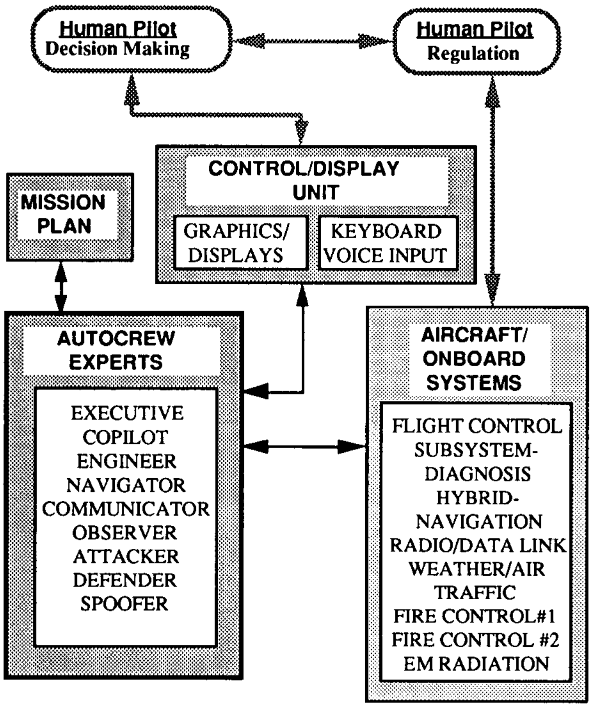


Figure 4. AUTOCREW Configuration with Pilot/Aircraft Interface (adapted from [52]).

AUTOCREW was developed by defining each member expert system as a *knowledge base*, according to the following principles:

- Divide each knowledge base into major task groups: time-critical, routine, and mission-specific.
- Order the task groups from most to least time-critical to quicken the inference engine's search.
- Break major tasks into sub-tasks according to the detail necessary for communicating system functions.
- Identify areas of cooperation between knowledge bases.

The five main task groups for each crew member are: tasks executed during attack on the aircraft, tasks executed during emergency or potential threat, tasks ordered by the EXECUTIVE, tasks executed on a routine basis, and mission-specific tasks. Routine and mission-specific tasks are executed on each cycle; emergency tasks are executed only when the situation warrants.

Operation of AUTOCREW was simulated to obtain comparative expert-system workloads for two mission scenarios: inbound surface-to-air missile attack and human pilot incapacitation [52]. Results are presented for five mission phases and three emergency conditions in Fig. 5. The *Rule Fraction* is the ratio of number of rules fired during the mission phase for the specified AUTOCREW member to the total number of mission phase rules in all AUTOCREW knowledge bases. The *Parameter Fraction* is the ratio of number of tasks performed during the mission phase for the specified AUTOCREW member to the total number of mission phase tasks in all knowledge bases. Such comparisons are helpful in allocating computer resources to component functions. Additional perspectives on intelligent flight management functions can be obtained from the literature on decision-making by teams, as in [57-59]. Alternate approaches to aiding the pilot in emergencies are given in [60, 61].

In addition to the overall AUTOCREW system, a functioning NAVIGATOR sensor-management expert system was developed. As shown in a later section, knowledge acquisition for the system presents an interesting challenge, because traditional methods (e.g., domain-expert interviews) do not provide sufficiently detailed information to design the system [62].

Reasoning Under Uncertainty

Rule-based control systems must make decisions under uncertainty. Measurements are noisy, physical systems are subject to random disturbances, and the environment within which decisions must be made is ambiguous. For procedural systems, the formalism of optimal state estimation provides a rigorous and useful means of handling uncertainty [63]. For declarative systems, there are a number of methods of uncertainty management, including probability theory, Dempster-Shafer theory, possibility theory (fuzzy logic), certainty factors, and the theory of endorsements [64].

Bayesian belief networks [65], which propagate event probabilities up and down a causal tree using Bayes's rule, have particular appeal for intelligent control applications because they deal with probabilities, which form the basis for stochastic optimal control. We have applied Bayesian networks to aiding a pilot who may be flying in the vicinity of hazardous wind shear [66]. Figure 6 shows a network of the causal relationships among meteorological phenomena associated with microburst wind shear, as well as temporal and spatial information that could affect the likelihood of microburst activity. A probability of occurrence is associated with each node, and a conditional probability based on empirical data is assigned to each arrow. The probability of encountering microburst wind shear is the principal concern; however, each time new evidence of a particular phenomenon is obtained, probabilities are updated throughout the entire tree. In the process, the estimated likelihood of actually encountering the hazardous wind condition on the plane's flight path is refined. Unlike other applications of hypothesis testing, the threshold for advising a go-around during landing or an abort prior to takeoff is a very low probability -- typically 0.01 or less -- as the consequences of actually encountering a strong microburst are severe and quite possibly fatal.

The safety of aircraft operations near microburst wind shear will be materially improved by forward-looking Doppler radar, which can sense variations in the wind speed. Procedural functions that can improve the reliability of the wind shear expert system include extended Kalman filtering of the velocity measurements at incremental ranges ahead of the aircraft [67].

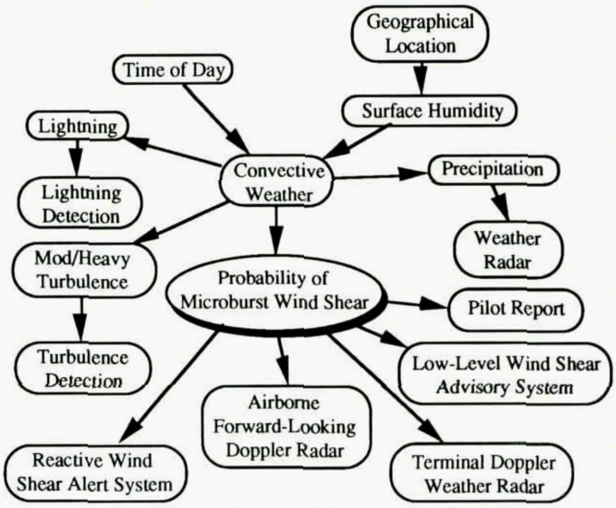


Figure 6. Bayesian Belief Network to Aid Wind Shear Avoidance (adapted from [67]).

Figure 7 shows the inevitable tradeoff between the probability of missed detection ($\mathbb{P}_{MD}$) and the probability of false warning ($\mathbb{P}_{FW}$) in processing such measurements. Monte Carlo simulations are used to define the detection statistics. In the present application, the penalty for false warning is less than for missed detection; however, too many false warnings could cause the crew to lose confidence in the expert system and would have a negative effect on transport operations, so the detection threshold must be chosen carefully.

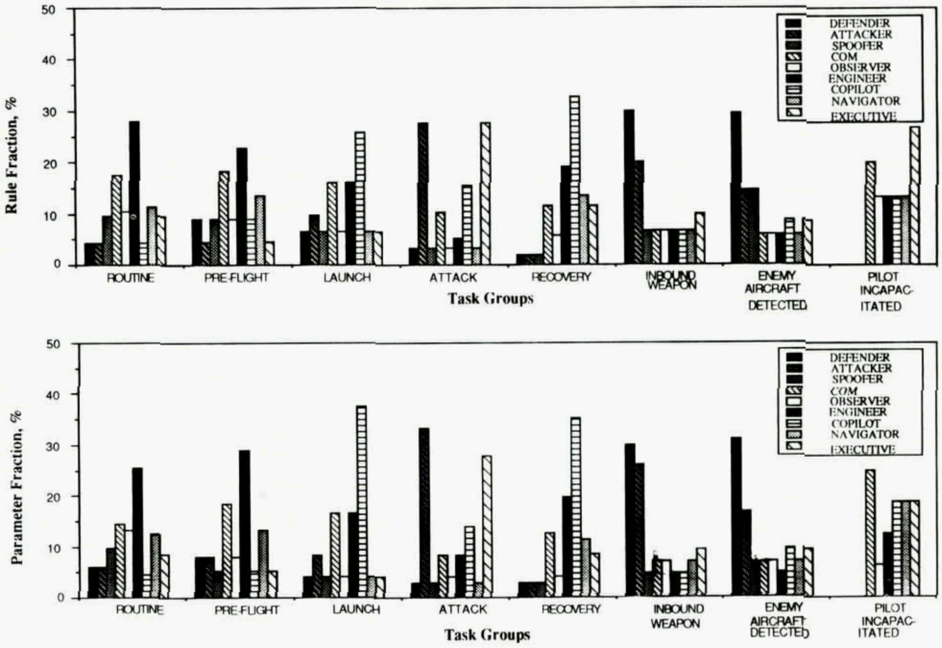


Figure 5. AUTOCREW Relative Workload for Mission-Specific and Emergency Scenarios [52].

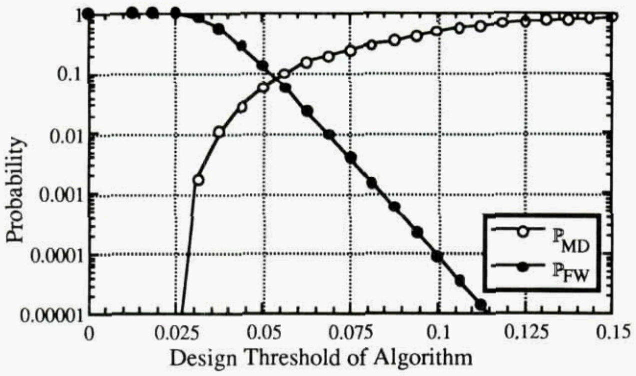


Figure 7. Statistical Analysis of Detection Probability [67].

Probabilistic reasoning of a different sort has been applied to a problem in automotive guidance that may have application in future Intelligent Vehicle/Highway Systems [68-70]. Intelligent guidance for headway and lane control on a highway with surrounding traffic is based on *worst-plausible-case decision-making*. It is assumed that the intelligent automobile (IA) has imaging capability as well as on-board motion sensors; hence, it can deduce the speed and position of neighboring automobiles. Each automobile is modeled as a simple discrete-time dynamic system, and estimates of vehicle states are propagated using extended Kalman filters [63]. There are limits on the performance capabilities of all vehicles, and IA strategy is developed using time-to-collide, braking ratios, driver aggressiveness, and desired security factors. Plausible guidance commands are formulated by minimizing a cost function based on these factors [70]. A general layout of the logic is shown in Fig. 8, and a partial decision tree for lateral guidance is presented as Fig. 9. Both normal and emergency expert systems govern the process, supported by procedural calculations for situation assessment, traffic prediction, estimation, and control. Guidance commands are formulated by minimizing a cost function based on these factors [70].

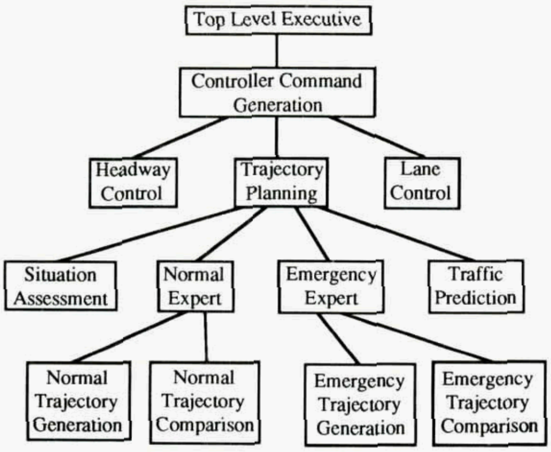


Figure 8. Intelligent Guidance for Automotive Headway and Lane Control [69].

Alternate plausible strategies for each neighboring automobile are extrapolated, with predictions of both means and covariances. Expected values of *plausibility, belief interval, and hazard functions* are calculated, scores for feasible IA actions are computed, and the best course of action is decided, subject to aggressiveness and security factors, as suggested by Fig. 9. Deterministic and Monte Carlo simulations are conducted to demonstrate system performance and to fine-tune logical parameters.

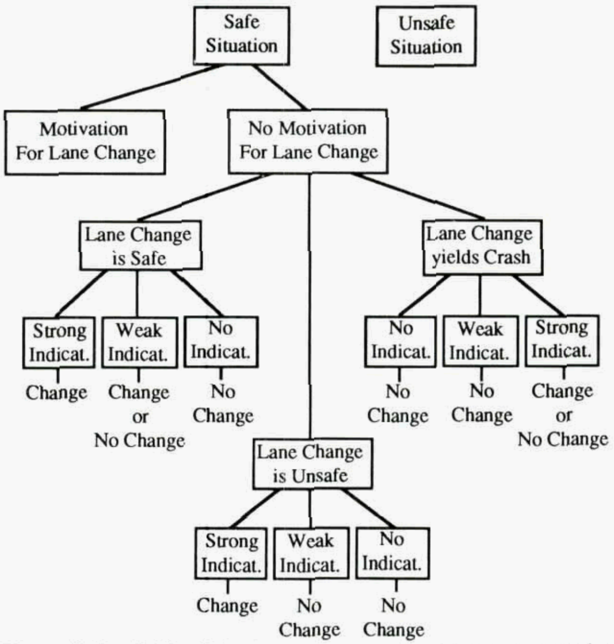


Figure 9. Partial Decision Tree used to Model Lateral Behavior in Intelligent Automotive Control [69].

Each of the expert systems discussed in this section performs deduction in a cyclical fashion based on prior logical structures, prior knowledge of parameters, and real-time measurements. It is clear that intelligent flight control systems must deal with unanticipated events, but it is difficult to identify aeronautical applications where on-line declarative learning is desirable. Nevertheless, off-line induction is needed to formulate the initial declarative system and perhaps (in a manner reminiscent of REM Sleep) to upgrade declarative logic between missions.

Inducing Knowledge in Declarative Systems

In common usage, "learning" may refer a) to collecting inputs and deducing outputs and b) to inducing the logic that relates inputs and outputs to specific tasks. Here, we view the first process as the normal function of the intelligent system and the second as "learning." Teaching an expert system the rules and parameters that generalize the decision-making process from specific knowledge is the inverse of expert-system operation. Given all possible values of the parameters, what are the rules that connect them? Perhaps the most common answer is to interview experts in an attempt to capture the logic that they use, or failing that, to study the problem intensely so that one becomes expert enough to identify naturally intelligent solutions. These approaches can be formalized [71, 72], and they were the ones used in [67] and [68]. Overviews of alternatives for induction can be found in [45, 46, 73, 74].

Two approaches are considered in greater detail. The first is called *rule recruitment* [75], and it involves the manipulation of "dormant rules" (or *rule templates*). This method was applied in the development of an intelligent failure-tolerant control system for a helicopter. Each template possesses a fixed premise-action structure and refers to parameters through *pointers*. Rules are constructed and incorporated in the rule base by defining links and modifying parameter-rule-association lists. Learning is based on Monte Carlo simulations of the controlled system with alternate failure scenarios. Learned parameter values then can be defined as "fuzzy functions" [76] contained in rule premises. For example, a typical rule reads, "IF Indicator 1 is near A and Indicator 2 is near B, THEN there is a good chance that Forward Collective Pitch Control is biased from nominal by an amount near C and that Failure Detection Delay is near D."

The second approach is to *construct classification or decision trees* that relate attributes in the data to decision classes [52]. The problem is to develop an Expert Navigation-Sensor Management System (NSM) that selects the best navigation aids from available measurements. Several aircraft paths were simulated, and the corresponding measurements that would have been made by GPS, Loran, Tacan, VOR, DME, Doppler radar, air data, and inertial sensors were calculated, with representative noise added. The simulated measurements were processed by extended Kalman filters to obtain optimal state estimates in 200 simulations. Using the root-sum-square error as a decision metric, Analysis of Variance (ANOVA) identifies the factors that make statistically significant contributions to the decision metric, and the Iterative Dichotomizer #3 (ID3) algorithm [77-79] extracts rules from the training set by inductive inference. The ID3 algorithm quantifies the *entropy content* of each attribute, that is, the information gained by testing the attribute at a given decision node. It uses an information-theoretic measure to find a splitting strategy that minimizes the number of nodes required to characterize the tree. Over 900 examples were used to develop the NSM decision tree.

NSM performance was assessed at nearly 500 points on two trajectories that differed from the training set. NSM correctly assessed the error class for each navaid type (±1 error class) most of the time (see Fig. 5 of [52]), and differences between NSM and optimal navigation solutions were found to be minimal.

PROCEDURAL SYSTEMS

Most guidance, navigation, and control systems fielded to date are procedural systems using sequential algorithms on sequential processors. Although optimality of a cost function is not always a necessary or even sufficient condition for a "good" system, linear-optimal stochastic controllers provide a good generic structure for discussion. They are presented in state-space form, they contain separate control and estimation functions, and they provide an unusual degree of design flexibility. The optimal regulator effectively produces an approximate stable inverse in providing desired response. The nonlinear-inverse-dynamic controller is a suitable design alternative in some cases.

Control and Estimation

We assume that a nominal (desired) flight path is generated by higher-level intelligence, such as the human pilot or declarative machine logic, or as a stored series of waypoints. The procedural system must follow the path, $\mathbf{x}^*(t)$ in $t_o < t < t_f$. Control is exercised by a digital computer at time intervals of Δt. The n-dimensional state vector perturbation at time t_k is $\mathbf{x}_k$, and the m-dimensional control vector perturbation is $\mathbf{u}_k$. The discrete-time linear-quadratic-Gaussian (LQG) control law is formed as [63],

$$\mathbf{u}_k = \mathbf{u}^*_k - \mathbf{C}_B[\hat{\mathbf{x}}_k - \mathbf{x}^*_k] = \mathbf{C}_F\mathbf{y}^*_k - \mathbf{C}_B\hat{\mathbf{x}}_k \qquad (1$$

$\mathbf{y}^*_k$ is the desired value of an output vector $(\overset{\Delta}{=} \mathbf{H}_x\mathbf{x}_k + \mathbf{H}_u\mathbf{u}_k)$, and $\hat{\mathbf{x}}_k$ is the *Kalman filter* estimate, expressed in two steps:

$$\hat{\mathbf{x}}_k(-) = \Phi\hat{\mathbf{x}}_{k-1}(+) + \Gamma\mathbf{u}_{k-1}$$
$$\hat{\mathbf{x}}_k \overset{\Delta}{=} \hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + \mathbf{K}[\mathbf{z}_k - \mathbf{H}_{obs}\hat{\mathbf{x}}_k(-)] \qquad (2$$

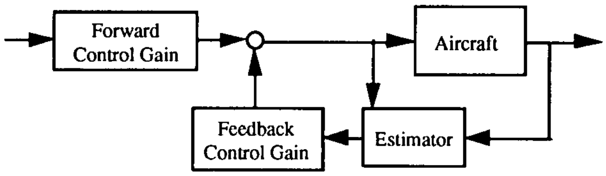
The forward and feedback control gain matrices are $\mathbf{C}_F$ and $\mathbf{C}_B$, Φ and Γ are state-transition and control-effect matrices that describe the aircraft's assumed dynamics, the estimator gain matrix is $\mathbf{K}$, and the measurement vector, $\mathbf{z}_k$, is a transformation of the state through $\mathbf{H}_{obs}$. The gains $\mathbf{C}_B$ and $\mathbf{K}$ result from solving two Riccati equations that introduce tradeoffs between control use and state perturbation and between the strengths of random disturbances and measurement error. $\mathbf{C}_F$, which provides proper steady-state command response, is an algebraic function of $\mathbf{C}_B$, Φ, Γ, and $\mathbf{H}_{obs}$. All of the matrices may vary in time, and it may be necessary to compute $\mathbf{K}$ on-line. In the remainder, it is not essential that $\mathbf{C}_B$ and $\mathbf{K}$ be optimal (i.e., they may have been derived from eigen-structure assignment, loop shaping, etc.), although the LQR gains guarantee useful properties of the nominal closed-loop system [63].

The control structure provided by eq. 1 and 2 is quite flexible. It can represent a scalar feedback loop if $\mathbf{z}$ contains one measurement and $\mathbf{u}$ one control, or it can address measurement and control redundancy with $\mathbf{z}$ and $\mathbf{u}$ dimensions that exceed the dimension of the state $\mathbf{x}$. It also is possible to incorporate reduced-order modeling in the estimator. Assuming that Φ and Γ have the same dimensions as the aircraft's dynamic model ($n \times n$ and $n \times m$), *the baseline estimator introduces n^{th}-order compensation in the feedback control loop.*

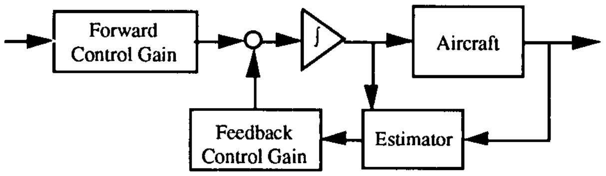
The weights of the quadratic control cost function can be chosen not only to penalize state and control perturbations but to produce *output weighting*, *state-rate weighting*, and *implicit model following*, all without modifying the dynamic model [63]. *Integral compensation*, *low-pass filter compensation*, and *explicit model following* can be obtained by augmenting the system model during the design process, increasing the compensation order and producing the control structures shown in Fig. 10.

These cost weighting and compensation features can be used together, as in the proportional-integral/implicit-model-following controllers developed in [80]. Implicit model following is especially valuable when an ideal model can be specified (as identified in flying qualities specifications and standards [5, 6]), and integral compensation provides automatic "trimming" (control that synthesizes $\mathbf{u}^*_k$ corresponding to $\mathbf{x}^*_k$ to achieve zero steady-state command error) and low-frequency robustness. Combining integral and filter compensation produces controllers with good command tracking performance and smooth control actions, as demonstrated in flight tests [81-83]. The LQG regulator naturally introduces an *internal model of the controlled plant*, a feature that facilitates control design [84]. It produces a stable approximation to the *system inverse*, which is at the heart of achieving desired command tracking.

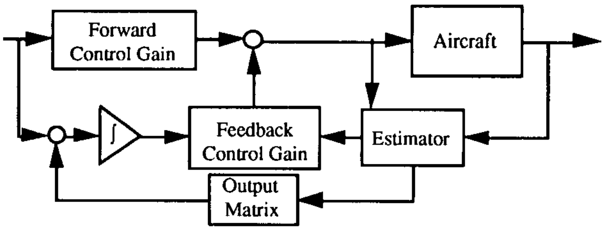
The estimator in the feedback loop presents an efficient means of dealing with uncertainty in the measurements, in the disturbance inputs, and (to a degree) in the aircraft's dynamic model. If measurements are very noisy, the estimator gain matrix $\mathbf{K}$ is "small," so that the filter relies on extrapolation of the system model to estimate the state. If disturbances are large, the state itself is more uncertain, and $\mathbf{K}$ is "large," putting more emphasis on the measurements. Effects of uncertain parameters can be approximated as "process noise" that increases the importance of measurements, leading to a higher $\mathbf{K}$. If the system uncertainties are constant but unknown biases or scale factors, a better approach is to augment the filter state to estimate these terms directly. Parametric uncertainty introduces nonlinearity; hence, an *extended Kalman filter* must be used [63].



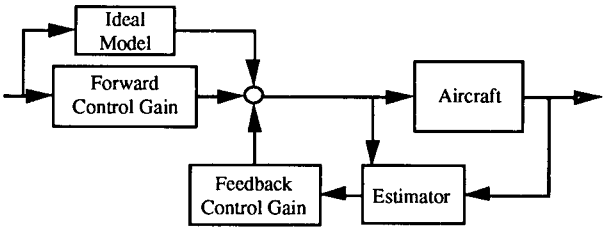
a) Linear-Quadratic-Gaussian (LQG) Regulator.



b) Proportional-Filter LQG Regulator.



c) Proportional-Integral LQG Regulator.



d) Explicit-Model-Following LQG Regulator.

Figure 10. Structured Linear-Quadratic-Gaussian Regulators.

Stability and Performance Robustness

Controlled system *robustness* is the ability to maintain satisfactory stability and performance in the presence of parametric or structural uncertainties in either the aircraft or its control system. All controlled systems must possess some degree of robustness against operational parameter variations. The inherent stability margins of certain algebraic control laws (e.g., the linear-quadratic (LQ) regulator [63, 85-87]) may become vanishingly small when dynamic compensation (e.g., the estimator in a linear-quadratic-Gaussian (LQG) regulator) is added [88]. Restoring the robustness to that of the LQ regulator typically requires increasing estimator gains (within practical limits) using the loop-transfer-recovery method [89].

Subjective judgments must be made in assessing the need for robustness and in establishing corresponding control system design criteria, as there is an inevitable tradeoff between robustness and nominal system performance [90]. The designer must know the normal operating ranges and distributions of parameter variations, as well as the specifications for system operability with failed components, else the final design may afford too little robustness for possible parameter variations or too much robustness for satisfactory nominal performance. Robustness traditionally has been assessed deterministically [91, 92]; gain and phase margins are an inherent part of the classical design of single-input/single-output systems, and there are multi-input/multi-output equivalents based on singular-value analysis (e.g., [93]). A critical difficulty in applying these techniques is relating singular-value bounds on return-difference and inverse-return-difference matrices to real parameter variations in the system.

Statistical measures of robustness can use knowledge of potential variations in real parameters. The *probability of instability* was introduced in [94] and is further described in [95, 96]. The *stochastic robustness* of a linear, time-invariant system, is judged using Monte Carlo simulation to estimate the probability distributions of closed-loop eigenvalues, given the statistics of the variable parameters in the system's dynamic model. The probability that one or more of these eigenvalues have positive real parts is the scalar measure of robustness, a figure of merit to be minimized by control system design. Because this metric can take one of only two values, it has a *binomial distribution*; hence, *the confidence intervals associated with estimating the metric from simulation are independent of the number or nature of the uncertain parameters* [95].

Considerations of performance robustness are easily taken into account in *Stochastic Robustness Analysis* (SRA). Systems designed using a variety of robust control methods (loop transfer recovery, H_∞ optimization, structured covariance, and game theory) are analyzed in [97], with attention directed to the probability of instability, probability of settling-time exceedence, probability of excess control usage, and tradeoffs between them. The analysis uncovers a wide range of system responses and graphically illustrates that gain and phase margins are not good indicators of the probability of instability[8]. This also raises doubts about the utility of singular values, as they are multivariable equivalents of the classical robustness metrics. Incorporating SRA into the design of an LQG regulator with implicit model following and filter compensation leads to designs that have high levels of stability and performance robustness [98]. The reason for improvement is that SRA measures the actual effects of parameter variations on stability and performance rather than incremental changes in the nominal margins.

Adaptation and Tolerance to Failures

Adaptation always has been a critical element of stability augmentation. Most aircraft requiring improved stability undergo large variations in dynamic characteristics on a typical mission. Gain scheduling and control interconnects initially were implemented mechanically, pneumatically, and hydraulically; now the intelligent part is done within a computer, and there is increased freedom to use sophisticated scheduling techniques that approach full nonlinear control [81, 99]. It becomes feasible not only to schedule according to flight condition but to account for differences in individual aircraft. Flight control systems that adapt to changes due to wear and exposure and that report changes for possible maintenance action can now be built.

Tolerance to system failures, such as plant alterations, actuator and sensor failures, computer failure, and power supply/transmission failure, is an important issue. Multiple failures can occur, particularly as a consequence of physical damage, and they may be intermittent. Factors that must be considered in designing failure-tolerant controls include: allowable performance degradation in the failed state, criticality and likelihood of the failure, urgency of response to failure, tradeoffs between correctness and speed of response, normal range of system uncertainty, disturbance environment, component reliability vs. redundancy, maintenance goals, system architecture, limits of manual intervention, and life-cycle costs [43].

One approach to failure tolerance is *parallel redundancy*: two or more control strings, each separately capable of satisfactory control, are implemented in parallel. A *voting* scheme is used for redundancy management. With two identical channels, a comparator can determine whether or not control signals are identical; hence, it can detect a failure but cannot identify which string has failed. Using three identical channels, the control signal with the middle value can be selected (or voted), assuring that a single failed channel never controls the plant. In any voting system, it remains for additional logic to declare unselected channels failed. Given the vectorial nature of control, this declaration may be equivocal, as middle values of control-vector elements can be drawn from different strings.

Parallel redundancy can protect against control-system component failures, but it does not address failures of plant components. *Analytical redundancy* provides a capability to improve tolerance to failures of both types. The principal functions of analytical redundancy are *failure detection*, *failure identification*, and *control-system reconfiguration*. These functions use the control computer's ability to compare expected response to actual response, inferring component failures from the differences and changing either the structure or the parameters of the control system as a consequence [47].

---

[8] Real parameter variations affect not only the magnitude and relative phase angle of the system's Nyquist contour but its shape as well [63]. Therefore, the points along the contour that establish gain and phase margin (i.e., the corresponding Bode-plot frequencies) are subject to change.

Procedural adaptation and failure-tolerance features will evolve outward, to become more declarative in their supervision and more reflexive in their implementation. Declarative functions are especially important for differentiating between normal and emergency control functions and sensitivities. They can work to reduce trim drag, to increase fatigue life, and to improve handling and ride qualities as functions of turbulence level, passenger loading, and so on. Gain-scheduling control can be viewed as *fuzzy control*, suggesting that the latter has a role to play in aircraft control systems [100-102]. Reflexive functions can be added by computational neural networks that approximate nonlinear multivariate functions or classify failures.

Nonlinear Control

Aircraft dynamics are inherently nonlinear, but aerodynamic nonlinearities and inertial coupling effects are generally smooth enough in the principal operating regions to allow linear control design techniques to be used. Control actuators impose hard constraints on operation because their displacements and rates are strictly limited. Nonlinear control laws can improve control precision and widen stability boundaries when flight must be conducted at high angles or high angular rates and when the control-actuator limits must be challenged.

The general principles of nonlinear inverse control are straightforward [103]. Given a nonlinear system of the form,

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u} \qquad (3$$

where $\mathbf{G}(\mathbf{x})$ is square ($m = n$) and non-singular, the control law

$$\mathbf{u} = -\mathbf{G}^{-1}\mathbf{f}(\mathbf{x}) + \mathbf{G}^{-1}\mathbf{v} \qquad (4$$

inverts the system, since

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})[-\mathbf{G}^{-1}\mathbf{f}(\mathbf{x}) + \mathbf{G}^{-1}\mathbf{v}] = \mathbf{v} \qquad (5$$

where $\mathbf{v}$ is the command input to the system.

In general, $\mathbf{G}(\mathbf{x})$ is not square ($m \neq n$); however, given an m-dimensional output vector,

$$\mathbf{y} \overset{\Delta}{=} \mathbf{H}\mathbf{x} \qquad (6$$

it is possible to define a nonlinear feedback control law that produces *output decoupling* of the elements of $\mathbf{y}$ or their derivatives such that $\mathbf{y}^{(d)} = \mathbf{v}$. The vector $\mathbf{y}^{(d)}$ contains *Lie derivatives* of $\mathbf{y}$,

$$\mathbf{y}^{(d)} = = \mathbf{f}^*(\mathbf{x}) + \mathbf{G}^*(\mathbf{x})\mathbf{u} \qquad (7$$

where d is the *relative degree* of differentiation required to identify a direct control effect on each element of $\mathbf{y}$. $\mathbf{G}^*(\mathbf{x})$ and $\mathbf{f}^*(\mathbf{x})$ are components of the Lie derivatives, and $\mathbf{G}^*(\mathbf{x})$ is guaranteed to be structurally invertible by the condition that defines relative degree [104]. The decoupling control law then takes the form

$$\mathbf{u} = -[\mathbf{G}^*(\mathbf{x})]^{-1}\mathbf{f}^*(\mathbf{x}) + [\mathbf{G}^*(\mathbf{x})]^{-1}\mathbf{v} \qquad (8$$

The control law is completed by feeding $\mathbf{y}$ back as appropriate to achieve desired transient response and by pre-filtering $\mathbf{v}$ to produce the desired command response [105]. Because the full state is rarely measured and measurements can contain errors, it may be necessary to estimate $\mathbf{x}$ with an extended Kalman filter, substituting $\hat{\mathbf{x}}$ for $\mathbf{x}$ in control computations.

Evaluating $\mathbf{G}^*(\mathbf{x})$ and $\mathbf{f}^*(\mathbf{x})$ requires that a full, d-differentiable model of aircraft dynamics be included in the control system; hence the statement of the control law is simple, but its implementation is complex (Fig. 11). Smooth interpolators of the aircraft model (e.g., cubic splines) are needed. Feedforward neural networks with sigmoidal activation functions are infinitely differentiable, providing a good means of representing this model on-line and allowing adaptation [106, 107].

There are limitations to the inverse control approach [108]. The principal concerns are pointwise singularity of $\mathbf{G}^*(\mathbf{x})$, the effects of control saturation, and the presence of the nonlinear equivalent of non-minimum-phase (NMP) zeros in aircraft dynamics. The command vector (eq. 6) has a direct effect on the definition of $\mathbf{G}^*(\mathbf{x})$. In [105], the singular points of $\mathbf{G}^*(\mathbf{x})$ are found to be outside the flight envelope of the subject aircraft for all command vectors. When saturation of a control effector occurs, the control dimension must be reduced by one; hence, the command vector is redefined to exclude the least important element of $\mathbf{y}$ in [105]. The command vector is returned to original dimension when the control effector is no longer saturated.

Whether or not NMP zero effects are encountered depends on the command-vector definition and on the physical model. Some command-vector definitions for aircraft control produce no NMP zeros [105]. When NMP zeros occur in conventional aircraft models, they are due to small force effects (e.g., lift due to elevator deflection and pitch rate); it may be possible to neglect them, eliminating the problem.

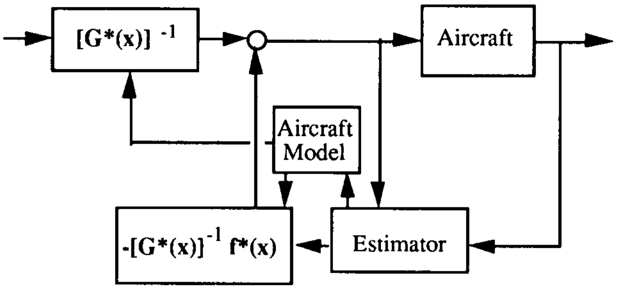


Figure 11. Decoupling Nonlinear-Inverse Control Law.

REFLEXIVE SYSTEMS

Inner-loop control is a reflexive (though not necessarily linear) function. To date, most inner loops have been designed as procedural control structures; computational neural networks may extend prior results to facilitate nonlinear control and adaptation. Neural networks can be viewed as *nonlinear generalizations of sensitivity, transformation, and gain matrices*. Consequently, compensation dynamics can be incorporated by following earlier models and control structures. Nonlinear proportional-integral and model following controllers, as well as nonlinear estimators, can be built using computational neural networks.

Computational Neural Networks

Computational neural networks are motivated by input-output and learning properties of biological neural systems, though in mathematical application the network becomes an abstraction that may bear little resemblance to its biological model. Computational neural networks consist of *nodes* that simulate the *neurons* and *weighting factors* that simulate the *synapses* of a living nervous system. The nodes are nonlinear basis functions, and the weights contain knowledge of the system. Neural networks are good candidates for performing a variety of reflexive functions in intelligent control systems because they are potentially very fast (in parallel hardware implementation), they are intrinsically nonlinear, they can address problems of high dimension, and they can learn from experience. From the biological analogy, the neurons are modeled as switching functions that take just two discrete values; however, "switching" may be softened to "saturation," not only to facilitate learning of the synaptic weights but to admit the modeling of continuous, differentiable functions.

The neural networks receiving most current attention are static expressions that perform one of two functions. The first is to *approximate multivariate functions* of the form

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) \qquad (9$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are input and output vectors and $\boldsymbol{h}(\bullet)$ is the (possibly unknown) relationship between them. Neural networks can be considered *generalized spline functions* that identify efficient input-output mappings from observations [109, 110]. The second application is to *classify attributes*, much like the decision trees mentioned earlier. (In fact, decision trees can be mapped to neural networks [111].) The following discussion emphasizes the first of these two applications.

An N-layer *feedforward neural network* (FNN) represents the function by a sequence of operations,

$$\mathbf{r}^{(k)} = \mathbf{s}^{(k)}[\mathbf{W}^{(k-1)}\mathbf{r}^{(k-1)}] \triangleq \mathbf{s}^{(k)}[\boldsymbol{\eta}^{(k)}]\ , \ k = 1 \text{ to } N \qquad (10$$

where $\boldsymbol{y} = \mathbf{r}^{(N)}$ and $\boldsymbol{x} = \mathbf{r}^{(0)}$. $\mathbf{W}^{(k-1)}$ is a matrix of weighting factors determined by the learning process, and $\mathbf{s}^{(k)}[\bullet]$ is an activation-function vector whose elements normally are identical, scalar, nonlinear functions $\sigma_i(\eta_i)$ appearing at each node:

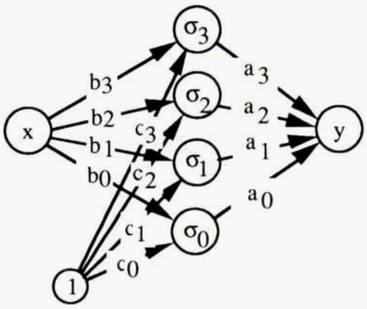
$$\mathbf{s}^{(k)}[\eta^{(k)}] = [\sigma_1(\eta_1{}^{(k)}) \ldots \sigma_n(\eta_n{}^{(k)})]^T \qquad (11$$

One of the inputs to each layer may be a unity threshold element that adjusts the bias of the layer's output. Networks consisting solely of linear activation functions are of little interest, as they merely perform a linear transformation $\boldsymbol{H}$, thus limiting eq. 9 to the form, $\boldsymbol{y} = \boldsymbol{Hx}$.

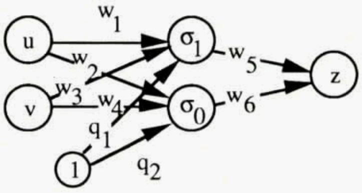
Figure 12 represents two simple feedforward neural networks. Each circle represents an arbitrary, scalar, nonlinear function $\sigma_i(\bullet)$ operating on the sum of its inputs, and each arrow transmits a signal from the previous node, multiplied by a weighting factor. A scalar network with a single hidden layer of four nodes and a unit threshold element (Fig. 12a) is clearly parallel, yet its output can be written as the series

$$y = a_0\sigma_0(b_0x + c_0) + a_1\sigma_1(b_1x + c_1) + a_2\sigma_2(b_2x + c_2) + a_3\sigma_3(b_3x + c_3) \qquad (12$$

illustrating that parallel and serial processing may be equivalent.



a) Single-Input/Single-Output Network.



b) Double-Input/Single-Output Network.

Figure 12. Two Feedforward Neural Networks.

Consider a simple example. Various nodal activation functions, σ_i, have been used, and there is no need for each node to be identical. Choosing $\sigma_0(\bullet) = (\bullet)$, $\sigma_1 = (\bullet)^2$, $\sigma_2 = (\bullet)^3$, $\sigma_3 = (\bullet)^4$, eq. 9 is represented by the truncated power series,

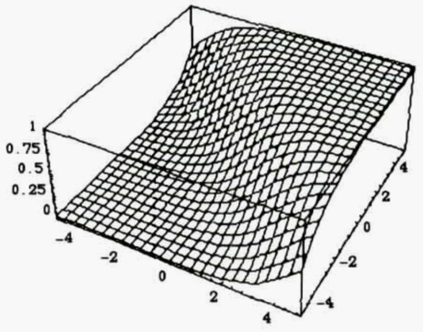
$$y = a_0(b_0x + c_0) + a_1(b_1x + c_1)^2 + a_2(b_2x + c_2)^3 + a_3(b_3x + c_3)^4 \qquad (13$$

and it is clear that network weights are redundant (i.e., that the (**a**, **b**, **c**) weighting factors are not independent). Consequently, more than one set of weights could produce the same functional relationship between x and y. Training sessions starting at different points could produce different sets of weights that yield identical outputs. This simple example also indicates that the unstructured feedforward network may not have compact support (i.e., its weights may have global effects) if its basis functions do not vanish for large magnitudes of their arguments.

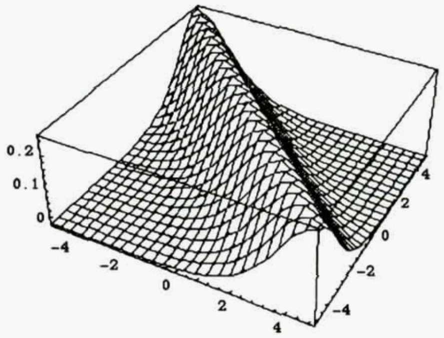
The *sigmoid* is commonly used as the artificial neuron. It is a saturating function defined variously as $\sigma(\eta) = 1/(1 + e^{-\eta})$ for output in (0,1) or $\sigma(\eta) = (1 - e^{-2\eta})/(1 + e^{-2\eta}) = \tanh \eta$ for output in (-1,1). Recent results indicate that any continuous mapping can be approximated arbitrarily closely with sigmoidal networks containing a single hidden layer (N = 2) [112, 113]. Symmetric functions like the *Gaussian radial basis function* ($\sigma(\eta) = e^{-\eta^2}$) have better convergence properties for many functions and have more compact support as a consequence of near-orthogonality [109, 114]. Classical *B-splines* [115] could be expressed in parallel form, and it has been suggested that they be used in multilayered networks [116]. Adding hidden layers strengthens the analogy to biological models, though additional layers are not necessary for approximating continuous functions, and they complicate the training process.

In control application, neural networks perform functions analogous to gain scheduling or nonlinear control. Consider the simple two-input network of Fig. 12b. The scalar output and derivative of a single sigmoid with unit weights are shown in Fig. 13. If u is a fast variable and v is a slow variable, choosing the proper weights on the inputs and threshold can produce a gain schedule that is approximately linear in one region and nonlinear (with an inflection point) in another. More complex surfaces can be generated by increasing the number of sigmoids. If u and v are both fast variables, then the sigmoid can represent a generalization of their nonlinear interaction.

For comparison, a typical radial basis function produces the output shown in Fig. 14. Whereas the sigmoid has a preferred input axis and simple curvature, the RBF admits more complex curvature of the output surface, and its effect is more localized. The most efficient nodal activation function depends on the general shape of the surface to be approximated. There may be cases best handled by a mix of sigmoids and RBF in the same network.

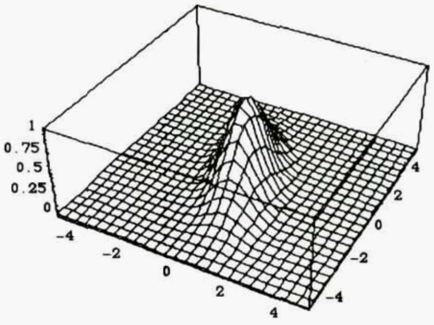


a) Sigmoid.

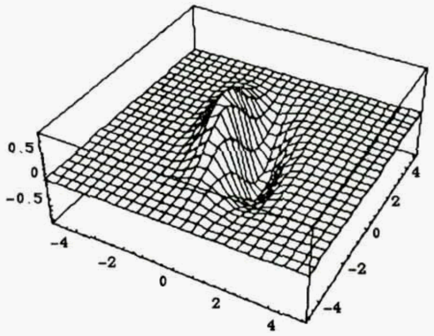


b) x-Derivative of Sigmoid.

Figure 13. Example of Sigmoid Output with Two Inputs.



a) Radial Basis Function (RBF).



b) x-Derivative of RBF.

Figure 14. Example of Radial Basis Function Output with Two Inputs.

The cerebellar model articulation controller (CMAC) is an alternate network formulation with somewhat different properties but similar promise for application in control systems [117, 118]. The CMAC performs table look-up of a nonlinear function over a particular region of function space. CMAC operation can be split into two mappings. The first maps each input into an *association space A*. The mapping generates a *selector vector* $\boldsymbol{a}$ of dimension n_A, with c non-zero elements (usually ones) from overlapping *receptive regions* for the input. The second mapping, R, goes from the selector vector $\boldsymbol{a}$ to the scalar output y through the weight vector $\boldsymbol{w}$, which is derived from training:

$$y = \boldsymbol{w}^T\boldsymbol{a} \qquad (14$$

Training is inherently local, as the extent of the receptive regions is fixed. The CMAC has quantized output, producing "staircased" rather than continuous output. A recent paper proposes to smooth the output using B-spline receptive regions [119].

The FNN and CMAC are both examples of *static networks*, that is, their outputs are essentially instantaneous: given an input, the speed of output depends only on the speed of the computer. *Dynamic networks* rely on stable resonance of the network about an equilibrium condition to relate a fixed set of initial conditions to a steady-state output. Bidirectional Associative Memory (BAM) networks [120] are nonlinear dynamical systems that subsume Hopfield networks [121], Adaptive-Resonance-Theory (ART) networks [122], and Kohonen networks [123]. Like FNN, they use binary or sigmoidal neurons and store knowledge in the weights that connect them; however, the "neural circuits" take time to stabilize on an output. While dynamic networks may operate more like biological neurons, which have a *refractory period* between differing outputs, computational delay degrades aircraft control functions.

Although neural networks performing function approximation may gain little from multiple hidden layers, networks used for classification typically require multiple layers, as follows from the ability to map decision trees to neural networks [111]. The principal values of performing such a mapping are that it identifies an efficient structure for parallel computation, and it may facilitate incremental learning and generalization.

Neural networks can be applied to *failure detection and identification* (FDI) by mapping data patterns (or *feature vectors*) associated with failures onto detector/identification vectors (e.g., [124-126]). To detect failure, the output is a scalar, and the network is trained (for example) with "1" corresponding to failure and "0" corresponding to no failure. To identify specific failures, the output is a vector, with a training value of "1" in the i^{th} element corresponding to the i^{th} failure mode and zeros elsewhere. For M failure modes, either M neural networks with scalar outputs are employed or a single neural network with M-vector output is used; there are evident tradeoffs related to efficiency, correlation, and so on. The data patterns associated with each failure may require *feature extraction*, pre-processing that transforms the input time series into a feature vector. In [124], this was done by computing two dozen Fourier coefficients of the input signal in a moving temporal window. As an alternative, the feature vector could be specified as a *parity vector* [127], and the neural network could be used for the decision-making logic in FDI. When assessing the efficiency of neural-network FDI logic, feature extraction must be considered part of the total process.

Reflexive Learning and Adaptation

Training neural networks involves either supervised or unsupervised learning. In *supervised learning*, the network is furnished typical histories of inputs and outputs, and the training algorithm computes the weights that minimize fit error. FNN and CMAC require this type of training. In *unsupervised learning*, the internal dynamics are self-organizing, tuning the network to home on different cells of the output *semantic map* in response to differing input patterns [128]. Dynamic networks train rapidly and are suitable for pattern matching, as in speech or character recognition. The remaining discussion focuses on supervised learning, which is more consistent with control functions.

Backpropagation learning algorithms for the elements of $\mathbf{W}^{(k)}$ typically involve a *gradient search* (e.g., [129, 130]) that minimizes the mean-square output error

$$\mathcal{E} = \left[\mathbf{r}_d - \mathbf{r}^{(N)}\right]^T\left[\mathbf{r}_d - \mathbf{r}^{(N)}\right] \qquad (15$$

where $\mathbf{r}_d$ is the desired output. For each input-output example presented to the network, the gradient of the error with respect to the weight matrix is calculated, and the weights are updated by

$$\mathbf{W}^{(k)}_{new} = \mathbf{W}^{(k)}_{old} + \beta\mathbf{r}^{(k-1)}\left[\mathbf{d}^{(k)}\right]^T \qquad (16$$

β is the learning rate, and $\mathbf{d}$ is a function of the error between desired and actual outputs. For the output layer, the error term is

$$\mathbf{d}^{(N)} = \mathbf{S}'\left[\mathbf{W}^{(N-1)}\mathbf{r}^{(N-1)}\right]\left(\mathbf{r}_d - \mathbf{r}^{(N)}\right) \qquad (17$$

where the prime indicates differentiation with respect to $\mathbf{r}$. For interior layers, the error from the output layer is propagated from the output error using

$$\mathbf{d}^{(k)} = \mathbf{S}'\left[\mathbf{W}^{(k-1)}\mathbf{r}^{(k-1)}\right]\left[\mathbf{W}^{(k-1)}\right]^T\mathbf{d}^{(k-1)} \qquad (18$$

Search rate can be modified by adding momentum or conjugate-gradient terms to eq. 16.

The CMAC network learning algorithm is similar to back-propagation. The weights and output are connected by a simple linear operation, so a learning algorithm is easy to prescribe. Each weight contributing to a particular output value is adjusted by a fraction of the difference between the network output and the desired output. The fraction is determined by the desired learning speed and the number of receptive regions contributing to the output.

Learning speed and accuracy for FNN can be further improved using an *extended Kalman filter* [106, 107, 131]. The dynamic and observation models for the filter are

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \mathbf{q}_{k-1} \qquad (19$$

$$\mathbf{z}_k = \mathbf{h}(\mathbf{w}_k, \mathbf{r}_k) + \mathbf{n}_k \qquad (20$$

where $\mathbf{w}_k$ is a vector of the matrix $\mathbf{W}_k$'s elements, $\mathbf{h}(\bullet)$ is an observation function, and $\mathbf{q}_k$ and $\mathbf{n}_k$ are noise processes. If the network has a scalar output, then $\mathbf{z}_k$ is scalar, and the extended Kalman filter minimizes the fit error between the training hypersurface and that produced by the network (eq. 15). It has been found that the fit error can be dramatically reduced by considering the *gradients* of the surfaces as well [106, 107]. The observation vector becomes

$$\mathbf{z}_k = \begin{bmatrix} \mathbf{h}(\mathbf{w}_k,\mathbf{r}_k) \\ \dfrac{\partial \mathbf{h}}{\partial \mathbf{r}}(\mathbf{w}_k,\mathbf{r}_k) \end{bmatrix} + \mathbf{n}_k \qquad (21$$

with concomitant increase in the complexity of the filter. The relative significance given to function and derivative error during training can be adjusted through the measurement-error covariance matrix used in filter design.

Recursive estimation of the weights is useful when smooth relationships between fit errors and the weights are expected, when the weight-vector dimension is not high, and when local minima are global. When one of these is not true, it may speed the computation of weights to use a *random search*, at least until convergent regions are identified. Such methods as *simulated annealing* or *genetic algorithms* can be considered (and the latter has philosophic appeal for intelligent systems) [132-134]. The first of these is motivated by statistical mechanics and the effects that repeated heating and cooling have on the ground states of atoms (which are analogous to the network weights). The second models the reproduction, crossover, and mutation of biological strings (e.g., chromosomes, again analogous to the weights), in which only the fittest combinations survive.

Statistical methods can go hand-in-hand with SRA to train *robust neural networks*. Following [98], the random search could be combined with Monte Carlo variation of system parameters during training, numerically minimizing the *expected value of fit error* rather than a deterministic fit error.

Problems that may be encountered in neural network training include proper choice of the input vector, local vs. global training, speed of learning and forgetting, generalization over untrained regions, and trajectory-dependent correlations in the training sets. We envision an aerodynamic model that spans the entire flight envelope of an aircraft, including post-stall and spinning regions. The model contains six neural networks with multiple inputs and scalar outputs, three for force coefficients and three for moment coefficients (for example, the pitch moment network takes the form $C_m = g(\mathbf{x},\mathbf{u})$, where $\mathbf{x}$ represents the state and $\mathbf{u}$ the control). If input variables are not restricted to those having plausible aerodynamic effect, false correlations may be created in the network; hence, attitude Euler angles and horizontal position should be neglected, while physically meaningful terms like elevator deflection, angle of attack, pitch rate, Mach number, and dynamic pressure should be included [107].

The aircraft spends most of its flying time within normal mission envelopes. Unless it is a trainer, the aircraft does not enter post-stall and spinning regions; consequently, on-line network training focuses on normal flight and neglects extreme conditions. This implies not only that networks must be pre-trained in the latter regions but that normal training must not destroy knowledge in extreme regions while improving knowledge in normal regions. Therefore, radial basis functions appear to be a better choice than sigmoid activation functions for adaptive networks.

Elements of the input vector may be strongly correlated with each other through the aircraft's equations of motion; hence, networks may not be able to distinguish between highly correlated variables (e.g., pitch rate and normal acceleration). This is problematical only when the aircraft is outside its normal envelope. Pre-training should provide inputs that are rich in frequency content, that span the state and control spaces, and that are as uncorrelated as possible. Generalization between training points may provide smoothness, but it does not guarantee accuracy.

Control Systems Based on Neural Networks

Neural networks can find application in logic for control, estimation, system identification, and physical modeling. In addition to work already referenced, additional examples can be found in [135-140].

Figure 15a illustrates an application in which the neural network forms the aircraft model for a nonlinear-inverse control law. The aircraft model of Fig. 11 is implemented with a neural network that is trained by a dedicated (weight) extended Kalman filter (the thick gray arrow indicating training). The extended Kalman filter for state estimation is expanded to estimate histories of forces and moments as well as the usual motion variables.

It is possible to conduct supervised learning on-line without interfering with normal operation because the state Kalman filter produces both the necessary inputs and the desired outputs for the

the network training algorithm. There is no need to provide an ideal control response for training, as the form of the control law is fixed. Procedural and reflexive functions are combined in this control implementation, under the assumption that the direct expression of inversion is the most efficient approach.

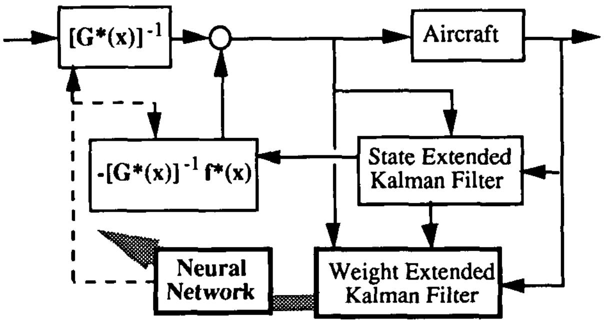
Figure 15b shows a logical extension in which the inverse control law is implemented by neural networks. Inversion is an implicit goal of neural-network controllers [135, 136], and the formal existence of inversion networks has been explored [141]. Although Fig. 15b implies that the inversion networks are pre-trained and fixed, they, too, can be trained with the explicit help of the network that models the system [136]. Here it is assumed that the control networks have been pre-trained, as no desired output has been specified.

If a desired control output is specified (Fig. 15c), then the formal model of the aircraft is no longer needed. The control networks learn implicit knowledge of the aircraft model. Referring to Fig. 10 and eq. 1 and 2, control and estimation gains, state-transition and control-effect matrices, and measurement transformations can be implemented as static neural networks with either off-line or on-line learning.

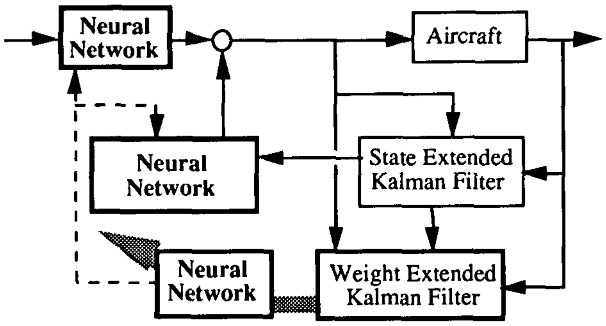
It can be useful to divide control networks into separate feedback and forward parts, as this may facilitate training to achieve design goals. A feedback neural network has strongest effect on homogeneous modes of motion, while a forward neural network is most effective for shaping command (forced) response. This structure is adopted in [139], where the forward and feedback networks are identified as *reason* and *instinct networks*. In pre-training, it is plausible that the feedback network would be trained with initial condition responses first, to obtain satisfactory transient response. The forward network would be trained next to achieve desired steady states and input decoupling. A third training step could be the addition of command-error integrals while focusing on disturbance inputs and parameter variations in training sets.

Once baselines have been achieved, it could prove useful to admit limited coupling between forward and feedback networks to enable additional nonlinear compensation. In adaptive applications, networks would be pre-trained with the best available models and scenarios to establish satisfactory baselines; on-line training would slowly adjust individual systems to vehicle and mission characteristics.

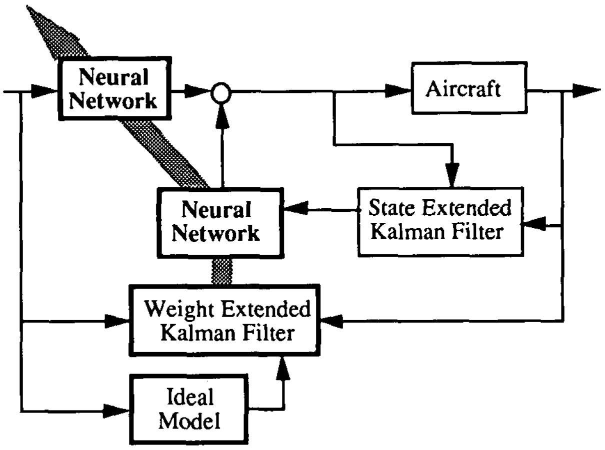
Including the integral of command-vector error as a neural network input produces a *proportional-integral* structure [140], while placing the integrator beyond the network gives a *proportional-filter* structure (Fig. 10). The principal purpose of these structures is, as before, to assure good low- and high-frequency performance in a classical sense. Extension of neural networks to state and weight filters is a logical next step that is interesting in its own right as a means of more nearly optimal nonlinear estimation.



a) Neural Network for Modeling and Adaptation.



b) Neural Networks for Modeling, Adaptation, and Control.



c) Neural Networks for Control Alone.

Figure 15. Adaptive Control Structures Using Neural Networks.

CONCLUSION

Intelligent flight control systems can do much to improve the operating characteristics of aircraft. An examination of cognitive and biological models for human control of systems suggest that there is a declarative, procedural, and reflexive hierarchy of functions. Top-level aircraft control functions are analogous to conscious and preconscious thoughts that are transmitted to lower-level subsystems through subconscious, neural, and reflex-like activities. Human cognition and biology also suggest models for learning and adaptation, not only during operation but between periods of activity.

The computational analogs of the three cognitive/biological paradigms are expert systems, stochastic controllers, and neural networks. Expert systems organize decision-making efficiently, stochastic controllers optimize estimation and control, and neural networks provide rapid, nonlinear, input-output functions. It appears that many functions at all levels could be implemented as neural networks. While this may not always be necessary or even desirable using sequential processors, mapping declarative and procedural functions as neural networks may prove most useful as a route to new algorithms for the massively parallel processors of the future.

ACKNOWLEDGMENTS

This work has been supported by the Federal Aviation Administration and the National Aeronautics and Space Administration under Grant No. NGL 31-001-252 and by the Army Research Office under Contract No. DAAL03-89-K-0092.

REFERENCES

1. *The Papers of Wilbur and Orville Wright*, Vol. 1, 1898-1905; Vol. 2, 1906-1948.
2. Harper, R. P., Jr., and Cooper, G. E., Handling Qualities and Pilot Evaluation, *J. Guid., Cont., Dyn.*, Vol. 9, No. 5, Sept.-Oct. 1986, pp. 515-529.
3. Moorhouse, D. J., "The History and Future of U.S. Military Flying Qualities Specification," *AIAA Aero. Sci. Mtg.*, New Orleans, Jan. 1979.
4. Anderson, R. O., "Flying Qualities Yesterday Today and Tomorrow," *AIAA Atmos. Flgt. Mech. Conf.*, Danvers, MA, Aug. 1980.
5. --, "Military Specification, Flying Qualities of Piloted Airplanes, " MIL-F-8785C, WPAFB, OH, Nov. 1980.
6. Hoh, R. H., *et al*, Proposed MIL Standard and Handbook - Flying Qualities of Air Vehicles, AFWAL-TR-82-3081, WPAFB, OH, Nov. 1982.
7. Tustin, A., "The Nature of the Operator's Response in Manual Control and Its Implications for Controller Design," *Proc. IEE*, Vol. 94, Part IIA, 1947, pp. 190-202.
8. McRuer, D. T., "Development of Pilot-In-The-Loop Analysis," *J. Aircraft*, Vol. 10, No. 9, Sept. 1973, pp. 515-524.
9. Kleinman, D. L., Baron, S., and Levison, W., "An Optimal Control Model of Human Response," *Automatica*, Vol. 9, No. 3, May 1970, pp. 357-383.
10. McRuer, D. T., *et al*, "New Approaches to Human Pilot/Vehicle Dynamic Analysis," AFFDL-TR-67-150, WPAFB, OH, Feb. 1968.
11. Stengel, R. F., and Broussard, J. R., "Prediction of Pilot-Aircraft Stability Boundaries and Performance Contours, *IEEE Trans. Syst., Man, Cyber.*, Vol. SMC-8, No. 5, May 1978, pp. 349-356.
12. Stark, L., *Neurological Control Systems*, Plenum Press, New York, 1968.
13. Stone, J. R., and Gerken, G. J., "The Prediction of Pilot Acceptance for a Large Aircraft," *Proc. 1973 Joint Auto. Cont. Conf.*, Columbus, OH, June 1973, pp. 807-809.
14. Hollister, W. M., ed., *Improved Guidance and Control Automation at the Man-Machine Interface*, AGARD-AR-228, Neuilly-sur-Seine, Dec. 1986.
15. Hartman, B. O., ed., *Higher Mental Functioning in Operational Environments*, AGARD-CP-181, Neuilly-sur-Seine, Apr. 1986.
16. Patton, R. M., Tanner, T. A., Jr., and Swets, J. A., *Applications of Research on Human Decisionmaking*, NASA SP-209, Wash., DC, 1970.
17. Perkins, C. D., "Development of Airplane Stability and Control Technology," *J. Aircraft*, Vol. 7, No. 4, Jul.-Aug., 1970, pp. 290-301.
18. McRuer, D., Ashkenas, I., and Graham, D., *Aircraft Dynamics and Automatic Control*, Princeton U. Press, Princeton, 1973.
19. Blakelock, J. H., *Automatic Control of Aircraft and Missiles*, J. Wiley & Son, New York, 1991.
20. McLean, D., *Automatic Flight Control Systems*, Prentice-Hall, NY, 1990.
21. Bryson, A. E., Jr., Desai, M. N., and Hoffman, W. L., "The Energy State Approximation in Performance Optimization of Supersonic Aircraft," *J. Aircraft*, Vol. 6, No. 6, Nov.-Dec., 1969, pp. 481-487.
22. Erzberger, H., McLean, J. D., and Barman, J. F., "Fixed-Range Optimum Trajectories for Short Haul Aircraft, NASA TN D-8115, Washington, DC, Dec. 1975.
23. Stengel, R. F., and Marcus, F. J., "Energy Management for Fuel Conservation in Transport Aircraft," *IEEE Trans. Aero. Elec. Sys.*, Vol. AES-12, No. 4, July 1976, pp. 464-470.
24. Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA, New York, 1987.
25. Widnall, W. S., "Lunar Module Digital Autopilot, *J. Space Rockets*, Vol. 8, No. 1, Jan. 1971, pp. 56-62.
26. Stengel, R. F., "Manual Attitude Control of the Lunar Module," *J. Space Rockets*, Vol. 7, No. 8, Aug 1970, pp. 941-948.
27. Lambregts, A. A., "Integrated System Design for Flight and Propulsion Control Using Total Energy Principles," AIAA Paper No. 83-2561, New York, Oct. 1983.
28. Wagner, S. M., and Rothstein, S. W., " Integrated Control and Avionics for Air Superiority: Computational Aspects of Real-Time Flight Management," *AIAA Guid., Nav., Cont. Conf.*, Boston, Aug. 1989, pp. 321-326.
29. Harris, W. H., and Levey, J. S., ed., *New Columbia Desk Encyclopedia*, Columbia U. Press, New York, 1975.
30. Sternberg, R. J., "Human Intelligence: The Model is the Message," *Science*, Vol. 230, Dec. 6, 1985, pp. 1111-1118.
31. Searle, J. R., "Is the Brain's Mind a Computer Program?" *Scient. Amer.*, Vol. 262, No. 1, Jan. 1990, pp. 26-31.
32. Penrose, R., *The Emperor's New Mind*, Penguin Books, New York, 1989.
33. Kolata, G., "Does Gödel's Theorem Matter to Mathematics?" *Science*, Vol. 218, Nov. 19, 1982, pp. 779-780.
34. Churchland, P. M., and Churchland, P. S., "Could a Machine Think?" *Scient. Amer.*, Vol. 262, No. 1, Jan. 1990, pp. 32-37.
35. Minsky, M., "Re: penrose," NetNews communication, May 13, 1992.
36. Turing, A., "Computing Machinery and Intelligence," *Mind*, Vol. 59, Oct. 1950, pp. 433-460.
37. Kihlstrom, J. F., "The Cognitive Unconscious," *Science*, Vol. 237, Sept. 18, 1987, pp. 1445-1452.
38. Anderson, J. R., *Language, Memory, and Thought*, Erlbaum, Hillsdale, NJ, 1976.
39. Hinto, D. E., and Anderson, J. A., *Parallel Models of Associative Memory*, Erlbaum, Hillsdale, NJ, 1981.
40. Mitchell, J., ed., *The Random House Encyclopedia*, Random House, New York, 1990.
41. Tulving, E., and Schacter, D. L., "Priming and Human Memory Systems," *Science*, Vol. 247, Jan. 10, 1990, pp. 301-306.
42. Blakeslee, S., "Scientists Unraveling Chemistry of Dreams," The New York Times, Jan. 7, 1992, pp. C1,C10.
43. Stengel, R. F., "Intelligent Failure-Tolerant Control," *IEEE Cont. Sys. Mag.*, Vol. 11, No. 4, June 1991, pp. 14-23.
44. Waldrop, M. M., "Causality, Structure, and Common Sense," *Science*, Vol. 237, Sept. 11, 1987, pp. 1297-1299.
45. Cohen, P. R., and Feigenbaum, E. A., *The Handbook of Artificial Intelligence*, William Kaufmann, Los Altos, CA, 1982.
46. Charniak, E., and McDermott, D., *Introduction to Artificial Intelligence*, Addison-Wesley, Reading, MA, 1985.
47. Handelman, D. A., and Stengel, R. F., "Combining Expert System and Analytical Redundancy Concepts for Fault-Tolerant Flight Control," *J. Guid., Cont., Dyn.*, Vol. 12, No. 1, Jan.-Feb. 1989, pp. 39-45.
48. Handelman, D. A., and Stengel, R. F., "An Architecture for Real-Time Rule-Based Control," *Proc. Amer. Cont. Conf.*, Minneapolis, MN, June 1987, pp. 1636-1642.
49. Huang, C. Y., and Stengel, R. F., "Failure Model Determination in a Knowledge-Based Control System," *Proc. Amer. Cont. Conf.*, Minneapolis, MN, June 1987, pp. 1643-1648.
50. Frankovich, K., Pedersen, K., and Bernsteen, S., "Expert System Applications to the Cockpit of the '90s," *IEEE . Aero. Elec. Sys. Mag.*, Jan. 1986, pp. 13-19.
51. Belkin, B. L., and Stengel, R. F., "Cooperative Rule-Based Control Systems for Aircraft Navigation and Control," *Proc. IEEE Conf. Dec. Cont.*, Los Angeles, Dec. 1987, pp. 1934-1940.
52. Belkin, B. L., and Stengel, R. F., "Systematic Methods for Knowledge Acquisition and Expert System Development, "*IEEE Aero. Elec. Sys. Mag.*, Vol. 6, No. 6, June 1991, pp. 3-11.
53. Belkin, B. L., and Stengel, R. F., "AUTOCREW: A Paradigm for Intelligent Flight Control," to appear in *An Introduction to Intelligent and Autonomous Control*, P. Antsaklis and K. Passino, ed., Kluwer, Norwell, MA.
54. Maxwell, K. J., Davis, J. A., "Artificial Intelligence Implications for Advanced Pilot/Vehicle Interface Design," AIAA 84-2617, 1984.
55. Leavitt, C. A., and Smith, D. M., "Integrated Dynamic Planning in the Pilot's Associate," *Proc. AIAA Guid., Nav., Cont. Conf.*, Boston, Aug. 1989, pp. 327-331.

56. Broadwell, M., Jr., and Smith, D. M., "Associate Systems Technology Issues," *Proc. AIAA Guid., Nav., Cont. Conf.*, New Orleans, Aug. 1991, pp. 1451-1457.
57. Miao, X., Luh, P. B., Kleinman, D. L., and Castanon, D. A., "Distributed Stochastic Resource Allocation in Teams," *IEEE Trans. Syst., Man, Cyber.*, Vol. 21, No. 1, Jan.-Feb. 1991, pp. 61-69.
58. Kapsouris, P., Serfaty, D., Deckert, J. C., Wohl, J. G., and Pattipati, K. R., Resource Allocation and Performance Evaluation in Large Human-Machine Organizations," *IEEE Trans. Syst., Man, Cyber.*, Vol. 21, No. 3, May-June. 1991, pp. 521-531.
59. Mallubhatla, R., Pattipati, K. R., Kleinman, D. L., and Tang, Z. B., "A Model of Distributed Team Information Processing under Ambiguity," *IEEE Trans. Syst., Man, Cyber.*, 21 (4), July-Aug 1991, pp. 713-725.
60. Wagner, E., "On-Board Automatic Aid and Advisory for Pilots of Control-Impaired Aircraft," *Proc. AIAA Guid., Nav., Cont. Conf.*, Boston, Aug. 1989, pp. 306-320.
61. Anderson, B. M., Cramer, N. L., Lineberry, M., Lystad, G. S., and Stern, R. C., "Intelligent Automation of Emergency Procedures in Advanced Fighter Aircraft," *Proc. IEEE Conf. Art. Intell. App.*, Silver Spring, MD, Dec. 1984, pp. 496-501.
62. Berning, S., Glasson, D. P., and Pomarede, J. L., "Knowledge Engineering for the Adaptive Tactical Navigator," *Proc. IEEE Nat'l. Aero. Elec. Conf.*, Dayton, May 1988, pp. 1266-1273.
63. Stengel, R. F., *Stochastic Optimal Control: Theory and Applications*, J. Wiley & Sons, New York, 1986.
64. Ng, K.-C., and Abramson, B., "Uncertainty Management in Expert Systems," *IEEE Expert*, Vol. 5, No. 2, Apr. 1990, pp. 29-47.
65. Pearl, J., *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, Palo Alto, CA, 1988.
66. Stratton, D. A., and Stengel, R. F., "Probabilistic Reasoning for Intelligent Wind Shear Avoidance," *J. Guid., Cont., Dyn.*, Vol. 15, No. 1, Jan-Feb 1992, pp. 247-254.
67. Stratton, D. A., and Stengel, R. F., "Real-Time Decision Aiding: Aircraft Guidance for Wind Shear Avoidance," AIAA 92-0290, Reno, Jan. 1992.
68. Niehaus, A., and Stengel, R. F., "An Expert System for Automated Highway Driving," *IEEE Cont. Sys. Mag.*, 11 (3), Apr 1991, pp. 53-61.
69. Niehaus, A., and Stengel, R. F., "Rule-Based Guidance for Vehicle Highway Driving in the Presence of Uncertainty, " *Proc. Amer. Cont. Conf.*, Boston, June 1991, pp. 3119-3124.
70. Niehaus, A., and Stengel, R. F., "Probability-Based Decision Making for Automated Highway Driving, " *Proc. Veh. Nav. Info. Sys. '91 Conf.*, SAE 912869, Dearborn, MI, Oct. 1991, pp. 1125-1136.
71. Perez, M., Gemoets, L., and McIntyre, R. G., "Knowledge Extraction Methods for the Development of Expert Systems," *Knowledge Based System Applications for Guidance and Control*, AGARD-CP-474, Apr. 1991, pp. 26-1 to 26-10.
72. Ryan, P. M., and Wilkinson, A. J., "Knowledge Acquisition for ATE Diagnosis," *IEEE Aero. Elec. Sys. Mag.*, July 1986, pp. 5-12.
73. Weiss, S. M., and Kulikowski, C. A., *Computer Systems That Learn*, Morgan Kaufmann, San Mateo, CA, 1991.
74. Safavian, S. R., and Landgrebe, D., "A Survey of Decision Tree Classifier Methodology," *IEEE Trans. Syst., Man, Cyber.*, Vol. 21, No. 3, May-June 1991, pp. 660-674.
75. Handelman, D. A., and Stengel, R. F., "Rule-Based Mechanisms of Learning for Intelligent Adaptive Flight Control," *Proc. Amer. Cont. Conf.*, Atlanta, June 1988, pp. 208-213.
76. Tong, R., "A Control Engineering Review of Fuzzy Systems," *Automatica*, Vol. 13, No. 6, Nov. 1977, pp. 559-569.
77. Quinlan, J. R., "Discovering Rules by Induction from Large Collections of Samples," in *Expert Systems in the Micro Electronic Age*, D. Michie, ed., Edinburgh U. Press, Edinburgh, 1979, pp. 169-201.
78. Thompson, B., and Thompson, W., "Finding Rules in Data," *Byte*, Nov. 1986, pp. 149-158.
79. Durkin, J., "Induction... ," *AI Expert*, Apr. 1992, pp. 48-53.
80. Huang, C. Y., and Stengel, R. F., "Restructurable Control Using Proportional-Integral Implicit Model Following," *J. Guid., Cont., Dyn.*, Vol. 13, No. 2, Mar.-Apr. 1990, pp. 303-309.
81. Stengel, R. F., Broussard, J. R., and Berry, P., "Digital Flight Control Design for a Tandem-Rotor Helicopter," *Automatica*, Vol. 14, No. 4, July 1978, pp. 301-311.
82. Foxgrover, J. A., *Design and Flight Test of a Digital Flight Control System for General Aviation Aircraft*, Princeton U. M.S.E. Thesis, MAE 1559-T, June 1982.
83. Broussard, J. R., "Design, Implementation and Flight Testing of PIF Autopilots for General Aviation Aircraft," NASA CR-3709, Washington, DC, July 1983.
84. Garcia, C. E., and Morari, M., "Internal Model Control. 1. A Unifying Review and Some New Results," *I&EC Proc. Des. & Devel.*, Vol. 21, 1982, pp. 308-323.
85. Kalman, R. E., "When is a Linear Control System Optimal?" *ASME J. Basic Eng.*, Vol. 86, Mar 1964, pp. 51-60.
86. Anderson, B. D. O., and Moore, J. B., *Linear Optimal Control*, Prentice Hall, Englewood Cliffs, NJ, 1971.
87. Lehtomaki, N. A., Sandell, N. R., and Athans, M., "Robustness Results in Linear-Quadratic-Gaussian Based Multivariable Control Designs," *IEEE Trans. Auto. Cont.*, Vol. AC-26, No. 1, Feb. 1981, pp. 75-93.
88. Doyle, J. C., "Guaranteed Margins for LQG Regulators," *IEEE Trans. Auto. Cont.*, Vol. AC-23, No. 4, Aug 1978, pp. 756-757.
89. Doyle, J. C., and Stein, G., "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," *IEEE Trans. Auto. Cont.*, Vol. AC-26, No. 1, Feb 1981, pp. 4-16.
90. Safonov, M. G., Laub, A. J., and Hartmann, G. L., "Feedback Properties of Multivariable Systems: The Role and Use of the Return Difference Matrix," *IEEE Trans. Auto. Cont.*, AC-26 (1), Feb 1981, pp. 47-65.
91. Dorato, P., ed., *Robust Control*, IEEE Press, New York, 1987.
92. Dorato, P., and Yedavalli, R. K., ed., *Recent Advances in Robust Control*, IEEE Press, New York, 1990.
93. Doyle, J. C., "Analysis of Feedback Systems with Structured Uncertainties," *IEE Proc.*, Vol. 129, Part D, No. 6, pp. 242-250, Nov. 1982.
94. Stengel, R. F., "Some Effects of Parameter Variations on the Lateral-Directional Stability of Aircraft," *J. Guid., Cont.*, Vol. 3, No. 2, Apr. 1980, pp. 124-131.
95. Stengel, R. F., and Ryan, L., "Stochastic Robustness of Linear-Time-Invariant Control Systems," *IEEE Trans. Auto. Cont.*, Vol. 36, No. 1, Jan. 1991, pp. 82-87.
96. Ray, L. R., and Stengel, R. F., "Application of Stochastic Robustness to Aircraft Control," *J. Guid., Cont., Dyn.*, Vol. 14, No. 6, Nov.-Dec. 1991, pp. 1251-1259.
97. Stengel, R. F., and Marrison, C. I., "Robustness of Solutions to a Benchmark Control Problem," *Proc. Amer. Cont. Conf.*, Boston, June 1991, pp. 1915-1916. (to appear *J. Guid., Cont., Dyn.*)
98. Stengel, R. F., and Marrison, C. I., "Stochastic Robustness Synthesis for a Benchmark Problem," *Proc. Amer. Cont. Conf.*, Chicago, June 1992.
99. Stengel, R. F., Berry, P. W., and Broussard, J. R., "Command Augmentation Control Laws for Maneuvering Aircraft," AIAA 77-1044, New York, Aug 1977.
100. Sugeno, M. "An Introductory Survey of Fuzzy Control," *Info. Sci.*, Vol. 36, 1985, pp. 59-83.
101. Chand, S., and Chiu, S., "Robustness Analysis of Fuzzy Control Systems with Application to Aircraft Roll Control," *Proc. Guid., Nav., Cont. Conf.*, New Orleans, Aug. 1991, pp. 1676-1679.
102. Steinberg, M., "Potential Role of Neural Networks and Fuzzy Logic in Flight Control Design and Development," AIAA 92-0999, Washington, DC, Feb. 1992.
103. Singh, S. N., and Rugh, W. J., "Decoupling in a Class of Nonlinear Systems by State Feedback, *ASME J. Dyn. Syst. Meas. Cont.*, Series G, Vol. 94, Dec. 1972, pp. 323-329.
104. Isidori, A., *Nonlinear Control Systems*, Springer-Verlag, Berlin, 1989.
105. Lane, S. H., and Stengel, R. F., "Flight Control Design Using Non-linear Inverse Dynamics," *Automatica*, 24 (4), July 1988, pp. 471-483.
106. Linse, D. J., and Stengel, R. F., "A System Identification Model for Adaptive Nonlinear Control," *Proc. Amer. Cont. Conf.*, Boston, June 1991, pp. 1752-1757.
107. Linse, D. J., and Stengel, R. F., "Identification of Aerodynamic Coefficients Using Computational Neural Networks," AIAA 92-0172, Washington, DC, Jan. 1992.
108. Sentoh, E., and Bryson, A. E., Jr., "Inverse and Optimal Control for Desired Outputs," *J. Guid., Cont., Dyn.*, Vol. 15, No. 3, May-June 1992, pp. 687-691.
109. Poggio, T., and Girosi, F., "Regularization Algorithms for Learning That Are Equivalent to Multilayer Networks," *Science*, Vol. 247, No. 4945, Feb 23, 1990, pp. 978-982.
110. Linse, D. J., and Stengel, R. F., "Neural Networks for Function Approximation in Nonlinear Control, *Proc. Amer. Cont. Conf.*, San Diego, May 1990, pp. 675-679.
111. Sethi, I. K., "Entropy Nets: From Decision Trees to Neural Networks," *Proc. IEEE*, Vol. 78, No. 10, Oct. 1990, pp. 1605-1613.
112. Funahashi, K.-I., "On the Approximate Realization of Continuous Mappings by Neural Networks," *Neural Networks*, Vol. 2, 1989, pp. 183-192.
113. Cybenko, G., "Approximation by Superposition of a Sigmoidal Function," *Math. Cont., Sig., Sys.*, Vol. 2, No. 4, 1989, pp. 303-314.
114. Holcomb, T., and Morari, M., "Local Training for Radial Basis Function Networks: Towards Solving the Hidden Unit Problem," *Proc. Amer. Cont. Conf.*, June 1991, pp. 2331-2336.
115. Cox, M. G., "Data Approximation by Splines in One and Two Independent Variables," in *The State of the Art in Numerical Analysis*, A. Iserles and M. J. D. Powell, ed., Clarendon, Oxford, 1987, pp. 111-138.
116. Lane, S. H., Flax, M. G., Handelman, D. A., and Gelfand, J. J., "Multi-Layered Perceptrons with B-Spline Receptive Field Functions," to appear in *Neural Information Processing Systems*, Morgan Kaufmann, Palo Alto.

117. Albus, J. S., "A New Approach to Manipulator Control: The Cerebellar Model Articulation Controller (CMAC)," *ASME J. Dyn. Sys., Meas., Cont.*, Vol. 97, Sept. 1975, pp. 220-227.
118. Miller, W. T., "Sensor-Based Control of Robotic Manipulators Using a General Learning Algorithm," *J. Robot. Auto.*, Vol. RA-3, No. 2, Apr. 1987, pp. 157-165.
119. Lane, S. H., Handelman, D. A., and Gelfand, J. J., "Theory and Development of Higher-Order CMAC Neural Networks," *IEEE Cont. Sys. Mag.*, Vol. 12, No. 3, Apr. 1992, pp. 23-30.
120. Kosko, B., "Bidirectional Associative Memories," *IEEE Trans. Syst., Man, Cyber.*, Vol. 18, No. 1, Jan.-Feb. 1988, pp. 49-60.
121. Hopfield, J. J., "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proc. Nat'l. Acad. Sci.*, Vol. 79, Apr. 1982, pp. 2554-2558.
122. Cohen, M. A., and Grossberg, S., "Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks," *IEEE Trans. Syst., Man, Cyber.*, Vol. SMC-13, No. 5, Sept.-Oct. 1983, pp. 815-826.
123. Kohonen, T., "Adaptive, Associative, and Self-Organizing Functions in Neural Computing," *Appl. Opt.*, Vol. 26, No. 23, Dec. 1987, pp. 4910-4918.
124. Naidu, S., Zafiriou, E., and McAvoy, T., "Use of Neural Networks for Sensor Failure Detection in a Control System," *IEEE Cont. Sys. Mag.*, Vol. 10, No. 3, Apr. 1990, pp. 49-55.
125. Watanabe, K., *et al*, "Incipient Fault Diagnosis of Chemical Processes via Artificial Neural Networks," *AIChE J.*, Vol. 35, No. 11, Nov. 1989, pp. 1803-1812.
126. Sorsa, T., Koivo, H. N., and Koivisto, H., "Neural Networks in Process Fault Diagnosis," *IEEE Trans. Syst., Man, Cyber.*, Vol. 21, No, 4, July-Aug. 1991, pp. 815-825.
127. Chow, E. Y., and Willsky, A. S., "Analytical Redundancy and the Design of Robust Failure Detection Systems," *IEEE Trans. Auto. Cont.*, Vol. AC-29, No. 7, July 1984, pp. 603-614.
128. Kohonen, T., "The Self-Organizing Map," *Proc. IEEE*, Vol. 78, No. 9, Sept. 1990, pp. 1464-1480.
129. Rumelhart, D., Hinton, G., and Williams, R., "Learning Internal Representations by Error Propagation," *Parallel Distributed Processing: Explorations in the Microstructure of Cognitions, Vol. 1: Foundations*, D. Rumelhart and J. McClelland, ed., MIT Press, Cambridge, 1986.
130. Werbos, P. J., "Backpropagation Through Time: What It Does and How to Do It," *Proc. IEEE*, Vol. 78, No. 10, Oct. 1990, pp. 1550-1560.
131. Singhal, S., and Wu, L., "Training Feed-Forward Networks with the Extended Kalman Algorithm," *Proc. Int'l. Conf. Acous., Speech, Sig. Proc.*, Glasgow, May 1989, pp. 1187-1190.
132. Levin, E., Tishby, N., and Solla, A. A., "A Statistical Approach to Learning and Generalization in Layered Neural Networks," *Proc. IEEE*, Vol. 78, No. 10, Oct. 1990, pp. 1568-1574.
133. Davis, L., ed., *Genetic Algorithms and Simulated Annealing*, Morgan Kaufmann, Palo Alto, 1987.
134. Bilbro, G. L., and Snyder, W. E., "Optimization of Functions with Many Minima," *IEEE Trans. Syst., Man, Cyber.*, Vol. 21, No. 4, July-Aug. 1991, pp. 840-849.
135. Goldenthal, W., and Farrell, J., "Application of Neural Networks to Automatic Control," *Proc. AIAA Guid., Nav., Cont. Conf.*, Portland, OR, Aug. 1990, pp. 1108-1112.
136. Nguyen, D. H., and Widrow, B., "Neural Networks for Self-Learning Control Systems," *IEEE Cont. Sys. Mag.*, Vol. 10, No. 3, Apr. 1990, pp. 18-23.
137. Narendra, K. S., and Parthasarathy, K., "Identification and Control of Dynamical Systems Using Neural Networks," *IEEE Trans. Neural Networks*, Vol. 1, No. 1, Mar. 1990, pp. 4-27.
138. Fadali, M., *et al*, "Minimum-Time Control of Robotic Manipulators Using a Back Propagation Neural Network," *Proc. Amer. Cont. Conf.*, San Diego, May 1990, pp. 2997-3000.
139. Nagata, S., Sekiguchi, M., and Asakawa, K., "Mobile Robot Control by a Structured Hierarchical Neural Network," *IEEE Cont. Sys. Mag.*, Vol. 10, No. 3, Apr. 1990, pp. 69-76.
140. Troudet, T., Garg, S., and Merrill, W. C., "Neural Network Application to Aircraft Control System Design," *Proc. Guid., Nav., Cont. Conf.*, Aug. 1991, pp. 993-1009.
141. Gu, Y.-L., "On Nonlinear System Invertibility and Learning Approaches by Neural Networks," *Proc. Amer. Cont. Conf.*, San Diego, May 1990, pp. 3013-3017.

X-29: LONGITUDINAL INSTABILITY AT HIGH ANGLE-OF-ATTACK

by

Dr Lawrence A. Walchli
Wright Laboratory
Wright-Patterson AFB OH 45433-6553
United States

ABSTRACT

Relaxed static stability (RSS) was chosen as one of the primary technologies to be flight demonstrated on the forward swept wing X-29 aircraft. Development experiences and performance benefits of this technology in the high angle-of-attack (AOA) regime of flight are described. Flight test results validate the X-29's wind tunnel database and the updated piloted simulation is used for parameter variations to thoroughly explore the potential performance of an aircraft with high levels of static instability.

INTRODUCTION

The X-29 integrates several different technologies into one airframe as depicted in Figure 1. The aeroelastically tailored composite wing covers cause the forward swept wing to twist as it deflects, successfully delaying wing divergence. The thin supercritical airfoil, coupled with the discrete variable camber produced by the double-hinged full span flaperons, provide optimum wing performance at all flight conditions. The aircraft was designed to be 35 percent statically unstable by adding a close-coupled, variable incidence canard, without which the wing-body combination would be near-neutrally stable. The canard, which has an area about 20 percent of the wing area, produces lift and its downwash delays flow separation at the wing root. The three-surface pitch control--the canard, flaperon, and strake flap--is used by the digital fly-by-wire flight control system to control an otherwise unflyable unstable vehicle. The success of the X-29 really rests with the integration of these technologies into a single synergistic configuration built for drag reduction in turning flight.

Two X-29 aircraft were designed and built. The first entered flight testing in December 1984 and concluded in December 1988, completing 242 flights and over 200 flight hours. A primary objective of Ship #1 testing was to validate, evaluate, and quantify the benefits of RSS at subsonic and supersonic speeds below 20 degrees AOA. Ship #2, which was modified to allow high AOA testing, began flying in May 1989. Its spin chute was designed to assist the pilot in regaining control in the event of a departure from controlled flight. Control surface tutorial lights mounted in the cockpit assist in this task. The flight control system software was significantly modified in order to best utilize the various surfaces in controlling this highly unstable aircraft in a post-stall environment. One g envelope expansion was completed to 67 degrees AOA and ten degrees sideslip. Accelerated entry high AOA expansion allowed all-axis maneuvering to 45 degrees. The inherent high-lift capability of this unstable configuration resulted in coordinated rolls to instantaneous rates of 70 degrees per second under approximately 2 g conditions at 30 degrees AOA. The military utility of this vehicle is in a class by itself.

AIRCRAFT AND FLIGHT CONTROL SYSTEM DESCRIPTION

Two essentially identical X-29s were designed and built by Grumman Aerospace Corporation, Bethpage, New York. To reduce overall program costs, the Air Force supplied several major components of the aircraft to Grumman. These included the F-5A forebody and nosegear; F-16 main gear, actuators, airframe-mounted accessory drive and emergency power unit; F-18 F404 engine; SR-71 HDP5301 flight control computers; and F-14 accelerometers and rate gyros. Use of these time-proven components also increased the reliability of the flight vehicle.

The X-29 flight control system (FCS) is a triplex digital fly-by-wire system with triplex analog backup (as shown in Figure 2). The fail-op/fail-safe system used MIL-F-8785C and MIL-F-9490D specifications as design guides. Flying quality design goals were Level I for the primary digital mode and Level II for the analog back-up mode.

Normal aircraft operation is accomplished through the normal digital (ND) mode with its associated functional options such as automatic camber control (ACC), manual camber control (MCC), speed stability, precision approach control, and direct electrical link. ND also contains options in its gain tables for power approach, up-and-away, and degraded operation.

The normal digital mode has a pitch rate control law with gravity vector compensation, driving a discrete ACC system. This mode is gain-scheduled as a function of Mach number and altitude and incorporates a sophisticated redundancy management system allowing fail-op/fail-safe flight. MCC is a pilot-selected, fixed flaperon sub-mode of ND normally used for landing.

The analog reversion (AR) mode is the back-up flight control system, designed to bring the aircraft safely back to base. The AR mode provides a highly reliable, dissimilar control mode to protect against generic digital control failures. It incorporates functions similar to those of the ND mode. AR contains no longitudinal trim capability or pitch loop gain compensation with dynamic pressure while the aircraft is on the ground. In all other aspects, it performs like the ND control system.

The Ship #2 flight control laws were modified to permit all-axis maneuvering to 40 degrees AOA, and pitch-only maneuvering to as high as 70 degrees AOA. Below 10 degrees, the control laws are identical to those last flown on Ship #1. Between 10 and 20 degrees, the high AOA modifications are faded in until above 20 degrees they are fully functional.

The high AOA changes are fairly simple. A spin prevention logic is active above 50 degrees or below minus 25 degrees AOA with increasing yaw rate. The logic increases the authority of both the rudder pedals and lateral stick and disconnects all other lateral/directional feedbacks. Besides the spin prevention logic, an aileron-to-rudder interconnect provides for better roll coordination at high AOA. Also assisting in roll coordination is a rate-of-sideslip feedback to the rudder. Since substantial wing rock was predicted for the X-29 above 30 degrees angle-of-attack, a high gain roll rate-to-aileron feedback loop has been added to compensate for the unstable rolling moment coefficient due to roll rate. For a more detailed description of the control system, see Reference 1.

RELAXED STATIC STABILITY (RSS)

Longitudinal static instability has long been recognized as having the potential for improving the overall performance of a high performance fighter aircraft. Subsonically, RSS allows for improved agility through rapid "g" onset and nose pointing. Care must be taken in the design of the FCS to provide enough control power to arrest any motions initiated by the pilot and overcome the inertial coupling of the aircraft, or in other words, design robustness. Trim drag in the supersonic regime is decreased through reduced control surface deflection requirements. Further, in the case of a canarded vehicle such as the X-29, the aircraft requires positive lift to trim at high speed. This in turn can improve the overall supersonic drag polar.

Since static stability in pitch increases with both Mach number and angle-of-attack, it is easy to see the juggling act which a designer faces when designing an aircraft suitable for air superiority as well as high speed penetration. The supersonic regime requires moderate to neutral static stability for best performance.

Figure 3 clearly shows that to maintain a reasonable level of longitudinal stability (positive two percent) during supersonic flight, the low speed range inherits a high level of static instability (negative 35 percent).

It is projected that the air superiority role of future high performance fighters will encompass all areas of the flight envelope including low speed, very high AOA. It is in this region that RSS produces the most serious demands on the aircraft flight control system. Not only does the FCS need to offset the severe instability, but it must also have enough robustness to allow the pilot to demand changes in flight path or orientation as required in combat.

Figure 4 depicts the longitudinal control power typical of a fighter operating at low speed, high angle-of-attack. The respective inflection points for pitch authority occur nominally above 30 degrees AOA for all current fighters. The critical value of nose down pitch authority (Cm*) is a function of flight condition, airframe parameters, and the tendency of the aircraft to inertially couple in pitch.

Because an aircraft operating at high angle-of-attack can suffer a large sink rate, a small value of Cm* may not be sufficient for a timely recovery to low AOA flight. In the event that maximum nose-down pitching moment actually becomes positive, significant problems with departure from controlled flight and/or deep stall occur. Throughout the design process, particular attention is focused on avoiding pitch control deficiencies. Wind tunnel data is used to predict pitch requirements of the flight control system, but often this data is inadequate in the predictive process. FCS robustness is the most satisfactory answer to provide adequate margins for error in control power, surface rates, and bandwidths.

EVOLUTION OF RSS FOR THE X-29

The primary design point for the X-29 was the transonic region of flight. This low AOA condition was used to identify the desired level of relaxed static stability. Figure 5 depicts transonic wind tunnel results from a test dedicated to explore RSS on this close-coupled canard plus forward swept wing configuration. These results suggested an optimum location for the aircraft center of gravity of negative 45 percent of the mean aerodynamic chord of the wing. A margin of safety was added and the FCS designers were faced with stabilizing a 35 percent unstable vehicle. And, as they say, the rest is history!

The longitudinal high AOA control laws (architecture shown in Figure 6) were based on wind tunnel data that predicted the X-29's longitudinal stability and control characteristic trends would be as shown in Figure 7 (as reproduced from Reference 2). The "neutral controls" curve implies that the basic X-29 airframe is unflyable without stability augmentation. The large, all-movable canards and the strake flaps provide powerful pitching moments throughout the flight envelope. The inherent static instability is helpful in initiating a rapid pitch maneuver, and the large pitch authority allows easy capture of a pitch attitude to end the maneuver and re-establish aircraft trim.

Free-flight wind tunnel tests provided information on the dynamics of the X-29. A simplified control law was employed (Reference 2) which used both angle-of-attack and pitch-rate feedback to the canard. The test model showed good stability and controllability over a 13 to 40 degree angle-of-attack range. The level of pitch stability augmentation was varied during the test in order to bound the region of predicted good flying qualities for the X-29.

FLIGHT TEST RESULTS

Stability and Control

The actual nose-up pitching capability of the X-29 proved to be superior to predictions. Because the FCS specifically used the canard in a light to moderate lifting position to maintain maximum nose-down margin, nothing was learned about the maximum nose-up capability of the aircraft. The 1 g pitch-ups to very high angles-of-attack did provide data on the total pitching moment coefficient, and this data showed clearly that the aircraft exhibited more nose-up pitching moment at a given AOA and canard position than was predicted.

The nose-down pitching moment flight data is shown in Figure 8. As with the pitch-up data, there is a general shift in the curves towards more positive Cm's. The model used for predictions was based on high Reynolds number wind tunnel data, about 1.3 million per chord. Flight Reynolds number ranged from about 13 million at 200 KEAS to 5 million at 1 g flight conditions. The flight data was actually better predicted if lower Reynolds number data was used. Presumably, this peculiarity resulted from particular tunnel or model anomalies and has, as yet, not been pursued. The pitch-down flight data clearly shows that the "trouble zone" for Cm* as described in Figure 4 occurred for the X-29. The nose-down pitching moment was less than half of predicted at AOA above 50 degrees at aft c.g.

Following the guidelines of Nguyen and Foster (Reference 3), Figure 9 shows that the X-29's Cm* is less than desirable at very high angles-of-attack while operating at aft center of gravity. For this reason, highest AOA test points were acquired early in a given flight, since the aircraft center of gravity moves aft as fuel is consumed. It must be emphasized here that no loss of control was encountered in flight. It was simply deemed wise not to tempt fate by operating at extreme AOA under heavy weight conditions.

During maneuvering flight, velocity vector rolls produced inertial coupling in the pitch axis. Cm* then had to be treated as the minimum nose-down pitching moment required to overcome <u>all</u> pitch-up moments. With the inertial coupling terms being in the nose-up direction for the X-29, further restrictions were needed for a safe pitch recovery. Figure 10 provides data for coordinated stability axis rolls, assuming the roll and yaw rates as required for coordination. Figure 11 adds increased yaw rates as produced in an uncoordinated roll. Note that for a coordinated roll of 20 degrees per second, the X-29 had enough pitch-down power to overcome the nose-up coupling at centers of gravity ahead of about 452 inches. For additional yaw rate in an uncoordinated 20 degree per second roll, the combined coupling produces a Cm* requirement greater than available at a center of gravity of only 450 inches. From a different perspective, a coordinated 20 degree per second roll required a nose-down moment coefficient of -0.05; a simular uncoordinated roll more than doubled the requirement to -1.3.

From a longitudinal stability and control viewpoint, the X-29 exhibited excellent maneuvering characteristics below 50 degrees AOA, using a simple pitch rate command FCS. The wind tunnel predictions adequately predicted trends in stability and control parameters, but were unsuccessful in predicting accurate Cm* values. The aircraft was operated safely several times to about 55 degrees in 1 g flight, and once to 67 degrees. However, it was deemed prudent to limit maneuvering to AOA below 50 degrees.

Aero Performance

What does RSS do to the lift and drag of the aircraft? To answer this question, a member of the test team, Joe Krumenacker of Grumman, "flew" the X-29 on the flight validated simulator at various levels of instability. The simulator model uses a simplified thrust and ram drag model based on specific General Electric F404-400 data as a function of flight condition and power setting. Figures 12-15 show the simulator results for the three chosen levels of instability. The curves represent trimmed conditions at 0.4 Mach and 30,000 feet, with fixed center gravity. The aircraft was actually flown in the range of -23 to -36 percent static margin.

Figure 12 shows the canard schedule which results from changing the level of RSS. The curves for -15 and -35 percent static margin clearly show that with the chosen mechanization of the three-surface pitch control, the canard is always a lifting surface. But, in the extreme case of -45 percent margin, the canard stops lifting at about 22 degrees AOA because both the flaperon and strake flap have saturated. The canard must then produce a nose-down moment to assure the trim of the aircraft.

The drag coefficients are shown in Figure 13. The lowest drag at this low Mach number condition results from the lowest level of RSS. The trim drag increases with decreasing RSS (Figure 5). However, at -15 percent RSS, neither the strake nor the flap on the X-29 produces much lift; therefore, its overall drag is low. At an RSS of -35 percent, both surfaces saturate trailing edge down by an AOA of 25 degrees in order to balance the aircraft. Induced drag reaches a maximum. For an RSS of -45 percent, the canard stops lifting and induced drag begins to fall off.

The lift curve, Figure 14, shows quite clearly that the optimum configuration for this aircraft is with an RSS of -35 percent. Maximum lift coefficient occurs about 40 degrees AOA.

The drag polar shown in Figure 15 for the negative 35 percent static margin is excellent in both shape and magnitude. Both the flight control laws and the FCS mechanization were successful in providing this aero performance for the X-29.

Can we improve the X-29? Two other members of the test team, Paul Pellicano of Grumman and Bill Gillard of Wright Laboratory, ran a parameter variation study on the simulator to gain insight on the significance of the ACC schedule in determining the aero performance of the X-29. The study was done for the most forward center of gravity flown, 445 inches. This location translates to a negative static margin of 23 percent. With the ACC scheduled as it was, none of the three longitudinal control surfaces was heavily loaded. To increase the aero performance, the flaperon was set to its maximum lifting position, 20 degrees trailing edge down. The strake was then fixed at three positions, zero degrees and 30 degrees up and down. Figure 16 provides the resulting trimmed canard positions. It is intuitively obvious that to force more lift from the flap and strake, the canard must lift more to keep the aircraft balanced, and so the canard schedule is shifted trailing edge down.

Figure 17 shows the significant drag polar improvement achieved by forcing all control surfaces to be more heavily loaded at this forward center of gravity condition. The resulting turn rate capability of the aircraft is significantly enhanced.

CONCLUDING REMARKS

The X-29 has successfully demonstrated that a very large negative static margin can be safely flown on a fighter aircraft and provide the benefits predicted from ground test results. The total pitch-up performance exceeded expectations up to the 1 g, 67 degree AOA condition that was flown. However, this "natural ability" of the aircraft to increase its pitch attitude created some concern above 50 degrees AOA in generating enough nose-down moment to recover to low angle-of-attack. This forced an aft center of gravity prohibition on the aircraft above 50 degrees AOA beyond fuselage station 450 inches (-29 percent static margin). With "only" 29 percent RSS, the X-29 was able to perform coordinated velocity vector rolls at 50 degrees AOA of 20 degrees per second.

Performance studies were done on the flight validated X-29 simulator using both control surface scheduling and level of RSS as variables. The results showed that a negative static margin of about 35 percent was near optimum. However, the ACC canard schedule was designed at a fixed RSS (-30 percent). A more efficient scheduling of the control surfaces would have included adjusting the surface positions as a function of RSS.

REFERENCES

1. Pellicano, et al, "X-29 High Angle-of-Attack Flight Test Procedures, Results, and Lessons Learned," Society of Flight Test Engineers 21st Symposium, August 1990.
2. Croom, Mark A., et al, "High-Alpha Flight Dynamics Research on the X-29 Configuration Using Dynamic Model Test Techniques," Aerospace Technology Conference and Exposition, Anaheim CA, 3-6 October 1988.
3. Nguyen, L. T. and Foster, J. V., "Development of a Preliminary High Angle-of-Attack Nose-Down Pitch Control Requirement for High-Performance Aircraft," NASA Tech Memorandum 101684, NASA Langley Research Center, February 1990.

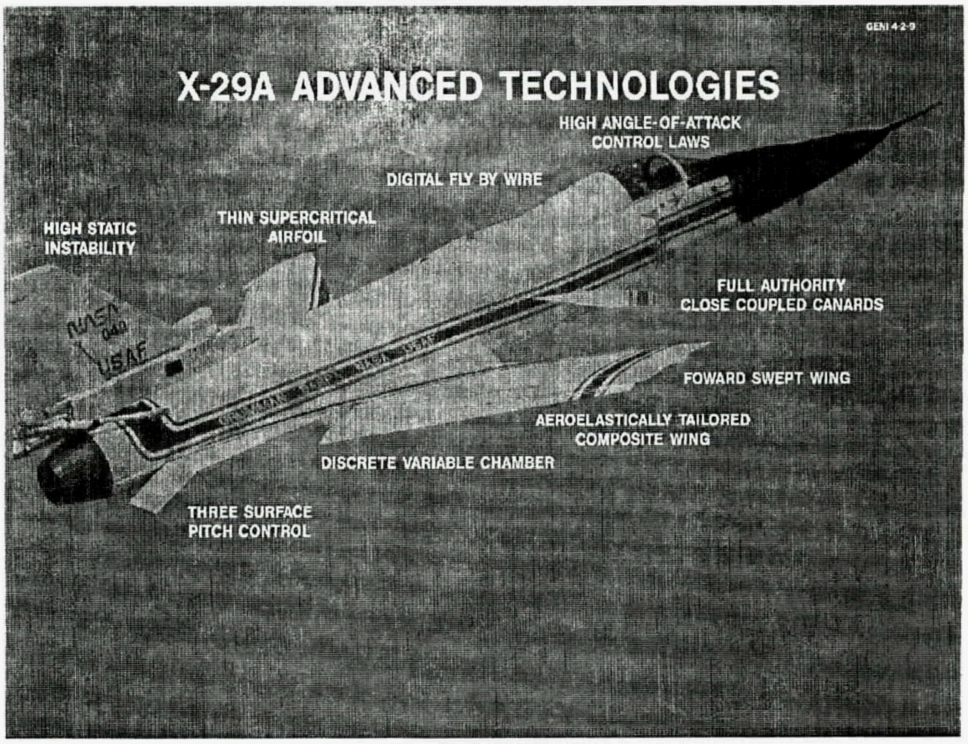


FIGURE 1. X-29 TECHNOLOGIES

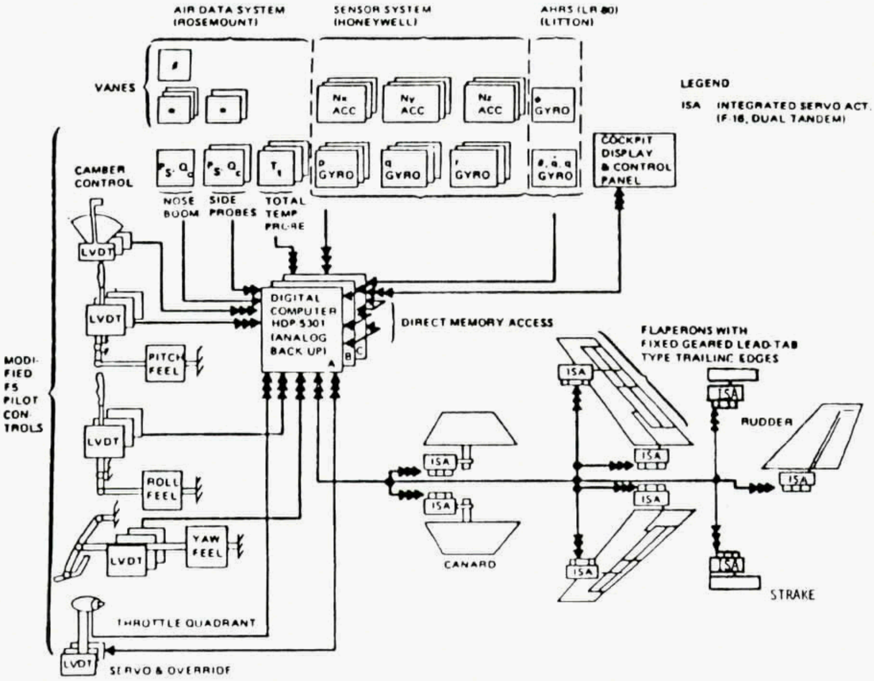


FIGURE 2. X-29 FLIGHT CONTROL SYSTEM

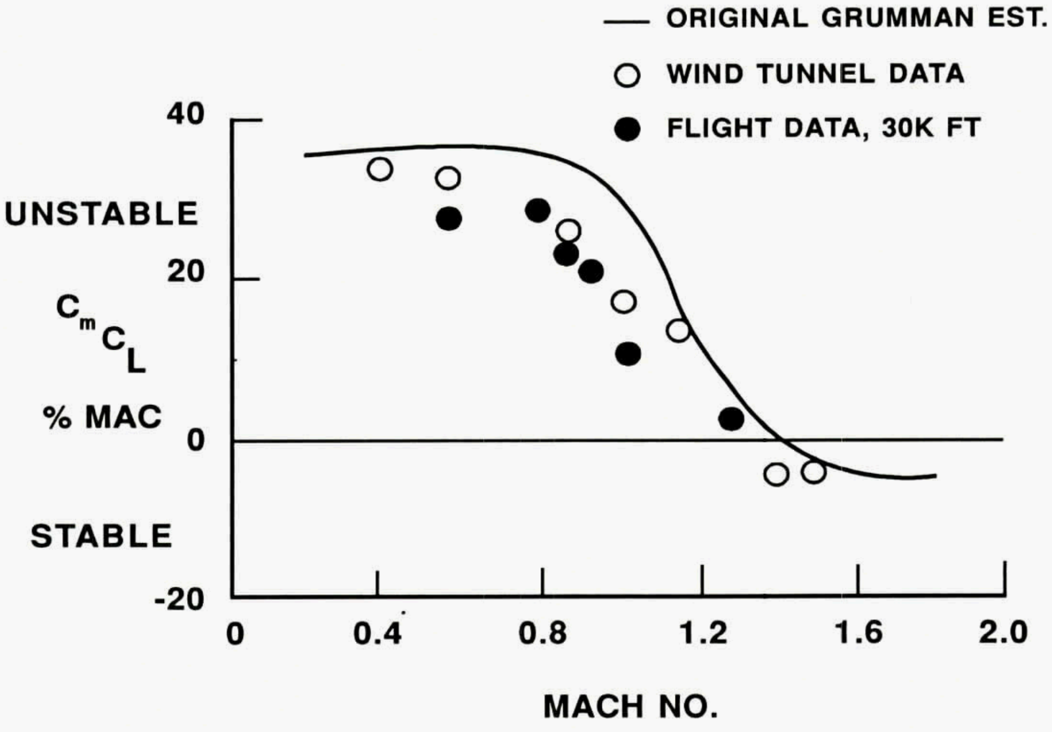


FIGURE 3. X-29 VARIATION OF LONGITUDINAL STATIC STABILITY WITH MACH NO.

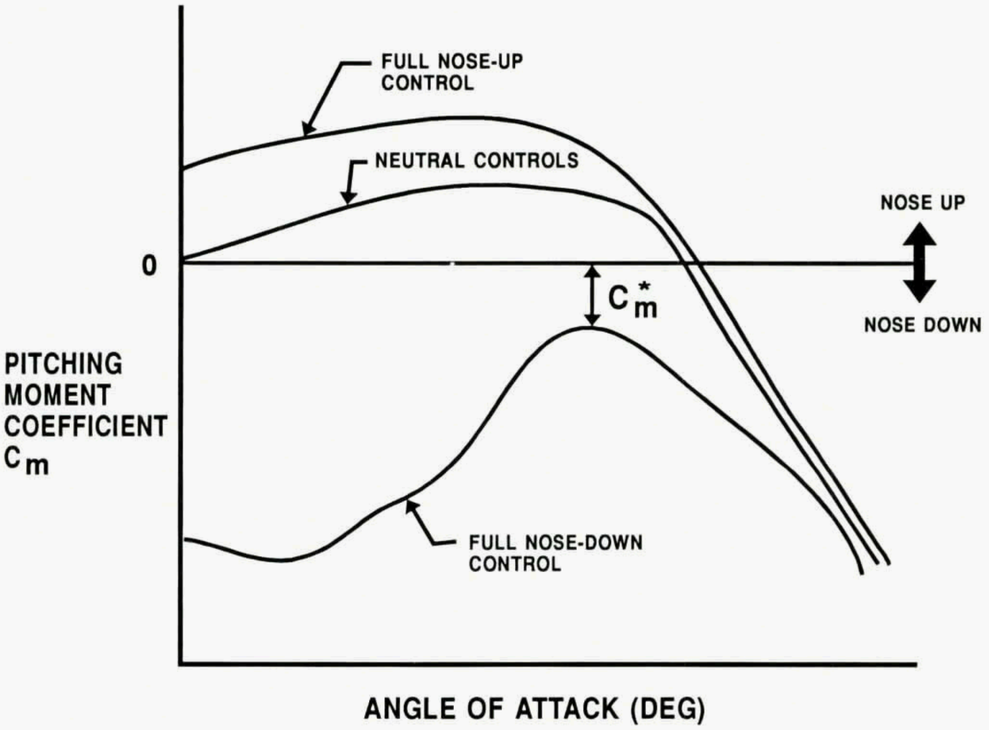


FIGURE 4. GENERIC RSS PITCHING MOMENT CHARACTERISTICS

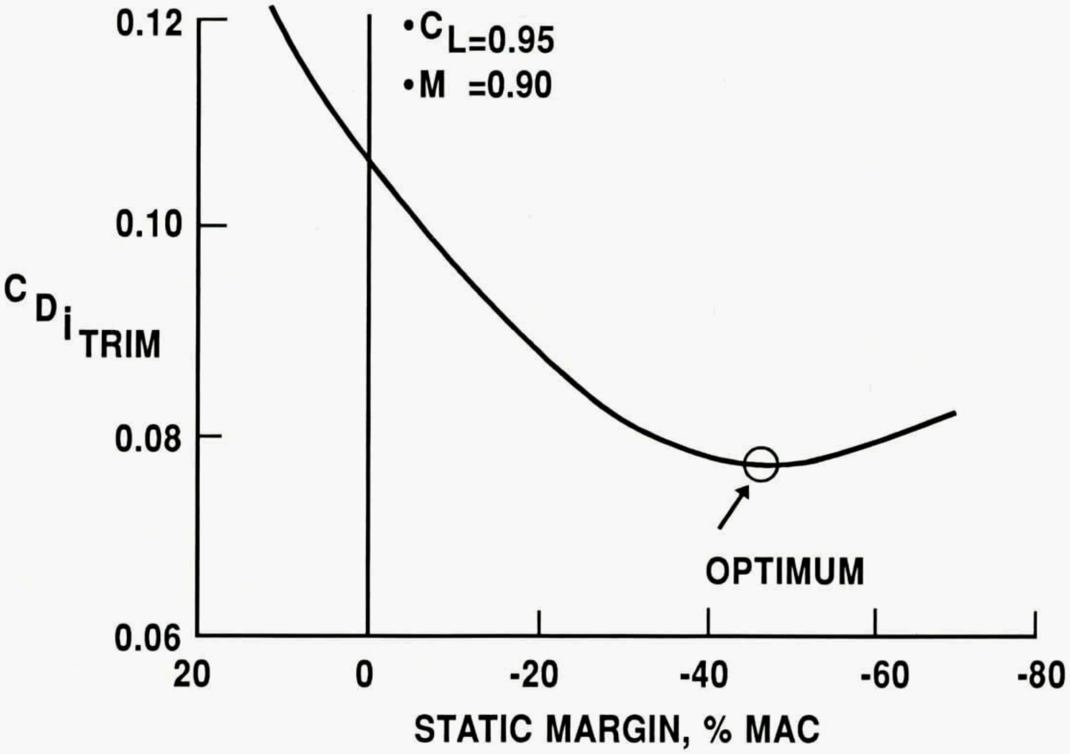


FIGURE 5. EFFECT OF RSS ON MANEUVER TRIM DRAG

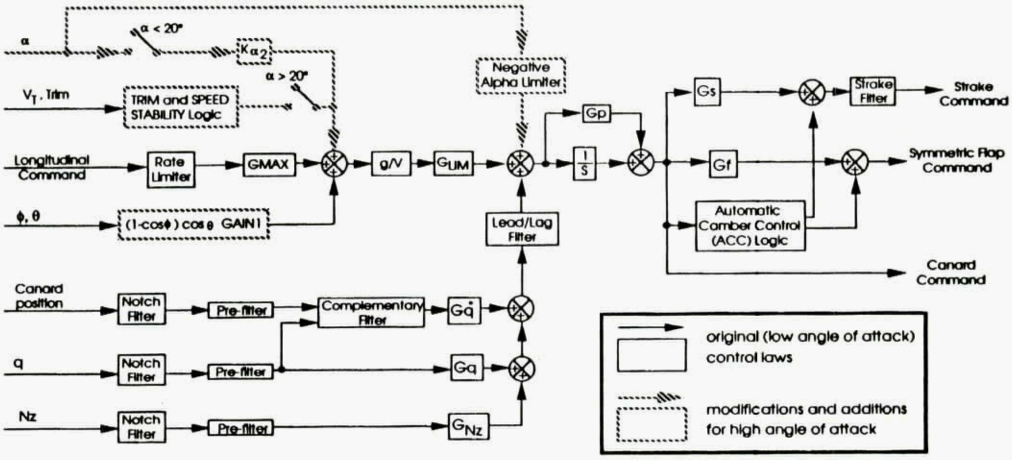


FIGURE 6. X-29 HIGH AOA LONGITUDINAL CONTROL LAW DIAGRAM

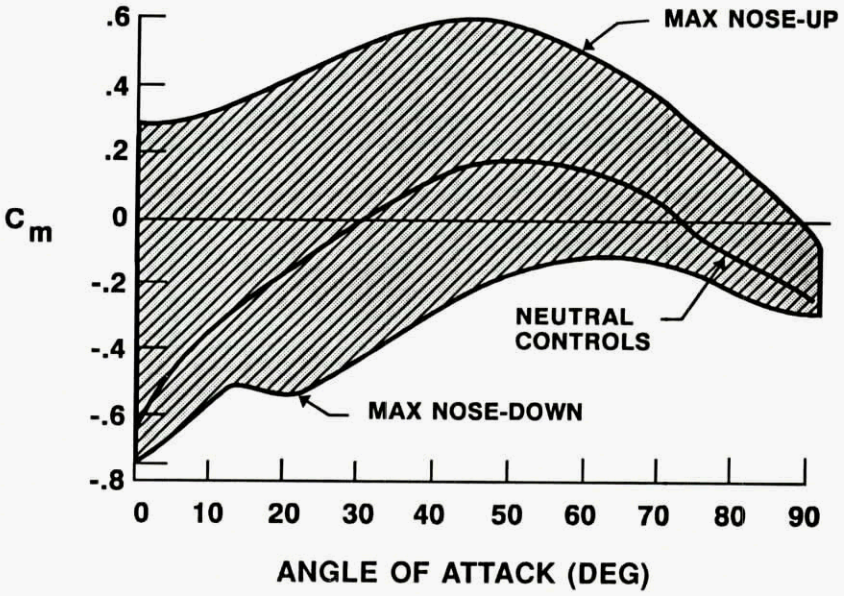


FIGURE 7. X-29 LONGITUDINAL STABILITY CHARACTERISTICS

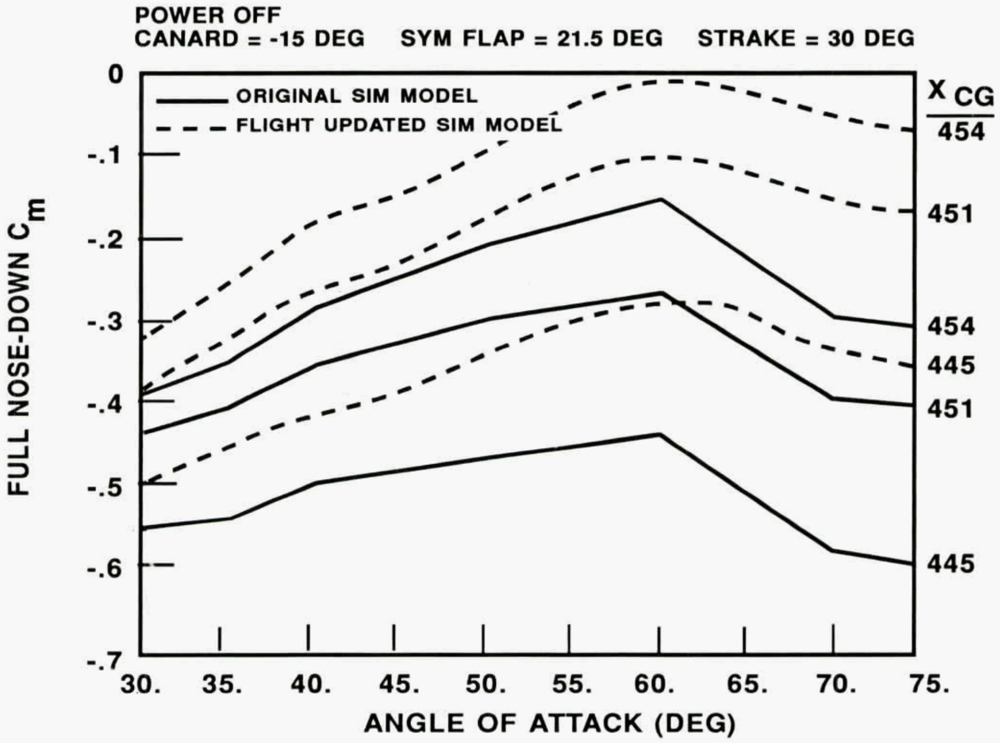


FIGURE 8. X-29 NOSE-DOWN PITCHING MOMENT CAPABILITY

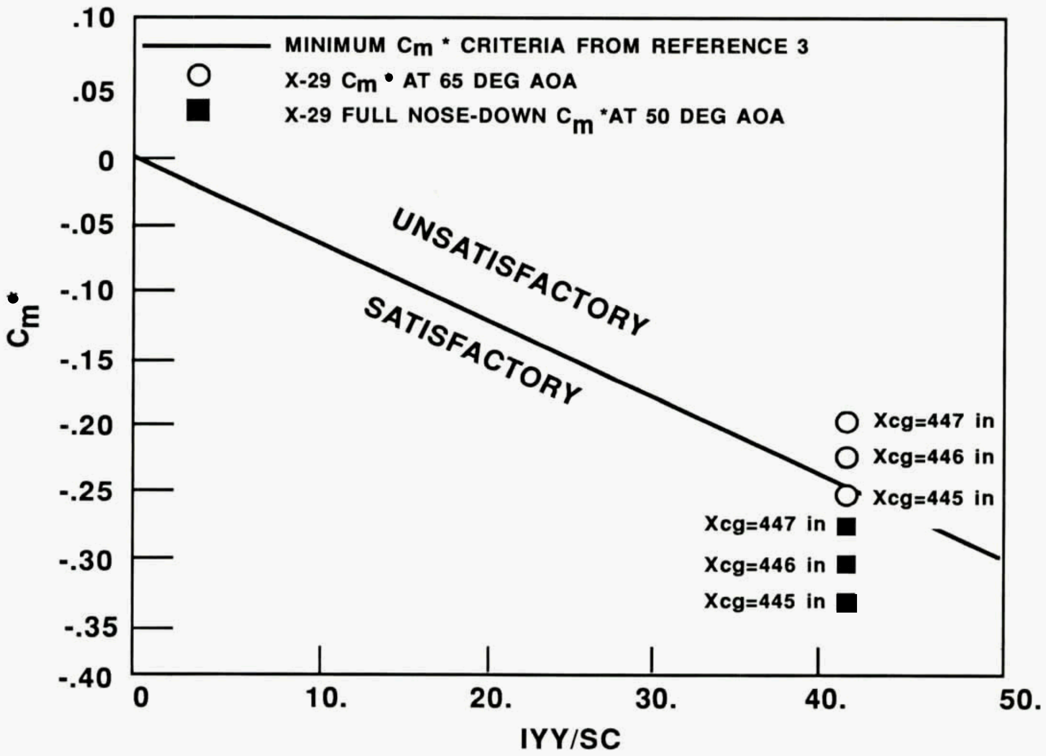


FIGURE 9. HANDLING QUALITY CRITERIA FOR MINIMUM NOSE-DOWN PITCHING MOMENT

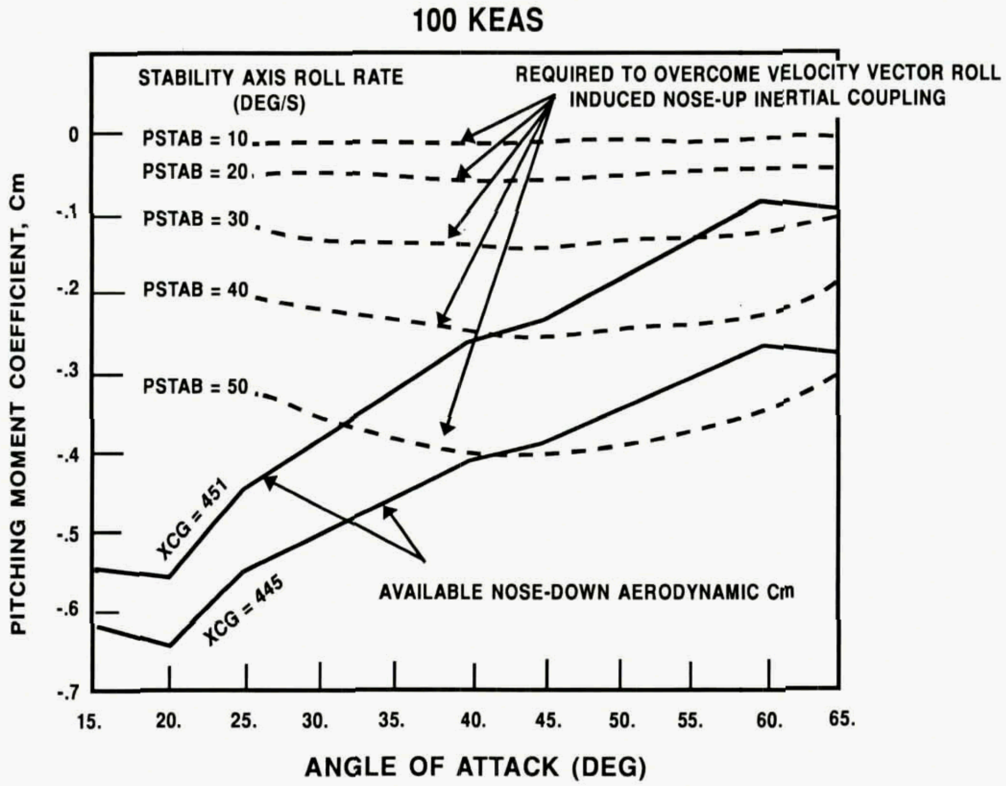


FIGURE 10. INERTIAL COUPLING DURING COORDINATED VELOCITY VECTOR ROLLS

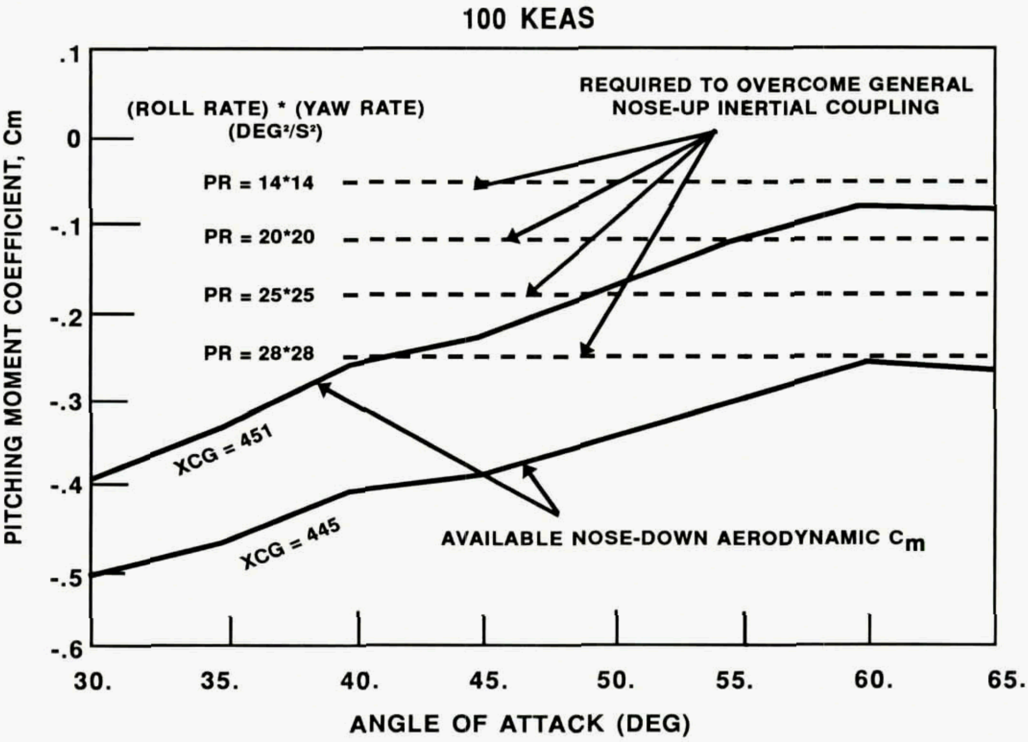


FIGURE 11. INERTIAL COUPLING DURING UNCOORDINATED VELOCITY VECTOR ROLLS

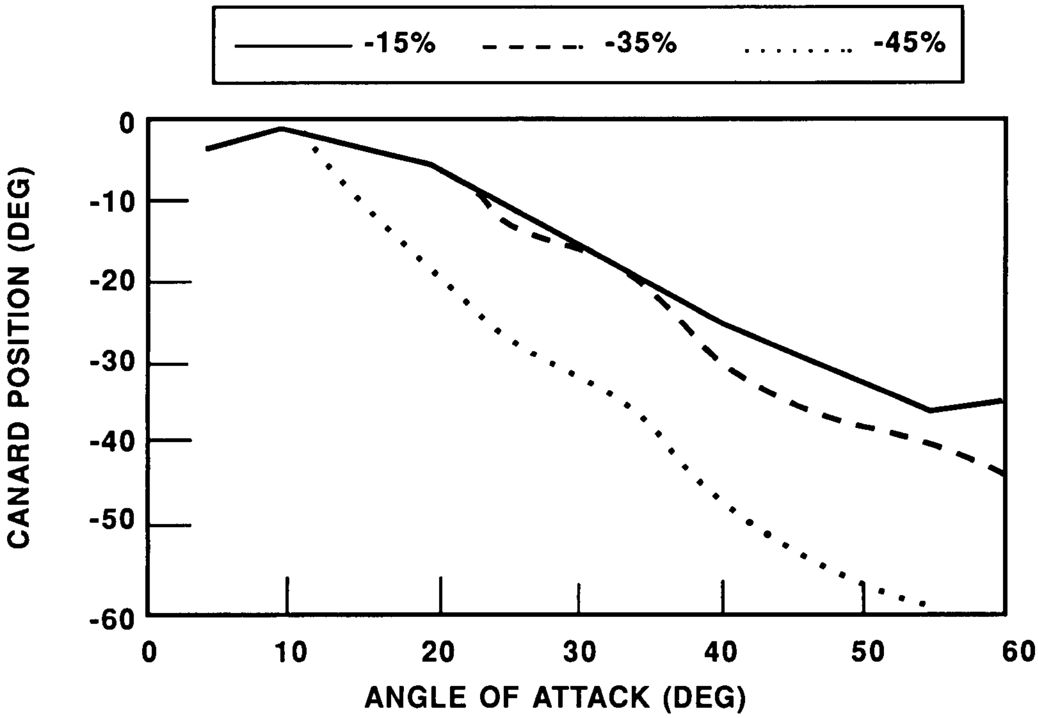


FIGURE 12. EFFECT OF RSS ON X-29
CANARD SCHEDULE

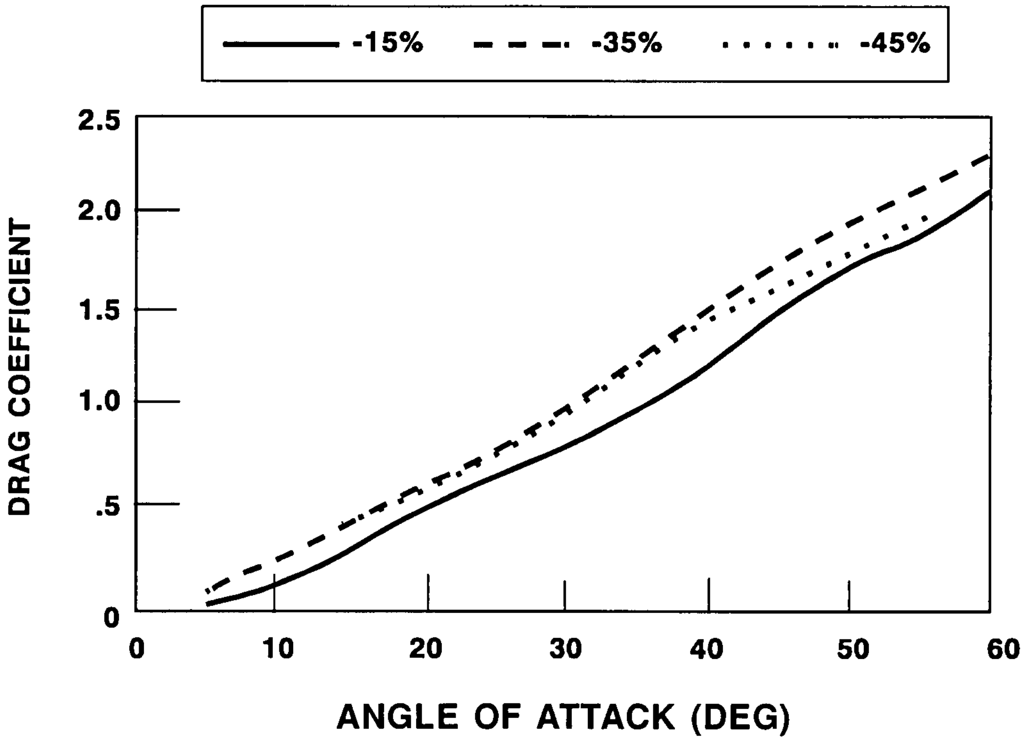


FIGURE 13. EFFECT OF RSS ON X-29 DRAG

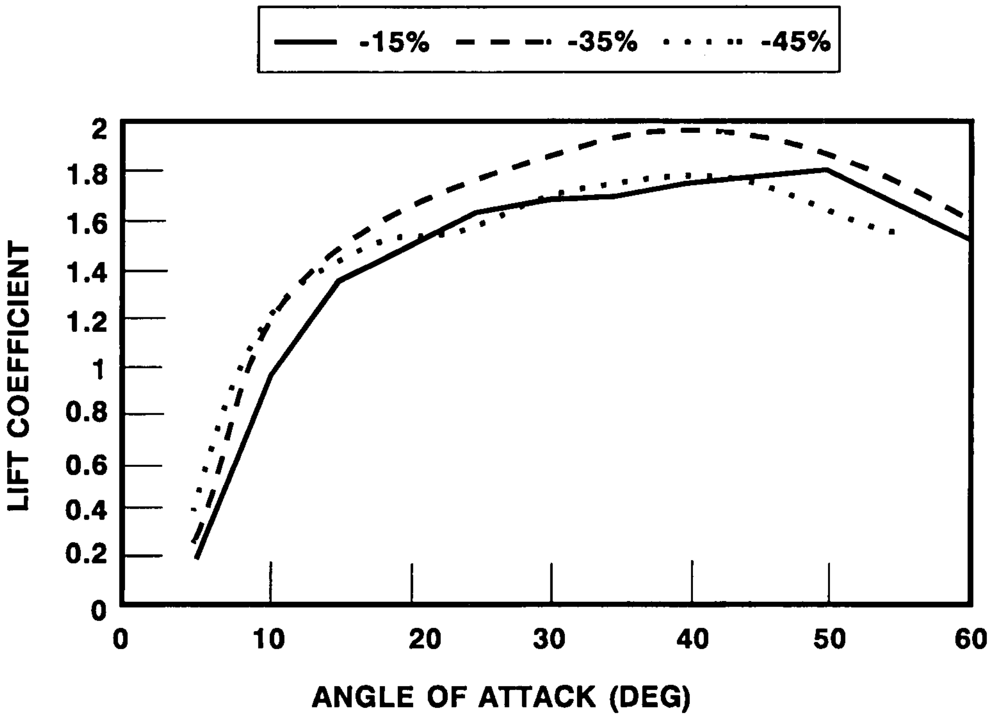


FIGURE 14. EFFECT OF RSS ON X-29 LIFT CURVE

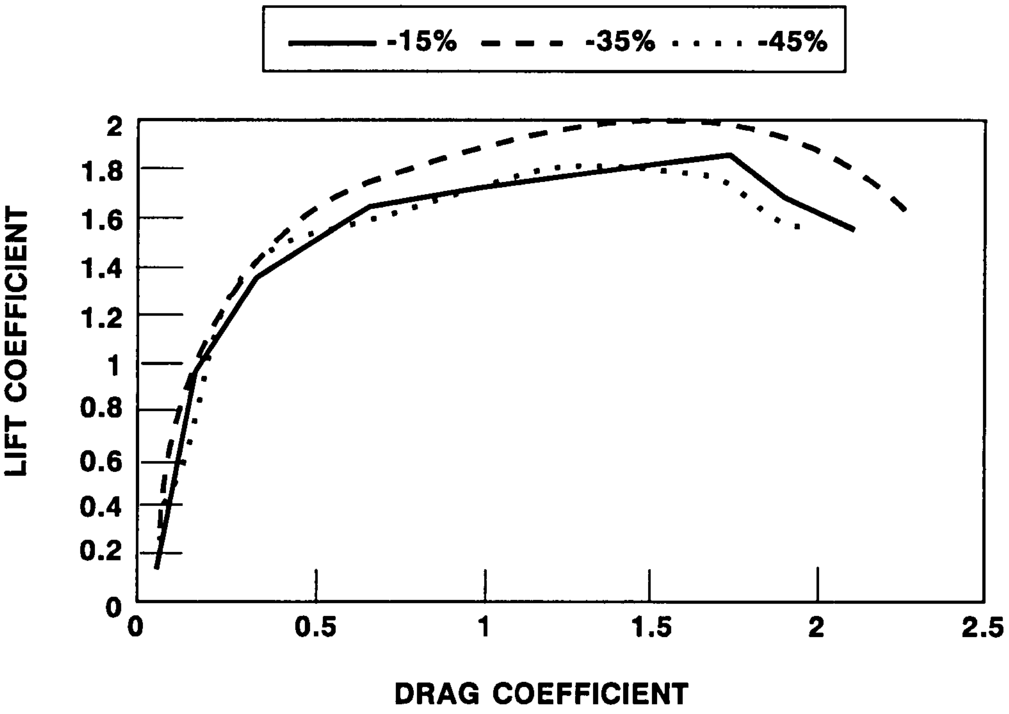


FIGURE 15. EFFECT OF RSS ON X-29 DRAG POLAR

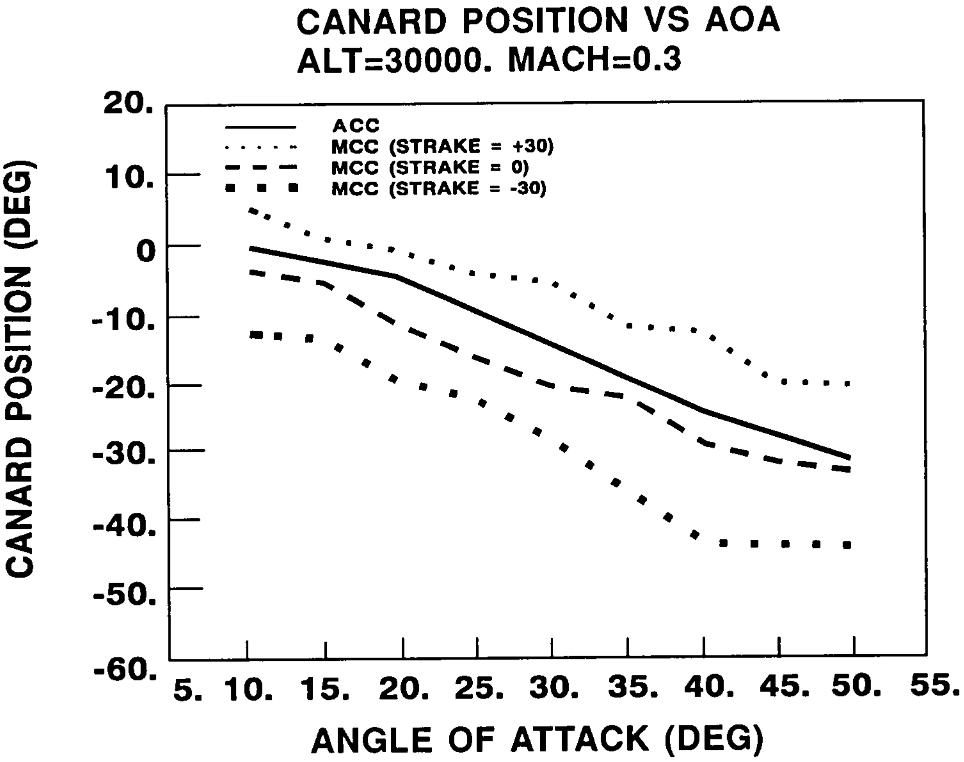


FIGURE 16. X-29 CANARD SCHEDULE VARIATION

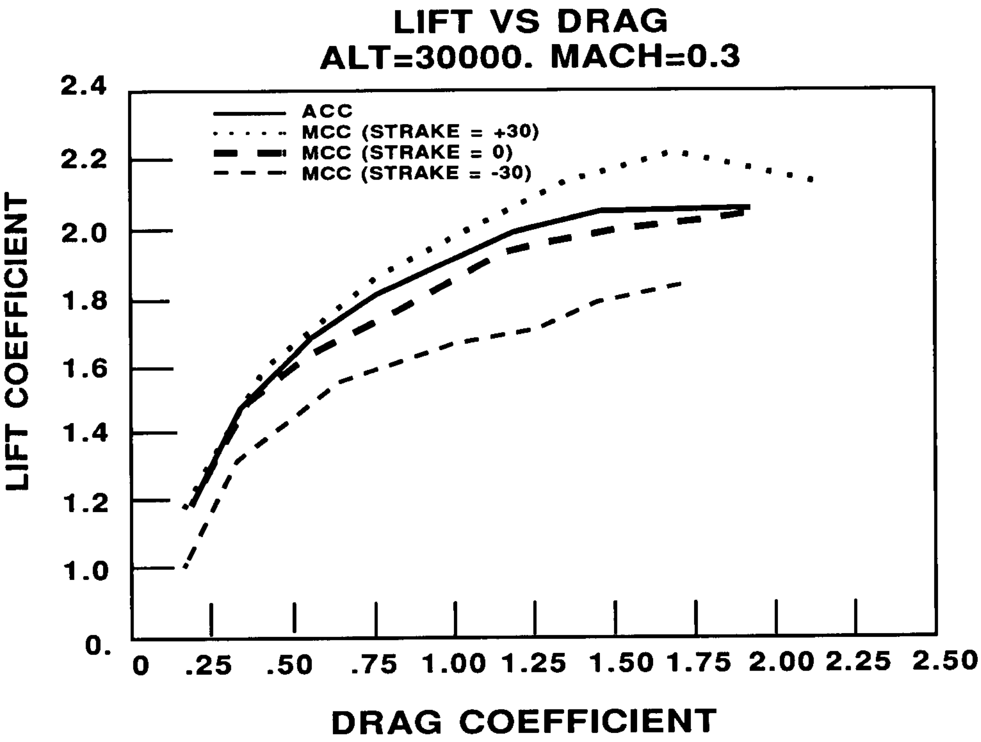


FIGURE 17. CANARD SCHEDULE EFFECT ON X-29 DRAG POLAR

Discussion

SESSION III - QUESTIONS & ANSWERS (PAPERS 13, 15, 16, 17, 18)

PAPER 13: J E JENKINS

Question:

Can you briefly explain what you meant by the bad effect of using higher-order derivatives being to introduce extraneous roots?

Answer:

The stability derivative approach can be shown to be equivalent to the first representing the aerodynamic reactions by indicial responses and summing these through the superposition (convolution) integral. Second, the superposition integral is approximated by an asymptotic expansion (valid for sufficiently slow motions). The terms of the expansion have a one-to-one correlation with terms of the Taylor series expansion of the aerodynamic forces with respect to the motion variables (stability derivatives). Thus, when this representation is put into the equations of motion, the effect is to change the characteristic equation for the system from a transcendental equation to a polynomial. Increasing the number of terms retained in the expansion increases the polynomial degree and thus the number of roots. Therefore, extraneous roots (not belonging to the transcendental equation) can be introduced. Some of these will be in the right-half plane and the solution will "blow-up".

PAPER 15: C H HOUPIS

Question:

How do you choose W, the squaring down matrix?

Answer:

By an understanding of the physical nature of the plant to be controlled. Currently positive and/or negative or zero values are used for the elements of W -- a "trial and error" approach. Research needs to be done on developing an "optimum" technique for choosing these elements.

Question:

How do you generate the bonds on Yii/δii for MIMO systems?

Answer:

As seen from Eq (11) the output is composed of two components: a "tracking" and a "disturbance" output. Tracking and disturbance bands are obtained for each component, based on the specifications (see References 8 & 10).

Question:

How complex are the compensations G and F? Do they cancel the non-minimum phase poles in P?

Answer:

The order of the compensations are very reasonable when the nominal plant is the starting point for synthesizing the loop transmission function. Right-half-plane poles are not cancelled.

PAPER 16: T SADEGHI

Question:

Have you done any work regarding real-time implementations of the Bierman estimator, since there is a large computational overhead?

Answer:

No. We are in the process of putting together a real-time simulator with flight control laws in external and dedicated R3000 CPUs with quadruple redundancy.

Question:

When does γ drop below 0.5, does this influence the rank of ß?

Answer:

γ drops below 0.5 when both surfaces of a pair have been failed or damaged. When $\gamma = 0$ (i.e., both surfaces in a pair failed), then rank of ß is droped by 1. $\gamma = 0$ or $\gamma < 0.5$ represents very rare failure scenarios and often the multiple failure cases.

Question:

How efficient is the pseudo inverse computation?

Answer:

Using the Moore-Penrose algorithm outlined by Karmarker has proven to be very reliable and fast.

PAPER 17: R F STENGEL

Question:

How difficult is it to generate $\mathbf{G}^*(\mathbf{x})$ for an aircraft system? Does it have to match $\mathbf{G}(\mathbf{x})$ closely?

Answer:

The original system equation is

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}$$

In general, $\mathbf{x}$ has dimension (n X 1), $\mathbf{u}$ has dimension (m X 1), and $\mathbf{G}(\mathbf{x})$ is non-square (n X m). Consequently, $\mathbf{G}(\mathbf{x})$ is not strictly invertible, although various pseudo-inverses can be defined. Given an m-dimensional output vector,

$$\mathbf{y} = \mathbf{H}\mathbf{x}$$

it is possible to define an output derivative equation,

$$\mathbf{y}^{(d)} == \mathbf{f}^*(\mathbf{x}) = \mathbf{G}^*(\mathbf{x})\mathbf{u}$$

for which $\mathbf{G}^*(\mathbf{x})$ is square (m X m) and invertible. Equations for $\mathbf{G}^*(\mathbf{x})$ are given for alternate definitions of the output matrix $\mathbf{H}$ in Ref. 105 (where $\mathbf{G}^*(\mathbf{x})$ is called $\mathbf{B}^*(\mathbf{x})$ and $\mathbf{H}$ is called $\mathbf{C}$). These relationships are straightforward; however, on-line evaluation is based on a full, diferentiable, nonlinear model of aircraft aerodynamics. Unless $\mathbf{H}$ is the identity matrix and m = n, $\mathbf{G}(\mathbf{x})$ and $\mathbf{G}^*(\mathbf{x})$ are quite different, both in dimension and numerical definition; therefore, they do not match each other in any obvious way.

Question:(about the film)

I have been surprised at the amount of numerical data which appears on the HUD. This reminds me of a report I read, some 5 years ago, on an F-18 accident. One conclusion of the report was that the attention of the pilot was called by numerical values (here the IAS, because he experienced a minor problem on the afterburner). The report justifies its recommendation by the following fact: to read numerical values you need the foveal vision which is controlled by the brain, and the information goes to it; it needs 1/10 to 1/5 of a second; the "analog-type" data are extracted by the peripheral vision and do not "mobilize" the brain. I am prepared to share this conclusion; however, the pilot has to know its speed!

Answer:

Many tests of human control reveal that experienced pilots perform well with a minimal amount of displayed information and, where possible, that this information should be displayed in analog fashion. The pilots find additional data to be extraneous and distracting. Less experienced pilots can benefit from explicit display of more information, but as their skills increase, they too tend to favor sparse, analog displays. Furthermore, it can be shown that there is a workload/opinion tradeoff between display and control-system complexity for a fixed task: given a more complex controller, the pilot performs as well with simpler displays. It appears to be desirable to reduce the pilot's need to perform compensatory tracking tasks, particularly those that can be performed as well or better by automatic systems. When that is not possible, information display should be tailored to the likely skill levels of the human operator.

PAPER 18: L A WALCHLI

Question:

You mentioned an instability of 35% for the X-28. High instability has aerodynamic advantages, but requires high control power and high rate activators. Others suggest 15-20% instability is closer to optimum. Can you comment?

Answer:

The X-29 is a technology demonstrator and was specifically designed to explore the limits of relaxed static stability. The aircraft was configured to safely fly -35% margin. It produced significant performance benefits - reduced drag, high lift, and excellent manoeuvrability. However, the integration of RSS with the specific X-29 technology suite produced these results. An RSS of -15 to -20% on some other configuration may be optimal for that specific configuration.

Question:

You showed jets or blowing for control at High AoA. How does it compare with thrust vectoring?

Answer:

The Vortex Flow Control (VFC) concept is being flown on the high AoA X-29 testbed as a proof-of-concept experiment. To date it is performing well and is supplying the data necessary to incorporate this new control effector into the flight control system. Multi-axis thrust vectoring is more mature and may be superior. VFC uses a lightweight system, a small amount of engine bleed air through an accumulator, is not totally dependent on engine operation (emergency consideration) and should be simpler to incorporate into the flight control system.

COMMENT ON PAPER 17:

My comment is that with all the new technology, why do we still crash aircraft? There was the SAAB Viggen and recently the YF-22 in the USA. There is a large data base of what pilots like and what they do not like. I see experiment. To date it is performing well and is supplying the data necessary to incorporate this new control effector into the flight control system. Multi-axis thrust vectoring is more mature and may be superior. VFC uses a lightweight system, a small amount of engine bleed air through an accumulator, is not totally dependent on engine operation (emergency consideration) and should be simpler to incorporate into the flight control system.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	**2. Originator's Reference** AGARD-R-789 1|11	**3. Further Reference** ISBN 92-835-0702-9 2|8	**4. Security Classification of Document** UNCLASSIFIED/ UNLIMITED

5. Originator	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 Rue Ancelle, 92200 Neuilly sur Seine, France 1|14
6. Title	STABILITY IN AEROSPACE SYSTEMS 1|18
7. Presented at	the Guidance and Control Panel Workshop held at ONERA, Mauzac (Toulouse), France, 23rd—25th June 1992.

8. Author(s)/Editor(s) Various	**9. Date** February 1993
10. Author's/Editor's Address Various	**11. Pages** 242

12. Distribution Statement	There are no restrictions on the distribution of this document. Information about the availability of this and other AGARD unclassified publications is given on the back cover.

13. Keywords/Descriptors

Boundary layer transition
Chaos
Flight control
Instability
Non-linear system
Robustness
Stability

14. Abstract

This volume contains the 18 unclassified papers, presented at the Guidance and Control Panel Workshop held at ONERA, Centre de Fauga — Mauzac (Toulouse), France, 23rd—25th June 1992.

The papers were presented covering the following headings:
- Fundamental Aspects of Stability with Examples;
- Basic Theoretical Aspects and Chaos;
- Applications to Aerospace Techniques.

AGARD Report 789
Advisory Group for Aerospace Research and Development, NATO
STABILITY IN AEROSPACE SYSTEMS
Published February 1993
242 pages

This volume contains the 18 unclassified papers, presented at the Guidance and Control Panel Workshop held at ONERA, Centre de Fauga — Mauzac (Toulouse), France, 23rd—25th June 1992.

The papers were presented covering the following headings:
- Fundamental Aspects of Stability with Examples;
- Basic Theoretical Aspects and Chaos;
- Applications to Aerospace Techniques.

ISBN 92-835-0702-9

AGARD-R-789

Boundary layer transition
Chaos
Flight control
Instability
Non-linear system
Robustness
Stability

AGARD Report 789
Advisory Group for Aerospace Research and Development, NATO
STABILITY IN AEROSPACE SYSTEMS
Published February 1993
242 pages

This volume contains the 18 unclassified papers, presented at the Guidance and Control Panel Workshop held at ONERA, Centre de Fauga — Mauzac (Toulouse), France, 23rd—25th June 1992.

The papers were presented covering the following headings:
- Fundamental Aspects of Stability with Examples;
- Basic Theoretical Aspects and Chaos;
- Applications to Aerospace Techniques.

ISBN 92-835-0702-9

AGARD-R-789

Boundary layer transition
Chaos
Flight control
Instability
Non-linear system
Robustness
Stability

AGARD Report 789
Advisory Group for Aerospace Research and Development, NATO
STABILITY IN AEROSPACE SYSTEMS
Published February 1993
242 pages

This volume contains the 18 unclassified papers, presented at the Guidance and Control Panel Workshop held at ONERA, Centre de Fauga — Mauzac (Toulouse), France, 23rd—25th June 1992.

The papers were presented covering the following headings:
- Fundamental Aspects of Stability with Examples;
- Basic Theoretical Aspects and Chaos;
- Applications to Aerospace Techniques.

ISBN 92-835-0702-9

AGARD-R-789

Boundary layer transition
Chaos
Flight control
Instability
Non-linear system
Robustness
Stability

AGARD Report 789
Advisory Group for Aerospace Research and Development, NATO
STABILITY IN AEROSPACE SYSTEMS
Published February 1993
242 pages

This volume contains the 18 unclassified papers, presented at the Guidance and Control Panel Workshop held at ONERA, Centre de Fauga — Mauzac (Toulouse), France, 23rd—25th June 1992.

The papers were presented covering the following headings:
- Fundamental Aspects of Stability with Examples;
- Basic Theoretical Aspects and Chaos;
- Applications to Aerospace Techniques.

ISBN 92-835-0702-9

AGARD-R-789

Boundary layer transition
Chaos
Flight control
Instability
Non-linear system
Robustness
Stability

AGARD

NATO OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE

FRANCE

Télécopie (1)47.38.57.99 · Télex 610 176

DIFFUSION DES PUBLICATIONS
AGARD NON CLASSIFIEES

Aucun stock de publications n'a existé à AGARD. A partir de 1993, AGARD détiendra un stock limité des publications associées aux cycles de conférences et cours spéciaux ainsi que les AGARDographies et les rapports des groupes de travail, organisés et publiés à partir de 1993 inclus. Les demandes de renseignements doivent être adressées à AGARD par lettre ou par fax à l'adresse indiquée ci-dessus. *Veuillez ne pas téléphoner.* La diffusion initiale de toutes les publications de l'AGARD est effectuée auprès des pays membres de l'OTAN par l'intermédiaire des centres de distribution nationaux indiqués ci-dessous. Des exemplaires supplémentaires peuvent parfois être obtenus auprès de ces centres (à l'exception des Etats-Unis). Si vous souhaitez reçevoir toutes les publications de l'AGARD, ou simplement celles qui concernent certains Panels, vous pouvez demander à être inclu sur la liste d'envoi de l'un de ces centres. Les publications de l'AGARD sont en vente auprès des agences indiquées ci-dessous, sous forme de photocopie ou de microfiche.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Fachinformationszentrum,
Karlsruhe
D-7514 Eggenstein-Leopoldshafen 2

BELGIQUE
Coordonnateur AGARD-VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Directeur du Service des Renseignements Scientifiques
Ministère de la Défense Nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Board
Ved Idraetsparken 4
2100 Copenhagen Ø

ESPAGNE
INTA (AGARD Publications)
Pintor Rosales 34
28008 Madrid

ETATS-UNIS
National Aeronautics and Space Administration
Langley Research Center
M/S 180
Hampton, Virginia 23665

FRANCE
O.N.E.R.A. (Direction)
29, Avenue de la Division Leclerc
92322 Châtillon Cedex

GRECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PAYS-BAS
Netherlands Delegation to AGARD
National Aerospace Laboratory NLR
P.O. Box 90502
1006 BM Amsterdam

PORTUGAL
Portuguese National Coordinator to AGARD
Gabinete de Estudos e Programas
CLAFA
Base de Alfragide
Alfragide
2700 Amadora

ROYAUME UNI
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE
Millî Savunma Başkanlığı (MSB)
ARGE Daire Başkanlığı (ARGE)
Ankara

Le centre de distribution national des Etats-Unis (NASA/Langley) ne détient PAS de stocks des publications de l'AGARD.
D'éventuelles demandes de photocopies doivent être formulées directement auprès du NASA Center for Aerospace Information (CASI) à l'adresse suivante:

AGENCES DE VENTE

NASA Center for
Aerospace Information (CASI)
P.O. Box 8757
BWI Airport, Maryland 21240
United States

ESA/Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris
France

The British Library
Document Supply Division
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume Uni

Les demandes de microfiches ou de photocopies de documents AGARD (y compris les demandes faites auprès du CASI) doivent comporter la dénomination AGARD, ainsi que le numéro de série d'AGARD (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Veuiller noter qu'il y a lieu de spécifier AGARD-R-nnn et AGARD-AR-nnn lors de la commande des rapports AGARD et des rapports consultatifs AGARD respectivement. Des références bibliographiques complètes ainsi que des résumés des publications AGARD figurent dans les journaux suivants:

Scientifique and Technical Aerospace Reports (STAR)
publié par la NASA Scientific and Technical
Information Division
NASA Headquarters (NTT)
Washington D.C. 20546
Etats-Unis

Government Reports Announcements and Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 22161
Etats-Unis

(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

SPS

Imprimé par Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

AGARD
NATO OTAN
7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE
FRANCE
Telefax (1)47.38.57.99 · Telex 610 176

DISTRIBUTION OF UNCLASSIFIED AGARD PUBLICATIONS

AGARD holds limited quantities of the publications that accompanied Lecture Series and Special Courses held in 1993 or later, and of AGARDographs and Working Group reports published from 1993 onward. For details, write or send a telefax to the address given above. *Please do not telephone.*

AGARD does not hold stocks of publications that accompanied earlier Lecture Series or Courses or of any other publications. Initial distribution of all AGARD publications is made to NATO nations through the National Distribution Centres listed below. Further copies are sometimes available from these centres (except in the United States). If you have a need to receive all AGARD publications, or just those relating to one or more specific AGARD Panels, they may be willing to include you (or your organisation) on their distribution list. AGARD publications may be purchased from the Sales Agencies listed below, in photocopy or microfiche form.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD — VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Director Scientific Information Services
Dept of National Defence
Ottawa, Ontario K1A 0K2

DENMARK
Danish Defence Research Board
Ved Idraetsparken 4
2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92322 Châtillon Cedex

GERMANY
Fachinformationszentrum
Karlsruhe
D-7514 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
See Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 90502
1006 BM Amsterdam

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Portuguese National Coordinator to AGARD
Gabinete de Estudos e Programas
CLAFA
Base de Alfragide
Alfragide
2700 Amadora

SPAIN
INTA (AGARD Publications)
Pintor Rosales 34
28008 Madrid

TURKEY
Milli Savunma Başkanlığı (MSB)
ARGE Daire Başkanlığı (ARGE)
Ankara

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
National Aeronautics and Space Administration (NASA)
Langley Research Center
M/S 180
Hampton, Virginia 23665

The United States National Distribution Centre (NASA/Langley) does NOT hold stocks of AGARD publications.
Applications for copies should be made direct to the NASA Center for Aerospace Information (CASI) at the address below.

SALES AGENCIES

NASA Center for
Aerospace Information (CASI)
P.O. Box 8757
BWI Airport, Maryland 21240
United States

ESA/Information Retrieval Service
European Space Agency
10, rue Mario Nikis
75015 Paris
France

The British Library
Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
United Kingdom

Requests for microfiches or photocopies of AGARD documents (including requests to CASI) should include the word 'AGARD' and the AGARD serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Note that AGARD Reports and Advisory Reports should be specified as AGARD-R-nnn and AGARD-AR-nnn, respectively. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
published by NASA Scientific and Technical
Informat
NASA H
Washing
United S

Government Reports Announcements and Index (GRA&I)
published by the National Technical Information Service