NASA/TP–2007-213474

# Synthesis from Design Requirements of a Hybrid System for Transport Aircraft Longitudinal Control

## Volume I

*Charles S. Hynes and Gordon H. Hardy*
*Ames Research Center, Moffett Field, California*

*Lance Sherry*
*Honeywell International Inc., Phoenix, Arizona*

December 2007

NASA/TP–2007-213474

# Synthesis from Design Requirements of a Hybrid System for Transport Aircraft Longitudinal Control

## Volume I

*Charles S. Hynes and Gordon H. Hardy*
*Ames Research Center, Moffett Field, California*

*Lance Sherry*
*Honeywell International Inc., Phoenix, Arizona*

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

**December 2007**

## Acknowledgments

We thank Michael Heymann for suggesting the investigation of validity regions in state space, and for many discussions. Charles E. Billings, William Bjorkman, James Franklin, Charles W. (Bill) Harper, Feng Lin, George Meyer, and Everett Palmer reviewed the report and provided useful comments on both content and presentation. We are grateful to them all.

Available from:

NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
(703) 487-4650

# TABLE OF CONTENTS

## TABLE OF CONTENTS (continued)

## TABLE OF CONTENTS (continued)

## LIST OF FIGURES

## LIST OF FIGURES (continued)

## LIST OF TABLES

## LIST OF ACRONYMS

| | |
|---|---|
| AFS | Autoflight system |
| ATC | Air traffic control |
| CDU | Control display unit |
| EAS | Equivalent airspeed |
| EGT | Exhaust gas temperature |
| EPR | Exhaust pressure ratio |
| FAA | Federal Aviation Administration |
| FMS | Flight management system |
| ILS | Instrument landing system |
| MCP | Mode control panel |
| QSRA | Quiet short-haul research aircraft |
| TAS | True airspeed |
| VNAV | vertical navigation |

## LIST OF SYMBOLS

### English Symbols

| | | | |
|---|---|---|---|
| a | = | a (H) | Sonic velocity, ft/sec |
| $a_{NLIM}$ | | | Normal acceleration limit, 3 ft/sec$^2$ |
| AR | ≡ | $b^2/S$ | Wing aspect ratio, dimensionless |
| b | | | Wing span, ft |
| $C_L$ | | | Lift coefficient, dimensionless |
| $C_{DP}$ | = | $C_{DP}$ (M) | Parasite drag coefficient, dimensionless |
| dH/dt | | | Rate of change of height H, ft/sec |
| $(dH/dt)_0$ | | | Threshold rate of change of height H, 250 ft/min |
| dV/dt | | | Rate of change of airspeed V, ft/sec$^2$ |
| $d\gamma/dt$ | | | Rate of change of flightpath angle $\gamma$, rad/sec |
| D | | | Drag, lb |
| e | = | e (M) | Span efficiency, dimensionless |
| g | | | Acceleration of gravity, ft/sec$^2$ |
| H | | | Height in Standard Atmosphere, ft |
| $H_0$ | | | Altitude deviation threshold, ft |
| $\Delta H$ | | | Altitude error, ft |
| $I_Y$ | | | Pitch moment of inertia, slug/ft$^2$ |

## LIST OF SYMBOLS (continued)

### English Symbols (continued)

| | | | |
|---|---|---|---|
| $K_V$ | | | Longitudinal acceleration limiter parameter, dimensionless |
| L | | | Lift, lb |
| M | = | V/a | Mach number, dimensionless |
| $M_Y$ | | | Pitching moment, ft-lb |
| N | | | Engine rpm (revolutions per minute), percent |
| $P_0$ | | | Standard sea-level barometric pressure, 2116.22 lb/ft$^2$ |
| q | | | Rate of change of pitch angle $\theta$, rad/sec |
| S | | | Aircraft wing area, ft$^2$ |
| T | | | Thrust, lb |
| $T/\delta$ | = | f (N, M) | Corrected thrust, lb |
| V | | | Airspeed, ft/sec |
| $V_E$ | = | $V\sqrt{\sigma}$ | Equivalent airspeed, ft/sec |
| W | | | Aircraft weight, lb |
| W/S | | | Aircraft wing loading, lb/ft$^2$ |

### Greek Symbols

| | |
|---|---|
| $\alpha$ | Angle of attack, rad |
| $\gamma$ | Flightpath angle, rad |
| $\Delta\gamma$ | Flightpath angle increment, 3 deg |
| $\gamma_{POT}$ | Trim (equilibrium) flightpath angle, rad |
| $\gamma_{SPEED}$ | Flightpath angle for airspeed control, rad |
| $\delta \equiv p/p_0 \equiv \delta\ (H)$ | Ambient pressure ratio, dimensionless |
| $\delta_{ELEV}$ | Elevator deflection, rad |
| $\delta_T$ | Throttle position, rad |
| $\theta$ | Pitch angle, rad |
| $\rho_0$ | Standard sea-level density, 0.00237691 sl/ft$^3$ |
| $\sigma \equiv \rho/\rho_0 \equiv \sigma\ (H)$ | Ambient density ratio, dimensionless |
| $\phi$ | Bank angle, rad |

## LIST OF SYMBOLS (continued)

### Subscripts

| | |
|---|---|
| CMD | Commanded |
| EXP | Exponential law |
| LIM | Limited |
| MAX | Maximum |
| MIN | Minimum |
| MC | Minimum control airspeed |
| PAR | Parabolic law |
| POT | Potential (equilibrium or trim) conditions |
| REF | Reference |
| SAFE | Safety envelope limit |
| TGT | Target |

### Logical Symbols

| | | | | | |
|---|---|---|---|---|---|
| Equivalence | ≡ | Negation | ¬ | Implication | ⇒ |
| Logical OR | ∪ | Logical AND omitted | | | |

## LIST OF SYMBOLS (continued)

### Logical Conditions

$VT1 \equiv (V_{TGT} < V_{MIN})$

$VT2 \equiv (V_{TGT} > V_{MAX})$

$VT3 \equiv (V_{TGT} \leq V_{MIN\ DRAG})$

$V1 \equiv (V \leq V_{MIN\ DRAG})$

$V2 \equiv (V < V_{TGT})$

$V3 \equiv (V = V_{TGT})$

$V4 \equiv (V > V_{TGT})$

$G1 \equiv (\gamma = 0)$

$GT1 \equiv (\gamma_{TGT} < \gamma_{SAFE})$

$GT2 \equiv (\gamma_{TGT} < \gamma_{POT\ TGT})$

$GT3 \equiv (\gamma_{TGT} = \gamma_{POT\ TGT})$

$GT4 \equiv (\gamma_{TGT} > \gamma_{POT\ TGT})$

$GT5 \equiv (\gamma_{TGT} \leq \gamma_{POT\ MIN})$

$GT6 \equiv (\gamma_{TGT} \geq \gamma_{POT\ MAX})$

$Q \equiv (\gamma_{TGT} \leq \gamma_{SPEED\ MIN})$

$P \equiv (\gamma_{TGT} \geq \gamma_{SPEED\ MAX})$

$DC1 \equiv (\gamma_{POT\ MIN} < 0)$

$PF1 \equiv (\gamma_{POT\ MAX} = \gamma_{POT\ MIN})$

$TC1 \equiv (\gamma_{POT\ MAX} < 0)$

$TC2 \equiv (\gamma_{POT\ MAX} = 0)$

$TS1 \equiv (\gamma \leq \gamma_{SPEED\ MIN})$

$TS2 \equiv (\gamma \geq \gamma_{SPEED\ MAX})$

$TT1 \equiv (\gamma_{POT\ TGT} = \gamma_{POT\ MIN})$

$TT2 \equiv (\gamma_{POT\ TGT} = \gamma_{POT\ MAX})$

$TT3 \equiv (\gamma_{POT\ TGT} < \gamma_{POT\ MIN})$

$TT4 \equiv (\gamma_{POT\ TGT} > \gamma_{POT\ MAX})$

$\neg PP \equiv (\text{PRIORITY} \equiv \text{SPEED})$

$PP \equiv (\text{PRIORITY} \equiv \text{PATH})$

$PE\ (V) \equiv \neg\ [P\ TT1]\ \neg\ [Q\ TT2]$

$PE\ (\gamma) \equiv \neg\ (GT6\ V1\ TS2)$

$H1 \equiv (H_{TGT} < H_{SAFE})$

$H2 \equiv (H_{TGT} > H_{MAX})$

$NE1 \equiv \neg\ [(GT2 \cup GT3)\ V4]\ \neg\ [GT2\ V3]$

$NE2 \equiv \neg\ [(GT32 \cup GT4)\ V2]\ \neg\ [GT4\ V3]$

# EXECUTIVE SUMMARY

The avionic systems for flight management and flight control on board current transport aircraft have become quite complex, resulting in heavy penalties on development time and cost. These systems have evolved during more than 20 years of ad hoc development, and now comprise very large codes characterized by functional duplication, so that a single flight task can often be accomplished in several different ways. Weak functional coordination at the design level can result, under unusual circumstances, in unsafe or inappropriate system actions. Furthermore, simulation is an inadequate method for finding and eliminating such system behavior, because it is impractical to simulate all possible cases. Excessively complex mode structure can result in pilot mode confusion, preventing the flight crew from identifying unsafe or inappropriate system actions in time to avoid potentially catastrophic consequences.

Recent advances in automation theory now enable an improved approach that ensures proper coordination of system elements. New design principles are derived from long-established criteria for aircraft design and certification. These criteria are based on more than 40 years of operating experience with jet transport aircraft, and specify functional requirements for aerodynamic, structural, and propulsion design. However, no such broad functional criteria have previously been applied to the design of avionic systems.

This report presents a detailed application to longitudinal control. Qualitative analysis based on the governing differential equations shows that longitudinally the aircraft is capable of only three kinds of dynamical behavior during manually controlled flight. Basic automated control modes are developed that correspond to each of the three kinds of behavior, and mode selection logic is then synthesized in such a way as to guarantee that the complete system satisfies behavioral design principles specified beforehand.

Formal system validation is based on rigorous proof of system behavioral properties. This formal validation supplements informal validation achieved by means of piloted simulation, but, unlike simulation, guarantees that unsafe or inappropriate system actions are ruled out by the system design.

Principled design results in a drastically simplified mode structure. This simplified mode structure is matched to an optimized division of cockpit duties between human crewmembers and the automated system, and can be expected to eliminate pilot mode confusion.

To summarize, major design goals are (1) system design integrity based on proof of correctness at the design level, (2) significant simplification and cost reduction in system development and certification, and (3) improved operational efficiency, with significant alleviation of human-factors problems encountered by pilots in current transport aircraft.

This report provides for the first time a firm technical basis for criteria governing design and certification of avionic systems for transport aircraft. It should be of primary interest to designers of next-generation avionic systems.

# SYNTHESIS FROM DESIGN REQUIREMENTS OF A HYBRID SYSTEM FOR TRANSPORT AIRCRAFT LONGITUDINAL CONTROL

Charles S. Hynes,[1] Gordon H. Hardy,[1] and Lance Sherry[2]

Ames Research Center

## SUMMARY

Hybrid systems are dynamical systems that consist of both continuous and discrete elements. This report presents a new method for synthesizing such systems from design requirements, and applies the method to the design of a system for longitudinal control of transport aircraft. The resulting system satisfies general requirements for safety and effectiveness specified a priori, enabling formal validation of the complete system to be achieved.

The design process begins by describing three primitive (basic) modes for continuous control of flightpath and airspeed that correspond to the three fundamental kinds of dynamical behavior characterizing transport aircraft. For each mode, the dynamical behavior specified by the governing differential equations is discretized by relating it geometrically to the aircraft performance envelope. Validity conditions for each mode are derived from safety and effectiveness requirements imposed a priori, and the three primitive modes are then combined to form a supermode that constitutes the lowest-level hybrid system. Within this lowest-level supermode, discrete mode control logic is synthesized directly from the validity conditions. Although synthesis is rigorous, heuristic arguments based on general notions of maximum effectiveness and logical dominance simplify the process, with the help of extensive use of elementary symbolic logic. In cases where aircraft performance limitations preclude capture of path or speed targets, the exact nature of the limitation is annunciated for resolution at higher levels within the mode hierarchy (or, ultimately, by the human crew), with assurance of logical completeness. Extension to higher-level supermodes is described, with a detailed altitude-control example.

The potential contributions of formal validation to system design integrity are discussed, and analysis of several transport aircraft accidents and incidents selected to serve as examples leads to the suggestion that new airworthiness criteria based on general safety and effectiveness properties might improve operational safety. The design of the human/machine cockpit interface is discussed briefly, and several heuristic guidelines for system and interface design are proposed. Regarding the pilot's mode selection as a command given to the system enables analysis from the human-factors viewpoint of the same accidents and incidents studied previously from the system design perspective, and leads to recommendations that cockpit interface design be integrated into the system design process, and that formal validation be extended to include the cockpit interface.

[1] Retired, NASA Ames Research Center, Moffett Field, CA 94035.

[2] Formerly at Honeywell International Inc., Flight Control Systems Design, Phoenix, AZ; currently at George Mason University, 4400 University Drive, MS 4A6, Fairfax, VA 22030.

The present results may contribute by example to the development of more general theoretical methods for synthesis of hybrid systems. They should also be of practical interest to designers of next-generation transport aircraft avionic systems.

# PART I
# BACKGROUND

## INTRODUCTION

Hybrid systems are dynamical systems that consist of both continuous and discrete elements. For example, autopilots with multiple control laws, each specified by a different mode, are hybrid systems in which the modes constitute the discrete elements. This report presents a new method for synthesizing the discrete elements of such systems (that is, the mode selection logic) directly from design requirements, and applies this method to the design of a system for longitudinal control of transport aircraft. The resulting system satisfies general requirements for safety and effectiveness specified a priori, enabling formal validation of the complete system to be achieved. Compared with the systems installed in current transport aircraft, this new method for mode synthesis leads to drastic simplification of mode structure, and has potential for streamlining system development and certification, reducing costs, and eliminating pilot mode confusion.

In order to provide necessary background, this section discusses the flight management systems on board current transport aircraft. It shows that the complexity of these large, reactive hybrid systems imposes heavy penalties on development time and cost. It then reviews methods for verification and validation of these systems. The simulation-based methods currently in use are limited by the impossibility of testing all cases, but formal mathematical methods have potential for avoiding that limitation. Recent advances in automation theory now enable an improved design approach, which is the subject of this report.

The problem of predicting theoretically the dynamical behavior of hybrid systems is discussed in detail, and it is shown how lack of predictability in safety-critical systems can lead to "automation surprises" with potentially catastrophic consequences. Although theoretical prediction of dynamical behavior is mathematically intractable for arbitrarily specified hybrid systems, the problem can in principle be solved by inversion: starting from certain safety and effectiveness properties specified a priori, the mode selection logic can be synthesized so as to force the completed system to satisfy the specified behavioral properties.

After this brief review of automation theory, the design goals of the synthesis method and the technical approach are described, the specific objectives and scope of this work are presented, and the section concludes with a detailed presentation of the plan of the report.

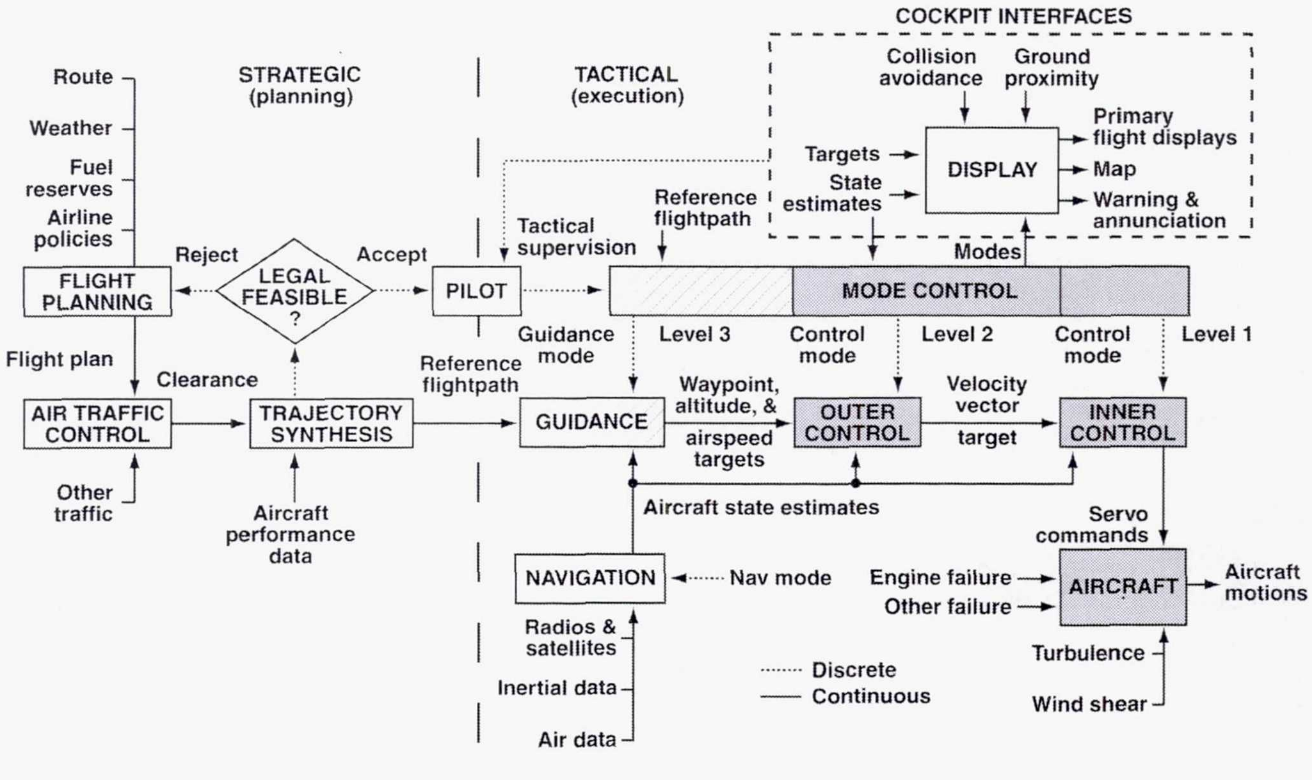
### Vehicle Management System

The block diagram of figure 1 illustrates operation of a transport aircraft by its human crew within the air traffic control (ATC) system. The illustrated system is termed a Vehicle Management System to distinguish it from the flight management systems (FMS) on board current transport aircraft, although the two systems share their major elements in common. In figure 1, solid lines represent continuous signals, broken lines represent discrete signals, and blocks represent transformations of signals from input to output. The diagram as a whole can be regarded as a simplified representation of the process by which the flight plan on the left side is transformed into the aircraft motions on the right side.

## Pilot Roles

The human pilot is represented in two roles, the strategic (planning) role on the left of the dashed vertical line, and the tactical (executive) role on its right. On the left (strategic) side, the pilot makes up the flight plan, accounting for weather, route, fuel reserves, airline policies, and the like, and proposes the flight plan to ATC. The pilot's strategic activities are knowledge-based (Rasmussen, 1976) because they require cause-and-effect understanding of whole disciplines such as meteorology and airline flight operations. After accounting for other air traffic so as to provide separation, ATC issues a clearance to the aircraft. On board the aircraft, the clearance must be checked for legality and feasibility before acceptance, and an executable trajectory must be synthesized that specifies the reference flightpath in space and time.

When the ATC clearance is accepted and the reference flightpath determined, the pilot's tactical role (right side of diagram) comes into play. The pilot's tactical activities involved in executing the flight plan and monitoring progress of the flight are rule-based, and could in principle be automated. In practice, much of this lower-level supervisory activity consists of selecting appropriate modes for the navigation, guidance, and flight control functions of the automated system, and monitoring



Pilot is shown in strategic and tactical supervisory roles. Shaded blocks indicate system elements treated in detail by this report. Cross-hatching indicates proposed application guidelines.

Figure 1. Vehicle management system.

displays and annunciations. (Direct manipulation of the flight controls during manual flight, which is a skill-based activity that is discussed later, is not illustrated.)

### Functional Description

The navigation system combines onboard sensor data with radio and satellite receiver data to form aircraft state estimates, which provide feedback of aircraft motions. The guidance system compares the actual trajectory of the aircraft based on state estimates with the reference trajectory, and generates local targets for waypoint position, altitude, and airspeed that are intended to null trajectory errors. The outer-loop control system accepts these local targets and generates a target velocity vector intended to ensure capture of waypoint, altitude, and airspeed targets. The target velocity vector is displayed to the flight crew and also injected into the inner-loop control system. The inner-loop control system accepts the target velocity vector and generates commands to the control surface and throttle servos of the aircraft that are intended to ensure capture of the target velocity vector. The resulting aerodynamic and propulsive forces and moments then act on the aircraft to generate its motions. These motions are displayed to the crew and fed back to the guidance and control systems by means of the state estimates. Aircraft position and altitude are also fed back to ATC by means of ground-based radar (not illustrated).

The actions of the navigation, guidance, and control systems are subject to the laws (that is, transformational algorithms) specified by the navigation, guidance, and control modes, which are selected partially by the pilot (tactical role) and partially by the automated system. The three hierarchical levels of mode control illustrated by the diagram are explained in detail later, and should be ignored for the present.

### Environmental Influences

Figure 1 illustrates environmental influences of several kinds that act on the aircraft system. The tactical (right) side of the diagram shows that atmospheric turbulence and wind shear act on the aircraft itself, disturbing its motion directly. Engine failure is considered an environmental influence, because the failed engine is external to the remaining elements of the system. Likewise, failures within other aircraft systems (hydraulic system, electrical system, and the like) are external failures to which the Vehicle Management System is required to react (fig. 1). Similarly, warnings from the collision avoidance system and the ground proximity warning system are considered environmental influences, because these warnings also represent external conditions to which the aircraft system is required to react.

On the strategic (left) side of the diagram, factors that could force revision of the flight plan are exemplified by deterioration of weather conditions at the destination airport, depletion of fuel reserves owing to unexpected headwinds, or a change of route to avoid storms. Such factors also are considered environmental influences to which the aircraft system must react. To summarize, the aircraft system must be regarded as reacting dynamically to more or less continuous external disturbances.

### Implementation Software

The code that implements the flight management systems on board current transport aircraft consists of about 2 to 2.5 million source-code instructions, and doubles in size about every 4 years (Sherry and McCrobie, 1998). About 20% of this code consists of computer modeling of continuous system elements for state estimation, trajectory synthesis, guidance, flight control, and cockpit display. The other 80% of the code consists of discrete decision logic. The complexity of this large, reactive hybrid system imposes heavy penalties on development time and cost.

In general, the continuous elements of the system implement mathematically based algorithms. For these continuous elements, existing knowledge of numerical analysis provides an adequate basis for design provided that it is correctly applied, and the computational behavior of the resulting continuous routines can be checked independently by well-understood input-output test methods. Furthermore, many of these routines can be reused.

On the other hand, design of the discrete decision logic relies largely on heuristics that lack a firm mathematical basis. Furthermore, software architecture has evolved during more than 20 years of ad-hoc development; current mode structures are vehicle-dependent, and are characterized by multiply-overlapping functions. Under these circumstances, it is not surprising that most industry effort required for development and for verification and validation of implementation code is expended on the discrete elements of the system. The next section discusses current industry trends for verification and validation.

## Verification and Validation

Flight qualification of safety-critical systems requires both verification and validation. Verification seeks to determine that the system implementation satisfies its specification, while validation examines whether the specification itself is correct. The industry approach to verification and validation has been based on massive testing of the completed system, and Federal Aviation Administration (FAA) certification has required inspection of the results of these tests. Development and certification have become increasingly expensive and time-consuming as system complexity has grown, and the industry is now seeking more efficient alternatives to labor-intensive manual coding and testing.

### Formal Verification

One alternative method uses decision tables as the fundamental specification document for the discrete decision logic, which as already noted comprises about 80% of the implementation code in current flight management systems. When the specification is complete, the implementation code can be generated directly from the decision table by an algorithm that has itself been formally verified. By this means, the verification process can, in principle, be automated, with very significant reductions in development time and cost. However, it seems likely that new certification criteria will be needed to enable regulatory approval.

### Formal Validation

If the trend toward formal verification develops as anticipated, it may be expected that the potential for formal validation will be explored next. Validation of safety-critical aircraft systems depends ultimately on the judgment of human pilots. The industry approach to validation has been based on the designer's interpretation of expert pilot judgments in formulating design specifications, and has required extensive piloted simulation for evaluation of the complete system in all expected operational contexts. Such informal validation is subject to the same limitations as informal verification. In practice, simulation cannot test all possible paths through codes as complex as those on board current transport aircraft, and cannot explore all possible operational situations.

Formal validation has potential for augmenting expert pilot judgment by eliminating any system behavior that is in conflict with general requirements for safety and for effective operation. These general requirements must be based on flight hazards identified a priori (Leveson, 1995), and on control tasks necessary for flight operations within the ATC system. Just as for formal verification, it seems likely that new certification criteria will be needed in order to take full advantage of this new approach to validation. Formal validation could be based on theoretical prediction of the dynamical behavior of the system, but no method is presently available for predicting the dynamical behavior of an arbitrarily specified reactive hybrid system; the discussion turns next to this mathematical problem.

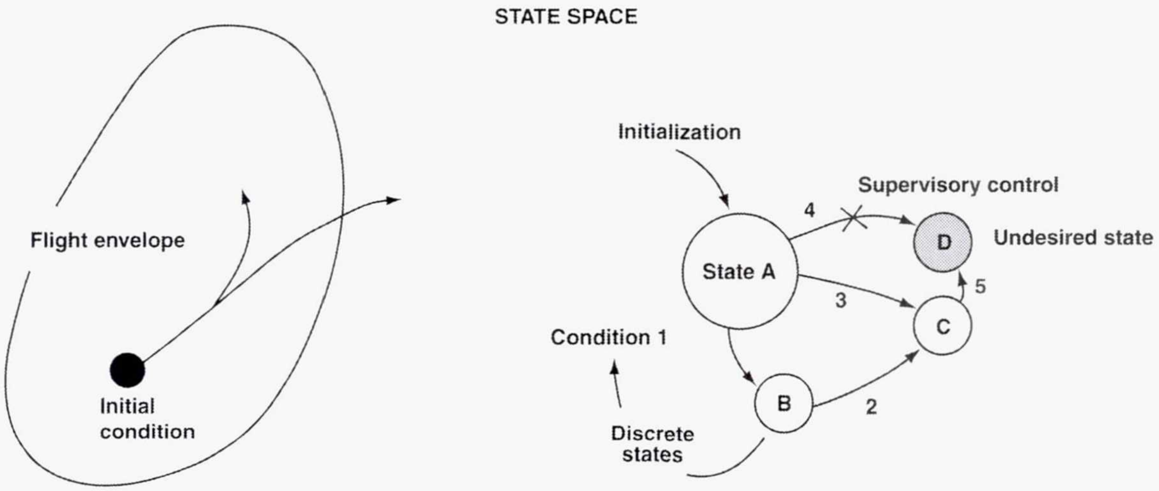
## Automation Theory

### Continuous Dynamical Systems

The dynamical behavior of an aircraft can be represented mathematically by a system of continuous differential equations that involve functions of state variables (that is, motion quantities such as position and velocity) and their time derivatives. The variations of these state variables with time (termed dynamical trajectories of the system) are illustrated by figure 2(a), which shows two possible trajectories determined by different applications of control or different environmental influences such as atmospheric turbulence. It can be seen that, starting from the same initial condition, one of these trajectories remains within the safe flight envelope, while the other, potentially catastrophic trajectory exceeds the boundaries of the safe envelope.

### Discrete Dynamical Systems

The various task-related modes of the automated control system of an airplane (for example, climb mode, altitude and heading hold modes, autoland mode, etc.) can be represented mathematically by a discrete dynamical system, as illustrated by figure 2(b). In figure 2(b), the state variables of the



a) Continuous system.

b) Discrete system.

Figure 2. Continuous and discrete systems.

discrete system (that is, the system modes) are represented by circles, and transitions between modes are represented by arrows. Each possible transition is governed by an associated logical condition: a transition occurs if and only if its associated logical condition is true. Starting from a known initial state (fig. 2(b)), the path taken through the state space defines the dynamical trajectory of the discrete system. It can be seen that different dynamical trajectories are possible, depending on the truth or falsity of the conditions involved.

For example, let State A (the initial state) correspond to the OFF condition, let State B correspond to engagement of the altitude hold mode, and let Condition 1 be set to TRUE if the pilot presses the button for selection of the altitude hold mode. Then the state transition diagram of figure 2(b) has the interpretation that the system transitions from OFF (State A) to ALTITUDE HOLD ENGAGED (State B) when the pilot presses the selection button (Condition 1 TRUE). Other conditions could be set TRUE or FALSE automatically, instead of being determined directly by pilot selection. Dependence of conditions on discrete states constitutes discrete feedback (fig. 2(b)).

Entry into certain states could result in undesirable or potentially catastrophic consequences, as illustrated by the shaded State D in figure 2(b). For example, State D could correspond to retraction of the landing gear with the aircraft on the ground, or to total electrical failure during an instrument approach, or (in a military context) to accidental launch of a missile. Fortunately, a well-developed theory of supervisory control enables the transition paths leading to such undesired states to be identified systematically and blocked off (fig. 2(b)). However, all such states must be identified and enumerated.

### Hybrid Dynamical Systems

A dynamical system that contains coupled continuous and discrete elements, as illustrated by figure 3, is termed a hybrid system. It can be seen that selection of control modes A, B, C, or D, each of which is characterized by a different control law, determines different trajectories for the continuous system (left side of figure 3). Therefore, the continuous system is coupled to the discrete system by means of the control laws characterizing the various modes.

Furthermore, because the truth or falsity of the conditions governing transitions within the discrete system depends in general on flight conditions (that is, on the state variables of the continuous system), perturbations in the continuous trajectories determine different trajectories for the discrete system (right side of figure 3) as the aircraft reacts to environmental influences. Therefore, the discrete system is coupled to the continuous system by means of the logic governing mode transitions. It follows that each of the two systems is coupled to the other.

### Dynamical Behavior

The dynamical behavior of continuous systems governed by differential equations is well-known. The theory of linear differential equations is mathematically complete, and the nonlinear differential equations describing aircraft motions can be linearized for small perturbations in the neighborhood of any desired operating point. Therefore, prediction of the dynamical behavior of the aircraft can be based on broad theoretical knowledge of the properties of the solutions of the governing differential equations, such as stability or instability, without reliance on numerical solution of the differential equations via simulation. Agreement between theoretical prediction, simulation, and experimental flight test results is generally satisfactory for transport aircraft.

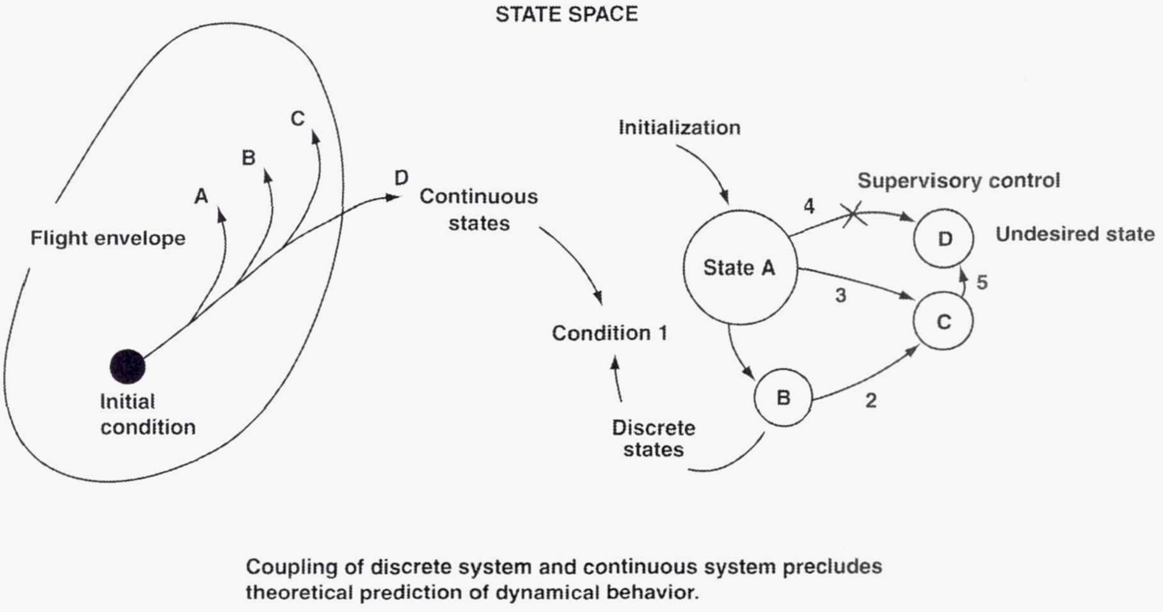


Figure 3. Hybrid system.

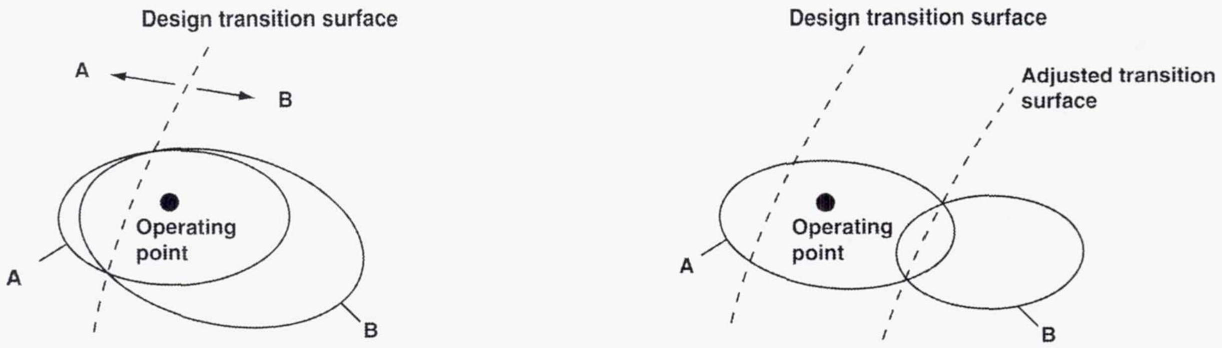
Prediction of dynamical behavior is more difficult for discrete systems than for continuous systems. When one trajectory is known for a continuous system such as an aircraft, physical continuity of the forces acting on the aircraft ensures that all neighboring trajectories must be similar in form. For discrete systems physical continuity plays no role, and even the notion of neighboring trajectories loses all significance. Knowledge of system behavior along one trajectory implies nothing for behavior along other trajectories, so that all combinatorial possibilities must be considered. Nevertheless, computational methods are available for discrete systems that enable all possible trajectories to be traced and enumerated, even for quite complex systems.

When a continuous system such as an aircraft is coupled to a discrete system consisting of several automated control modes to form a hybrid system like the flight management systems installed in current transport aircraft, the infinite-dimensional nature of the states characterizing the continuous system precludes enumeration. Furthermore, the continuous system introduces variable dynamical time delays that complicate the logical conditions governing state transitions within the discrete system. Therefore, the dynamical behavior of an arbitrarily specified hybrid system cannot be predicted theoretically by any known method.

## Automation Surprises

Further insight into the potentially catastrophic consequences of this lack of predictability can be gained by analyzing the behavior of the simplified hybrid system illustrated by figure 4, which shows the regions of validity in state space for two modes denoted as Mode A and Mode B. A formal definition of validity is presented later. Briefly, any continuous-system trajectory that lies within the illustrated boundaries is considered safe, and any excursion outside those limits is considered potentially catastrophic. These regions of mode validity must be regarded as bubbles floating in multidimensional state space. Furthermore, these bubbles move about in the state space and change shape as the system reacts to environmental influences.

STATE SPACE


Design transition surface
A
B
A
Operating point
B
Design transition surface
Adjusted transition surface
A
Operating point
B

To avoid invalid operation, transition surface must lie within intersection of regions of validity.

a) Normal conditions.

b) Abnormal conditions.

Figure 4. Regions of validity.

The diagram on the left (fig. 4(a)) illustrates a normal condition. At the operating point shown, which could lie anywhere on one of the trajectories illustrated on the left side of figure 3, both Mode A and Mode B are valid, and either could be selected safely. The selection logic is indicated by the transition surface, which shows that Mode A is selected for points to the left of the transition surface, and Mode B is selected for points to its right. For the operating point shown, Mode B would be selected. It is clear that, for a safe transition, the transition surface must lie within the intersection of the two regions of validity.

The diagram on the right (fig. 4(b)) illustrates an abnormal condition. It can be seen that the region of validity for Mode A remains the same as in figure 4(a), but that the region of validity for Mode B has moved to the right and diminished in size owing to system reaction to changes in environmental conditions, such as engine failure or wind shear. As a result, the operating point no longer lies within the region of validity for Mode B. If the transition surface in figure 4(b) remains specified as in figure 4(a), then Mode B will be selected even though it is invalid under the changed environmental conditions.

Such an invalid mode selection could result in automated system actions that are unsafe, or totally inappropriate for the intended operation, actions that take the human operators completely by surprise. A concrete example of an invalid mode selection that caused an aircraft accident is presented later in the section "An Example Illustrating Invalid Mode Selection." In order to avoid such potentially catastrophic mode selections, the transition surface must be adjusted as the system reacts to environmental influences so that it always lies within the intersection of the two regions of validity, as shown at the right of figure 4(b).

However, such adjustment of the transition surfaces in a complex reactive system is a very difficult problem, because it would require that many independent entities be maintained in strict correspondence under dynamically varying conditions. A simpler approach, which is taken in this report, is to eliminate separate transition surfaces entirely by making the regions of validity themselves the agency of mode selection.

### Design Methods

In the absence of any method for theoretical prediction of the dynamical behavior of reactive hybrid systems, system designers have been forced to take a computational approach via simulation, as already explained. The example illustrated by figure 4 shows that, because simulation of all environmental conditions is impossible, it cannot be determined by means of a purely computational approach whether an existing system satisfies requirements that ensure valid transitions in general. The resulting lack of behavioral predictability under exceptional conditions has been recognized as a serious human-factors deficiency (Wiener, 1989; Billings, 1996).

There remains the possibility of a theoretical attack on the inverse design problem. This inverse approach seeks to synthesize (that is, to deduce by a mathematical process) the specification of a hybrid system in such a way as to ensure that the system satisfies a set of dynamical behavior properties imposed a priori that rule out unsafe or inappropriate behaviors. This approach seems especially attractive for the design of transport aircraft flight management systems, for which the a priori specification of behavioral properties can take advantage of criteria for design and certification of jet transports that are based on more than 40 years of experience with both civil and military operation. This report presents a method for synthesizing mode selection logic for transport aircraft flight management systems so as to satisfy general safety and effectiveness properties specified a priori, enabling formal validation to be achieved.

## Design Goals and Technical Approach

### Design Goals

Major design goals are (1) system design integrity based on proof of logical correctness at the design level, (2) significant simplification and cost reduction in system development and certification, and (3) improved operational efficiency, with significant alleviation of certain human-factors problems encountered by pilots in current transport aircraft.

### Technical Approach

The technical approach to be taken leads toward these goals through development of a system architecture capable of formal verification and validation, which would replace much of the software testing currently required for certification with logical proof of algorithmic properties for all safety-critical system elements, and would enable comprehensive validation. Because such an architecture must of necessity make use of a drastically simplified mode structure, it may be expected to alleviate human-factors problems relating to mode confusion (Hughes and Dornheim, 1995).

### Synthesis Method

Following a suggestion of Michael Heymann (Heymann, personal communication, 1994), the present approach concentrates on investigating the region of state space for each mode that corresponds to valid operation of the system (fig. 4). Based on these well-defined regions, the state space is discretized geometrically, and the conditions for validity of each mode are tabulated. This table must be inverted (interchanging tabulated quantities with their arguments) to provide the basis for a mode selection strategy. However, it is found that naive inversion leads to a combinatorial explosion of cases. To avoid this explosion, heuristic strategies are used to partition the validity table, and the partitioned table is then inverted to obtain a logically complete enumerated list of relevant cases. Each case is related geometrically to a well-defined region of the state space, and selection criteria determine mode selection uniquely for each such region. The complete system satisfies general safety and effectiveness properties specified a priori. Behavioral properties are summarized by general theorems that enable formal validation of the complete system to be achieved. The potential for generalization of the method to other hybrid systems will be discussed later.

### Logical Consistency

A fully axiomatic treatment (Rushby, 1995) would seek to examine formally the logical consistency of the specified safety and effectiveness properties, which is not attempted here. It might seem that, in case of conflict between safety and mission effectiveness, in airline service the latter should always be sacrificed in favor of safety. But this formulation is too simple to deal with situations in which violation of effectiveness could in itself give rise to a safety hazard. For example, if a wind shear that overwhelmed aircraft performance should be encountered during final approach, avoiding the a priori hazard of stalling might seem to require maintaining the conventional generous airspeed margins above stalling speed, even if this caused the aircraft to settle below the approach path. But it is now well-accepted that the safest policy for traversing wind shear requires holding the approach path to avoid striking the ground short of the runway, and sacrificing airspeed margin, if necessary, right down to stalling speed. More subtle examples of cases in which violation of effectiveness could create a hazard can be found in the accidents and incidents selected for discussion in appendix G. In several of these cases, the hazard was compounded by poor annunciations to the human crew, compromising their potential capability for resolving problems not known to the designers a priori.

Recognizing the limitations of our knowledge in this human-factors area, the present work does not attempt a complete axiomatic treatment. Instead, it deals on a case-by-case basis with logical conflicts between safety and effectiveness, such as performance limitations that preclude capture of altitude or airspeed targets. In each case, the exact nature of the limitation is annunciated for resolution at higher levels within the system hierarchy, or, ultimately, by the human crew. Assurance that all such cases of conflict are certain to be identified is provided by a guarantee of logical completeness.

## Specific Objectives

The primary objective of the present work is to enable formal validation by developing a practical design procedure, one useful to industry designers in the near term, for synthesis of hybrid systems that satisfy general safety and effectiveness properties specified a priori.

A secondary objective is to realize the potential of formal validation for improved operational safety, with particular concern for the design integrity of the complete system, including the human/machine cockpit interface, and to provide a firm technical basis for certification criteria.

A tertiary objective is to contribute to the development of more general theoretical methods for synthesizing hybrid systems directly from design requirements.

## Scope

This report treats only the problem of transport aircraft longitudinal control, which is more interesting than lateral-directional control because two longitudinal response parameters (flightpath angle and airspeed) are controlled by two control parameters (pitch and thrust). Furthermore, the longitudinal problem is complicated by abrupt saturation of the thrust control. In contrast, lateral-directional control in coordinated flight with small sideslip angles involves only one response parameter (heading) and one control parameter (bank angle) not subject to abrupt saturation. For simplicity, no failures are treated in detail in this report, with the exception of abrupt failure of one engine, as previously explained. However, methods that enable detailed treatment of other failures are discussed briefly.

The elements of the Vehicle Management System that are treated in this report are indicated by the shaded blocks in figure 1. The continuous elements treated are the outer-loop control, the inner-loop control, and the aircraft itself together with its control surface and throttle servos. The design of these continuous elements uses the nonlinear inverse control concept developed theoretically by Meyer (Meyer and Cicolani, 1981) and applied by Franklin (Franklin, Hynes, Hardy, Martin, and Innis, 1986) to the flight control system of a NASA research aircraft.

As indicated by the shading in figure 1, the discrete elements of the Vehicle Management System treated in detail in this report consist of the mode control logic for the first two levels of the mode hierarchy, which correspond functionally to the autoflight systems installed in current transport aircraft. As indicated by the cross-hatching in figure 1, guidelines are presented in this report for simplified development of guidance functions and associated third-level supermodes, which correspond functionally to current vertical navigation (VNAV) systems.

The navigation function and the corresponding elements for lateral-directional control can be treated by applying the synthesis methodology developed in this report. It can be seen that, on the right (tactical) side of the vertical dashed line in figure 1, the two remaining elements are, first, the cockpit interfaces grouped within the display function, and second, the human crew's role in tactical supervision. Comments on human-factors research needed to model these elements, which could enable complete integration of the human-automation system, can be found in the section "Future Work."

## Plan of Report

### Contents of Report

The report is divided into two volumes and three major parts. Part I ("Background") begins with the "Introduction," and presents a brief overview of the formulation and solution of the synthesis problem in the section "Overview of Design Method."

Next, in the section "Aircraft Model," the aircraft model is described in detail. The dynamical response of the aircraft to longitudinal control is developed from the governing differential equations, and related geometrically to the forms of trajectories in the flightpath-airspeed plane.

Part II ("Design Synthesis") begins with the section "Flight Control System," in which the mode hierarchy is described, and the continuous elements of the flight control system are treated in detail. The lowest-level elements in the mode hierarchy, which are termed primitive modes (that is, not composite), relate directly to the dynamical behavior of the aircraft. The higher-level elements relate to the control tasks required for operation within the ATC system, such as capturing and holding an altitude, tracking the approach path, and the like. Three primitive modes for automated longitudinal control are defined, and their stability properties are analyzed and related geometrically to certain regions of the flightpath-airspeed plane.

In the section "Safety Requirements," flight safety hazards are identified a priori, and are associated with certain boundary contours in the flightpath-airspeed plane. Safety margins are then applied to these boundaries to define the safe flight envelope. Flight envelope protection is discussed, and envelope protection modes are defined.

In the section "Effectiveness Requirements," the concept of effectiveness of the automated control system for its intended purpose is developed, and related to the performance capability of the aircraft. (The concept of effectiveness plays a role in the synthesis problem similar to that of liveness in computer science.) Complete effectiveness is defined to require capture of altitude, airspeed, and flightpath targets. Partial effectiveness is defined, and requirements for annunciation of violations of effectiveness are specified.

In the section "Mode Validity," validity is defined to require both safety and effectiveness. Validity is analyzed for each mode based on the differential equations specified by the control law, and is related geometrically to the stability regions of the flightpath-airspeed plane studied previously. The conditions for validity of each mode are summarized by a validity table, which completes the formulation of the synthesis problem.

Solution of this problem for the lowest-level hybrid system is discussed in detail in the section "Synthesis of Path/Speed Command Supermode." Synthesis based on naive inversion of the validity table is shown to be intractably complex. Heuristic strategies based on general notions of maximum effectiveness, best approximation, and logical dominance are adopted, and are used to partition the validity table, accounting for logical dependencies. The partitioned table is then inverted to obtain a logically complete enumerated list of relevant cases. Each case is related geometrically to a well-defined region of the flightpath-airspeed plane, and mode selection criteria determine mode selection uniquely for each such region; statecharts specifying mode selection logic are constructed directly from the mode selection tables. The complete system satisfies the safety and effectiveness properties specified a priori.

In the section "Synthesis of Altitude Command Supermode," extension of the synthesis method to the second level of the mode hierarchy is described, with detailed application to altitude control. System properties are summarized by general theorems that enable formal validation of the complete system to be achieved.

The section "Other Second-Level and Third-Level Supermodes" presents guidelines for simplified development of other second-level and third-level supermodes based on modifications to the second-level Altitude Command supermode, and shows how the complete Vehicle Management System could be used in airline service. This section completes the second major part of the report, "Design Synthesis."

The last major part of the report, Part III, "Consequences," deals with assessment of its consequences for airworthiness and certification, and for human-factors design of cockpit interfaces. The section "Airworthiness, Certification, and Cockpit Interface Design Issues" discusses those issues in the context of several transport aircraft accidents and incidents. These accidents and incidents are selected to show the broad applicability of airworthiness principles based on the safety and effectiveness properties developed in this report. Simple but plausible analysis shows that, although the causes of these accidents and incidents differ greatly in detail, each of them resulted from some violation of safety and effectiveness properties. (In anthropomorphic terms, each of the automated systems involved exhibited a lack of trustworthiness in carrying out its assigned task that would have been recognized as unacceptable had the same task been assigned to a human crewmember.) Furthermore, these problems have not been alleviated by experience, because the most recently designed systems are characterized by violations of safety and effectiveness properties similar to those encountered in previous design generations. The discussion leads to the suggestion that formal validation might significantly improve operational safety. Cockpit interface design is discussed briefly, and several heuristic guidelines are proposed.

The next section, "Future Work," discusses the possibility of modeling the lower levels of the human pilot's activities involved in tactical supervision, enabling formal validation to be extended to include the complete human machine cockpit interface. The potential for generalization of the synthesis method developed in this report to other hybrid systems such as highly automated industrial plants for petrochemicals and nuclear power is discussed in some detail.

Finally, the "Conclusions" are presented, and the "References" and "Figures" complete the report. The main report is presented in "Volume One."

Volume Two presents seven appendices, which are intended to make the main report accessible to readers with various backgrounds.

Appendix A presents a tutorial review of the U. S. Standard Atmosphere, with atmospheric parameters tabulated at six selected altitudes for reference.

Appendix B presents a tutorial review of transport aircraft design and operation. Dynamic models with three different levels of fidelity are developed to support the present work; however, these models may also be useful for future studies. Dynamic simulation is implemented by means of a widely available spreadsheet, and several flying qualities criteria are illustrated by time histories of aircraft motions in response to elevator steps and pulses.

Appendix C presents a rigorous development of the aircraft equations of motion derived by elementary methods in appendix B, and generalizes those equations for symmetric (coordinated) turning flight.

Appendix D presents a brief tutorial review of elementary propositional logic, and provides a list of theorems in symbolic logic for reference.

Appendix E discusses the application of elementary propositional logic to real-time embedded systems; statechart semantics are summarized, and modifications are proposed that facilitate implementation within a clock-based sequential machine.

Appendix F develops methodology for generating formal proofs, and applies it to provide rigorous proof of several theorems that summarize the dynamical behavior of the transport aircraft and control system described by the main report, enabling formal validation of the complete system to be achieved. Independently verified regulator properties that ensure eventual capture of altitude and airspeed targets are taken as axioms of the formal system, ignoring time-dependent detail. This methodology enables application of ordinary (static) propositional logic to the dynamical system. Currently available codes for automated hypothesis testing can provide the basis for a theorem-proving tool, and it is suggested that such tools could play a crucial role in future system development.

Appendix G presents brief statements summarizing the aspects of several selected transport aircraft accidents and incidents that are considered relevant to issues of system design.

### Recommendations for Readers

Readers desiring only an overview of the design procedure should read the section, "Overview of Design Method," and then skip to the "Conclusions." Readers interested in the design and operation of the complete system, but not in the details of the method of logical synthesis, should skip the section "Synthesis of Path/Speed Command Supermode" and the synthesis discussion in the section "Synthesis of Altitude Command Supermode." Readers interested primarily in the human-factors aspects of formal validation should read the sections "Overview of Design Method," "Airworthiness, Certification, and Cockpit Interface Design Issues," "Future Work," and "Conclusions."

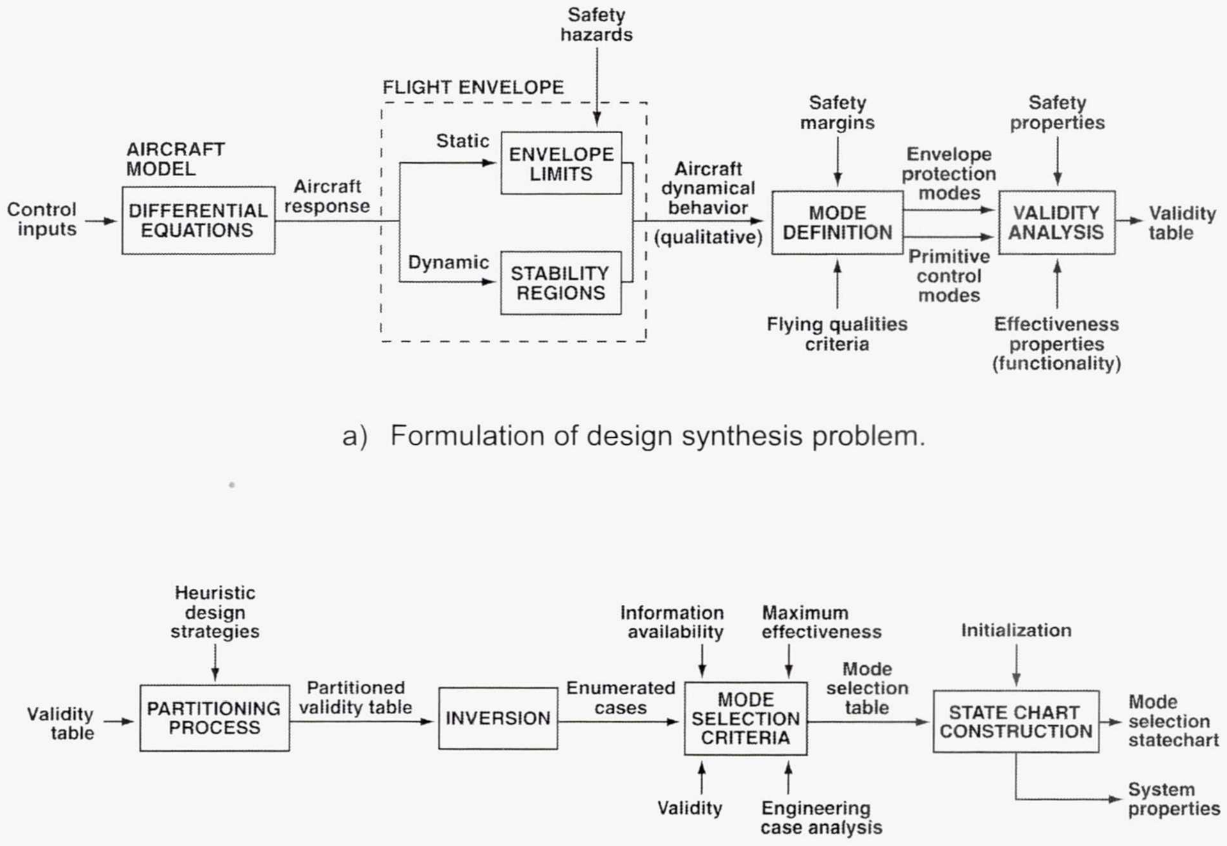
Readers desiring additional background in transport aircraft design and operation can consult the tutorial treatments in appendices A, B, and C, and the summary of selected accidents and incidents in appendix G. Readers desiring additional background in formal logic and its application to real-time computation within embedded systems can consult appendices D and E. Specialists in formal logic and automated hypothesis testing who are interested in the contribution that their expert knowledge can make to avionic system design should read the sections "Overview of Design Method," "Synthesis of Path/Speed Command Supermode," "Synthesis of Altitude Command Supermode," and appendix F. The formal proofs contained in appendix F are essential for formal validation of the complete system, but it is expected that only readers familiar with formal logic and others with a vital interest in formal validation will need to follow these proofs in detail; the introductory discussion of proof methodology should be of more general interest.

## OVERVIEW OF DESIGN METHOD

This section provides a brief, qualitative overview of the design method that is treated in detail in this report. Because this section provides an overview of selected sections of the report, references to tables 2, 3, 4, 6, and 7, and to figures 6 and 7 do not appear here in numerical sequence. The design procedure, which is illustrated by the block diagram of figure 5, can be summarized by the following steps (terms in *italics* refer to the elements of figure 5):

### Formulation of Synthesis Problem

1. The *aircraft model* for longitudinal motions in the plane of symmetry is developed, and the governing differential equations are derived. The aircraft motion is described, like that of any rigid body, by six ordinary differential equations. As a consequence of the assumed bilateral symmetry of both the aircraft itself and the surrounding aerodynamic flow field, those six equations separate into two uncoupled sets of three equations each, one set describing the *longitudinal* motions, which are the subject of this report, and the other set describing the *lateral-directional* motions, which are not treated.



a) Formulation of design synthesis problem.

b) Solution of design synthesis problem.

Figure 5. Design Method.

The aircraft model is further simplified by the assumption that elevator control is sufficiently powerful and pitch response sufficiently rapid to generate any desired lift variation. This assumption eliminates the pitching moment equation from the three-degree-of-freedom model, reducing the aircraft model to two degrees of freedom. Because rotational pitch dynamics are eliminated from the model together with the pitching moment equation, the two-degree-of-freedom model treats the aircraft as a point mass concentrated at its center of gravity. A detailed justification for use of this simplified model is presented, based on response calculations that demonstrate the excellence of the approximation for representative transport aircraft (appendix B).

The response parameters of the two-degree-of-freedom model (eqs. (1) and (2)) are flightpath angle (climb angle) and airspeed, and the control parameters are pitch and thrust. The state space for the continuous differential equations consists, therefore, of the two-dimensional flightpath-airspeed plane. Furthermore, flightpath angle and airspeed are also the selected output variables.

2. The differential equations are solved for the steady state, and the *flight envelope* diagram is constructed (fig. 8) by cross-plotting flightpath angle versus airspeed with contours of pitch and thrust as parameters. The physical limits of the flight envelope *(envelope limits*), which are *quasi-static*, correspond to minimum thrust, maximum thrust, minimum airspeed, and maximum airspeed. Minimum airspeed is determined in normal symmetric flight by the stalling speed, and in abnormal engine-out flight by the minimum control speed. Maximum airspeed is determined by structural limits at low altitude and by compressibility (Mach) limits at high altitude.

   During manual flight, the pilot controls pitch with elevator and engine revolutions per minute (rpm) with throttle. It can be seen from the flight envelope chart of figure 8 that the steady-state solution of the differential equations for any combination of control inputs is determined by the intersection of the corresponding pitch and thrust contours. The desired solution is termed the target operating point.

3. The *dynamic response* of the aircraft to control is related to its *flight envelope* as follows. The accelerations acting on the aircraft can be determined graphically from the flight envelope diagram by noting the position of the point corresponding to the current path and speed of the aircraft relative to the pitch and thrust control contours (fig. 9), which is equivalent to constructing the direction field of solutions to the differential equations. By stepwise numerical integration of the accelerations, each solution of the differential equations can be mapped into a unique path in the path-speed plane. (Representative paths connecting initial and final points are illustrated in the flight envelope diagrams of figure 10.) Thus, in principle, complete information about the *dynamical behavior* of the aircraft is contained geometrically in the flight envelope diagram.

4. Analysis of the differential equations shows that, qualitatively, only three kinds of aircraft behavior need to be considered. If both pitch and thrust are available for control (termed Type (i) control), then both flightpath angle (climb angle) and airspeed can be controlled independently. If thrust is fixed (saturated) so that only pitch is available for control, then either airspeed can be controlled at the expense of flightpath angle (termed Type (ii) control), or else flightpath angle can be controlled at the expense of airspeed (termed Type (iii) control). By analyzing the

stability properties of the solutions corresponding to each of the three kinds of dynamical behavior, the flight envelope is partitioned geometrically into seven discrete regions (fig. 10), within each of which the dynamical behavior of the aircraft is qualitatively similar. Physically, the boundaries separating these *stability regions* correspond to thrust control saturation and to speed for minimum drag.

These boundaries are treated as mathematically sharp. (Uncertainties contributed by noisy measurements or external disturbances remain to be dealt with during implementation.) With sharp boundaries, logical completeness of the classification scheme is established by inspection of the flight envelope diagram to ensure that all points in the plane lie in one discrete region or another, so that no point exists in the state space that is not identified by its region.

5. *Primitive control modes* are defined for the automated system (fig. 11) that correspond qualitatively to the three kinds of dynamical behavior, and control laws are specified quantitatively to satisfy applicable *flying qualities criteria* for manual control.

6. Flight *safety hazards* are identified a priori, and *safety margins* providing protection against these hazards are applied to the physical limits of the flight envelope to define safe envelope limits (fig. 13). Envelope protection is discussed, and *envelope protection modes* are defined.

7. General *effectiveness properties* that the complete system is required to satisfy to assure capture of flightpath and airspeed targets are specified, and are related to the performance capability of the aircraft. Steady-state values of flightpath angle and airspeed must lie within the aircraft flight envelope. Complete effectiveness, partial effectiveness, and normal effectiveness are defined, and requirements for annunciation of effectiveness violations are specified.

8. Mode validity is defined to require that both safety and effectiveness properties hold. For each primitive mode, a *validity analysis* based on the differential equations specified by the control law is performed, and the region in state space (the flightpath-airspeed plane) is determined that corresponds to valid operation of the aircraft in that mode. In figure 10(a), only Region I is valid for Type (i) control. In figure 10(b), all regions except Regions VI and VII are valid for Type (ii) control. In figure 10(c), only Regions I, II, III, and IV are valid for Type (iii) control.

   Discrete validity conditions are enumerated, and these conditions are related geometrically to the discrete stability regions determined previously (fig. 10). The validity conditions are summarized by a table (table 2) that gives the validity of each mode as a function of the instantaneous measured values of path and speed, and of the desired (target) values of these parameters. This *validity table* completes the formulation of the synthesis problem.

## Solution of Synthesis Problem

9. In principle, the mode selection logic for the primitive modes can be synthesized by inversion of the validity table (table 2), interchanging arguments with tabulated quantities. Before inversion, validity conditions are tabulated, with modes and parameters as arguments. After inversion, truth values would be tabulated with modes and validity conditions as arguments (that is, the inverted

table would be a truth table for table 2). However, brute-force inversion based on enumeration of all cases is shown to be intractably complex.

In table 2, there are 10 binary conditions. (These binary conditions can be seen more clearly by expressing them in terms of symbolic logic, as in table 3.) It may be seen that there are 5 binary conditions that determine the validity of the $\gamma$-V Command mode (first column of table 3), 7 conditions that determine the validity of the V Command mode (second column), and 7 conditions that determine the validity of the $\gamma$ Command mode (third column). If these conditions were all independent, there would be $2^5 = 32$ combinations requiring evaluation for assessment of the validity of the $\gamma$-V Command mode, $2^7 = 128$ combinations for the V Command mode, and $2^7 = 128$ combinations for the $\gamma$ Command mode. After accounting for duplication of conditions between columns, it is found that there are actually 10 independent conditions in table 3. It follows that, for assessment of the validity of all three primitive modes by inversion of the whole table, there would be $2^{10} = 1024$ combinations requiring enumeration. Furthermore, a mode selection policy would have to be formulated specifying a unique mode selection for each of these 1024 combinations. Therefore, it must be expected that a naive, brute-force approach based on enumeration of all cases would be found intractably complex.

10. *Heuristic design strategies* for mode selection and for input screening of target values (table 4) are developed, based on general notions of maximum effectiveness, best approximation, and logical dominance. When adopted, these strategies can be regarded as part of the functional specification if desired.

11. The validity table (table 2) is *partitioned* with the help of the heuristic strategies and extensive use of elementary symbolic logic, accounting for logical dependencies.

12. The *partitioned validity table* is inverted to generate a logically complete list of *enumerated cases* (table 6). Each case corresponds to a single combination of conditions that determine the validity of each mode, and is related geometrically to a well-defined region of the flightpath-airspeed plane.

13. *Mode selection criteria* are applied to each enumerated case to generate the *mode selection table* (table 7). In some cases *validity* considerations alone are sufficient to determine mode selection. In some cases for which no modes are valid, the *information available* is nevertheless sufficient to determine mode selection; otherwise, *engineering analysis* is applied on a *case-by-case* basis to determine the least adverse choice. In cases for which multiple modes are valid, mode selection is based on *maximizing effectiveness*. The resulting mode selection table (table 7) determines mode selection uniquely for each enumerated case and for the geometric region (fig. 10) to which that case corresponds.

14. The *mode selection statechart* (fig. 15) is *constructed* directly from the mode selection table, specifying mode *initialization* according to a simplified initialization strategy developed separately.

15. After completing steps 1–14, *system properties* are summarized by concise statements of system behavioral properties that enable formal validation of the complete system to be achieved. (For the Altitude Command supermode, which is functionally similar to the Flight Level Change mode in current transport aircraft, the dynamical behavior of the synthesized system is summarized by five theorems (appendix F) that demonstrate satisfaction of the general safety and effectiveness properties imposed a priori.)

## An Example Illustrating Invalid Mode Selection

Partitioning of the two-dimensional state space into well-defined regions of mode validity, as just discussed, has laid the groundwork for a concrete example illustrating the potentially catastrophic consequences of invalid mode selection, which is presented next. This example supplements the abstract discussion of automation surprises presented in the "Introduction."

### Performance Envelope

Figure 6 illustrates the performance envelope for a representative transport aircraft. It can be seen that the upper broken contour represents the variation of maximum flightpath angle (climb angle) with airspeed with all engines operating at maximum thrust, and the solid contour (center of diagram) represents its variation with one engine inoperative. The operating point shown (open circle) is typical for the initial climb after takeoff. Comparison with figure 10(a), shows that the initial operating point lies in Region I.

### Engine Failure After Takeoff

Now assume that an engine fails abruptly, and that maximum available thrust is applied to the remaining engines. The maximum possible climb angle is therefore reduced to the level corresponding to the solid contour, and the thrust control becomes saturated. Since by definition Type (i) control requires unsaturated thrust, Type (i) control is not available, and either Type (ii) control (that is, speed control) or Type (iii) control (that is, path control) must be selected.

### System Behavior

Comparison with figure 10(c) shows that just after engine failure the operating point lies in Region V, even though the operating point itself remains unchanged during the brief time interval during which the failure is assumed to take place—it is the maximum thrust contour that changes, not the operating point.

In Region V, Type (iii) control is invalid. If Type (iii) control is selected, the aircraft trajectory will follow the horizontal path marked P in the diagram (fig. 6): path is controlled at the expense of speed, which decreases steadily until penetration of the minimum-control-speed boundary (left side of diagram) causes potentially catastrophic loss of control.

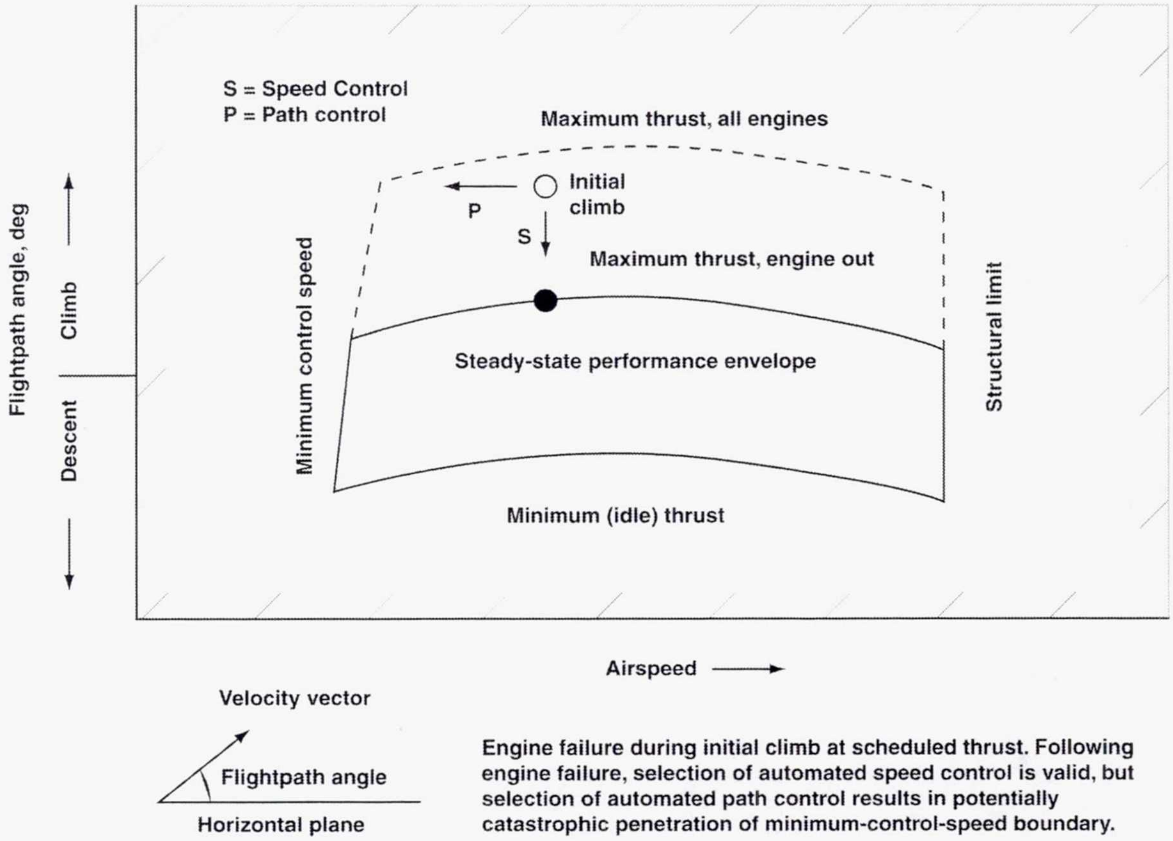
On the other hand, if Type (ii) control is selected, the aircraft trajectory will follow the vertical path marked S in the diagram (fig. 6): speed is controlled at the expense of path, which decreases steadily until equilibrium is reached as shown (filled circle) on the maximum-thrust contour. Certification criteria require that the resulting climb gradient must provide adequate obstacle clearance, ensuring a safe outcome when speed control is maintained.

## Discussion

It is clear that Type (ii) control should be selected after engine failure. The resulting operating point has two desirable properties: first, it represents the best approximation to the initially selected operating point that is physically realizable following engine failure; and second, it results in the steepest climb gradient achievable at the selected airspeed. Nevertheless, selection of Type (iii) control has caused at least one aircraft accident under similar conditions. Avoiding such catastrophic mode selections presents a challenge to designers of automated systems.

The method of synthesizing mode selection logic that is developed in this report avoids selection of invalid modes whenever possible. If none of the available modes is valid, case analysis based on enumeration enables selection of the least adverse choice. If several modes are valid, the selection is made so as to maximize effectiveness. This strategy results in best physically realizable approximations when the aircraft is subject to performance limitations of various kinds.

In the engine failure case discussed in this example, it should be noted that correct mode selection need depend only on mode validity (which is based on general identification of geometrical stability regions within the state space), and does *not* require specific identification of engine failure. With this validity-based mode selection strategy, it is recognition that the target climb angle exceeds the



**Engine failure during initial climb at scheduled thrust. Following engine failure, selection of automated speed control is valid, but selection of automated path control results in potentially catastrophic penetration of minimum-control-speed boundary.**

Figure 6. Engine failure during initial climb.

performance capability of the aircraft that triggers reversion to Type (ii) control, not recognition of engine failure. Therefore, any other failure that caused a similar loss of performance (perhaps a hydraulic fault that prevented landing-gear retraction) would result in a similarly safe outcome—all that is required is onboard determination of validity regions within the state space (that is, determination of performance-envelope limits) with sufficient accuracy. This strategy stands in sharp contrast to other design methods that attempt specific identification of each possible abnormal condition in order to enable appropriate mode selection.

## AIRCRAFT MODEL

The aim of the following discussion is to enable a qualitative understanding of "what the automated control system is trying to do, and how it does it" that is accessible to readers without a detailed mathematical grasp of the properties of nonlinear differential equations. The necessary background follows: (1) elementary differential calculus for definition of derivative; (2) elementary integral calculus for definition of integral, and for approximate evaluation of definite integrals by trapezoidal rule; and (3) elementary physics for formulation of differential equations according to Newton's second law.

Because the differential equations governing the aircraft motion are solved numerically by the control system, and because this motion is driven by pilot control inputs and by atmospheric disturbances whose forms are unknown until they are encountered in real time, classical methods of solving differential equations for known forcing functions are of little help in understanding control system actions and aircraft dynamic responses. (Dynamical responses to representative simplified pilot control inputs are calculated numerically and discussed in detail in appendix B.) Fortunately, the qualitative understanding of aircraft dynamical behavior that is essential for this report is not difficult, and indeed is already familiar to pilots.

### Definitions and Assumptions

Readers unfamiliar with aeronautical terminology would no doubt find unintelligible the summary statement that the mathematical aircraft model to be developed will be based on *path-frame* description of the motion of a *rigid* aircraft with negligible *thrust inclination* during *straight, symmetric, wings-level flight* in *still air* over a *flat, nonrotating Earth.* The discussion must therefore begin by explaining the technical definitions of the italicized terms and the simplifying assumptions they imply. A more expanded treatment can be found in appendix B.

#### Symmetric

Not only is the aircraft structure bilaterally symmetric, but the aerodynamic flow fields on each side of the aircraft plane of symmetry are mirror images of each other. The small asymmetric rotational effects owing to like-rotation engines are neglected. The aircraft velocity vector lies in the plane of symmetry.

Transport aircraft operate in symmetric flight except during crosswind landing and in case of engine failure. The effects of thrust asymmetry owing to engine failure are discussed later.

### Straight

The projection of the aircraft trajectory in the ground plane (its *track*) is a straight line. For straight, symmetric flight, the wings of a symmetric aircraft must be level. Results are generalized later to include symmetric (coordinated) turning flight (appendix C).

### Rigid Aircraft

By definition, the distances between each pair of points in a rigid body must remain fixed, but transport aircraft cannot be considered structurally rigid (indeed, they are quite flexible). Nevertheless, the frequencies characterizing elastic structural deformations (vibration) usually lie so far above those for rigid-body motions of the whole aircraft that the airflow can be considered to act upon the deformed aircraft. In that case (the *quasi-static* assumption), the aircraft can be treated as a rigid body provided that aerodynamic force and moment parameters are suitably modified to incorporate the effects of structural deformation. With the quasi-static assumption, aircraft motions (like those of any other rigid body) can be described by six ordinary differential equations involving six degrees of freedom.

Movable control surfaces are assumed to be balanced and irreversibly actuated, so that no additional differential equations are required for description of their motions. Gyroscopic effects of rotating engine machinery are neglected.

### Still Air

The static atmospheric environment is specified by the Standard Atmosphere (appendix A), which describes the variations of ambient temperature, pressure, air density, and sonic velocity with height above the surface of the Earth. The effects of atmospheric disturbances such as steady (horizontal) *wind*, steady (vertical) *draft*, *wind shear*, and *atmospheric turbulence* are discussed later.

### Flat, Nonrotating Earth

The gravitational force at the surface of the Earth is assumed to remain fixed independent of height, and small Coriolis and centrifugal ("weightlessness") accelerations are neglected. The spherical Earth is approximated by a tangent horizontal plane, with the point of tangency directly below the aircraft (local tangent plane approximation). These approximations are valid for control of subsonic transport aircraft.

## Reference Frames

### Body Frame

The origin of the *body frame* is located at the aircraft center of gravity. Its longitudinal X axis is aligned with the aircraft reference fuselage axis (often parallel to the cabin floor). The lateral Y axis is orthogonal to the plane of symmetry and positive toward the right wing. The vertical Z axis is positive downward in the plane of symmetry, forming a right-hand orthogonal triad with the X and Y axes. The pitch angle $\theta$ measures the angular elevation of the X axis relative to the local horizontal plane (fig. 7(a)). It is essential to describe the rotational motions of the aircraft in the body frame to avoid appearance of the moments and products of inertia as time-varying parameters in the equations of motion.

### Path Frame

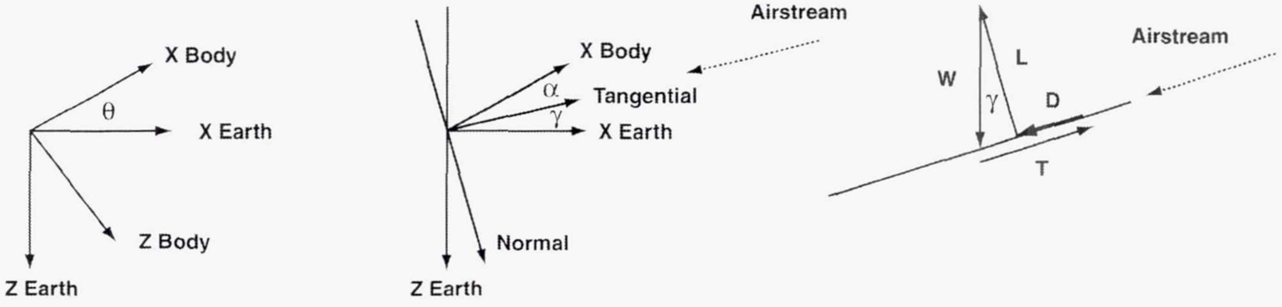
The origin of the *path frame* is located at the aircraft center of gravity. Its tangential *T* axis is aligned with the aircraft velocity vector. In symmetric flight, the lateral *L* axis is orthogonal to the plane of symmetry, and coincides with the Y axis of the body frame. The normal *N* axis is positive downward in the plane of symmetry, forming a right-hand orthogonal triad with the *T* and *L* axes. The flight-path angle $\gamma$ measures the angular elevation of the *T* axis relative to the local horizontal plane. In straight, symmetric flight the angle of attack $\alpha$ is equal to the angular difference $(\theta - \gamma)$ (fig. 7(b)); symmetric (coordinated) turning flight is discussed later.

Pilots perceive these angular relationships from inside the aircraft, of course, and express them by saying that the pitch angle describes where the aircraft is pointing, the flightpath angle describes where it is going, and the angle of attack is the difference between where the aircraft is pointing and where it is going. Except during takeoff and landing, the pilot's primary control task can be regarded as control of the velocity vector in such a way as to realize the mission objectives summarized by the flight plan; pitch control is then considered a means by which control of the velocity vector can be achieved.

It will be shown next that the aerodynamic lift and drag forces and also the engine thrust force are aligned with the axes of the path frame, so that only the gravity force requires trigonometric transformation. Therefore, choice of the path frame for the translational motions is convenient from two viewpoints: first, the pilot's task is closely related to the path-frame motion parameters (flightpath angle and airspeed); and second, the force description and the resulting equations are simplified.

### Lift, Drag, and Thrust

The propulsive force (*gross thrust*) acts in the plane of symmetry; its component along the *T* axis is termed the *net thrust*. *Thrust inclination* relative to the *T* axis is negligible for conventional transport aircraft, together with interference effects between engine efflux and external aerodynamic flow. Therefore, the total engine thrust is assumed to act along the *T* axis. By definition, the aerodynamic *drag* force acts in the negative *T* (streamwise) direction, and the *lift* force L acts upward in the negative *N* direction (fig. 7(c)). Under the *quasi-steady flow* assumption (appendix B), the aerodynamic forces are assumed to depend only on the (vectorial) velocity of the aerodynamic flow relative to the aircraft.



a) Earth frame and body frame.

b) Path frame aligned with velocity vector.

c) Force equilibrium.

Figure 7. Reference frames.

## Equations of Motion

Under the assumptions just discussed, the aircraft motion is described, like that of any rigid body, by six ordinary differential equations. As a consequence of the assumed bilateral symmetry of both the aircraft itself and the surrounding aerodynamic flow field, those six equations separate into two uncoupled sets of three equations each, one set describing the *longitudinal* motions, which are the subject of this report, and the other set describing the *lateral-directional* motions, which are not treated. The set of three longitudinal equations comprises (1) the *longitudinal force equation* (sometimes termed the *streamwise force equation*), which describes translation along the *T* axis of the path frame, (2) the *normal force equation* (sometimes termed the *lift equation*), which describes translation along the *N* axis of the path frame, and (3) the *pitching-moment* equation, which describes rotation about the Y axis of the body frame.

### Three-Degree-of-Freedom Model

These three *equations of motion* are obtained by application of Newton's second law to the rigid aircraft. An elementary derivation can be found in appendix B, and a rigorous but nonelementary derivation of the two force equations making use of vector-matrix methods is presented in appendix C. The three longitudinal equations of motion are as follows (fig. 7(c)):

$$\frac{1}{g}\frac{dV}{dt} = \frac{T - D}{W} - \sin\gamma \tag{1}$$

$$\frac{V}{g}\frac{d\gamma}{dt} = \frac{L}{W} - \cos\gamma \tag{2}$$

$$I_Y \frac{dq}{dt} = M_Y\,(q, d\alpha/dt, \alpha, V, \delta_{ELEV}) \tag{3}$$

where the notation is given in the "List of Symbols."

The quantity (1/g)(dV/dt) on the left side of equation (1) is the (normalized) longitudinal acceleration, and the quantity (V/g)(dγ/dt) on the left side of equation (2) is the (normalized) normal acceleration, which is proportional to curvature of the flightpath in the plane of symmetry. The lift and drag forces are related to angle of attack and airspeed by functions (appendix B) of the form

$$L = L\,(\alpha, V) \quad \text{and} \quad D = D\,(L, V) \tag{4}$$

The angle of attack is related to pitch angle and flightpath angle by the kinematic equation

$$\alpha = \theta - \gamma \tag{5}$$

which is valid for straight, symmetric, wings-level flight, as already noted.

It can be seen that equation (1) is coupled to equation (2) by the $\gamma$ term, and to equation (3) by the lift and drag variations of equation (4). Equation (2) is coupled to equation (3) by the lift variation (4) and the kinematic relation (5), and to equation (1) by the airspeed factor V on the left side of equation (2). Equation (3) is coupled to the other equations by the functional dependence of pitching moment on pitch rate, angle-of-attack rate, angle of attack, and airspeed, as indicated on the right side of equation (3). Therefore, the three differential equations form a coupled nonlinear dynamical system. Detailed numerical calculations of the response of this coupled system to elevator impulses and steps are presented in appendix B.

### Two-Degree-of-Freedom Model

The pitching moment $M_Y$ on the right side of equation (3) is proportional to the elevator deflection $\delta_{ELEV}$, which the pilot controls directly during manual flight by means of the cockpit control column. The pilot shapes the column control input as necessary to obtain the desired pitch response, despite the unwanted coupling that results from the dependence of the pitching moment $M_Y$ upon the angle of attack $\alpha$ and its rate of change $d\alpha/dt$ and upon the airspeed V (eq. (3)). This desired pitch response then results in the desired angle-of-attack response (eq. (5)) and the desired lift (eq. (4)). The lift determines the normal acceleration according to equation (2). Indeed, one of the oldest and most important flying qualities criteria is the specification of column force required to generate an incremental normal acceleration of 1g.

The time integral of normal acceleration then determines the flightpath response according to equation (2), which is coupled with equation (1), as already noted. Dynamic response calculations for a representative transport (appendix B) show that flightpath follows pitch with a time lag in the order of 2 sec that characterizes the aircraft. Flightpath response to an elevator (column) impulse stabilizes within 3 sec (the *short term*); airspeed responds much more slowly.

Therefore, provided that the elevator control is sufficiently powerful and the pitch response sufficiently rapid (that is, pitch response bandwidth is sufficiently wide) relative to the flightpath response, the pitch angle itself can be regarded as the controlled quantity instead of the elevator (or column) deflection. It then follows from equations (4) and (5) that any desired variation of lift can be generated by means of pitch control. This assumption is justified for transport aircraft designed to meet applicable flying-qualities criteria for manual control, because pitch bandwidth and control power are much higher than necessary for low-bandwidth automated control of flightpath and airspeed, especially when modest acceleration limits are imposed for passenger comfort. (High-bandwidth tasks such as automatic landing are not considered in this report.)

This assumption that pitch control can generate any desired lift variation eliminates the pitching moment equation, reducing the aircraft model to two degrees of freedom. Because rotational dynamics are eliminated from the model together with equation (3), the two-degree-of-freedom model treats the aircraft as a point mass concentrated at its center of gravity. Appendix B presents a detailed justification for choice of this two-degree-of-freedom model based on response calculations that demonstrate the excellence of the approximation for representative transport aircraft.

To summarize, the two-degree-of-freedom model consists of the longitudinal force equation (1), the normal force equation (2), and the functional relations (4) that describe the lift force L and the drag force D. As a first step toward understanding the dynamical behavior of the aircraft, the steady-state solution of these equations is studied in the next section.

## Flight Envelope

For steady (trimmed) flight conditions, the acceleration terms on the left sides of equations (1) and (2) must vanish, so that the equations become (as illustrated by figure 7(c))

$$\sin \gamma = \frac{T - D}{W} \tag{1a}$$

and

$$\cos \gamma = \frac{L}{W} \tag{2a}$$

Denoting the trimmed (potential) flightpath angle by $\gamma_{POT}$, equations (1a) and (2a) become

$$\sin \gamma_{POT} = \frac{T - D}{W} \tag{1b}$$

and

$$\cos \gamma_{POT} = \frac{L}{W} \tag{2b}$$

For flight in still air, $\gamma_{POT}$ is also an aerodynamic flightpath angle that is referred to the air mass, as are of course the aerodynamic forces L and D. The parameter $\gamma_{POT}$ can also be regarded as a control parameter that plays the same role as thrust. (The parameter $\gamma_{POT}$ is the same as that denoted by $\gamma_{TRIM}$ in appendix B.)

### Lift and Drag Characteristics

Appendix B shows how the lift and drag forces are reduced to nondimensional coefficients whose characteristic variations can be studied in canonical form. Briefly summarizing the results, the *lift coefficient* $C_L$ and the *drag coefficient* $C_D$ are defined by the equations

$$L = \left(\frac{\rho V^2}{2}\right) S\ C_L(\alpha, M) \qquad D = \left(\frac{\rho V^2}{2}\right) S\ C_D(\alpha, M) \tag{B-6d}$$

In the equations (B-6d), the *Mach number* M is defined by M = V/a where a denotes the sonic velocity, and the quantity $\rho V^2/2$ is termed the *dynamic pressure*, where ρ denotes atmospheric density.

At low Mach numbers below M = 0.3, the effects of Mach number are unimportant, so that M can be dropped from the functional relations (B-6d). Appendix B shows that the lift coefficient $C_L$ increases linearly with angle of attack, reaching its maximum value at the *stalling angle.* At higher Mach numbers, the lift-curve slope $dC_L/d\alpha$ increases slightly, and the maximum value of the lift coefficient decreases as shock waves form on the aircraft wings.

Appendix B shows that the drag coefficient $C_D$ varies parabolically with angle of attack, and therefore with lift coefficient (that is, the *drag polar* curve is parabolic). At Mach numbers above M = 0.6 the drag coefficient increases with Mach number, slowly at first and then more abruptly as shock waves form on the aircraft.

In steady flight, the lift L must equilibrate the normal component of aircraft weight W according to equation (2a). Appendix B shows that the variation of the drag D with airspeed V determined by equations (2a) and (B-6d) is quadratic, taking on its minimum value at an intermediate airspeed termed the *speed for minimum drag*, and increasing at both lower and higher airspeeds.

### Engine Thrust Characteristics

Appendix B shows that, at altitudes below 15,000 ft, maximum available engine thrust decreases slowly with increasing airspeed and Mach number. As altitude increases, the slope dT/dM of this thrust variation decreases; at cruising altitudes, the available thrust becomes nearly independent of Mach number.

### Performance Envelope

The aircraft performance envelope for steady flight (that is, for flight conditions in which both the longitudinal acceleration dV/dt and the normal acceleration V ($d\gamma/dt$) vanish) is determined by the variation of the trimmed flightpath angle $\gamma_{POT}$ with airspeed or Mach number, which is found by substituting into equations (1b) and (2b) the lift, drag, and thrust functions just discussed. Appendix B shows that the resulting quadratic variation of $\gamma_{POT}$ with airspeed is given by the equation

$$\sin\gamma_{POT} - \frac{1}{\pi\,AR\,e}\,\frac{W/S}{\rho_0 V_E^{\,2}/2}\sin^2\gamma_{POT} = \frac{T}{W} - \frac{C_{DP}}{W/S}\rho_0 V_E^{\,2}/2 - \frac{1}{\pi\,AR\,e}\,\frac{W/S}{(\rho_0 V_E^{\,2}/2)} \qquad \text{(B-10d)}$$

where the notation is given in the "List of Symbols."

**Sample calculation**– Equation (B-10d) gives the variation of the equilibrium (steady-state) flightpath angle $\gamma_{POT}$ with equivalent airspeed $V_E$, with the thrust T (or, alternatively, the engine rpm N) as parameter. It can be seen that, at a known height H in the Standard Atmosphere (appendix A) and at a selected equivalent airspeed $V_E$ and with known aircraft parameters T/W, W/S, $C_{DP}$, AR, and e, equation (B-10d) is quadratic in sin $\gamma_{POT}$ and can be solved for $\gamma_{POT}$ in closed form. The following parameter values are representative for a transport aircraft climbing with maximum thrust at sea level at the minimum-drag airspeed (table 1, appendix B):

| | | | |
|---|---|---|---|
| W/S = 150 lb/ft$^2$ | $C_{DP}$ = 0.0150 | AR = 7.19 | e = 0.83 |
| V = 487.9 ft/sec | $V_E$ = 289.1 kt | M = 0.437 | T/W = 0.1765 |

With these parameter values, equation (B-l0d) becomes

$$\sin \gamma_{POT} - 0.028\,286 \sin^2 \gamma_{POT} = 0.1200 \qquad \text{(B-10e)}$$

The numerical solution for sin $\gamma_{POT}$ is 17.6766 ±17.5562, or, because the property sin (•) ≤ 1 rules out the positive sign,

$$\sin \gamma_{POT} = 0.1204 \qquad \gamma_{POT} = 6.92 \text{ deg}$$

Because $\gamma_{POT}$ is small, the exact result differs trivially from the approximation obtained by neglecting the small $\sin^2 \gamma_{POT}$ term. At this climb angle and airspeed, the lift coefficient, the drag coefficient, the lift/drag ratio, and the rate of climb are found for reference to be

$$C_L = 0.53 \qquad C_D = 0.0298 \qquad L/D = 17.7 \qquad dH/dt = 3524 \text{ ft/min}$$

**Thrust contours**– Repeating the calculation over the airspeed range of interest and plotting $\gamma_{POT}$ against $V_E$ or M (eq. (B-7m), appendix B) defines the contour for maximum thrust, which forms the upper boundary of the flight envelope. The contour for minimum (idle) thrust forms the lower boundary. Contours for intermediate values of thrust can be added as desired to complete the diagram for the aircraft performance envelope (Innis, Holzhauser, and Quigley, 1970).

**Pitch contours**– A second set of contours corresponding to constant pitch angle can be constructed in the following way. Solve equation (2b) and equation (B-6d) for the lift coefficient

$$C_L = C_W \cos \gamma_{POT} = [(W/S)/(\rho V^2/2)] \cos \gamma_{POT} \qquad \text{(B-10f)}$$

which determines the lift coefficient for any point ($V_E$, $\gamma_{POT}$) within the flight envelope, because $\gamma_{POT}$ is given by the ordinate and the dynamic pressure ($\rho V^2/2$) is determined by the abscissa $V_E$ (or M). Now, the lift coefficient is related to angle of attack by a linear equation, which can be solved to find the angle of attack corresponding to the selected point. With both the flightpath angle $\gamma_{POT}$ and the angle of attack $\alpha$ known, the pitch attitude $\theta$ can be found from the following equation:

$$\theta = \alpha + \gamma \qquad (5)$$

which is valid for wings-level flight as already noted.

**Flight envelope chart**– A representative flight envelope is illustrated by figure 8. The left-hand boundary corresponds to the stalling speed. The right-hand boundary is determined by structural limits at low altitude, and by compressibility (Mach) effects at high altitude. The upper and lower boundaries are determined by the maximum and minimum thrust limits, respectively, as noted previously. Because at any fixed throttle setting (constant engine rpm), thrust decreases with speed (eqs. (B-8f) and (B-8g), appendix B), the thrust contours take on their maximum values of $\gamma_{POT}$ at airspeeds slightly below the speed for minimum drag.

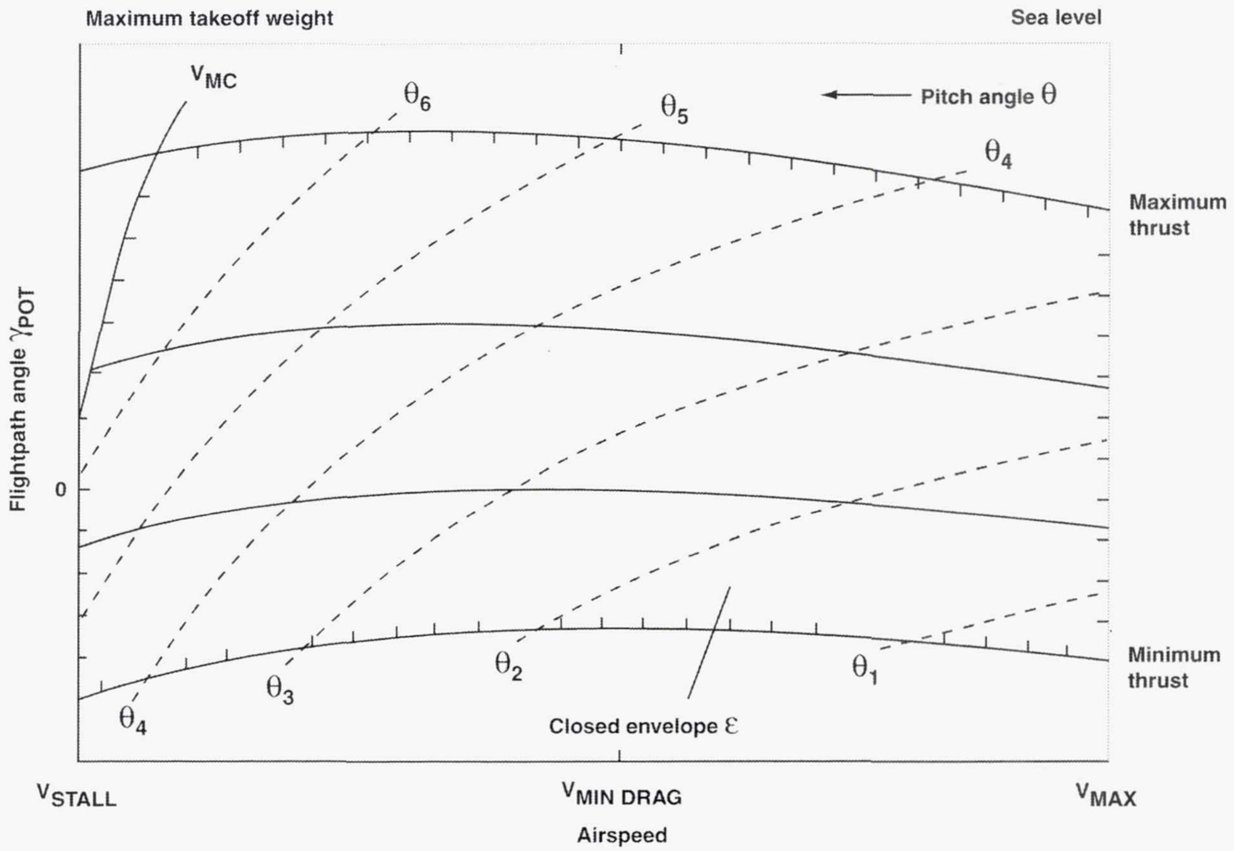


Figure 8. Aircraft performance envelope.

The flight envelope illustrated by figure 8 applies to an aircraft at maximum takeoff weight (wing loading W/S of 150 lb/ft$^2$) operating at sea level, as indicated on the diagram. Other parametric values of wing loading and altitude would result in flight envelope charts similar in form but differing in details. Various schemes are available for nondimensionalizing the parameters to enable a single chart to serve for all flight conditions (Taylor, 1974). Appendix B shows that the equivalent airspeed at which wing stall occurs depends only on aircraft weight (specifically, on wing loading), and is independent of altitude. Therefore, at low speed the equivalent airspeed, which is approximately the speed that is indicated by the standard cockpit airspeed indicator, is an especially convenient choice for the abscissa of the flight envelope chart. At high speed, the Mach number may be a more convenient choice for the abscissa.

**Onboard calculation**– However, it is not necessary to calculate the entire flight envelope on board the aircraft. The following discussion shows that knowledge of the upper flight envelope limit $\gamma_{POT\ MAX}$ and the lower limit $\gamma_{POT\ MIN}$ at the instantaneously prevailing (measured) airspeed is sufficient to determine mode selection for the flight control system to be developed. For that purpose, only the aircraft weight W, the drag D (estimated from the aircraft polar drag curve), and the maximum and minimum thrust limits are necessary (eq. (lb)). These parameters must be updated continuously by the onboard system. Correction of the flight envelope for the effects of nonstandard ambient conditions, of turning flight, of wind, draft, and wind shear, and of engine failure are discussed next.

### Nonstandard Atmospheric Conditions

The effects of nonstandard atmospheric conditions can be accounted for by using onboard measurements of ambient temperature and pressure in performance envelope calculations instead of Standard Atmosphere values.

### Symmetric (Coordinated) Turning Flight

In symmetric (coordinated) turning flight with near-zero sideslip, the increased lift required can be approximated by replacing the aircraft weight W with W/cos $\phi$ in the lift equation, where $\phi$ is the angle of bank (appendix C). Therefore, the effect of turning flight is the displacement of the constant-thrust contours downward and to the right, the same as for increased weight. To verify this effect, note that reducing the T/W term in equation (B-10d) reduces $\gamma_{POT}$ displacing the contours downward, while increasing V so as to hold the other terms constant at increased weight displaces the contours to the right.

The increased angle of attack required to generate the increased lift can be approximated by dividing the right-hand side of equation (5) by cos $\phi$. Both these approximations are valid over the range of bank angles encountered during controlled flight of transport aircraft (appendix C).

### Steady Horizontal Wind

Appendix B shows that the equations of motion remain unchanged during flight in steady horizontal wind, provided that flightpath angle and speed are measured with respect to the airmass. Therefore, the flight envelope also remains unchanged.

### Wind Shear and Vertical Draft

If the airmass accelerates instead of moving uniformly, or if the wind field varies with altitude so that the aircraft encounters changing winds during climb or descent, appendix B shows that an additional term involving the wind shear (time rate of change of tailwind velocity) appears in the longitudinal (streamwise) force equation. In an increasing tailwind, this wind-shear term acts to reduce airspeed in the same way as a drag increase. Because lift is proportional to the square of airspeed, the loss of lift owing to reduced airspeed results in downward acceleration, causing the aircraft to settle below its still-air path. Aircraft performance can be severely degraded. Steady downdrafts have a similar performance-degrading effect on the aircraft even though the airmass moves uniformly, because level flight relative to the Earth requires climbing flight relative to the airmass.

It follows that tailwind shear and downdraft displace the constant-thrust contours of the flight envelope downward toward smaller $\gamma$, and, conversely, headwind shear and updraft displace them upward. Contour displacement can be estimated from onboard measurements (Funabiki, Bando, Tanaka, Hynes, and Hardy, 1993). (Shear disturbances exceeding the FAA-specified threshold of 2 kt/sec must be detected, and must cause the control system to transition to a special wind-shear recovery mode, which is not treated in this report).

### Engine Failure

Engine failure causes a loss of thrust that displaces the maximum-thrust contour downward toward smaller $\gamma$, leaving the zero-thrust contour (lower envelope boundary) unchanged. Thrust asymmetry combines with limited rudder control power to define a sloping minimum-control boundary ($V_{MC}$

contour in figure 8) such that, at airspeeds below $V_{MC}$ aircraft controllability cannot be retained if maximum thrust is applied to the operating engines. Since an engine could fail at any time, flight below the minimum-control airspeed must be avoided in airline service. Actual thrust remaining after engine failure can be estimated on board from measurements of engine rpm and ambient conditions (appendix B).

## Aircraft Dynamical Behavior

### Pitch and Thrust Control

Specializing equations (1) and (2) for control of flightpath angle $\gamma$ and airspeed V by means of pitch angle and thrust and placing the control quantities on the right, equations (1) and (2) become

$$\frac{1}{g}\frac{dV}{dt} + \sin\gamma = \sin\gamma_{POT} \tag{1c}$$

and

$$\frac{V}{g}\frac{d\gamma}{dt} + \cos\gamma = \frac{L}{W} \tag{2c}$$

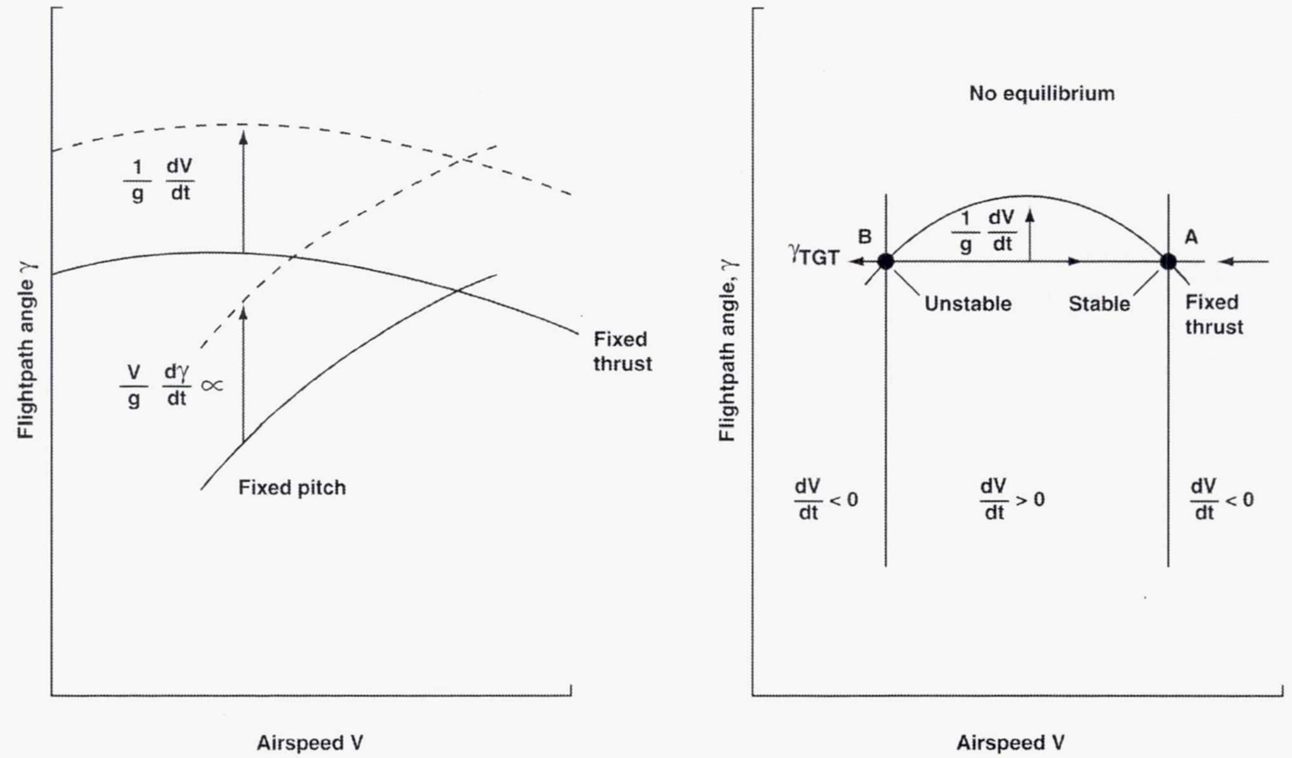
where $\gamma_{POT}$ is defined by equation (1b). The complete solution of equations (1c) and (2c) determines the dynamic response of the aircraft to pitch and thrust control, which is discussed next.

### Response Trajectories

The dynamic response of the aircraft to pitch and thrust control is closely related to the geometrical form of the flight envelope diagram. The discussion begins with an overview of design issues before taking up the dynamical behavior of the aircraft in detail.

**Overview**– The dynamic response of the aircraft to pitch and thrust control depends on the initial operating point within its performance envelope. Operating points are selected to realize the flight profile summarized by the flight plan (appendix B). Each operating point corresponds to a steady-state solution of the equations of motion for specified pitch and throttle control inputs, and the transitions from one operating point to the next correspond to transient solutions of the equations of motion. Flying-qualities criteria based on pilot evaluation define what is meant by well-shaped responses, which enable smooth, rapid capture of the desired (target) operating point during manual control. Automated systems should be designed to achieve response shapes similar to those desired for manual control, because from the flying-qualities viewpoint it can be taken as axiomatic that behavior that is difficult to anticipate is difficult to monitor.

**Steady-state solution**– During manual flight, the pilot controls pitch with elevator and engine rpm with throttle. It can be seen from the flight envelope chart of figure 8 that the steady-state solution of the force equations (lc) and (2c) for any combination of control inputs is determined by the intersection of the pitch and rpm (thrust) contours corresponding to those control inputs. The desired steady-state solution is termed the target operating point. By definition, the longitudinal and normal accelerations must vanish at each steady-state solution point. It will be shown next that the accelerations prevailing at any other point (that is, any point not corresponding to a steady-state solution) can be determined directly from the flight envelope diagram.



a) Longitudinal and normal acceleration variation with pitch and thrust control.

b) Longitudinal acceleration variation with airspeed at constant flightpath angle.

Figure 9. Speed stability.

**Determination of normal acceleration**– Equation (2c) shows that, at any airspeed V, the (normalized) normal acceleration (V/g) dγ/dt is equal to the difference between the normalized lift L/W and cos γ. In figure 9(a), L/W is related to the pitch angle contour corresponding to the instantaneous pitch angle in the following way. At any fixed speed V, the lift force L is proportional to angle of attack, and according to equation (5) the angle of attack is equal to the difference between pitch angle and flightpath angle. Thus at any speed V and flightpath angle γ, the normal acceleration is proportional to the vertical distance from the instantaneous flightpath angle γ upward to the pitch angle contour corresponding to the instantaneous pitch angle θ. Therefore, if the pitch angle contour is displaced upward from an equilibrium angle by an abrupt increase of pitch, the resulting normal acceleration is proportional to the incremental displacement Δθ (fig. 9(a)).

Mathematically, the variation of the lift coefficient $C_L$ with angle of attack α given by equation (B-6d) can be linearized (appendix B). Then, by combining equations (2c), (5), and (B-6d), it can be shown that the variation of normal acceleration (V/g) (dγ/dt) with pitch angle θ (while the airspeed V and the flightpath angle γ remain fixed at their initial values $V_0$ and $\gamma_0$) is determined approximately by the partial derivative

$$(\partial/\partial\theta)\,(V/g)\,(d\gamma/dt)_{V,\gamma} = (1/C_W)\,(\partial C_L/\partial\theta).$$

**Determination of longitudinal acceleration**– Since equilibration of normal forces takes place rapidly following any disturbance (appendix B), it is reasonable to assume that normal forces remain in equilibrium during the change of airspeed that follows an abrupt change of thrust, especially since thrust inclination to the longitudinal axis has already been assumed to be negligible. It then follows from equation (4) that the variation of drag with airspeed remains the same during such a speed change as in steady flight.

Under these conditions, equation (1c) shows that, for small $\gamma$, the normalized longitudinal acceleration (l/g)(dV/dt) is approximately equal to the increment $\gamma - \gamma_{POT}$. At any point, the normalized longitudinal acceleration is therefore equal to the vertical distance from the instantaneous flightpath angle $\gamma$ upward to the $\gamma_{POT}$ contour corresponding to the instantaneous thrust (fig. 9(a)). It follows that, if the thrust contour is displaced upward from an equilibrium setting by an abrupt increase of thrust, the resulting normalized longitudinal acceleration is equal to the incremental displacement $\Delta\gamma_{POT}$. Extensive use will be made of this geometric relationship.

**Direction field**– The determination of the longitudinal and normal accelerations at a point (V, $\gamma$) is equivalent to construction of the direction field for the differential equations (1c) and (2c) at that point. Starting from any initial point, each solution of the differential equations for specified pitch and thrust control inputs can be mapped into a unique trajectory in the (V, $\gamma$) plane by stepwise numerical integration of the longitudinal and normal accelerations over the direction field. Thus, in principle, complete information about the dynamic response of the aircraft is contained geometrically in the flight envelope diagram.

**Flying qualities criteria**– The final point of each trajectory (the steady-state solution) is determined by the intersection of the pitch and thrust contours corresponding to the specified control inputs, as already noted. In order to capture a specified target point, these control inputs should cause the aircraft to traverse a suitable trajectory from the initial point to the target point. Desirable trajectories result in rapid but smooth capture maneuvers, subject to constraints on normal and longitudinal acceleration and on control power. During manual control, the form of each capture trajectory is generated by the human pilot. Flying qualities criteria applicable to transport aircraft specify aircraft design characteristics that assure satisfactory dynamic response. These criteria are based on more than 40 years of operating experience with both civil and military jet transport aircraft (Anonymous, 1985).

## Dynamic Response

With the assumption that pitch control can generate any desired lift variation, as previously discussed, equation (2c) shows that the normal acceleration (V/g)(d$\gamma$/dt) can be controlled as desired, whatever the value of $\gamma$. By adjusting the lift L to compensate for variations in airspeed V, the flightpath rate d$\gamma$/dt can be controlled as desired, whatever the prevailing values of V and $\gamma$. For example, by making d$\gamma$/dt proportional to the path error relative to the target $\gamma$, smooth, rapid control of $\gamma$ without overshoot can be obtained independent of V

Similarly, equation (1) shows that the longitudinal acceleration dV/dt can be controlled as desired by varying the thrust T. By making adjustments in T to compensate for variations in flightpath $\gamma$ and drag D, dV/dt can be controlled as desired, whatever the prevailing values of $\gamma$ and D. For example, by making dV/dt proportional to speed error relative to the target V, smooth, rapid control of V without overshoot can be obtained independent of $\gamma$.

**Pitch and thrust control**– By combining control of the flightpath angle $\gamma$ by varying the lift L (that is, by varying pitch angle $\theta$) and control of the airspeed V by varying thrust T, both the flightpath angle $\gamma$ and the airspeed V can be controlled independently by simultaneous pitch and thrust control inputs. Geometrically, the final solution point in the flight envelope diagram (fig. 8) is then determined by the intersection of the specified pitch and thrust contours, which can be made to coincide with any target (V, $\gamma$) point within the steady-state flight envelope.

**Pitch control of airspeed with thrust saturated**– If the thrust T is saturated at maximum or minimum thrust, or fixed at any intermediate value so that only pitch is available for control, then equation (2c) shows that the flightpath angle $\gamma$ can be controlled by pitch just as before. Equation (1) shows that, if the flightpath angle $\gamma$ is regarded as the control variable replacing the thrust T (which is now fixed), then dV/dt can be controlled as desired by making adjustments in $\gamma$ to compensate for variations in drag D. For example, if dV/dt is maintained constant over a range of airspeed, then the locus of operating points lies on a contour parallel to the fixed-thrust contour, but displaced from it vertically by a distance equal to the normalized longitudinal acceleration (l/g)(dV/dt), as illustrated by figure 9(a) (with the arrow reversed in sign). In order to control airspeed by varying pitch, the flightpath angle $\gamma$ must be controlled so as to obtain the desired value of dV/dt in equation (1). The flightpath angle $\gamma$ therefore cannot be controlled independently.

Geometrically, the final solution point in figure 8 is then located at the intersection of the fixed-thrust contour with the vertical line corresponding to the target airspeed. By selecting the fixed (target) thrust appropriately, any target point within the steady-state flight envelope can be captured. (Strictly speaking, it is throttle setting that is maintained at a fixed value during manually controlled flight; thrust itself then exhibits some variation with airspeed. For simplicity, thrust will be treated as the control variable.) However, this report is concerned chiefly with thrust saturation at either maximum or minimum thrust.

**Pitch control of flightpath with thrust saturated**– It remains to study the control of the flightpath angle $\gamma$ by varying pitch while thrust remains fixed. In contrast to the control of airspeed by pitch just discussed, it will be shown that, when flightpath is controlled by pitch while thrust remains fixed, it is *not* possible to capture any target point within the steady-state flight envelope, but only those points that lie at or above the speed for minimum drag. This restriction results from considerations of speed stability when the flightpath is constrained, which is analyzed in the next section.

## Speed Stability With Thrust Saturated

If the flightpath angle $\gamma$ is controlled by varying pitch with thrust fixed, equation (2c) shows that flightpath rate d$\gamma$/dt can be controlled as desired, as noted previously. However, it is clear from equation (1) that, with the variation of $\gamma$ specified, the longitudinal acceleration dV/dt cannot be controlled independently when thrust is fixed. For simplicity, it is assumed that the flightpath angle $\gamma$ is controlled so as to capture a fixed $\gamma$ target. Geometrically, the final solution point in figure 8 is then located at the intersection of the specified fixed-thrust contour with the horizontal line corresponding to the target flightpath.

But inspection of figure 8 shows that the intersection of these contours does not determine a unique solution. Because of its concave-downward form, any fixed-thrust contour has two possible points of intersection with a horizontal line corresponding to a target flightpath, one at an airspeed above the speed for minimum drag and the other below it. This ambiguity can be resolved by studying the speed stability at each of these two points of equilibrium (fig. 9(b)), which illustrates longitudinal stability while maintaining lift equilibrium (that is, $d\gamma/dt = 0$).

**Stable airspeed equilibrium**– Considering first the speed stability in the neighborhood of the equilibrium point to the right of the speed for minimum drag (point A in figure 9(b)), it can be seen that at lower airspeeds to the left of point A where the target $\gamma$ lies below the fixed-thrust contour for $\gamma_{POT}$, the longitudinal acceleration is positive (eq. (1c)), so that the airspeed tends toward point A. Similarly, at higher airspeeds to the right of point A, the acceleration is negative, and again the airspeed tends toward point A. It is clear that point A is a point of stable equilibrium.

**Unstable airspeed equilibrium**– By a similar argument, point B is found to be a point of unstable equilibrium. Physically, this instability at point B results from its location below the speed for minimum drag. There, the total drag force is dominated by the induced drag, which increases with decreasing speed (appendix B). Thus the tangent $d\gamma/dV$ is positive along the fixed-thrust contour near point B.

**Geometrical regions**– Geometrically, it is clear from figure 9(b) that point A is an attractor point whose region of attraction includes all points to the right of point B on the flightpath target line through points A and B. If the target flightpath angle is increased (or the fixed target thrust is reduced), points A and B approach each other until they coincide at the speed for minimum drag. If the target flightpath angle is increased still farther, it lies above the target thrust contour everywhere. No steady-state solution then exists, and the longitudinal acceleration $dV/dt$ is negative for all points on the target flightpath line.

**Aircraft operating characteristics**– The consequences of these facts for aircraft operation are as follows. If the flightpath target lies below the fixed-thrust contour anywhere, so that some points of equilibrium exist, then, starting from any initial point on the flightpath target line to the right of point B, the aircraft will capture and hold point A no matter what airspeed target has been selected. On the other hand, starting from any initial point to the left of point B, the airspeed will diverge toward stalling speed. By selecting the fixed (target) thrust appropriately, any target point within the steady-state flight envelope can be captured, provided that it lies to the right of the speed for minimum drag. But the aircraft cannot capture a target point that lies within the steady-state envelope to the left of the speed for minimum drag (that is, a point of unstable equilibrium). Dynamic response calculations that illustrate these behaviors are presented and discussed in detail in appendix B.

If the aircraft attempts to capture a point that lies above the maximum-thrust contour everywhere, no steady-state solution exists because the target path exceeds the performance capability of the aircraft. The airspeed will then diverge toward stalling speed, whatever the initial speed. For example, this situation could arise following engine failure during high-altitude cruise, if the remaining thrust were insufficient for level flight. (The aircraft must then "drift down" to some lower altitude at which sufficient thrust becomes available.)

The instability characterizing control to a constrained path below the speed for minimum drag with fixed thrust has been known for a long time (Neumark, 1953). Some early barometric-hold autopilots could stall the aircraft if given sufficient pitch authority. Clearly, this potentially catastrophic behavior presents a challenge to designers of full-authority fly-by-wire systems to provide some means of ensuring safety.

On the other hand, operation at a point of stable equilibrium presents no inherent safety hazard even though the target airspeed is not captured. Overspeed must be avoided, of course, but for some purposes such as level-flight cruise, operation with fixed thrust (that is, with autothrottle off) may be preferred because all unwanted throttle activity owing to airspeed disturbances is suppressed. Some wandering of airspeed must then be accepted, but this can be reduced by permitting some altitude deviation ("soft" altitude control).

### Dynamical Behavior Theorem

The dynamical behavior of the aircraft resulting from control of flightpath angle and airspeed by means of pitch and thrust can be summarized by the following theorem, which results directly from the aircraft equations of motion:

> **Theorem**– If both pitch and thrust are available for control, then (i) flightpath angle and airspeed can be controlled independently. If the thrust is fixed so that only pitch is available for control, then either (ii) airspeed, or else (iii) flightpath angle can be controlled, but not both.

The aircraft is therefore capable of only three kinds of dynamical behavior:

> if thrust is available for control, then
> (i) path and speed can both be controlled;
> if thrust is fixed, then
> (ii) speed can be controlled at the expense of path,
> or else
> (iii) path can be controlled at the expense of speed.

It is clear from the statement of the theorem that Type (i) control is restricted to the interior of the steady-state performance envelope of the aircraft, which is the only region of the state space where thrust is available for control. The previous discussion showed that Type (ii) control is unrestricted, but Type (iii) control results in airspeed instability at airspeeds below the speed for minimum drag. Therefore, the theorem can be regarded from a mathematical control-theoretic viewpoint as a qualitative statement of controllability resulting from the form of the differential equations (l) and (2): if both pitch and thrust are available for control, then both airspeed and flightpath angle can be controlled throughout the region of interest; if thrust is fixed, then airspeed is controllable throughout the region of interest, but flightpath angle is controllable only within a subset of the region of interest.

These three kinds of behavior are familiar to pilots. Automated control modes that correspond to these three fundamental types of control are developed later in this report. Geometrically, the three types of control partition the flightpath-airspeed plane into several discrete regions, which are termed stability regions. Within each of these regions, the dynamical behavior of the aircraft is qualitatively similar. These geometric stability regions are examined next.

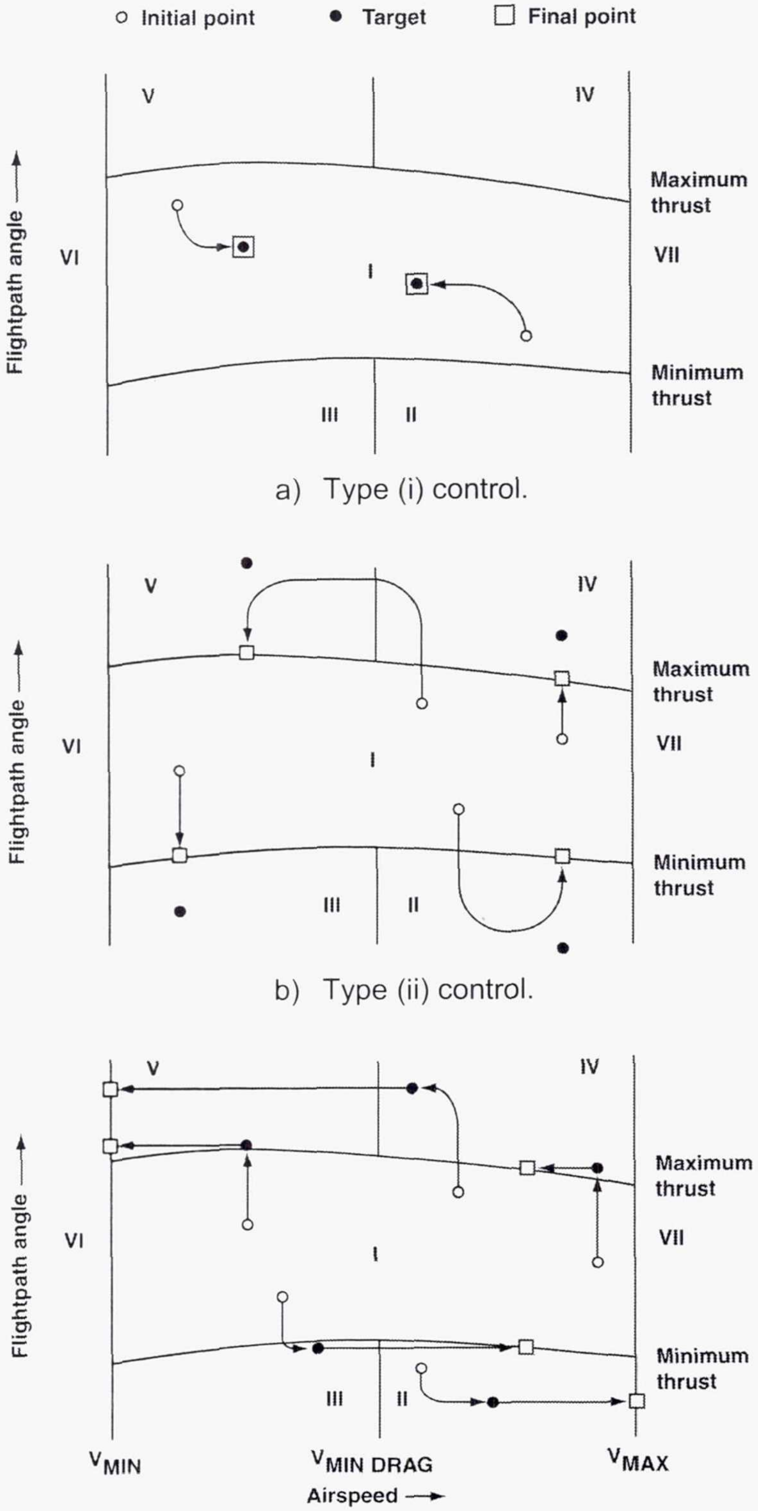
## Geometrical Stability Regions

Capture trajectories corresponding to the three types of control just discussed are illustrated by the three diagrams of figure 10.

**Type (i) control**– In figure 10(a), which illustrates Type (i) control, region I corresponds to the steady-state flight envelope. Two representative capture trajectories are illustrated. Because the flightpath response to pitch is much more rapid than the airspeed response to thrust (appendix B), as a rough approximation the initial trajectory is nearly vertical, so that capture of the path target is nearly complete before significant change in airspeed has taken place. The final portion is nearly horizontal, as capture of the speed target is completed with little change of flightpath. No Type (i) control is possible outside region I, because thrust is saturated there.

**Type (ii) control**– Type (ii) control is illustrated by figure 10(b). Representative capture trajectories when thrust is saturated at maximum thrust are illustrated in regions IV and V, and representative capture trajectories when thrust is saturated at minimum thrust are illustrated in regions II and III. Intermediate thrust settings are not illustrated. It will be seen that each final solution point is located at the intersection of the specified fixed-thrust contour with the vertical line corresponding to the target airspeed, as already noted. All these points are points of stable flightpath equilibrium, because the lift-curve slope is positive for all angles of attack below the stalling angle (appendix B). Target flightpaths lying in regions II, III, IV, or V are not captured because they lie outside the performance capability of the aircraft. Portions of the capture trajectories for which the longitudinal acceleration is constant are parallel to the specified fixed-thrust contours, as noted previously.

**Type (iii) control**– Type (iii) control is illustrated by figure 10(c), in which the maximum and minimum thrust contours are the same as in figure 10(b). It will be seen that each final solution point is located at the intersection of the specified fixed-thrust contour with the horizontal line corresponding to the target flightpath angle. The solution points lying along the lower boundary of region V are points of unstable equilibrium, from which negative airspeed excursions would cause divergence toward stalling speed. (Positive airspeed excursions into region I are of no particular concern, because there the thrust is not saturated and Type (i) control becomes possible.) Attempts to capture target flightpaths lying above the maximum-thrust contour everywhere in regions IV and V also lead to airspeed divergence toward stalling speed. No equilibrium solutions exist along these target flightpaths because they lie outside the performance capability of the aircraft. Thus for behavioral classification, the points of unstable equilibrium and of no equilibrium in region V can be grouped together, because they lead to the same airspeed divergence. Furthermore, target points in region IV that lie outside the performance capability of the aircraft need not be considered separately, because the airspeed divergence resulting from attempts to capture them leads inevitably into region V.



a) Type (i) control.

b) Type (ii) control.

c) Type (iii) control.

Figure 10. Aircraft dynamical behavior.

The final solution points lying above the speed for minimum drag (regions II and IV) are points of stable airspeed equilibrium. (Points of stable airspeed equilibrium with maximum thrust (region IV) are often used for high-altitude cruising flight with auto throttle disengaged.) It will be seen that attempts to capture steeply descending target flightpaths in the interior of region II can cause airspeed divergence toward overspeed. Points on the upper boundary of region III are points of unstable equilibrium. (Airspeed excursions from these points into region I are of no particular concern, because there the thrust is not saturated and Type (i) control becomes possible.) Positive airspeed excursions lead into region II where overspeed is possible, but even if a point of stable equilibrium is reached that avoids overspeed, the increase of airspeed can be excessive. In no case of Type (iii) control is the target airspeed captured.

**Summary**– To summarize, the stability regions I, II, III, IV, and V are discrete regions within each of which the dynamical behavior of the aircraft is qualitatively similar in response to Type (i), Type (ii), or Type (iii) control. For logical completeness, regions VI (underspeed) and VII (overspeed) are also defined; region I is defined as open with respect to thrust, and the others closed. It can then be verified by inspection of figure 10 that no point exists in the (V, $\gamma$) plane that does not lie in one of the seven identified regions.

# PART II
# DESIGN SYNTHESIS

## FLIGHT CONTROL SYSTEM

### Concept

The flight control system is a major element of the Vehicle Management System, as already mentioned. The following brief functional description is taken from a previous publication (Sherry, Youssefi, and Hynes, 1995) to summarize the concept.

### Manual Control

Augmented manual flight control modes have been developed that accept commands for horizontal track angle, vertical flightpath angle, and airspeed. During manual flight, the pilot commands track-angle acceleration with the wheel and flightpath angle rate with the column. The pedals command sideslip. With the controls centered, commanded track angle rate and flightpath angle are held fixed against external disturbances. For these modes, the high-reliability autothrottle remains engaged at all times during flight, accepting airspeed commands entered manually into the mode control panel. By minimizing aircraft response to unwanted external disturbances, this advanced control augmentation improves precision of control and reduces pilot workload (Franklin, Hynes et al., 1986). In particular, the high-reliability autothrottle alleviates difficulties with manual control of airspeed owing to the strongly "backsided" variation of drag with speed that characterizes certain transport aircraft during landing approach.

### Automated Control

During fully automated flight, a velocity-vector command generated by the guidance function is injected directly into the same input port of the flight control system as that used during manual flight. This interface between the guidance and flight control functions can be regarded as the specification of the three components of the aircraft velocity vector in a mixed Earth/airmass spherical coordinate frame. (In the present report, only the two longitudinal components, that is, flightpath angle and airspeed, are treated.) This structure satisfies two human-factors criteria. First, specification of the velocity vector is more meaningful for the pilot than the roll, pitch, and autothrottle loop closures often specified in current aircraft, because the velocity vector relates directly to the pilot's task. Servo loop closures are confined to the interior of the flight-control function, together with other vehicle-specific details. Second, use of the same control algorithms during fully automated flight as those used for manual flight provides the crew an intuitive understanding of "what the automated system is trying to do" (Billings, 1996). Thus specifying the velocity vector as the input command to the flight control system is a simple, natural choice that is desirable from several points of view.

This system is similar in concept to the Total Energy Control System developed by Lambregts (Lambregts, 1983), and successfully evaluated in flight experiments conducted on the NASA-Langley B-737 research aircraft in 1984. The present NASA-Ames implementation makes use of the nonlinear inverse control concept originated by Meyer (Meyer and Cicolani, 1981), and applied by Franklin to the NASA-Ames quiet short-haul research aircraft (QSRA) (Franklin, Hynes et al., 1986). This nonlinear inverse control system contains an internal model of the aircraft, which enables response characteristics such as column force per unit of normal acceleration (that is, per g) to be held constant over the entire flight envelope without explicit gain scheduling. A separate control system incorporating conventional pitch damping is used for ground operation.

The discussion of the flight control system now proceeds by describing the system structure. First, a definition is presented that formalizes the notion of mode so as to apply in general to all real-time process-control systems. A brief discussion of system structure then shows how the control modes to be described fit into their supporting hierarchical framework within the Vehicle Management System. These structural matters are the subjects of the next two sections, after which the design of the flight control system is examined in detail.

## Mode Definition

The design process must begin with careful consideration of exactly how the aircraft is to be operated under both normal and abnormal conditions. These considerations define the modes from the engineering design viewpoint. Mathematically, a mode is a set of actions that the machine can take; that is, a physical behavior. From the control-theoretic viewpoint, a mode is specified by its topological structure (for example, its block diagram) together with specific details of gains, limiters, and the like that determine its characteristic behavior. In the human-factors view of aircraft systems, the human pilot selects a mode in order to obtain its characteristic set of actions, so that mode selection can be regarded as an expression of the pilot's intentions. In the implementation software, a mode corresponds to a path through the code along which the commands for its characteristic actions are generated.

## Mode Hierarchy

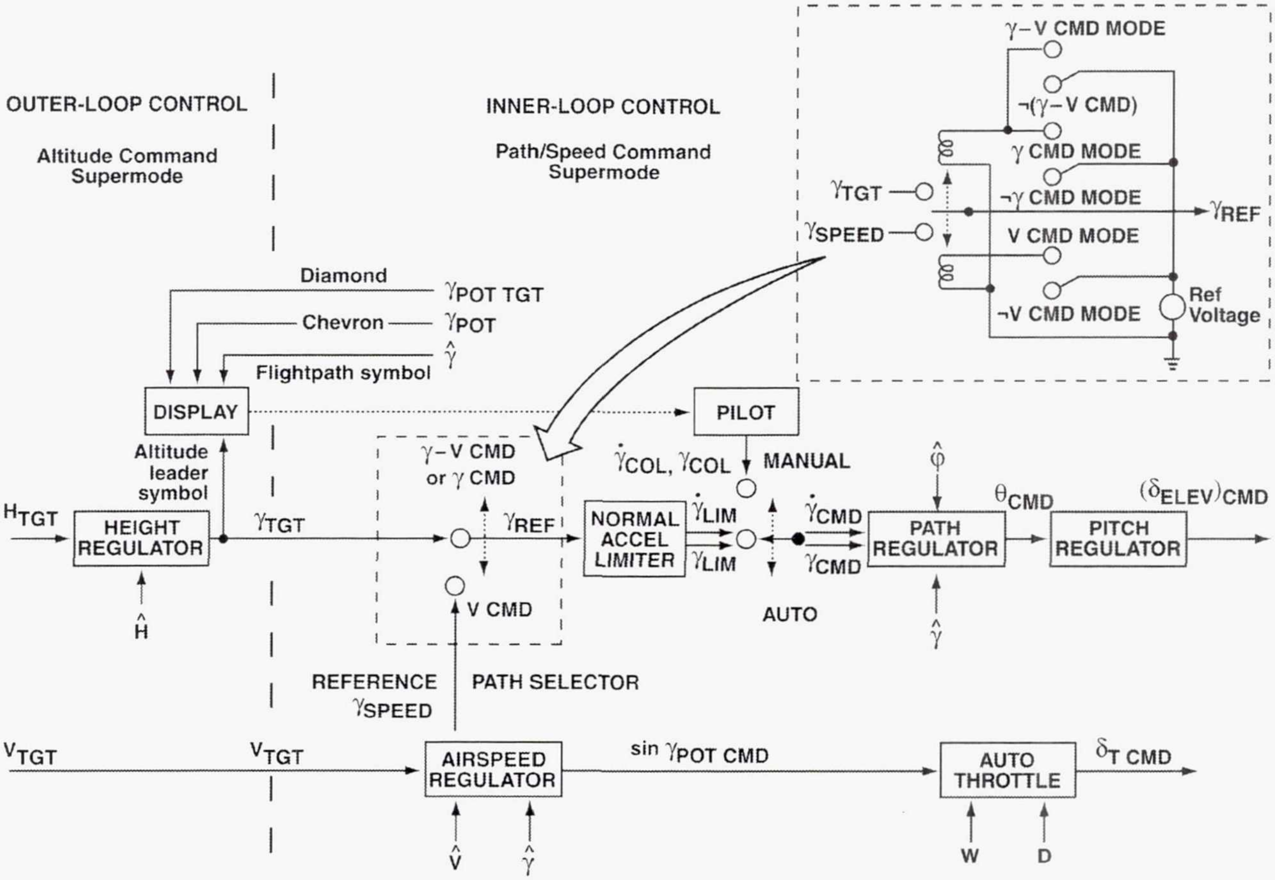
In the system to be described, the flight control modes are organized into a hierarchical structure with three levels, as illustrated by figure 1. In bottom-up order, the modes at the lowest level are termed primitive modes. They constitute elements of the higher-level supermodes, but do not themselves contain any subelements; that is, they are not composite. The three primitive longitudinal control modes to be described correspond to the three types of dynamical behavior of which the aircraft is capable, as already discussed in detail.

The second level of the mode hierarchy relates to the control tasks required for operation in the ATC system, such as capturing and holding an assigned altitude or tracking an approach glideslope. During manual operation, the crew defines the desired flightpath by setting the desired altitude and airspeed targets into the aircraft mode control panel. For example, in current aircraft the Flight Level Change mode occupies the second level of the mode hierarchy. On this second level, operational efficiency is determined entirely by the crew's manual selections.

The third and highest level of the mode hierarchy enables trajectory optimization for conservation of time or fuel by trajectory-synthesis algorithms (Erzberger, 1982) resident within the onboard flight-planning function. Readers can consult a previous publication (Sherry, Youssefi, and Hynes, 1995) for a more detailed description of the Vehicle Management System.

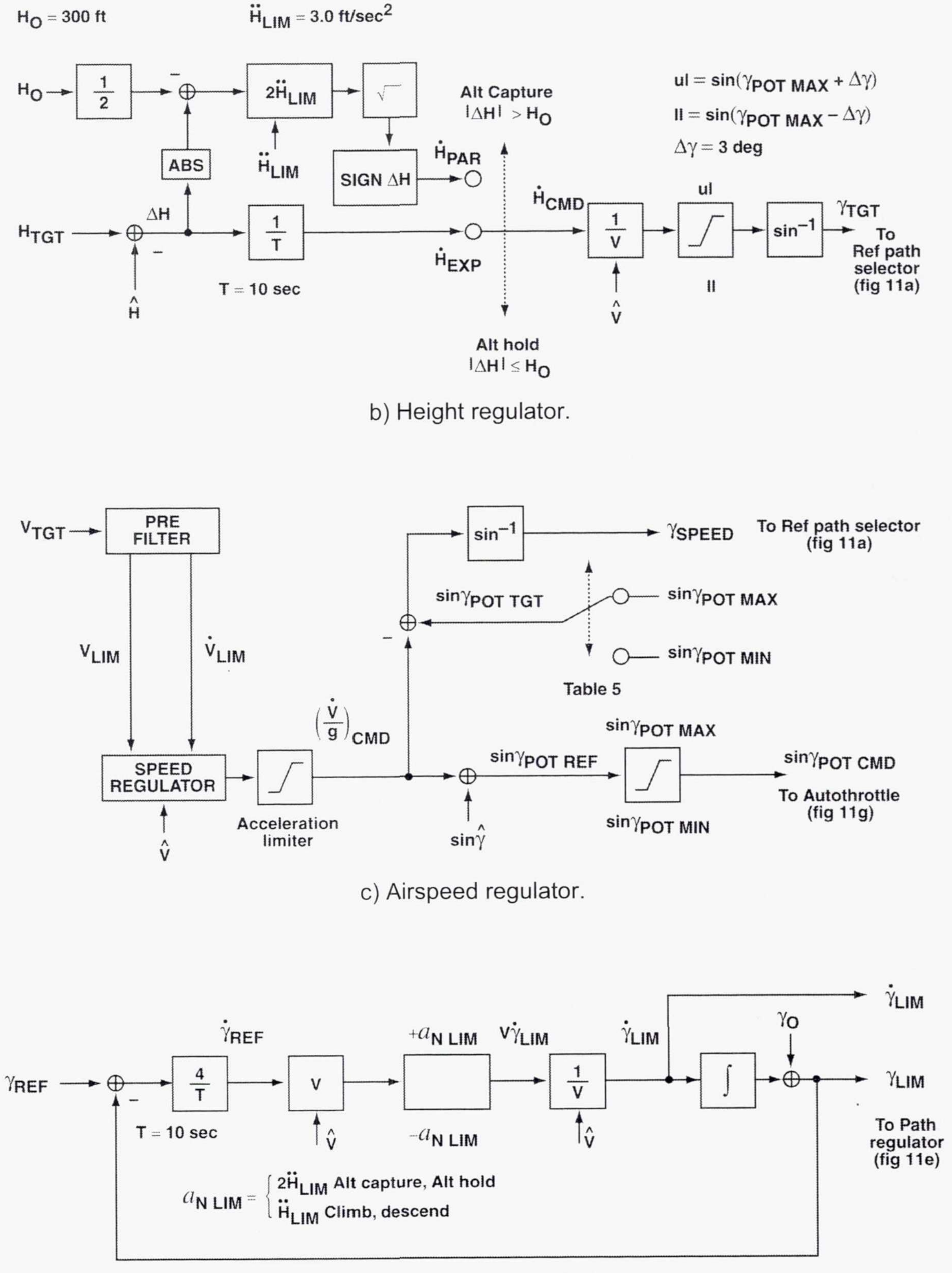
This report discusses the first level of the mode hierarchy in detail, and shows how the three primitive modes can be combined to form the lowest-level supermode, which is termed Path/Speed Command. The present work focuses on the design method and does not treat the complete mode hierarchy, although one second-level supermode termed Altitude Command is developed in detail in order to illustrate extension of the design method to higher level. Guidelines for extension to the third level of the mode hierarchy are presented briefly. The three primitive modes, termed $\gamma$-V Command, V Command, and $\gamma$ Command, are described in detail in the following sections. As noted previously, they correspond to the three fundamental kinds of dynamical behavior, namely Type (i) control, Type (ii) control, and Type (iii) control, respectively.

The complete longitudinal flight control system is illustrated by the block diagrams of figure 11. The altitude regulator function on the left of the broken vertical line in figure 11(a) is discussed in a later section, and should be ignored for the present. The three primitive modes to be described next are illustrated by the structure to the right of the broken vertical line. It can be seen that the external inputs are the target flightpath angle $\gamma_{TGT}$ and the target airspeed $V_{TGT}$.



a) Regulator structure.

Figure 11. Flight control system.

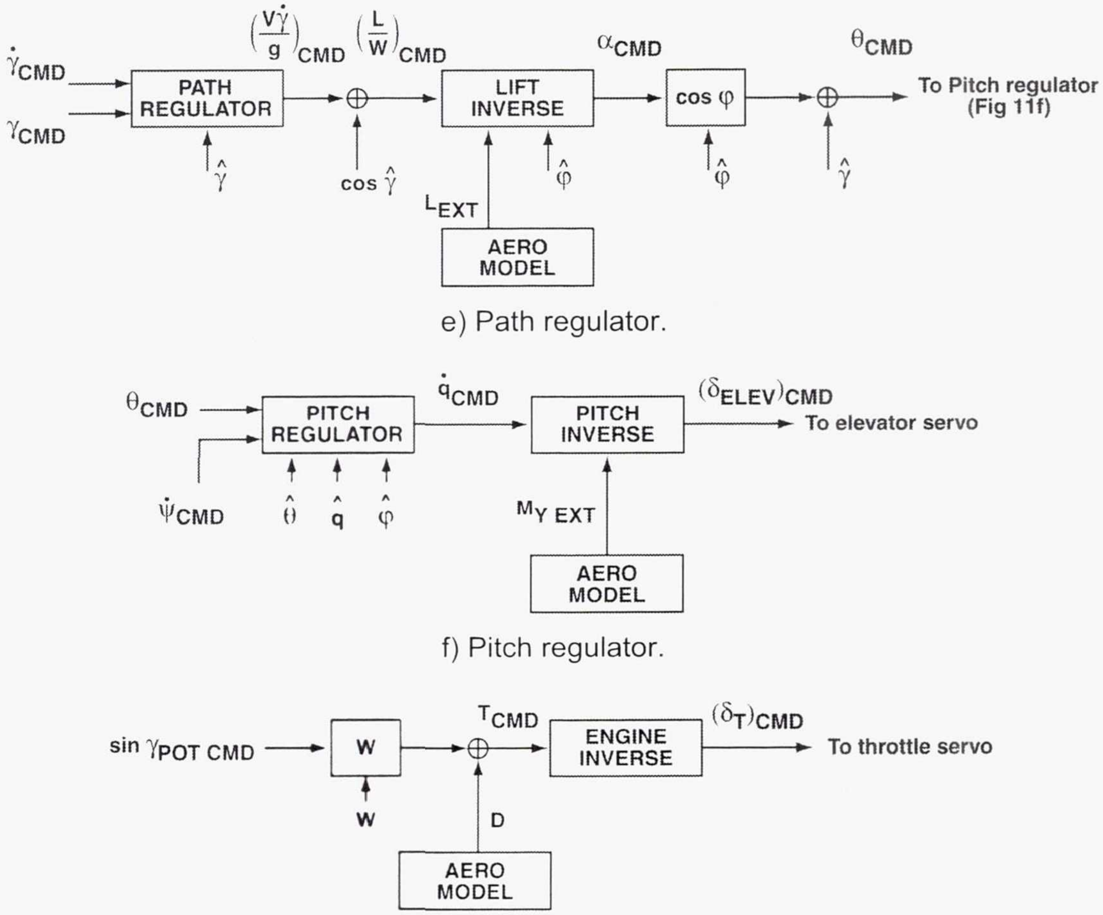


b) Height regulator.

c) Airspeed regulator.

d) Normal acceleration limiter.

Figure 11. Flight control systems (continued).



e) Path regulator.

f) Pitch regulator.

g) Autothrottle.

Figure 11. Flight control systems (concluded).

## γ-V Command Mode

### Control Law

The γ-V Command mode corresponds to Type (i) behavior, and requires that both pitch and thrust be available for control. The control law for the γ-V Command mode is obtained directly from the aircraft differential equations (1c) and (2c) by specializing them as follows.

In equation (1c), the desired (commanded) normalized longitudinal acceleration, denoted by $(1/g)(dV/dt)_{CMD}$, is generated by the airspeed regulator (fig. 11(c)), which compares the measured airspeed with the desired airspeed $V_{LIM}$ and generates a longitudinal acceleration command intended to null the airspeed error. If the airspeed regulator contains an integral term, the integrator must be clamped to prevent wind-up if the thrust response becomes rate-limited. For simplicity, it is assumed that the airspeed regulator contains only a proportional term. The rate-limited reference airspeed $V_{LIM}$ is obtained by passing the target airspeed $V_{TGT}$ through a pre-filter that limits its rate of change to a maximum of 2 kt/sec (about 0.1 g). Making use of the measured flightpath angle, equation (1c) is then solved for the desired potential flightpath angle, which is denoted by $\gamma_{POT\ CMD}$.

In equation (2c), the desired normal acceleration, denoted by $(V/g)(d\gamma/dt)_{CMD}$, is generated by the flightpath regulator (fig. 11(e)), which compares the measured flightpath angle with the desired flightpath angle $\gamma_{CMD}$ and generates a normal acceleration command intended to null the flightpath error. During automated flight, $\gamma_{CMD}$ is obtained by passing the target flightpath angle $\gamma_{TGT}$ through a rate limiter that limits normal acceleration to 0.1g for passenger comfort (figs. 11(a) and 11(d)). During manual flight, $\gamma_{CMD}$ is obtained from the time integral of column deflection, and is rate-limited by the human pilot (fig. 11(a)).

The control law for the γ-V Command mode, which is derived in the manner just described, is given by the equations

$$\frac{1}{g}\frac{dV}{dt}_{CMD} + \sin\gamma = \sin\gamma_{POT\ CMD} \qquad (6a)$$

and

$$\frac{V}{g}\frac{d\gamma}{dt}_{CMD} + \cos\gamma = \frac{L}{W}_{CMD} \qquad (6b)$$

Path errors are fed back to pitch control, and speed errors are fed back to thrust. This topological structure is the appropriate choice for transport aircraft with small thrust inclination. As already noted, the characteristics of the airspeed regulator and the flightpath regulator are selected to meet flying qualities criteria based on extensive operating experience with transport aircraft.

### Thrust Command

The thrust command is generated by specializing equation (lb) and solving for thrust, making use of the prevailing drag calculated from the onboard aircraft model:

$$T_{CMD} = W\sin\gamma_{POT\ CMD} + D \qquad (6c)$$

The parameter $\gamma_{POT\ CMD}$ is calculated from equation (6a). The commanded thrust $T_{CMD}$ must be physically realizable. The throttle servo command is then obtained by inverting the engine thrust characteristic, as illustrated by the autothrottle diagram of figure 11(g).

### Thrust Saturation

Thrust saturation must be absent in the γ-V Command mode, as previously noted. To avoid over-boosting the engines, some means for determining thrust saturation, such as observation of exhaust gas temperature (EGT), exhaust pressure ratio (EPR), or the like, must be available on board the aircraft. What is needed for mode selection is a mode-independent method of determining whether thrust would be saturated if the γ-V Command mode were selected under prevailing flight conditions. For simplicity, it will be assumed that maximum thrust $T_{MAX}$ and minimum thrust $T_{MIN}$ can be calculated from the onboard engine model with sufficient accuracy.

To aid the determination of thrust saturation, a reference thrust termed $T_{REF}$ is calculated by specializing equation (6c):

$$T_{REF} = W \sin \gamma_{POT\ REF} + D \tag{7a}$$

In equation (7a), $\gamma_{POT\ REF}$ is calculated from equation (6a), as shown in figure 11(c):

$$\frac{1}{g}\frac{dV}{dt}_{CMD} + \sin \gamma = \sin \gamma_{POT\ REF} \tag{7b}$$

The reference parameters $T_{REF}$ and $\gamma_{POT\ REF}$ differ from their command counterparts $T_{CMD}$ and $\gamma_{POT\ CMD}$ because, during an accelerating climb or a decelerating descent, $T_{REF}$ and $\gamma_{POT\ REF}$ can exceed the physically realizable values corresponding to the performance capability of the aircraft (eq. (1c)), whereas $T_{CMD}$ and $\gamma_{POT\ CMD}$ must be physically realizable. Therefore

$$T_{CMD} = T_{REF} \quad \text{when} \quad T_{MIN} \le T_{REF} \le T_{MAX} \tag{7c}$$

Thrust saturation corresponds to the conditions

$$T_{REF} \le T_{MIN} \qquad\qquad T_{REF} \ge T_{MAX} \tag{7d}$$

or, equivalently, to the conditions

$$\gamma_{POT\ REF} \le \gamma_{POT\ MIN} \qquad\qquad \gamma_{POT\ REF} \ge \gamma_{POT\ MAX} \tag{7e}$$

Equation (1b) is specialized as follows:

$$\sin \gamma_{POT\ MAX} = \frac{T_{MAX} - D}{W} \tag{7f}$$

$$\sin \gamma_{POT\ MIN} = \frac{T_{MIN} - D}{W} \tag{7g}$$

It should be noted that the parameter $\gamma_{POT\ MIN}$ is negative for transport aircraft. Furthermore, the condition $T_{MAX} > T_{MIN}$ holds except during total propulsion failure.

### Pitch Command

The derivation of the elevator servo command is illustrated for completeness by the block diagrams of figure 11(e) and 11(f), but it is not treated in any detail by this report because the pitch control assumption already made implies that the lift command is identically satisfied. It can be seen that the angle-of-attack command is obtained by solving equation (6b) for the lift function $L = L(\alpha)$ and then inverting, as illustrated by figure 11(e). The pitch command $\theta_{CMD}$ is then obtained by making use of the measured flightpath angle and bank angle (compare equation (5)). The pitch regulator and the inversion of the pitching acceleration command to obtain the elevator servo command are illustrated by figure 11(f).

### V Command Mode

The V Command mode corresponds to Type (ii) behavior. Thrust is fixed, so that only pitch is available for control, and airspeed is to be controlled at the expense of flightpath angle by feeding back airspeed error to pitch control. Again the control law is obtained by specializing the aircraft differential equations.

#### Thrust Command

As already noted, this report is chiefly concerned with thrust saturation at maximum or minimum thrust, but for generality it is assumed for the present that the thrust can be fixed at any physically realizable value $T_{TGT}$. In the V Command mode, the commanded thrust is then given by the equation

$$T_{CMD} = T_{TGT} \quad \text{when } T_{MIN} \leq T_{TGT} \leq T_{MAX} \tag{8a}$$

Equation (1b) is specialized as follows:

$$\sin \gamma_{POT\ TGT} = \frac{T_{TGT} - D}{W} \tag{8b}$$

#### Reference Thrust

In the V Command mode, the reference thrust is calculated just as it is for the γ-V Command mode (eqs. (7a) and (7b)).

#### Control Law

In equation (1c), the desired normalized longitudinal acceleration, denoted by $(l/g)(dV/dt)_{CMD}$, is generated by the airspeed regulator (fig. 11(c)), just as in the γ-V Command mode. With the thrust fixed at $T_{TGT}$, $\gamma_{POT}$ is fixed at $\gamma_{POT\ TGT}$ (eq. (8b)), and equation (1c) can be solved for the flightpath angle required to satisfy the longitudinal acceleration command. Denoting this flightpath angle by $\gamma_{SPEED}$, equation (1c) becomes

$$\frac{1}{g}\frac{dV}{dt}_{CMD} - \sin \gamma_{POT\ TGT} = - \sin \gamma_{SPEED} \tag{8c}$$

The derivation of $\gamma_{SPEED}$ is illustrated by the block diagram of figure 11(c). $\gamma_{SPEED}$ is then rate-limited and applied to the input of the path regulator to generate the normal acceleration command (figs. 11(a) and 11(e)).

Equation (6b) is solved for the lift command just as it is in the γ-V Command mode (fig. 11(e)):

$$\frac{V}{g}\frac{d\gamma}{dt}_{CMD} + \cos\gamma = \frac{L}{W}_{CMD} \tag{8d}$$

It should be noted that equations (8c) and (8d) are not solved simultaneously, because the immediately previous value of the drag D is used for the calculation of $\gamma_{POT\ TGT}$ in equation (8b).

### Longitudinal Acceleration Limiting

Following Lambregts (Lambregts, 1983), an acceleration limiter is incorporated within the airspeed regulator (fig. 11(c)) that limits the longitudinal acceleration command in a manner to be described. The acceleration limiter logic is independent of mode, but it is described here because its function is essential for the V Command mode. For clarity, the basic concept is presented first without refinements to be incorporated later.

**Concept**– The acceleration limiter is specified conceptually by the following equations:

$$(1/g)(dV/dt)_{CMD\ MIN} \le (1/g)(dV/dt)_{CMD} \le (1/g)(dV/dt)_{CMD\ MAX} \tag{9a}$$

$$(1/g)(dV/dt)_{CMD\ MAX} = K_V \sin\gamma_{POT\ MAX} \tag{9b}$$

$$(1/g)(dV/dt)_{CMD\ MIN} = K_V \sin\gamma_{POT\ MIN} \tag{9c}$$

where $0 \le K_V \le 1$.

The parameter $K_V$ determines the division of excess thrust between acceleration and climb, or between deceleration and descent. For example, assume that thrust is saturated at maximum thrust and that the target thrust $\gamma_{POT\ TGT}$ is set equal to $\gamma_{POT\ MAX}$, and denote the corresponding value of $\gamma_{SPEED}$ by $\gamma_{SPEED\ MAX}$. Then equation (8c) shows that $\gamma_{SPEED\ MAX}$ is given by the equation

$$\sin\gamma_{SPEED\ MAX} = \sin\gamma_{POT\ MAX} - \frac{1}{g}\frac{dV}{dt}_{CMD} \tag{9d}$$

Further assume that the commanded longitudinal acceleration $(l/g)(dV/dt)_{CMD}$ takes on its maximum value of $K_V \sin\gamma_{POT\ MAX}$ according to equations (9a) and (9b). With this assumption, equation (9d) becomes

$$\sin\gamma_{SPEED\ MAX} = (1 - K_V)\sin\gamma_{POT\ MAX} \tag{9e}$$

Equation (9e) shows that, if $K_V = 0$, then $\gamma_{SPEED\ MAX} = \gamma_{POT\ MAX}$, so that all available excess thrust $(T_{MAX} - D)/W = \sin\gamma_{POT\ MAX}$ is used for climb. If $K_V = 1$, then $\gamma_{SPEED\ MAX}$ vanishes, so that all excess thrust is used for level-flight acceleration. The positive upper bound on commanded acceleration (eq. (9b)) prevents an excessive positive acceleration command from causing the aircraft to descend, because the right-hand side of equation (9e) is nonnegative for all $K_V$ within its range $0 \leq K_V \leq 1$.

Similarly, the negative lower bound (eq. (9c)) prevents an excessive negative acceleration command from causing the aircraft to climb. If thrust is saturated at minimum thrust and the target thrust $\gamma_{POT\ TGT}$ is set equal to $\gamma_{POT\ MIN}$, and if the commanded longitudinal acceleration $(1g)(dV/dt)_{CMD}$ takes on its minimum value of $K_V \sin \gamma_{POT\ MIN}$, then it can be shown by a similar argument that equation (9d) becomes

$$\sin \gamma_{SPEED\ MIN} = (1 - K_V) \sin \gamma_{POT\ MIN} \tag{9f}$$

Because the condition ($\gamma_{POT\ MIN} < 0$) always holds for transport aircraft, as already noted, the right-hand side of equation (9f) is nonpositive for all $K_V$ within its range $0 \leq K_V \leq 1$.

**Pilot evaluation**– Evaluation of these limiter characteristics by means of piloted simulation has shown that the condition ($K_V = 1$) is unacceptable, because some definitely positive flightpath angle is required for climb. Similarly, a definitely negative flightpath angle is required for descent. Therefore, the parameter $K_V$ should be bounded away from unity by adopting the modified specification $0 \leq K_V < 1$. The nominal value of $K_V$ has been set tentatively to 0.7, based on pilot opinion obtained during an exploratory simulation carried out on the NASA-Ames Vertical Motion Simulator (Sherry, Youssefi, and Hynes, 1995).

**Performance degradation**– A second refinement is needed to deal with the possibility that the condition ($\gamma_{POT\ MAX} < 0$) might hold, which could occur, for example, following engine failure during high-altitude cruising flight if aircraft performance became so severely degraded that available thrust were insufficient for level flight. In that case, equation (9b) shows that the system would command anomalous negative acceleration unrelated to airspeed error, resulting in airspeed divergence toward stalling speed.

In order to prevent this potentially catastrophic behavior and ensure that some minimal positive acceleration capability is always retained, a lower bound of 0.00525 g = 0.1 kt/sec is imposed on the upper acceleration limit $(dV/dt)_{CMD\ MAX}$. This minimal acceleration capability is selected to coincide with the acceleration available at maximum cruising altitude (appendix B), which is discussed later. It is clear from equation (9b) that imposing this lower bound of $0.00525 \leq K_V \sin \gamma_{POT\ MAX}$ also requires that the parameter $K_V$ be bounded away from zero.

**Limiter properties**– With the two refinements just discussed, the final limiter specification can be stated as follows:

$$(1/g)(dV/dt)_{CMD\ MIN} \leq (1/g)(dV/dt)_{CMD} \leq (1/g)(dV/dt)_{CMD\ MAX} \tag{9a}$$

$$0.00525 \leq (1/g)(dV/dt)_{CMD\ MAX} = K_V \sin \gamma_{POT\ MAX} \tag{9g}$$

$$(1/g)(dV/dt)_{CMD\ MIN} = K_V \sin \gamma_{POT\ MIN} \tag{9h}$$

where $0 < K_V < 1$.

The limiter specified by conditions (9a), (9g), and (9h) is characterized by several properties that will be used extensively during later developments. If aircraft performance is normal, then the condition ($K_V \sin \gamma_{POT\ MAX} \geq 0.00525$) holds. In that case, equations (9e) and (9f) show that the conditions

$$\gamma_{SPEED\ MIN} < 0 \qquad \gamma_{SPEED\ MAX} > 0 \tag{9i}$$

must hold. Furthermore, as a consequence of the physical condition ($\gamma_{POT\ MIN} < \gamma_{POT\ MAX}$), which holds in the absence of total propulsion failure, the condition

$$\gamma_{SPEED\ MAX} > \gamma_{SPEED\ MIN} \tag{9j}$$

holds by definition of $\gamma_{SPEED}$.

If aircraft performance is degraded, so that the condition ($K_V \sin \gamma_{POT\ MAX} \geq 0.00525$) does not hold, then the condition ($\gamma_{SPEED\ MAX} > 0$) can hold only if the condition ($V > V_{TGT}$) holds, because in that case deceleration commanded by the airspeed regulator augments $\gamma_{SPEED\ MAX}$ according to equation (9d). With the specification $0 < K_V < 1$, equation (9f) shows that the condition ($\gamma_{SPEED\ MIN} < 0$) holds generally, because as already noted $\gamma_{POT\ MIN}$ is negative for transport aircraft.

### Pitch Command

In the V Command mode, the pitch command is calculated just as it is for the γ-V Command mode (figs. 11(e) and 11(f)).

## γ Command Mode

The γ Command mode corresponds to Type (iii) behavior. Thrust is fixed, so that only pitch is available for control, and flightpath angle is to be controlled at the expense of airspeed by feeding back path error to pitch control. As before, the control law is obtained by specializing the aircraft differential equations.

### Control Law

In equation (2c), the desired normal acceleration $(V/g)(d\gamma/dt)_{CMD}$ is generated by the flightpath regulator (fig. 11(e)), and with the measured flightpath angle γ equation (2c) is solved for $(L/W)_{CMD}$, just as it is in the γ-V Command mode.

In equation (1c), the normalized longitudinal acceleration $(1/g)(dV/dt)$ is determined by the instantaneous flightpath angle γ and the target value of $\gamma_{POT}$, which depends on the fixed target thrust (eq. (8b)). In the γ Command mode, there is no closed-loop control of airspeed.

The control law for the $\gamma$ Command mode is given by the equations

$$\frac{1}{g}\frac{dV}{dt} + \sin\gamma = \sin\gamma_{\text{POT TGT}} \tag{10a}$$

and

$$\frac{V}{g}\frac{d\gamma}{dt}_{\text{CMD}} + \cos\gamma = \frac{L}{W}_{\text{CMD}} \tag{10b}$$

### Pitch Command

In the $\gamma$ Command mode, the pitch command is calculated just as it is for the $\gamma$-V Command mode (figs. 11(e) and 11(f)).

### Thrust Command

In the $\gamma$ Command mode, the commanded thrust is given by equation (8a), just as it is in the V Command mode.

### Reference Thrust

In the $\gamma$ Command mode, the reference thrust is calculated just as it is for the $\gamma$-V Command mode (eqs. (7a) and (7b)).

## Dynamic Thrust Saturation

Now that the control laws for the three primitive control modes and the details of the longitudinal acceleration limiter have been specified, it is of interest to study the geometrical stability regions for each of the primitive modes that correspond to the regions of figure 10, accounting for thrust saturation under dynamic conditions (that is, including the contribution of the $(dV/dt)_{\text{CMD}}$ term according to equation (6a)). It will be shown that contours corresponding to dynamic thrust saturation will subdivide the (V, $\gamma$) plane.

### Thrust Saturation Contours

The thrust saturation contour corresponding to the variation of $\gamma_{\text{SPEED MAX}}$ with airspeed can be obtained as follows. Substitute the condition $\gamma_{\text{POT TGT}} = \gamma_{\text{POT MAX}}$ and the definition of' $\gamma_{\text{POT MAX}}$ (eq. (7f)) into equation (8c) to obtain the equation

$$\sin\gamma_{\text{SPEED MAX}} = \frac{T_{\text{MAX}} - D}{W} - \frac{1}{g}\frac{dV}{dt}_{\text{CMD}}. \tag{11a}$$

In equation (11a), the drag D should be evaluated under the assumption that the normal forces remain in equilibrium during speed changes, as previously discussed. The variation of $T_{\text{MAX}}$ with airspeed is known (appendix B), and the variation of the commanded longitudinal acceleration $(dV/dt)_{\text{CMD}}$ is defined by the speed regulator law (fig. 11(c)). It follows that the variation of $\gamma_{\text{SPEED MAX}}$ with airspeed" is determined by equation (11a). Similarly, replacing $T_{\text{MAX}}$ by $T_{\text{MIN}}$ in equation (11a), the variation of $\gamma_{\text{SPEED MIN}}$ with airspeed is given by the equation

$$\sin \gamma_{\text{SPEED MIN}} = \frac{T_{\text{MIN}} - D}{W} - \frac{1}{g}\frac{dV}{dt}_{\text{CMD}}. \tag{11b}$$

## Equivalent Conditions for Thrust Saturation

As already noted, thrust saturation is evaluated by comparing the reference thrust $T_{\text{REF}}$, which is calculated continuously, with the maximum thrust $T_{\text{MAX}}$ and the minimum thrust $T_{\text{MIN}}$ (eqs. (7a) – (7g)). It can be seen from equations (7a), (7f), and (7g) that this thrust comparison is equivalent to comparing $\gamma_{\text{POT REF}}$ with $\gamma_{\text{POT MAX}}$ and $\gamma_{\text{POT MIN}}$. Equivalent conditions for thrust saturation that are useful for study of the corresponding geometrical regions can be developed as follows.

Equation (7b) can be put in the form

$$\sin \gamma = \sin \gamma_{\text{POT REF}} - \frac{1}{g}\frac{dV}{dt}_{\text{CMD}} \tag{7b}$$

Restating equation (8c),

$$\sin \gamma_{\text{SPEED}} = \sin \gamma_{\text{POT TGT}} - \frac{1}{g}\frac{dV}{dt}_{\text{CMD}} \tag{11c}$$

Equation (11c) can be subtracted from equation (7b) to obtain the equation

$$\sin \gamma - \sin \gamma_{\text{SPEED}} = \sin \gamma_{\text{POT REF}} - \sin \gamma_{\text{POT TGT}} \tag{11d}$$

By substituting $\gamma_{\text{POT MAX}}$ and $\gamma_{\text{POT MIN}}$ for $\gamma_{\text{POT TGT}}$ in equation (11d), the following equations for $\gamma_{\text{SPEED MAX}}$ and $\gamma_{\text{SPEED MIN}}$ can be obtained:

$$\sin \gamma - \sin \gamma_{\text{SPEED MAX}} = \sin \gamma_{\text{POT REF}} - \sin \gamma_{\text{POT MAX}} \tag{11e}$$

and

$$\sin \gamma - \sin \gamma_{\text{SPEED MIN}} = \sin \gamma_{\text{POT REF}} - \sin \gamma_{\text{POT MIN}} \tag{11f}$$

Equations (11e) and (11f) show that the following conditions on measured flightpath angle $\gamma$ are equivalent to the thrust saturation conditions (7e) on $\gamma_{\text{POT REF}}$:

$$(\gamma \geq \gamma_{\text{SPEED MAX}}) \equiv (\gamma_{\text{POT REF}} \geq \gamma_{\text{POT MAX}}) \tag{11g}$$

and

$$(\gamma \leq \gamma_{\text{SPEED MIN}}) \equiv (\gamma_{\text{POT REF}} \leq \gamma_{\text{POT MIN}}) \tag{11h}$$

Extensive use will be made of the alternative thrust saturation conditions (11g) and (11h).

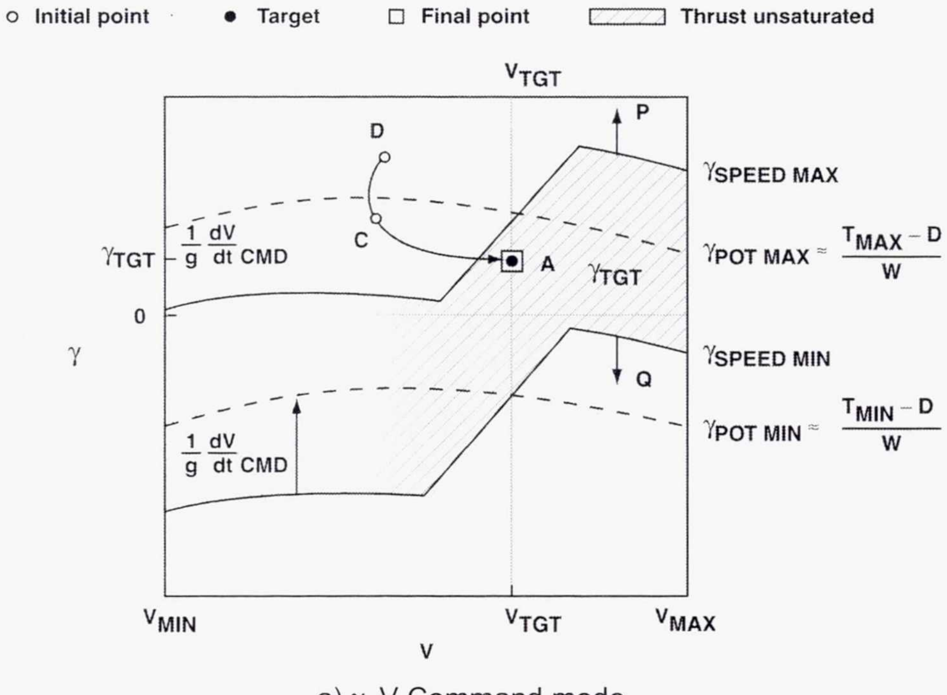
## Geometrical Stability Regions

The contours corresponding to $\gamma_{SPEED\ MAX}$ and $\gamma_{SPEED\ MIN}$ are illustrated by figure 12 for each of the three primitive control modes. In figure 12, the contour for $\gamma_{SPEED\ MAX}$ corresponds to the maximum instantaneous flightpath angle $\gamma$ that can be obtained without thrust saturation while simultaneously satisfying the longitudinal acceleration command. Similarly, the contour for $\gamma_{SPEED\ MIN}$ corresponds to the minimum instantaneous flightpath angle $\gamma$ that can be obtained without thrust saturation while simultaneously satisfying the longitudinal acceleration command.

It can be seen that the contours of $\gamma_{SPEED\ MAX}$ and $\gamma_{SPEED\ MIN}$ separate the (V, $\gamma$) plane into three regions that correspond to thrust saturation or to its absence, after accounting for the contribution of the commanded longitudinal acceleration $(dV/dt)_{CMD}$ to the total thrust required. The hatched central region corresponds to the absence of dynamic thrust saturation; elsewhere, the thrust control is saturated. It is convenient to designate the regions in which thrust would be saturated if $\gamma_{TGT}$ were captured by defining the conditions P and Q as follows:

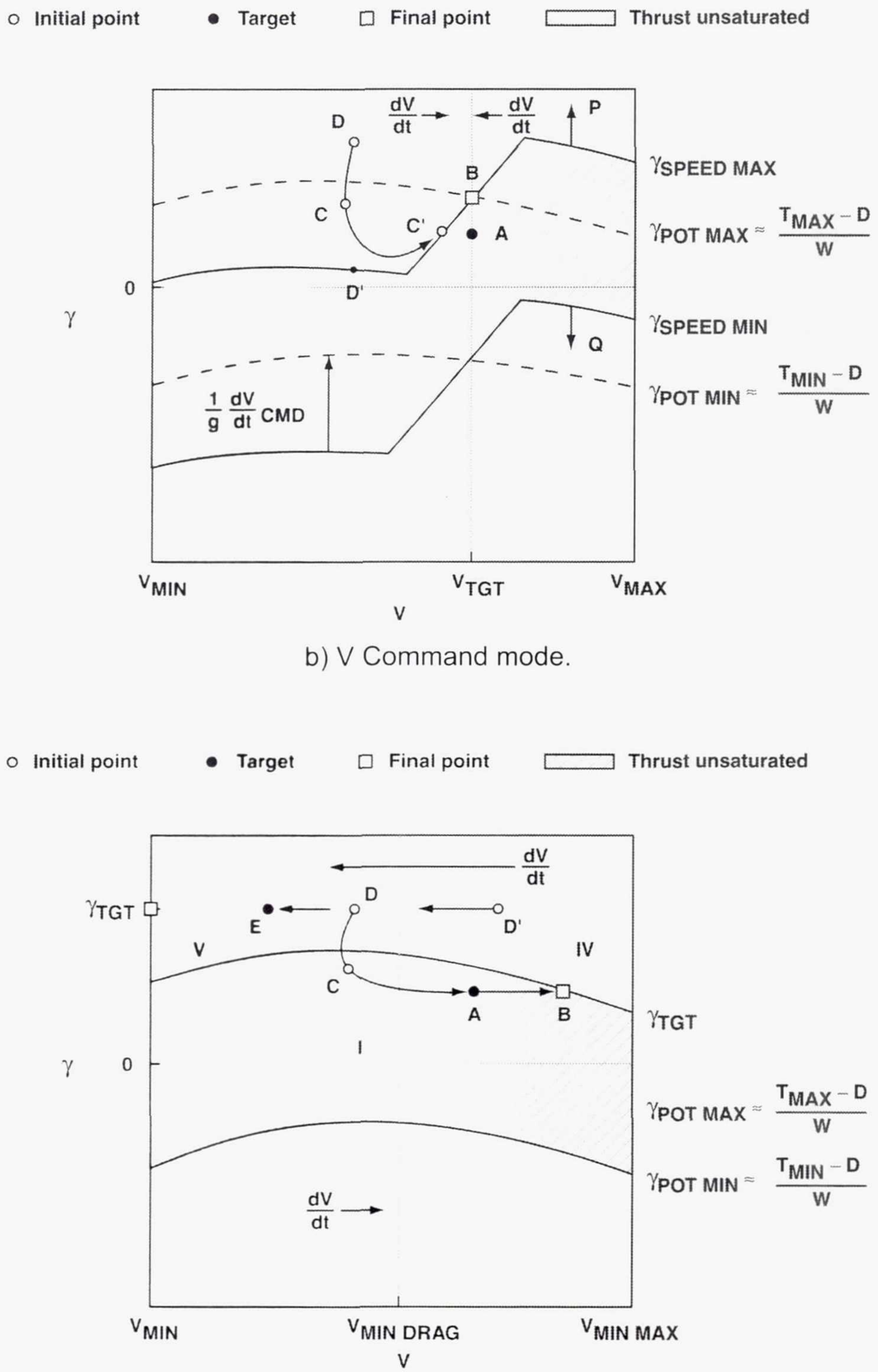
$$P \equiv (\gamma_{TGT} \ge \gamma_{SPEED\,MAX}) \qquad Q \equiv (\gamma_{TGT} \le \gamma_{SPEED\,MIN}) \tag{11i}$$

The regions in which $\gamma_{TGT}$ must lie when P and Q hold are marked on the diagrams of figures 12(a) and 12(b).



a) $\gamma$–V Command mode.

Figure 12. Geometric stability regions for primitive modes.



b) V Command mode.

c) γ Command mode.

Figure 12. Geometric stability regions for primitive modes (concluded).

In figures 12(a) and 12(b), contours of $\gamma_{\mathrm{SPEED\ MAX}}$ and $\gamma_{\mathrm{SPEED\ MIN}}$ for a representative target airspeed $V_{\mathrm{TGT}}$ are illustrated. The forms of these contours result from limiting of the longitudinal acceleration command, as discussed previously (eqs. (9a)–9(j)). For airspeeds near $V_{\mathrm{TGT}}$, the commanded longitudinal acceleration is proportional to the airspeed error (fig. 11(c)), so that the portions of the contours for $\gamma_{\mathrm{SPEED\ MAX}}$ and $\gamma_{\mathrm{SPEED\ MIN}}$ near $V_{\mathrm{TGT}}$ are steeply sloping straight lines. (Integral control within the airspeed regulator would distort their linear form to some extent.)

For larger airspeed errors, the acceleration limiter becomes active. In that case, because the acceleration command is limited to a constant fraction of $\gamma_{\mathrm{POT\ MAX}}$ or $\gamma_{\mathrm{POT\ MIN}}$ (eqs. (9a), (9g), and (9h)), the $\gamma_{\mathrm{SPEED\ MAX}}$ and $\gamma_{\mathrm{SPEED\ MIN}}$ contours lie roughly parallel to the maximum and minimum thrust contours (figs. 12(a) and 12(b)). It can be seen that the displacement of the $\gamma_{\mathrm{SPEED\ MAX}}$ contour from the $\gamma_{\mathrm{POT\ MAX}}$ contour is approximately equal, for small angles, to the normalized longitudinal acceleration command $(1/g)(dV/dt)_{\mathrm{CMD}}$, as required by equation (11c), and likewise for the displacement of the $\gamma_{\mathrm{SPEED\ MIN}}$ contour from the $\gamma_{\mathrm{POT\ MIN}}$ contour.

### γ-V Command Mode

In figure 12(a), the hatched central region that corresponds to absence of dynamic thrust saturation is available for operation in the γ-V Command mode. Capture of a representative target point ($V_{\mathrm{TGT}}$, $\gamma_{\mathrm{TGT}}$) is illustrated (point A). Outside the hatched region the thrust control is saturated, so that the γ-V Command mode cannot be engaged there. However, if operation in the γ-V Command mode were permitted with saturated thrust, the dynamical behavior of the system would be the same as it is in the γ Command mode, which is discussed shortly.

During operation in the γ-V Command mode, any target point can be captured provided that it lies within the steady-state performance envelope (region I, figure 10(a)), but this is not required of the initial point. As illustrated by figure 12(a), if the target lies at point A, it can be captured starting from points such as point D that lie outside the steady-state performance envelope. In this example, the γ Command mode would be selected for the initial part of the capture trajectory DCA, and the γ-V Command mode would be selected when the trajectory enters the central hatched region (fig. 12(a)).

### V Command Mode

In the V Command mode the thrust setting is fixed, as noted previously, but the physical significance of the hatched central region in figure 12(b) is the same as it is in figure 12(a): the hatched region is that within which the instantaneous operating point (V, γ) must lie to enable the longitudinal acceleration command to be satisfied with a physically realizable thrust setting.

During operation in the V Command mode, the flightpath target $\gamma_{\mathrm{TGT}}$ is ignored by the system, and $\gamma_{\mathrm{SPEED}}$ is captured instead. For example, suppose the instantaneous operating point (V, γ) lies initially at point D (fig. (12(b)). If $\gamma_{\mathrm{POT\ TGT}}$ is set to $\gamma_{\mathrm{POT\ MAX}}$ and the V Command mode is selected, the flightpath angle $\gamma_{\mathrm{SPEED\ MAX}}$ is captured. After path capture is complete, the (V, γ) operating point must lie on the $\gamma_{\mathrm{SPEED\ MAX}}$ contour at some point such as point C', with the acceleration directed toward the target airspeed (indicated by arrows on the diagram). In figure 12(b), point B indicates the point on the $\gamma_{\mathrm{SPEED\ MAX}}$ contour with coordinates ($V_{\mathrm{TGT}}$, $\gamma_{\mathrm{SPEED\ MAX}}$), which is the point finally captured. Point B is the best approximation to the target point (point A, as in figure 12(a)) that can be realized with maximum thrust when the target airspeed is specified as shown.

The points on the $\gamma_{\text{SPEED MAX}}$ contour to the left of point B correspond to an accelerating climb, and those to the right of point B correspond to a decelerating climb. Similarly, during minimum-thrust descent, the (V, $\gamma$) operating point lies on the $\gamma_{\text{SPEED MIN}}$ contour, with the acceleration directed toward the target airspeed.

### γ Command Mode

In the γ Command mode the thrust setting is fixed, just as it is in the V Command mode, but in the γ Command mode any target flightpath angle $\gamma_{\text{TGT}}$ can be captured without restriction to the steady-state flight envelope. Because there is no closed-loop control of airspeed, the airspeed target $V_{\text{TGT}}$ is ignored by the system, and no use is made of the commanded longitudinal acceleration $(dV/dt)_{\text{CMD}}$. The instantaneous longitudinal acceleration is determined from equation (l0a), which can be put in the form

$$(1/g)(dV/dt) = \sin \gamma_{\text{POT TGT}} - \sin \gamma \qquad (10a)$$

where γ is the instantaneous flightpath angle. Equation (l0a) shows that, in the γ Command mode, the sign of the longitudinal acceleration dV/dt is determined by comparing the instantaneous flightpath angle γ with the fixed target thrust $\gamma_{\text{POT TGT}}$.

During operation in the γ Command mode, the sign of the longitudinal acceleration for any selected value of $\gamma_{\text{TGT}}$ can be determined from figure 12(c). For example, suppose that the selected target ($V_{\text{TGT}}$, $\gamma_{\text{TGT}}$) lies at point A (fig. 12(c)), the initial operating point (V, γ) lies at point C, and $\gamma_{\text{POT TGT}}$ is set to $\gamma_{\text{POT MAX}}$. Then the longitudinal acceleration is positive at both point C and point A, as required by equation (10a). As shown by figure 12(c), the trajectory overshoots the target at point A and captures point B, where the longitudinal acceleration vanishes. Point B is the best approximation to the target (point A, as in figure 12(a)) that can be realized with maximum thrust when the target flightpath angle is specified as shown.

If the initial operating point (V, γ) lies at point D, the initial longitudinal acceleration is negative, becoming positive when the trajectory enters the hatched region as shown (fig. 12(c)). This example shows how the initial part of the trajectory DCA shown in figure 12(a) can be generated by selecting the γ Command mode while thrust is saturated.

If the γ Command mode is selected when the target point lies in region V, then the behavior of the system is quite different. For example, suppose that the selected target lies above the $\gamma_{\text{POT MAX}}$ contour (point E, figure 12(c)), and the initial operating point lies in region V (point D) or the upper part of region IV (point D'). In either case, the longitudinal acceleration (marked on the diagram by an arrow) is negative, as required by equation (l0a). The trajectory overshoots the target at point E, and the airspeed diverges toward stalling speed (fig. 12(c)). Therefore, when the target point lies in region V or the upper part of region IV above the $\gamma_{\text{POT MAX}}$ contour, selection of the γ Command mode would not be acceptable.

### Summary

Dynamic thrust saturation has been analyzed, and it has been shown how contours corresponding to dynamic thrust saturation subdivide the flightpath-airspeed plane into geometric regions (fig. 12). Several examples have illustrated the behavior of each of the three primitive modes for target points in each of those regions.

These examples suggest some of the dynamical issues that must be addressed during mode selection. It is essential to provide a comprehensive strategy that ensures appropriate system behavior when thrust saturates. Before a valid mode control logic can be developed, it is necessary to specify the requirements for safety and functionality that the system must satisfy, and to define explicitly what is meant by mode validity. These matters are discussed next.

## SAFETY REQUIREMENTS

Transport aircraft safety requirements are based on a priori identification of safety hazards, as are safety requirements in several other industries (Leveson, 1995). Safety margins intended to provide protection against these hazards are then applied to the ultimate physical operating limits of the aircraft to define a safe operating envelope within which all normal operation is required to take place. Thus the development of safety requirements for the longitudinal control system begins with a survey of known flight hazards that restrict the physical operational envelope of the aircraft.

### Flight Hazards

### Wing Stall

Wing stall occurs at the angle of attack for which the lift function takes on its maximum value (appendix B). The stalling airspeed depends on aircraft weight and on normal acceleration, and thus on the flight demonstration technique used during certification to establish the stalling speed experimentally. In some cases, available pitch control power may not be sufficient to permit demonstration of a well-defined stall. In others, safety hazards anticipated from wind tunnel tests (such as deep stall or pitch-up at high angles of attack) make it inadvisable to approach the stall closely during certification flight tests, so that certification is based on a declared maximum angle of attack that can be demonstrated safely. The consequences of exceeding such a declared limit are presumed to be catastrophic. Whatever the basis, firm values for stalling angle of attack and airspeed for all aircraft configurations and weights are determined during the certification process.

### Thrust Asymmetry

If engine failure occurs, the resulting thrust asymmetry generates yawing and rolling moments that can threaten transport aircraft controllability. Because control power decreases with decreasing airspeed, a minimum control airspeed exists below which controllability cannot be retained with maximum thrust applied to the operating engines. Engine-out controllability is assessed during the certification process, and minimum control airspeeds are determined for all aircraft configurations and weights. The consequences of violating minimum control-speed restrictions are presumed to be catastrophic.

### Overspeed

The maximum airspeed is determined at low altitude by structural limits such as excessive airloads or flutter, and at high altitudes by compressibility (Mach) effects that degrade controllability (for example, Mach buffet) as shock waves form on the aerodynamic surfaces of the aircraft. Just as for wing stall, maximum airspeed limits are determined for all aircraft configurations and weights during the certification process, and the consequences of exceeding those limits are presumed to be catastrophic.

### Ground Contact

At large airports suitable for transport-category aircraft, a survey plane is established at the approach and departure ends of each runway, and the locations and heights of all obstacles penetrating this survey plane (nominal gradient 2%) are marked on aeronautical charts. The gradients specified for all instrument approach procedures for each runway are selected to provide safe terrain clearance in the vicinity of the airport. Similar surveys determine safe minimum en-route altitudes for all airways throughout the National Airspace System. The consequences of contact with the ground other than the runway, or with an obstacle such as a building or a radio tower, are presumed to be catastrophic.

## Safety Margins

### Airspeed Margins

An airspeed margin equal to 30% of the stalling speed is applied to the stalling speed to determine the minimum airspeed to be used during final approach. Intentional operation below this reference approach speed is prohibited in airline service. Other airspeed margins are applied to the takeoff stalling speed and to the engine-out minimum control speed to establish the takeoff safety speed, and still others are subtracted from the maximum airspeed limits to determine the maximum safe operating airspeeds for all aircraft configurations and weights in both smooth and rough air.

### Safe Descent Margins

At low altitude, the maximum safe descent gradient during instrument approach is determined by the published approach procedure. At higher altitudes, operational limits on rate of descent are often imposed by airline operating policy to avoid excessive rates of descent that could lead to violation of safe altitude limits. At high altitudes where terrain clearance is not an issue, excessive rates of descent that could lead to overspeed should be avoided. These rate-of-descent limits can be expressed as flightpath angle limits on the descent angle $\gamma_{SAFE}$ that vary with airspeed.

### Safe Operating Envelope

Application of the airspeed margins and the safe descent margins to the physical envelope limits determine the safe envelope limits $V_{MIN}$, $V_{MAX}$, and $\gamma_{SAFE}$ within which the automated control system is required to operate (fig. 13). To provide protection against violation of safe operating envelope limits, envelope protection modes are developed in the next section.

### Discussion

The presence of these envelope protection modes should not be regarded as reducing the designer's obligation to ensure that the normal control modes operate safely. To the contrary, if the design of the normal control modes is valid, envelope protection should be invoked (activated) only in cases where aircraft performance limitations preclude other alternatives.

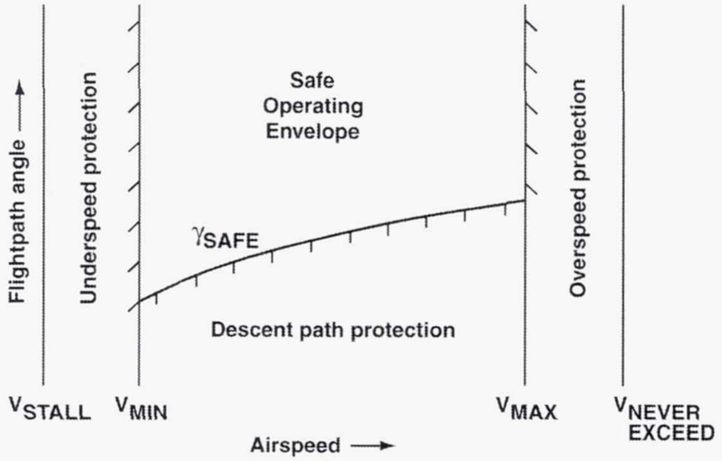


Figure 13. Safety envelope for Path/Speed Command supermode.

# Envelope Protection

## Envelope Protection Requirements

Envelope protection requirements can be stated as follows. Protection against underspeed should be provided by setting $V_{TGT}$ equal to $V_{MIN}$ plus a small tolerance and selecting the γ-V Command mode. Similarly, protection against overspeed should be provided by setting $V_{TGT}$ equal to $V_{MAX}$ less a small tolerance and selecting the γ-V Command mode. In both cases, if the thrust saturates, the V Command mode should be selected to give priority to airspeed.

Protection against penetration of the $\gamma_{SAFE}$ boundary should be provided by setting $\gamma_{TGT}$ equal to $\gamma_{SAFE}$ plus a small tolerance and selecting the γ-V Command mode. If the thrust saturates, the γ Command mode should be selected to give priority to path. In case of conflict between speed and path envelope violations, priority should be given to airspeed protection, except in the presence of strong wind shear. Shear disturbances exceeding the FAA-specified threshold of 2 kt/sec must be detected, and must cause the control system to transition to a special wind shear recovery mode that is not treated in this report.

In all cases where envelope protection is invoked, the maximum thrust limit should be increased as much as possible (takeoff or contingency thrust), and appropriate warnings must be annunciated. It is also desirable to increase the limits on normal and longitudinal acceleration imposed during fully automated flight for passenger comfort. The system should revert to normal operation if that again becomes valid.

## Envelope Protection Modes

The envelope protection logic developed to meet these requirements is specified by table 1, which is presented to illustrate representative control logic for envelope protection modes. Because the focus of this report is on the modes engaged during normal operation, no validity analysis is performed on the envelope protection mode logic, and no claims are made for its formal validation. A detailed description of condition-action decision tables like table 1 can be found in appendix D.

TABLE 1. SELECTION OF ENVELOPE PROTECTION MODES

| Condition | Underspeed Protection | Overspeed Protection | Descent Path Protection |
|---|---|---|---|
| $V_{MIN} \leq V \leq V_{MAX}$ | | | TRUE |
| $V < V_{MIN}$ | TRUE | | |
| $V > V_{MAX}$ | | TRUE | |
| $\gamma \geq \gamma_{SAFE}$ | | | |
| $\gamma < \gamma_{SAFE}$ | | | TRUE |
| **Action** | Set $V_{TGT}$ equal to $V_{MIN}$ + tolerance | Set $V_{TGT}$ equal to $V_{MAX}$ – tolerance | Set $\gamma_{TGT}$ equal to $\gamma_{SAFE}$ + tolerance |
| | Set thrust limit to contingency thrust<br>Increase acceleration limits | | |
| | Select $\gamma$-V Command mode. If thrust saturates: | | |
| | Select V Command mode | | Select<br>$\gamma$ Command mode |

The purpose of the tolerance parameters in table 1 is to provide hysteresis that prevents repetitive cycling of mode transitions (“chattering”). Appropriate values for these parameters should be selected during implementation. What exactly is meant by “giving priority to airspeed or to flight-path angle” is clarified in the next section.

## EFFECTIVENESS REQUIREMENTS

### Effectiveness Concept

In addition to the a priori safety requirements already discussed, the system must also be effective in performing its intended function. This notion is termed effectiveness. (In the design procedure to be presented, effectiveness plays a role similar to that of liveness in computer science.) The specification of desired system functionality in terms of general effectiveness properties is the task of this section.

For the longitudinal control system, the desired effectiveness can be stated as a general stability property in the following way. Starting at any initial operating point specified by $(V, \gamma)$ coordinates, abrupt change of target to a new point specified by $(V_{TGT}, \gamma_{TGT})$ coordinates lying within the closed flight envelope $\varepsilon$ (fig. 8) should cause the aircraft to capture and hold the new target smoothly, without exceeding (in smooth air) the acceleration limits imposed for passenger comfort. If the target lies outside the closed flight envelope $\varepsilon$, the aircraft should capture and hold a point within the closed envelope that provides an acceptable approximation to the target.

The desired safety properties require that the aircraft remain at all times within the envelope safety limits illustrated by figure 13, as previously discussed. These proposed statements of safety and effectiveness properties taken together restrict the effective targets to points within the steady-state (trimmed) performance envelope, and do not permit exploitation of dynamic maneuvers to provide enhanced performance. This limitation seems consistent with both the mission and the operating characteristics of transport aircraft, but likely would not be satisfactory for higher-performance aircraft.

If a single mode (for example, the $\gamma$-V Command mode) could provide the desired effectiveness throughout the closed envelope $\varepsilon$, that system would consist of a single continuous element, and the required stability property could be established by known control-theoretic methods. Such use of the $\gamma$-V Command mode would require avoidance of thrust saturation, as previously discussed. But thrust saturation cannot be avoided in practice, because efficient operation during climb and descent requires operation on the thrust saturation boundary (appendix B). Both the V Command mode and the $\gamma$ Command mode can deal with thrust saturation, but neither can capture both path and speed targets simultaneously. Therefore, none of the three primitive modes can provide the desired effectiveness throughout the closed envelope $\varepsilon$.

However, it will be shown that the desired effectiveness can be achieved by combining the three primitive modes into a higher-level entity that will be termed a supermode, with a well-defined strategy for choosing among the three primitive modes whenever a thrust saturation boundary is encountered. This supermode, which is termed Path/Speed Command, constitutes a hybrid system as previously defined. The desired safety and effectiveness properties must therefore be established for this hybrid system, a problem for whose solution no general theoretical method is known. It is hoped that the present study may contribute by example to the development of more general methods. The resulting design method should also be of practical interest to designers of next-generation transport aircraft avionic systems.

To aid in fixing ideas, the discussion of effectiveness will begin with a detailed example that illustrates use of the three primitive modes already discussed to perform a representative higher-level task. The task selected is that of capturing an altitude assigned by ATC; other representative tasks could have been chosen.

## Altitude Capture Example

It is assumed that an aircraft in level flight at 15,000 ft is cleared to climb to 35,000 ft (Flight Level 350). For efficiency, the aircraft is to climb at 250 kt, and then accelerate to 280 kt for cruise. To execute this clearance, the first task is to pull up from level flight into a steady climb, maintaining a constant normal acceleration of 0.1g during the pull-up. This pull-up task can be accomplished by the automated control system in the following way.

### Pull-up

Referring to figure 11(a), the altitude regulator function is illustrated by the structure to the left of the broken vertical line. In this example, the target altitude $H_{TGT}$ is changed abruptly from 15,000 ft to 35,000 ft while the measured altitude H remains at 15,000 ft. The altitude error $\Delta H$ is therefore

35,000 – 15,000 = 20,000 ft. It can be seen from the upper row of the block diagram of figure 11(b) that the commanded vertical velocity $(dH/dt)_{CMD}$ then becomes

$$(dH/dt)_{CMD} = \sqrt{2(3)\left[|\Delta H| - H_0/2\right]} = \sqrt{6\left[20{,}000 - 150\right]} = 345.1 \text{ ft/sec}$$

At 15,000 ft, an equivalent airspeed (EAS) of 250 kt corresponds to a true airspeed (TAS) of 531.9 ft/sec (appendix B). The sine of the commanded flightpath angle $\gamma_{TGT}$ is calculated to be

$$\sin \gamma_{TGT} = (dH/dt)_{CMD}/V = 345.1/531.9 = 0.6488 \qquad \gamma_{TGT} = 40.45 \text{ deg}$$

Clearly, the raw $\gamma_{TGT}$ calculated from the parabolic law greatly exceeds the performance capability of the aircraft. (It may be recalled that in the numerical example illustrating calculation of the flight envelope, the maximum climb angle $\gamma_{POT\ MAX}$ was found to be only 6.92 degrees at sea level with maximum thrust at 289 kt EAS). In practice, to ensure application of maximum thrust during climb, the flightpath angle $\gamma_{TGT}$ is set to an angle 3 degrees steeper than $\gamma_{POT\ MAX}$ (fig. 11(b)). This value of $\gamma_{TGT}$ and the airspeed target $V_{TGT}$ of 250 kt provide the inputs to the primitive mode structure to the right of the broken vertical line (fig. 11(a)). In the γ-V Command mode, the normal acceleration limiter limits the rate of change of flightpath angle to

$$d\gamma/dt = (g/V)(0.1) = (32.174/531.9)(0.1) = 0.00605 \text{ rad/sec} = 0.35 \text{ deg/sec}$$

and the aircraft pulls up at this constant rate, increasing thrust to maintain airspeed constant (fig. 11(c)). When the reference thrust $T_{REF}$ exceeds the maximum thrust $T_{MAX}$ the thrust command saturates (eq. (7d)). The γ-V Command mode used during the pull-up is then no longer viable, and one of the other two primitive modes must be engaged.

### Climb

It is assumed that, for efficiency, the climb to 35,000 ft is to be conducted with maximum thrust at the target airspeed of 250 kt EAS, accepting whatever climb angle results as thrust decreases with altitude. Such a climb procedure is representative of subsonic jet transport operations (appendix B). Clearly, it corresponds to operation in the V Command mode, which must therefore be selected when thrust saturation occurs, setting the target thrust $T_{TGT}$ equal to $T_{MAX}$ (figs. 11(a) and 11(c)).

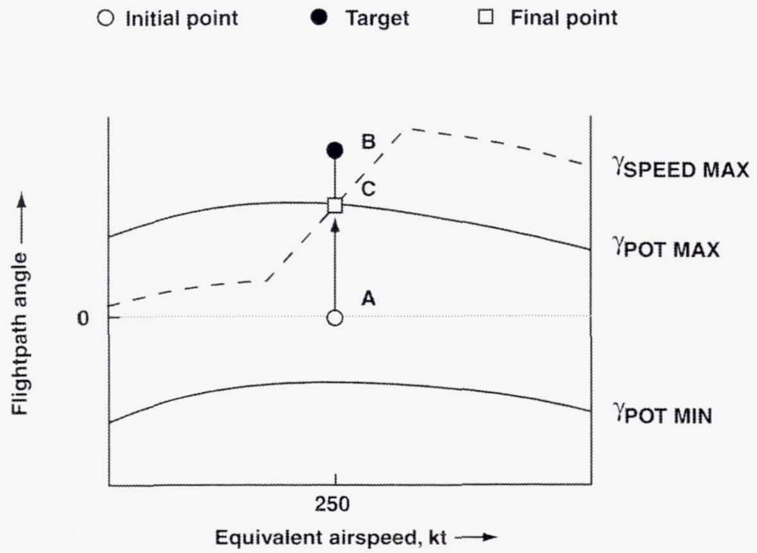
During the initial climb, the flightpath angle target $\gamma_{TGT}$ remains on its upper limit (fig. 11(b)), but as the aircraft approaches the target altitude of 35,000 ft and the altitude error is reduced, $\gamma_{TGT}$ is no longer limited, and it decreases smoothly to zero as the aircraft reaches the target altitude. But in the V Command mode in which the aircraft is operating during the climb, the system captures and holds $\gamma_{SPEED}$, ignoring $\gamma_{TGT}$ (fig. 11(a)). Therefore, to capture the target altitude smoothly, the γ-V Command mode must again be selected as soon as $\gamma_{TGT}$ becomes equal to or less than $\gamma_{SPEED\ MAX}$. The aircraft then pushes over to level flight as the altitude error is nulled.

### Altitude Capture

Up to this point, the altitude capture example has made use of only two of the three primitive modes. The aircraft must now accelerate to a cruising airspeed of 280 kt EAS at the top of the climb, the airspeed target being changed abruptly from 250 kt EAS to 280 kt EAS at the start of the push-over for altitude capture. The increased thrust required for longitudinal acceleration then maintains thrust saturation during the initial altitude capture. Capturing and tracking the flightpath angle target $\gamma_{TGT}$ with the thrust saturated requires use of the $\gamma$ Command mode (fig. 11(a)), which must be selected at the start of altitude capture. As the aircraft approaches the target airspeed, the commanded longitudinal acceleration is reduced until eventually the thrust is no longer saturated, and the $\gamma$-V Command mode can once again be engaged.

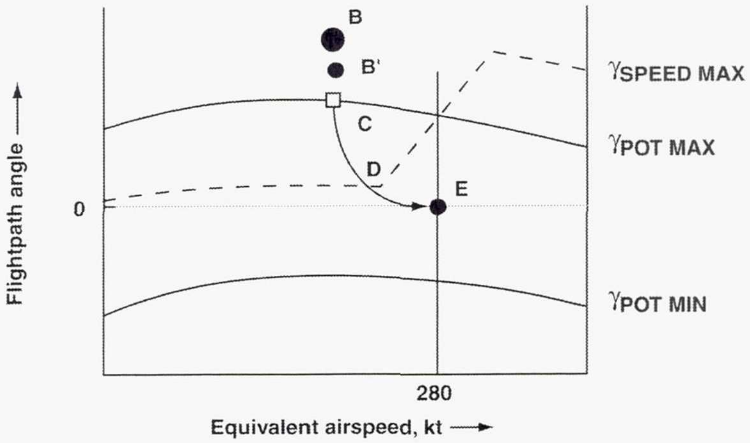
### Transition Paths

The transition paths in the (V, $\gamma$) plane corresponding to the altitude capture example are illustrated by the diagrams of figure 14, which are similar to those of figure 12. In figure 14(a), point A corresponds to the initial condition in level flight at 15,000 ft and 250 kt. Point B corresponds to the target point when the target altitude of 35,000 ft is first entered. The target flightpath angle $\gamma_{TGT}$ is limited to an angle 3 degrees above the maximum thrust contour at the target airspeed of 250 kt, as previously explained. During the pull-up, the rate-limited flightpath angle $\gamma_{LIM}$ traverses the path extending vertically upward from point A as the system, operating in the $\gamma$-V Command mode, attempts to capture the target point B. When the thrust saturates at point C, the V Command mode is selected, and the system then holds $\gamma_{SPEED}$ at 250 kt (point C) during the steady climb. The target flightpath angle $\gamma_{TGT}$ remains on its upper limit (point B) until the aircraft approaches the target altitude of 35,000 ft. When the limit is no longer active, the flightpath target (point B') moves downward toward point C as the altitude error is reduced (fig. 14(b)). When $\gamma_{TGT}$ coincides with point C, the airspeed target is changed to 280 kt, and the $\gamma$ Command mode is selected. During altitude capture, the flightpath and airspeed follow the path C-D-E, the thrust remaining saturated until point D is reached. At point D, the $\gamma$-V Command mode is engaged, and the altitude capture is completed in level flight at the cruising airspeed of 280 kt EAS (point E).



a) Pull-up.

Figure 14. Altitude capture example.



b) Push-over.

Figure 14. Altitude capture example (concluded).

As previously noted, there is no closed-loop control of airspeed while thrust is saturated along the path C-D. Nevertheless, the longitudinal acceleration is positive throughout the altitude capture, and the airspeed target is captured as expected. On the other hand, premature selection of the $\gamma$ Command mode before the target flightpath angle has reached point C (that is, before the aircraft has approached the target altitude closely enough) could result in operation in the $\gamma$ Command mode at point B'. In that case, the longitudinal acceleration would be negative during the initial portion of the altitude capture, contrary to the desired behavior.

The altitude capture example just presented has brought out several key points. In the next section, this example is generalized as a first step toward definition of requirements for a systematic mode selection strategy for the Path/Speed Command supermode.

## Path/Speed Command Supermode Specifications

### Path/Speed Priority

An essential element of this mode selection strategy is the priority to be placed on controlling path or controlling speed when thrust saturates, because only one of the two parameters can then be controlled. It is clear that operation in the $\gamma$ Command mode gives priority to path, because it holds path at the expense of speed. Similarly, operation in the V Command mode gives priority to speed. Thus the priority placed on path or speed determines the selection of either the $\gamma$ Command mode or the V Command mode within the Path/Speed Command supermode whenever thrust saturates.

But this priority decision must be made at a higher level within the mode hierarchy where sufficient information is available. The altitude capture example shows that the correct decision depends on the phase of flight (initial pull-up, climb, altitude capture, and the like), which is not known within the Path/Speed Command supermode. Therefore, the path/speed priority must be provided as an external input to the Path/Speed Command supermode that is specified by the higher-level entity in the mode hierarchy that invokes the Path/Speed Command supermode as a lower-level element. For the altitude capture task, this higher-level entity is a supermode termed Altitude Command; it is discussed in detail in a later section. The path/speed priority transmitted from the upper level to the lower can be regarded from a general system design viewpoint as control flowing downward.

### Mode Hierarchy

From the viewpoint of the higher-level supermode invoking the Path/Speed Command supermode, the higher-level supermode provides to the Path/Speed Command supermode three parameters: (1) a target flightpath angle, (2) a target airspeed, and (3) a choice of path or speed priority if thrust saturates. The Path/Speed Command supermode hides all information relative to the performance capability of the aircraft, its flight envelope, and the differential equations governing control of flightpath and airspeed, and annunciates to the invoking supermode the effectiveness to be expected in capturing the flightpath and airspeed targets. In the next section, these general notions are made precise by formal definition of three different classes of effectiveness.

### Effectiveness Definitions

The stability analysis carried out previously shows that, depending on which primitive mode is engaged, the Path/Speed Command supermode can exhibit three different kinds of behavior in capturing path and speed targets, which correspond to Type (i) control, Type (ii) control, or Type (iii) control.

> ***Complete effectiveness*** is defined as capture of both path and speed targets within small tolerances that account for expected disturbances, which characterizes Type (i) control. If thrust saturates, Type (i) control is not available, and either Type (ii) control or Type (iii) control must be selected. In that case, capture of the target that is given priority (either path or speed) is assured, but the other target is ignored by the system.
>
> ***Partial effectiveness*** is defined as capture of a point of stable equilibrium that provides an acceptable approximation to the target that is ignored.
>
> ***Normal effectiveness*** is defined as follows: If the target path and speed have not yet been captured, but the normal and longitudinal accelerations have the expected signs leading toward capture, this behavior is termed normal effectiveness.

It follows from their definitions that complete effectiveness implies partial effectiveness and is therefore logically dominant, whereas normal effectiveness is a weaker property that is useful chiefly in the short term.

Application of these definitions to each of the three primitive modes is described next. To enable effectiveness to play a useful role in mode selection, it is essential that each effectiveness property be defined in such a way that evaluation based on its definition is independent of the mode selected at the time the evaluation is made.

### Effectiveness of $\gamma$-V Command Mode

For the $\gamma$-V Command Mode, normal effectiveness holds immediately upon engagement, and complete effectiveness holds in the long term. These properties follow directly from the stability of the flightpath and airspeed regulators. Partial effectiveness does not apply to the $\gamma$-V Command mode, which can be engaged only in the absence of thrust saturation.

## Effectiveness of V Command Mode

In the V Command mode, priority is given to airspeed. The flightpath angle $\gamma_{\text{SPEED}}$ is captured, and subsequent capture of the target airspeed is assured by the stability properties of the path and speed regulators (figs. 11(c) and 11(e)). Therefore, normal effectiveness for the V Command mode applies only to capture of the airspeed target, for which it always holds because of the stability of the airspeed regulator. The flightpath target $\gamma_{\text{TGT}}$ is ignored, but partial effectiveness of the V Command mode requires that $\gamma_{\text{SPEED}}$ constitute an acceptable approximation to $\gamma_{\text{TGT}}$. How should this requirement be interpreted?

### Partial Effectiveness

If the flightpath target $\gamma_{\text{TGT}}$ lies in the region of thrust saturation above the $\gamma_{\text{SPEED MAX}}$ contour in figure 12(b) (that is, if the condition $P \equiv (\gamma_{\text{TGT}} \geq \gamma_{\text{SPEED MAX}})$ holds), then the best physically realizable approximation to $\gamma_{\text{TGT}}$ is the point on the $\gamma_{\text{SPEED MAX}}$ contour nearest $\gamma_{\text{TGT}}$ at the prevailing airspeed. For example, if point D (fig. 12(b)) has the coordinates $(V_D, \gamma_{\text{TGT}})$, then the best physically realizable approximation is the point D', which has coordinates $(V_D, \gamma_{\text{SPEED MAX}})$.

To realize the best approximation whenever the condition $P \equiv (\gamma_{\text{TGT}} \geq \gamma_{\text{SPEED MAX}})$ holds, the thrust target $\gamma_{\text{POT TGT}}$ should be set equal to $\gamma_{\text{POT MAX}}$, so that $\gamma_{\text{SPEED}}$ is equal to $\gamma_{\text{SPEED MAX}}$; no other physically realizable thrust setting could provide as good an approximation. In contrast, suppose that P holds, but $\gamma_{\text{POT TGT}}$ is *not* equal to $\gamma_{\text{POT MAX}}$. That case should be considered a violation of partial effectiveness of the V Command mode, because the best approximation is not realized; the same is true if Q holds, but $\gamma_{\text{POT TGT}}$ is not equal to $\gamma_{\text{POT MIN}}$ (fig. 12(b)).

Therefore, expressing logical negation by the word NOT and denoting partial effectiveness of the V Command mode symbolically as PE (V), a preliminary definition for PE (V) based on general notions of best approximation can be stated as follows:

$$\text{PE (V)} \equiv \text{NOT } [\text{P AND } (\gamma_{\text{POT TGT}} \neq \gamma_{\text{POT MAX}})] \text{ AND NOT } [\text{Q AND } (\gamma_{\text{POT TGT}} \neq \gamma_{\text{POT MIN}})]$$

To relate PE (V) to the geometrical regions illustrated by figure 12(b), its definition must be simplified. It is shown in appendix D by means of symbolic logic that the definition for PE (V) just stated can be simplified in the following three ways:

$$\text{If } (\gamma_{\text{POT TGT}} = \gamma_{\text{POT MAX}}) \text{ holds, then PE (V)} \equiv \text{NOT Q.}$$

$$\text{If } (\gamma_{\text{POT TGT}} = \gamma_{\text{POT MIN}}) \text{ holds, then PE (V)} \equiv \text{NOT P.}$$

$$\text{If } (\gamma_{\text{POT TGT}} < \gamma_{\text{POT TGT}} < \gamma_{\text{POT MAX}}) \text{ holds, then PE (V)} \equiv \text{NOT P AND NOT Q.}$$

(To verify the first statement without the use of symbolic logic, which otherwise will not be needed until later, assume that $(\gamma_{\text{POT TGT}} = \gamma_{\text{POT MAX}})$ holds. Then the definition of PE (V) becomes

$$\text{PE (V)} \equiv \text{NOT [P AND FALSE] AND NOT [Q AND TRUE].}$$

Because the proposition within the first right-hand bracket is FALSE whatever P may be, its negation NOT FALSE must be TRUE. The proposition within the second bracket reduces to Q, with negation NOT Q. With these simplifications, PE (V) becomes

$$\text{PE (V)} \equiv \text{TRUE AND NOT Q},$$

which reduces to the statement that PE (V) is equivalent to NOT Q, verifying the first result. The other two statements can be verified in a similar manner.)

These three simplified statements show that PE (V) is related to the regions illustrated by figure 12(b) in the following ways. When $\gamma_{POT\ TGT}$ is set equal to $\gamma_{POT\ MAX}$ and PE (V) is therefore equivalent to NOT Q, then PE (V) holds when $\gamma_{TGT}$ lies in the region above the $\gamma_{SPEED\ MIN}$ contour where NOT Q holds (fig. 12(b)). When $\gamma_{POT\ TGT}$ is set equal to $\gamma_{POT\ MIN}$ and PE (V) is therefore equivalent to NOT P, then PE (V) holds when $\gamma_{TGT}$ lies in the region below the $\gamma_{SPEED\ MAX}$ contour where NOT P holds (fig. 12(b)). When $\gamma_{POT\ TGT}$ is set equal to an intermediate thrust level and PE (V) is equivalent to NOT P AND NOT Q, then PE (V) holds when $\gamma_{TGT}$ lies in the central hatched region (fig. 12(b)) where NOT P AND NOT Q holds.

It is clear that, in general, violation of PE (V) would result from inappropriate setting of the target thrust $\gamma_{POT\ TGT}$. The conditions ($\gamma_{POT\ TGT} \neq \gamma_{POT\ MAX}$) and ($\gamma_{POT\ TGT} \neq \gamma_{POT\ MIN}$) are made more specific later.

## Effectiveness of γ Command Mode

In the γ Command mode, priority is given to path. Following engagement of the γ Command mode with a fixed-path target, capture of the selected path target is ensured by the stability property of the path regulator (fig. 11(e)). The airspeed target $V_{TGT}$ is ignored, but partial effectiveness of the γ Command mode requires that a point of stable airspeed equilibrium be captured that provides an acceptable approximation to $V_{TGT}$. What conditions must be imposed to ensure partial effectiveness of the γ Command mode?

### Partial Effectiveness

After path capture is complete, the instantaneous flightpath angle γ remains fixed at $\gamma_{TGT}$. It will be recalled from the previous discussion of the dynamical behavior of the aircraft that, when flightpath angle is constrained to a fixed angle such as $\gamma_{TGT}$ with thrust fixed, one point of stable airspeed equilibrium lies above the speed for minimum drag at the intersection of the horizontal line corresponding to $\gamma_{TGT}$ with the fixed-thrust contour (point A of figure 9(b)). Furthermore, this point of stable equilibrium will be captured if the operating point (V, γ) lies within its region of attraction, which consists of all points on the horizontal $\gamma_{TGT}$ line to the right of the point of unstable equilibrium (point B of figure 9(b)), as previously explained. Depending on the thrust setting $\gamma_{POT\ TGT}$, there are three cases; the case corresponding to maximum thrust is treated first.

**Maximum thrust**– When $\gamma_{POT\ TGT}$ is set equal to $\gamma_{POT\ MAX}$ and $\gamma_{TGT}$ lies in region I (point A of figure 12(c)), point B is a point of stable equilibrium. Its region of attraction consists of the points on the horizontal $\gamma_{TGT}$ line either within region I or within the lower part of region IV (fig. 12(c)). This region of attraction is defined by the condition ($\gamma_{TGT} < \gamma_{POT\ MAX}$) for region I and by the two conditions ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) and ($V > V_{MIN\ DRAG}$) for region IV.

Therefore, a preliminary definition for partial effectiveness of the γ Command mode (denoted by PE (γ)) based on the requirement that the operating point (V, γ) lie within the region of attraction of a point of stable equilibrium can be stated symbolically as follows:

$$PE\ (\gamma) \equiv (\gamma_{TGT} < \gamma_{POT\ MAX})\ \text{OR}\ [(\gamma_{TGT} \geq \gamma_{POT\ MAX})\ \text{AND}\ (V > V_{MIN\ DRAG})]$$

It is shown in appendix D by means of symbolic logic that this definition of PE (γ) can be simplified as follows:

$$PE\ (\gamma) \equiv (\gamma_{TGT} < \gamma_{POT\ MAX})\ \text{OR}\ (V > V_{MIN\ DRAG})$$

(This result can be verified without the use of symbolic logic, which otherwise is not needed until later, but that exposition is awkward and is not presented here.)

Violation of PE (γ), which requires annunciation to be discussed later, is expressed by its negation, using the word NOT:

$$\text{NOT}\ PE\ (\gamma) \equiv (\gamma_{TGT} \geq \gamma_{POT\ MAX})\ \text{AND}\ (V \leq V_{MIN\ DRAG})$$

It can be seen that, with this logical definition, PE (γ) holds if $\gamma_{TGT}$ lies in region I at the prevailing airspeed (point A of figure 12(c)), but is violated in region V, where the violation-defining conditions ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) and ($V \leq V_{MIN\ DRAG}$) both hold (point E of figure 12(c)). If the operating point (V, γ) should penetrate region V under these conditions, then the airspeed would diverge toward stalling speed, as explained previously. Thus the required annunciation of PE (γ) violation gives warning of potentially catastrophic behavior that can enable timely recovery action, avoiding invocation of underspeed protection.

**Discussion**– The preliminary logical definition for PE (γ) just stated appears faulty if $\gamma_{TGT}$ lies in the upper part of region IV above the $\gamma_{POT\ MAX}$ contour for all airspeeds within region IV (point D' of figure 12(c)). In that case, PE (γ) holds according to the stated definition, but no point of equilibrium can be captured because none exists, contradicting the defining property for PE (γ). The airspeed diverges toward stalling speed because the condition ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) holds for all airspeeds (fig. 12(c)). However, airspeed divergence presents no immediate threat to safe operation while the operating point remains within region IV. Furthermore, the operating point must eventually penetrate region V, generating a timely warning. Therefore, unnecessary complication can be avoided without compromising safety by allowing PE (γ) to hold everywhere in region IV. The preliminary definition is then satisfactory in the form stated.

This definition also appears faulty if the condition ($\gamma_{TGT} < \gamma_{POT\ MAX}$) holds at $V_{MAX}$, because in that case PE (γ) holds according to the definition, but no point of equilibrium can be captured because none exists below $V_{MAX}$ (fig. 12(c)). Overspeed could occur. However, if $\gamma_{TGT}$ lies in region I at $V_{MAX}$ (that is, if the condition ($\gamma_{POT\ MIN} < \gamma_{TGT} < \gamma_{POT\ MAX}$) holds at $V_{MAX}$), overspeed can be prevented by selecting the $\gamma$-V Command mode, which provides closed-loop control of both airspeed and path, as already noted. The $\gamma$-V Command mode is available for selection unless thrust saturates.

Therefore, the stated definition of PE ($\gamma$) is satisfactory provided that thrust remains unsaturated. The case in which $\gamma_{TGT}$ lies below region I is treated shortly.

**Minimum thrust**– When $\gamma_{POT\ TGT}$ is set equal to $\gamma_{POT\ MIN}$, a point of stable equilibrium (if one exists) lies on the $\gamma_{POT\ MIN}$ contour to the right of $V_{MIN\ DRAG}$ (fig. 12(c)). Overspeed could occur if the condition ($\gamma_{TGT} < Y_{POT\ MIN}$) holds at $V_{MAX}$, even though PE ($\gamma$) holds under the stated definition.

However, $\gamma_{TGT}$ is required to lie above the safe descent limit $\gamma_{SAFE}$, as previously discussed. At high altitudes where $\gamma_{SAFE}$ is not constrained by considerations of terrain clearance, airline operating policy often imposes a limit $(dH/dt)_{SAFE}$ on maximum rate of descent, which is related to $\gamma_{SAFE}$ by the equation $(dH/dt)_{SAFE} = V \sin \gamma_{SAFE}$. Therefore, overspeed can be prevented by choosing the safe descent limit so that the condition ($\gamma_{SAFE} \geq \gamma_{POT\ MIN}$) holds at $V_{MAX}$, and by enforcing the constraint ($\gamma_{TGT} \geq \gamma_{SAFE}$) on $\gamma_{TGT}$. These two conditions together imply that the condition ($\gamma_{TGT} \geq \gamma_{POT\ MIN}$) holds at $V_{MAX}$, which is sufficient to ensure that a point of stable equilibrium will be captured without the occurrence of overspeed (fig. 12(c)).

Therefore, the stated definition of PE ($\gamma$) remains satisfactory in this case also, provided that overspeed is prevented by a suitable choice of $\gamma_{SAFE}$. (At low altitudes, of course, the safe descent limit could be further restricted by terrain clearance considerations.)

**Intermediate thrust**– Finally, assume that thrust is fixed at an intermediate level, so that the condition ($\gamma_{POT\ MIN} < \gamma_{POT\ TGT} < \gamma_{POT\ MAX}$) holds. If the condition ($\gamma_{TGT} < \gamma_{POT\ MAX}$) also holds, then PE ($\gamma$) holds under the stated definition. However, airspeed could diverge toward stalling speed if the condition ($\gamma_{TGT} \geq \gamma_{POT\ TGT}$) holds, or toward overspeed if ($\gamma_{TGT} < \gamma_{POT\ TGT}$) holds. However, if $\gamma_{TGT}$ lies in region I (that is, if ($\gamma_{POT\ MIN} < \gamma_{TGT} < \gamma_{POT\ MAX}$) holds), airspeed divergence can be prevented by selecting the $\gamma$-V Command mode, which is available for selection unless thrust saturates. Therefore, the stated definition of PE ($\gamma$) is satisfactory in that case provided that thrust remains unsaturated.

If $\gamma_{TGT}$ lies below region I (fig. 12(c)) (that is, if ($\gamma_{TGT} < \gamma_{POT\ MIN}$) holds), the situation is the same as that discussed previously for minimum thrust: overspeed could occur, but can be prevented by a suitable choice of $\gamma_{SAFE}$. In that case, the stated definition of PE ($\gamma$) remains satisfactory in the form stated.

If $\gamma_{TGT}$ lies above region I (fig. 12(c)) (that is, if ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) holds), the situation is the same as that discussed previously for maximum thrust: the stated definition of PE ($\gamma$) is satisfactory in the form stated.

**Summary**– To summarize, analysis has shown that, with the stated definition

$$\text{NOT PE } (\gamma) \equiv (\gamma_{TGT} \geq \gamma_{POT\ MAX}) \text{ AND } (V \leq V_{MIN\ DRAG})$$

partial effectiveness holds in several situations in which airspeed diverges, and no point of stable equilibrium is captured. If $\gamma_{TGT}$ lies in region I, airspeed divergence can be prevented by selecting the $\gamma$-V Command mode unless thrust saturates. Furthermore, only saturation at maximum thrust is of concern, because at minimum thrust airspeed divergence toward overspeed can be prevented by a suitable choice of $\gamma_{SAFE}$. The same is true if $\gamma_{TGT}$ lies below region I. However, if thrust saturates, the $\gamma$-V Command mode is not available for recovery.

Therefore, to ensure that PE ($\gamma$) holds only in cases where a stable point of equilibrium is captured (the original definition), or else in cases where airspeed divergence can be prevented by some other means, the thrust saturation condition ($\gamma \geq \gamma_{SPEED\ MAX}$) will be conjoined to the previous definition of NOT PE ($\gamma$), so that the final definition can be stated as follows:

$$\text{NOT PE } (\gamma) \equiv (\gamma_{TGT} \geq \gamma_{POT\ MAX}) \text{ AND } (V \leq V_{MIN\ DRAG}) \text{ AND } (\gamma \geq \gamma_{SPEED\ MAX}).$$

### Normal Effectiveness

In the $\gamma$ Command mode, normal effectiveness always holds for capture of the flightpath target because of the stability property of the path regulator. Normal effectiveness may or may not hold for capture of the airspeed target, as shown by the altitude capture example.

By definition, normal effectiveness for capture of the airspeed target requires that the sign of the longitudinal acceleration be matched appropriately to that of the airspeed error. In the $\gamma$ Command mode, the longitudinal acceleration is given by equation (l0a):

$$(1/g)\ dV/dt = \sin \gamma_{POT\ TGT} - \sin \gamma \tag{10a}$$

When the flightpath target is captured, the flightpath angle $\gamma$ becomes equal to $\gamma_{TGT}$, and equation (l0a) becomes

$$(1/g)\ dV/dt = \sin \gamma_{POT\ TGT} - \sin \gamma_{TGT} \tag{12}$$

Equation (12) shows that in the medium and long term after path capture is complete, the longitudinal acceleration dV/dt is determined by comparing the target flightpath $\gamma_{TGT}$ with the target thrust $\gamma_{POT\ TGT}$. Similarly, the airspeed error is determined by comparing the prevailing airspeed V with the target airspeed $V_{TGT}$. Nine possible combinations require evaluation; they are summarized by the following table:

| | **DV/dt > 0** | **DV/dt = 0** | **DV/dt < 0** |
|---|---|---|---|
| | $\boldsymbol{\gamma_{TGT} < \gamma_{POT\ TGT}}$ | $\boldsymbol{\gamma_{TGT} = \gamma_{POT\ TGT}}$ | $\boldsymbol{\gamma_{TGT} > \gamma_{POT\ TGT}}$ |
| $\mathbf{V < V_{TGT}}$ | Case 1 | Case 4 | Case 7 |
| $\mathbf{V = V_{TGT}}$ | Case 2 | Case 5 | Case 8 |
| $\mathbf{V < V_{TGT}}$ | Case 3 | Case 6 | Case 9 |

It can be seen that normal effectiveness holds by definition in cases 1, 5, and 9:

| | | |
|---|---|---|
| 1. | $V < V_{TGT}$ | $DV/dt > 0$ |
| 5. | $V = V_{TGT}$ | $DV/dt = 0$ |
| 9. | $V > V_{TGT}$ | $DV/dt < 0$ |

In the other cases normal effectiveness is violated. To facilitate meaningful annunciation, it is convenient to group cases 2, 3, and 6 together by defining the property NOT NE1 as follows:

$$\text{NOT NE1} \equiv [(V = V_{TGT}) \text{ AND } (\gamma_{TGT} < \gamma_{POT\ TGT})] \text{ OR } [(V > V_{TGT}) \text{ AND } (\gamma_{TGT} \leq \gamma_{POT\ TGT})]$$

If the condition NOT NE1 holds, the system should annunciate the cautionary message “More drag.” Similarly, the property NOT NE2 is defined by grouping cases 4, 7, and 8 together:

$$\text{NOT NE2} \equiv [(V = V_{TGT}) \text{ AND } (\gamma_{TGT} > \gamma_{POT\ TGT})] \text{ OR } [(V < V_{TGT}) \text{ AND } (\gamma_{TGT} \geq \gamma_{POT\ TGT})]$$

If the condition NOT NE2 holds, the system should annunciate the cautionary message “More thrust.” With the violations NOT NE1 and NOT NE2 defined as stated, normal effectiveness can then be defined symbolically by the property NE, where

$$\text{NE} \equiv \text{NE1 AND NE2}$$

It should be noted that normal effectiveness is not essential, because its violation presents no immediate threat to continued safe operation in the $\gamma$ Command mode. Therefore, the required annunciations should be presented to the human crew as cautions (advisories), not as warnings.

### Annunciation Requirements

To assure design integrity, two kinds of requirements for annunciation of the effectiveness of the Path/Speed Command supermode must be imposed. In the first place, complete effectiveness, partial effectiveness, and normal effectiveness must be evaluated continuously, and violations must be annunciated over the whole mode hierarchy up to the top-level entity invoking the Path/Speed Command supermode, which specifically includes the human crew.

Secondly, the certainty that violations of effectiveness will be correctly identified in every case must be provided by a guarantee of logical completeness. The foundation for this logical completeness has already been laid by the previous demonstration that no point exists in the flightpath-airspeed plane that does not lie in one of the seven geometrical stability regions of figure 10. Therefore, definition of effectiveness over all seven regions ensures logical completeness.

This statement of annunciation requirements completes the specification of effectiveness for the Path/Speed Command supermode. In the next section, safety and effectiveness are combined to define mode validity.

# MODE VALIDITY

## Validity Concept

### Definition

Validity is defined to hold when both safety and effectiveness requirements are satisfied; at least partial effectiveness is required. Under this definition, invalid operation is inevitable if aircraft performance limitations preclude even partial effectiveness. For example, because envelope protection modes sacrifice effectiveness in favor of safety, by definition they are invalid relative to the purpose for which the Path/Speed Command mode is selected.

### Validity Regions

With validity defined, the next task is to determine the regions of state space that correspond to valid operation in each primitive mode, as suggested by Heymann (personal discussion). For the aircraft longitudinal control problem considered here, the relevant state space is that of the two controlled variables, that is, the $(V, \gamma)$ plane. To avoid an invalid mode transition, the transition boundary between each pair of candidate modes must lie within the intersection of their regions of validity, as explained by the previous discussion ("Introduction"). Furthermore, this condition must hold notwithstanding the effect upon the shape of each validity region of external influences (such as wind shear or engine failure) to which the hybrid system must react.

These regions of validity are determined in two stages. First, validity conditions for each primitive mode are determined directly from their governing differential equations. Second, these validity conditions are related to the geometrical stability regions studied previously (figs. 10 and 12).

## Validity of γ-V Command Mode

### Condition on $V_{TGT}$

The a priori safety requirement that prevailing airspeed remain within the limits $V_{MIN} \leq V \leq V_{MAX}$ imposes the same condition on the target airspeed:

$$V_{MIN} \leq V_{TGT} \leq V_{MAX} \tag{13a}$$

If the measured airspeed lies outside these safe envelope limits, envelope protection is invoked, as described previously, but these envelope protection modes are not treated in detail by this report. Thus for normal operation in the $\gamma$-V Command mode in the absence of envelope protection, it follows that both the target airspeed and the measured airspeed must lie within the safe envelope limits.

### Condition on $\gamma_{TGT}$

Similarly, the a priori safety requirement that the flightpath avoid penetration of the $\gamma_{SAFE}$ boundary imposes the same condition on the target flightpath:

$$\gamma_{TGT} \geq \gamma_{SAFE} \tag{13b}$$

If the measured flightpath penetrates the $\gamma_{SAFE}$ boundary, envelope protection is invoked, as described previously, but this envelope protection is not treated in detail by this report. Thus for normal operation in the $\gamma$-V Command mode in the absence of envelope protection, it follows that the target flightpath and the measured flightpath must both lie within the safe envelope limit.

### Condition on $\gamma$

For validity of the $\gamma$-V Command mode, thrust saturation must be avoided, as discussed previously. Thus the condition $T_{MIN} < T_{REF} < T_{MAX}$ must hold (eq. (7d)). As shown by equations (11g) and (11h), this is equivalent to the condition

$$\gamma_{SPEED\ MIN} < \gamma < \gamma_{SPEED\ MAX} \tag{14}$$

(To verify condition (14), substitute $T_{REF}$ from equation (7a), $T_{MAX}$ from equation (7f), and $T_{MIN}$ from equation (7g) into the condition ($T_{MIN} < T_{REF} < T_{MAX}$), cancel the drag D, divide out the weight W, and take inverse sines; then substitute the negations of conditions (11g) and (11h).)

The $\gamma$-V Command mode is valid if and only if conditions (13a), (13b), and (14) hold.

## Validity of V Command Mode

### Condition on $V_{TGT}$

In the V Command mode, the same validity condition on the target airspeed must hold as for the $\gamma$-V Command mode:

$$V_{MIN} \le V_{TGT} \le V_{MAX} \tag{13a}$$

### Condition on $\gamma_{SPEED}$

The safety condition (13b) is required to hold in general. In the V Command mode, however, $\gamma_{TGT}$ is replaced by $\gamma_{SPEED}$. Therefore, the a priori safety requirement that flightpath avoid penetration of the $\gamma_{SAFE}$ boundary imposes the same condition on $\gamma_{SPEED}$:

$$\gamma_{SPEED} \ge \gamma_{SAFE} \tag{15a}$$

It can be shown that the system to be developed makes no use of the V Command mode unless condition (15a) holds, provided that aircraft performance is normal (that is, the condition ($\gamma_{POT\ MAX} \ge 0$) holds). However, if aircraft performance is severely degraded (that is, if the condition ($\gamma_{POT\ MAX} < 0$) holds), then condition (15a) could be violated. In that case, no practical alternative exists that could prevent enforcement of the condition ($\gamma \ge \gamma_{SAFE}$) by invocation of descent path protection. Therefore, condition (15a) is not required to hold for validity of the V Command mode.

### Condition on PE (V)

The conditions discussed previously that determine the partial effectiveness of the V Command mode can be summarized symbolically by the statement that, for validity of the V Command mode, the property PE (V) must hold.

### Condition on $\gamma_{POT\ TGT}$

As previously discussed, in the V Command mode the target thrust must be set to some physically realizable value:

$$T_{MIN} \le T_{TGT} \le T_{MAX} \tag{15b}$$

This is equivalent to the condition

$$\gamma_{POT\ MIN} \le \gamma_{POT\ TGT} \le \gamma_{POT\ MAX} \tag{15c}$$

To verify condition (l5c), substitute $T_{TGT}$ from equation (8b), $T_{MAX}$ from equation (7f), and $T_{MIN}$ from equation (7g) into condition (15b), cancel the drag D, divide out the weight W, and take inverse sines.

### Path/Speed Priority

As previously discussed, for validity of the V Command mode, the path/speed priority must give priority to speed:

$$\text{PRIORITY} \equiv \text{SPEED} \tag{15d}$$

As previously discussed, normal effectiveness of the V Command mode applies only to capture of airspeed, for which it always holds. Therefore, no evaluation of normal effectiveness is required in the V Command mode. The V Command mode is valid if and only if conditions (13a), (13b), (l5c), and (l5d) hold, and partial effectiveness PE (V) holds.

## Validity of $\gamma$ Command Mode

### Condition on $V_{TGT}$

In the $\gamma$ Command mode, the same validity condition on the target airspeed must hold as for the $\gamma$-V Command mode and the V Command mode:

$$V_{MIN} \le V_{TGT} \le V_{MAX} \tag{13a}$$

### Condition on $\gamma_{TGT}$

In the $\gamma$ Command mode, the same validity condition on the target flightpath (eq. (13b)) must hold as for the $\gamma$-V Command mode and the V Command mode:

$$\gamma_{TGT} \ge \gamma_{SAFE} \tag{13b}$$

### Condition on PE($\gamma$)

The conditions discussed previously that determine the partial effectiveness of the $\gamma$ Command mode can be summarized symbolically by the statement that, for validity of the $\gamma$ Command mode, the property PE ($\gamma$) must hold.

For validity of the $\gamma$ Command mode, the normal effectiveness property NE is not required to hold, as noted previously, but violations must be annunciated as required.

### Condition on $\gamma_{\text{POT TGT}}$

In the $\gamma$ Command mode, the requirement that the target thrust be physically realizable imposes the same condition (eq. (l5c)) on $\gamma_{\text{POT TGT}}$ as for the V Command mode:

$$\gamma_{\text{POT MIN}} \leq \gamma_{\text{POT TGT}} \leq \gamma_{\text{POT MAX}} \tag{15c}$$

### Path/Speed Priority

As previously discussed, for validity of the $\gamma$ Command mode, the path/speed priority must give priority to path:

$$\text{PRIORITY} \equiv \text{PATH} \tag{16}$$

The $\gamma$ Command mode is valid if and only if conditions (13a), (13b), (15c) and (16) hold, and partial effectiveness PE ($\gamma$) holds.

## Summary of Validity Conditions

Collecting results, the validity conditions for all three primitive control modes are summarized by table 2.

### Geometrical Regions for Mode Validity

With qualifications to be pointed out, the validity conditions summarized by table 2 are simply related to the geometrical regions of the flightpath-airspeed plane illustrated by figures 10, 12, and 13, which were discussed previously. The two upper rows of table 2 show that the validity conditions of the target point ($V_{\text{TGT}}$, $\gamma_{\text{TGT}}$) apply to all three of the primitive modes. For validity, the target point must lie in region I, II, III, IV, or V (fig. 10); it cannot lie in region VI or VII, nor below the $\gamma_{\text{SAFE}}$ boundary (fig. 13). Additional requirements are as follows.

**$\gamma$-V Command mode**– For validity of the $\gamma$-V Command mode (first column of table 2), the instantaneous operating point (V, $\gamma$) must lie within the central hatched region of figure 12(a), within which the condition ($\gamma_{\text{SPEED MIN}} < \gamma < \gamma_{\text{SPEED MAX}}$) holds and thrust saturation is therefore absent.

**V Command mode**– For validity of the V Command mode (second column of table 2), the thrust target $\gamma_{\text{POT TGT}}$ must lie within region I or on its boundary (fig. 10(b)), and priority must be set to SPEED (not illustrated).

Furthermore, PE (V) is required to hold. The definition of PE (V) (table 2) shows that its truth value depends on P, Q, and $\gamma_{\text{POT TGT}}$. If $\gamma_{\text{POT TGT}}$ is set equal to $\gamma_{\text{POT MAX}}$, then the condition PE (V) ≡ NOT Q holds, restricting $\gamma_{\text{TGT}}$ to lie above the $\gamma_{\text{SPEED MIN}}$ contour (fig. 12(b)), as previously noted. If $\gamma_{\text{POT TGT}}$ is set equal to $\gamma_{\text{POT MIN}}$, then PE (V) ≡ NOT P holds, restricting $\gamma_{\text{TGT}}$ to lie below the $\gamma_{\text{SPEED MAX}}$ contour (fig. 12(b)), as previously noted. If $\gamma_{\text{POT TGT}}$ is set to some intermediate thrust level, then PE (V) ≡ NOT P NOT Q holds, restricting $\gamma_{\text{TGT}}$ to lie within the central hatched region of figure 12(b),

TABLE 2. VALIDITY CONDITIONS FOR PRIMITIVE MODES

| | **$\gamma$-V Command** | **V Command** | **$\gamma$ Command** |
|---|---|---|---|
| $V_{TGT}$ | $V_{MIN} \leq V_{TGT} \leq V_{MAX}$ | | |
| $\gamma_{TGT}$ | $\gamma_{TGT} \geq \gamma_{SAFE}$ | | |
| **PE(V),**<br>**PE ($\gamma$)** | | PE(V) | PE ($\gamma$) |
| $\gamma$<br>**OR**<br>$\gamma_{POT\ TGT}$ | $\gamma > \gamma_{SPEED\ MIN}$<br>$\gamma < \gamma_{SPEED\ MAX}$ | $\gamma_{POT\ MIN} \leq \gamma_{POT\ TGT} \leq \gamma_{POT\ MAX}$ | |
| **PRIORITY** | | SPEED | PATH |

**DEFINITIONS:**

PE (V) ≡ NOT [P AND ($\gamma_{POT\ TGT} \neq \gamma_{POT\ MAX}$)] AND NOT [Q AND ($\gamma_{POT\ TGT} \neq \gamma_{POT\ MIN}$)]

PE ($\gamma$) ≡ ($\gamma_{TGT} < \gamma_{POT\ MAX}$) OR ($V > V_{MIN\ DRAG}$) OR ($\gamma < \gamma_{SPEED\ MAX}$)

NOT PE ($\gamma$) ≡ ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) AND ($V \leq V_{MIN\ DRAG}$) AND ($\gamma \geq \gamma_{SPEED\ MAX}$)

P ≡ $\gamma_{TGT} \geq \gamma_{SPEED\ MAX}$    Q ≡ $\gamma_{TGT} \leq \gamma_{SPEED\ MIN}$

as previously noted. Because for validity $\gamma_{POT\ TGT}$ is required to be physically realizable (fourth row of table 2), there are no other cases in which PE (V) holds.

**$\gamma$ Command mode**– For validity of the $\gamma$ Command mode (third column of table 2), the thrust target $\gamma_{POT\ TGT}$ must lie within region I or on its boundary (fig. 10(c)), and priority must be set to PATH (not illustrated).

Furthermore, PE ($\gamma$) is required to hold. For simplicity, the region is determined within which PE ($\gamma$) is violated (that is, within which NOT PE ($\gamma$) holds). The definition of NOT PE ($\gamma$) (table 2) shows that its truth value depends on $\gamma_{TGT}$, V, and $\gamma$. It follows from the definition of NOT PE ($\gamma$) that, if NOT PE ($\gamma$) holds, then the conditions ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) and ($V \leq V_{MIN\ DRAG}$) must hold; that is, the point (V, $\gamma_{TGT}$) must lie within region V or on its boundary (fig. 12(c)).

Conversely, assume that the point (V, $\gamma_{TGT}$) lies within region V or on its boundary, so that the conditions ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) and ($V \leq V_{MIN\ DRAG}$) hold. Then for the usual case illustrated by figure 12 in which the condition ($V_{TGT} > V_{MIN\ DRAG}$) holds, ($V \leq V_{TGT}$) must hold. In that case, ($\gamma_{POT\ MAX} \geq \gamma_{SPEED\ MAX}$) holds by the definition of $\gamma_{SPEED\ MAX}$. But because ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) holds by assumption, the condition ($\gamma_{TGT} \geq \gamma_{SPEED\ MAX}$) must hold. Because path capture takes place rapidly, as explained previously, the analysis can be simplified further by examining thrust saturation only after path capture is complete (that is, after $\gamma$ becomes equal to $\gamma_{TGT}$). The condition ($\gamma_{TGT} \geq \gamma_{SPEED\ MAX}$) is then equivalent to the condition ($\gamma \geq \gamma_{SPEED\ MAX}$), so the conditions

($\gamma_{\text{TGT}} \geq \gamma_{\text{POT MAX}}$), ($V \leq V_{\text{MIN DRAG}}$), and ($\gamma \geq \gamma_{\text{SPEED MAX}}$) must all hold. It then follows from its definition that NOT PE ($\gamma$) holds.

This simplified analysis shows that PE ($\gamma$) is violated if and only if the point ($V$, $\gamma_{\text{TGT}}$) lies within region V or on its boundary (fig. 12(c)); in regions I, II, III, and IV, PE ($\gamma$) holds.

## SYNTHESIS OF PATH/SPEED COMMAND SUPERMODE

### The Synthesis Problem

The mode validity conditions summarized by table 2 complete the formulation of the synthesis problem (see the section "Overview of Design Method," steps 1–8). Its solution (see the section "Overview of Design Method," steps 9–15) requires that mode control logic for the Path/Speed Command supermode be synthesized in such a way that the required safety and effectiveness properties are established by construction.

As already noted, no general theoretical method is known for synthesizing such a hybrid system so as to ensure that its dynamical behavior meets requirements specified a priori. The behavior of such systems is usually assessed by simulation, but this is a logically incomplete process that cannot achieve formal verification or validation, as noted previously. In this section, a new method is presented for solution of the synthesis problem formulated by the discussion leading to the validity conditions of table 2. The first step toward synthesis of the required mode control logic is the inversion of table 2, interchanging its arguments with tabulated quantities.

Table 2 specifies the validity of each of the three primitive longitudinal control modes as a function of the truth values of logical conditions that depend on flightpath, airspeed, and thrust targets, on measured flightpath and airspeed, and on the path/speed priority parameter. In principle, what is needed for selection of a primitive mode is an inverted table (that is, a truth table for table 2) that shows, for each combination of logical conditions, which modes (if any) are valid, and which violations of effectiveness require annunciation to higher levels of the mode hierarchy. However, a naive attempt to construct such an inverted table by means of brute-force enumeration would lead to a combinatorial explosion of cases, as the following analysis shows.

#### Naive Inversion

After separating the validity conditions of table 2 into binary logical propositions that are either true or false, it is found that there are 5 binary conditions that determine the validity of the $\gamma$-V Command mode (first column of table 2), 7 binary conditions that determine the validity of the V Command mode (second column), and 7 binary conditions that determine the validity of the $\gamma$ Command mode (third column). If these binary conditions were all independent, there would be $2^5 = 32$ combinations of conditions requiring evaluation for assessment of the validity of the $\gamma$-V Command mode, $2^7 = 128$ combinations for the V Command mode, and $2^7 = 128$ combinations for the $\gamma$ Command mode.

After accounting for duplication of conditions between columns, it is found that there are actually 10 independent binary conditions in table 2. It follows that, for assessment of the validity of all three primitive modes by inversion of the whole table, there would be $2^{10} = 1024$ combinations requiring enumeration. Furthermore, a mode selection policy would have to be formulated specifying a unique mode selection for each of these 1024 combinations. Therefore, it must be expected that a naive, brute-force approach based on enumeration of all cases would be found intractably complex.

Nevertheless, the inversion problem can be made tractable by developing heuristic design strategies that enable partitioning of the table before inversion. This partitioning process bears a rough analogy to matrix inversion, which can often be simplified by partitioning the original matrix, inverting the partitioned elements, and re-assembling the results. To facilitate this development, it is convenient to express the validity conditions in terms of symbolic logic.

### Symbolic Logic

Readers unfamiliar with formal logic would no doubt find unintelligible the summary statement that *logical simplification of compound logical propositions* formed by *conjunction* and *disjunction* of *Boolean variables* is to be accomplished by application of *elementary logical theorems* established by *perfect induction* based on *truth tables*, leading to results expressed in the form of *condition-action decision tables* and *statecharts* characterized by *logical consistency and completeness*. Yet the solution of the synthesis problem to be described depends vitally on use of the tools of formal logic, in contrast to the formulation of the synthesis problem just completed, which has depended on continuous mathematics and on aeronautical knowledge. It is clear that solution of the synthesis problem will require a significant change in the reader's orientation, a mental "shifting of gears." Fortunately, only knowledge of elementary formal logic (on the level of an introductory course) is needed for the developments to follow.

At this point of the discussion, readers should refer to appendix D, which presents a brief tutorial review of elementary propositional logic, and to appendix E, which discusses its application to real-time computation within embedded systems. Familiarity with this background material is assumed in what follows, with specific references provided as necessary. In particular, appendix D contains a reference list of theorems in symbolic logic that are useful for simplifying compound logical propositions such as those in table 2; extensive use is made of these elementary theorems (Dromey, 1989; Bartee, 1985).

## Expression of Validity Conditions by Symbolic Logic

### Logical Notation

To express the validity conditions concisely, each validity condition is symbolized by a binary logical proposition denoted by an upper-case letter and number. For example, the validity condition ($V_{TGT} < V_{MIN}$) is denoted by the symbol VT1, and the validity condition ($V_{TGT} > V_{MAX}$) is denoted by the symbol VT2; mnemonically, VT2 is the second condition imposed on the Velocity Target, and GT6 is the sixth condition imposed on the Gamma Target.

For typographical convenience, logical relationships between these propositions are denoted by the following symbols:

| | | | | | |
|---|---|---|---|---|---|
| Equivalence | $\equiv$ | Negation | $\neg$ | Implication | $\Rightarrow$ |
| Logical OR | $\cup$ | Logical AND will be omitted | | | |

For example, the logical proposition "p implies q" is expressed symbolically as $p \Rightarrow q$, and its negation is expressed as $\neg (p \Rightarrow q)$. Similarly, the proposition "p OR q" is expressed as $p \cup q$, and the proposition "p AND q" is expressed as p q.

## Symbolic Logical Propositions

The validity conditions of table 2 are expressed symbolically, using the following notation to separate the conditions into binary form:

$$V1 \equiv (V \leq V_{MIN\ DRAG})$$

$$VT1 \equiv (V_{TGT} < V_{MIN}) \qquad VT2 \equiv (V_{TGT} > V_{MAX})$$

$$GT1 \equiv (\gamma_{TGT} < \gamma_{SAFE}) \qquad GT6 \equiv (\gamma_{TGT} \geq \gamma_{POT\ MAX})$$

$$Q \equiv (\gamma_{TGT} \leq \gamma_{SPEED\ MIN}) \qquad P \equiv (\gamma_{TGT} \geq \gamma_{SPEED\ MAX})$$

$$TS1 \equiv (\gamma \leq \gamma_{SPEED\ MIN}) \equiv (\gamma_{POT\ REF} = \leq \gamma_{POT\ MIN})$$

$$TS2 \equiv (\gamma \geq \gamma_{SPEED\ MAX}) \equiv (\gamma_{POT\ REF} \geq \gamma_{POT\ MAX})$$

$$TT1 \equiv (\gamma_{POT\ TGT} = \gamma_{POT\ MIN}) \qquad TT2 \equiv (\gamma_{POT\ TGT} = \gamma_{POT\ MAX})$$

$$TT3 \equiv (\gamma_{POT\ TGT} < \gamma_{POT\ MIN}) \qquad TT4 \equiv (\gamma_{POT\ TGT} > \gamma_{POT\ MAX})$$

$$\neg PP \equiv (PRIORITY \equiv SPEED) \qquad PP \equiv (PRIORITY \equiv PATH)$$

$$PF1 \equiv (\gamma_{POT\ MAX} = \gamma_{POT\ MIN})$$

A complete list of symbolic logical propositions can be found in the section "List of Symbols."

## Symbolic Validity Conditions

Making use of these binary propositions, the validity conditions of table 2 can be expressed in symbolic form by table 3.

## Conjunction of Symbolic Validity Conditions

Inspection of table 3 shows that, by conjoining the validity conditions for each of the three primitive modes, these conditions can be summarized compactly as follows:

### γ-V Command mode:

$\neg VT1\ \neg VT2\ \neg GT1$ Safe envelope limits

AND

$\neg TS1\ \neg TS2$ Absence of thrust saturation

**V Command mode:**

| | |
|---|---|
| $\neg$VT1 $\neg$VT2 $\neg$GT1 | Safe envelope limits |
| AND | |
| $\neg$PP | Speed priority |
| AND | |
| PE(V) | Partial effectiveness of V Command mode |

**$\gamma$ Command mode:**

| | |
|---|---|
| $\neg$VT1 $\neg$VT2 $\neg$GT1 | Safe envelope limits |
| AND | |
| PP | Path priority |
| AND | |
| PE($\gamma$) | Partial effectiveness of $\gamma$ Command mode |

The condition that the thrust target be physically realizable is satisfied by the heuristic strategy to be discussed next. For simplicity, it is not included in this summary.

TABLE 3. SYMBOLIC VALIDITY CONDITIONS

| | **$\gamma$-V Command** | **V Command** | **$\gamma$ Command** |
|---|---|---|---|
| $V_{TGT}$ | $\neg$VT1 $\neg$VT2 | | |
| $\gamma_{TGT}$ | $\neg$GT1 | | |
| **PE(V), PE($\gamma$)** | | PE(V) | PE($\gamma$) |
| **$\gamma$ OR $\gamma_{POT\ TGT}$** | $\neg$TS1 $\neg$TS2 | $\neg$TT3 $\neg$TT4 | |
| **PRIORITY** | | $\neg$PP | PP |

**DEFINITIONS:**

PE (V) $\equiv \neg$[P AND ($\gamma_{POT\ TGT} \neq \gamma_{POT\ MAX}$)] AND $\neg$ [Q AND ($\gamma_{POT\ TGT} \neq \gamma_{POT\ MIN}$)]

PE ($\gamma$) $\equiv$ (GT6 V1 TS2)

$\neg$PE ($\gamma$) $\equiv$ (GT6 V1 TS2)

P $\equiv$ ($\gamma_{TGT} \geq \gamma_{SPEED\ MAX}$)     Q $\equiv$ ($\gamma_{TGT} \leq \gamma_{SPEED\ MIN}$)

## Heuristic Design Strategies

As the first step toward partitioning the validity table (table 3) to simplify inversion, a heuristic strategy is developed that ensures that the safety conditions

$$\neg VT1 \equiv (V_{TGT} \geq V_{MIN}) \qquad \neg VT2 \equiv (V_{TGT} \leq V_{MAX}) \qquad \neg GT1 \equiv (\gamma_{TGT} \geq \gamma_{SAFE})$$

are satisfied by screening the target airspeed $V_{TGT}$ and the target flightpath $\gamma_{TGT}$ at the system input.

### Target Screening

According to table 3, the validity conditions ¬VT1, ¬VT2, and ¬GT1 are required for all three primitive modes. The compound condition (¬VT1 ¬VT2 ¬GT1) can be established by screening the flightpath and airspeed targets before initial mode engagement and whenever the safe envelope limits $V_{MIN}$, $V_{MAX}$, or $\gamma_{SAFE}$ are changed. The rationale for target screening is that allowing system operation with invalid targets is nonsensical.

Although continuous screening is necessary in practice to account for changing aircraft performance, the flightpath and airspeed targets must not be changed after initial engagement without pilot consent. Since the first-level Path/Speed Command supermode has no direct interface with the human crew, the need for such target changes must be annunciated to the invoking entity, which must then obtain pilot consent. The screening logic is specified by the following condition-action decision table (appendix D):

Table 4 shows that the safety conditions

$$\neg VT1 \equiv (V_{TGT} \geq V_{MIN}) \qquad \neg VT2 \equiv (V_{TGT} \leq V_{MAX}) \qquad \neg GT1 \equiv (\gamma_{TGT} \geq \gamma_{SAFE})$$

must hold after the screening process is complete. Therefore, the effect of the target screening strategy is the effective removal of those conditions from the validity table, because a condition known to hold can be dropped from the conjunction of conditions (Theorem 13, appendix D).

TABLE 4. TARGET SCREENING STRATEGY FOR PATH/SPEED COMMAND SUPERMODE

| $V_{TGT} < V_{MIN}$ | $V_{TGT} > V_{MAX}$ | $V_{MIN} \leq V_{TGT} \leq V_{MAX}$ | $\gamma_{TGT} < \gamma_{SAFE}$ | $\gamma_{TGT} \geq \gamma_{SAFE}$ |
|---|---|---|---|---|
| Set $V_{TGT}$ equal to $V_{MIN}$ | Set $V_{TGT}$ equal to $V_{MAX}$ | Accept $V_{TGT}$ | Set $\gamma_{TGT}$ equal to $\gamma_{SAFE}$ | Accept $\gamma_{TGT}$ |

### Unsaturated Thrust

After the safety conditions just discussed have been removed from the validity table (table 3) by target screening, it can be seen that in the absence of thrust saturation (that is, if the condition ¬TS1 ¬TS2 holds), the γ-V Command mode is valid. In that case, it should be selected.

The rationale for this mode selection strategy is that it maximizes system effectiveness. The $\gamma$-V Command mode provides complete effectiveness (that is, capture of both the flightpath target and the airspeed target), whereas either of the other two primitive modes could, at best, provide partial effectiveness (that is, capture of one of these targets, but not both). Logically, control of both path and speed dominates control of only one of them.

This heuristic decision to select the $\gamma$-V Command mode when thrust is unsaturated removes the first column from the validity table (table 3), leaving selection of either the V Command mode or the $\gamma$ Command mode as the only remaining possibilities. The same decision also removes the condition ($\gamma_{\text{POT MIN}} < \gamma_{\text{POT TGT}} < \gamma_{\text{POT MAX}}$) for absence of thrust saturation from the second and third columns, leaving only the two possibilities

$$\text{TT1} \equiv (\gamma_{\text{POT TGT}} = \gamma_{\text{POT MIN}}) \quad \text{OR} \quad \text{TT2} \equiv (\gamma_{\text{POT TGT}} = \gamma_{\text{POT MAX}})$$

The definition of PE(V) can now be simplified as follows:

$$\text{PE (V)} \equiv \neg(\text{P TT1}) \neg(\text{Q TT2})$$

The following heuristic strategy for setting the target thrust $\gamma_{\text{POT TGT}}$ determines which of the two possible thrust settings TT1 or TT2 is appropriate for operation in the V Command mode and the $\gamma$ Command mode.

### Target Thrust

When thrust is saturated, the thrust target should be set so as to provide the best physically realizable approximation to the reference thrust $\gamma_{\text{POT REF}}$. Therefore, the target thrust should be set to $\gamma_{\text{POT MIN}}$ when the thrust saturation condition TSI $\equiv$ ($\gamma_{\text{POT REF}} \leq \gamma_{\text{POT MIN}}$) holds, and to $\gamma_{\text{POT MAX}}$ when the condition TS2 $\equiv$ ($\gamma_{\text{POT REF}} \geq \gamma_{\text{POT MAX}}$) holds.

When thrust is unsaturated, the $\gamma$-V Command mode is selected, as already noted. In that case, the fixed target thrust is not used by the automated system, because it is used only when either the V Command mode or the $\gamma$ Command mode is selected. However, it is convenient to set the target thrust as it would be set by the best approximation rule just described, if thrust saturation resulted from capture of the flightpath target $\gamma_{\text{TGT}}$. This strategy avoids unnecessary violation of the partial effectiveness of the V Command mode by updating the target thrust appropriately whenever possible. In the absence of total propulsion failure (that is, when the condition ¬PF1 holds), maximum thrust remains computationally distinct from minimum thrust. The complete strategy for setting target thrust is specified by the following condition-action decision table (table 5):

TABLE 5. STRATEGY FOR SETTING TARGET THRUST

| **PF1** | **¬PF1** | | | | |
|---|---|---|---|---|---|
| | **TS1** | **TS2** | **¬TS1 ¬TS2** | | |
| | | | **Q** | **P** | **¬P ¬Q** |
| Set<br>TT1<br>TRUE | Set<br>TT1<br>TRUE | Set<br>TT2<br>TRUE | Set<br>TT1<br>TRUE | Set<br>TT2<br>TRUE | $\gamma_{\text{POT TGT}}$<br>not<br>updated |

If the condition ¬PFI ¬TS1 ¬TS2 ¬P ¬Q holds, table 5 shows that the thrust target $\gamma_{\text{POT TGT}}$ is not updated. In that case, the thrust target is not used by the flight control system. If the condition ¬PFI ¬TS1 ¬TS2 ¬P ¬Q holds during initialization, the thrust target $\gamma_{\text{POT TGT}}$ is initialized arbitrarily to $\gamma_{\text{POT MIN}}$.

Because it ensures that the physical realizability conditions ¬TT3 ≡ ($\gamma_{\text{POT TGT}} \geq \gamma_{\text{POT MIN}}$) and ¬TT4 ≡ ($\gamma_{\text{POT TGT}} \leq \gamma_{\text{POT MAX}}$) hold, the strategy for setting target thrust effectively removes them from the mode validity table (table 3).

## Partitioning and Inversion of Validity Table

### Partitioning

With the safety conditions (first two rows of table 3) removed by target screening, the first column removed by selection of the γ-V Command mode when thrust is unsaturated, and the conditions ¬TT3 and ¬TT4 (fourth row) removed by the strategy for setting target thrust, the validity table (table 3) has been greatly simplified by this partitioning. The reduced table that remains when thrust is saturated is now as follows:

| | **V Command** | **γ Command** |
|---|---|---|
| **Partial Effectiveness** | PE (V) | PE (γ) |
| **PRIORITY** | ¬PP | PP |

### Inversion

This reduced table can now be inverted by inspection. Only three binary conditions remain: PE (V), PE (γ), and PRIORITY. Therefore, there are $2^3 = 8$ combinations, which can be arranged as an enumerated list with 8 rows. This enumerated list forms the inverted decision table (table 6) that constitutes the solution to the mode selection problem when thrust is saturated; that is, table 6 is a truth table (appendix D) for the reduced validity table just discussed.

For each of these 8 combinations, a unique mode selection policy must be formulated. It can be seen that, in half the cases enumerated in table 6, neither the V Command mode nor the γ Command mode is valid. Nevertheless, in each of these cases, one of these modes must be selected. The next section turns to a discussion of mode selection criteria.

### Mode Selection Criteria

In cases 3, 4, 6, and 8, one of the two modes is valid and the other invalid. Therefore, the mode selection strategy can be based entirely on validity, and consists simply of avoiding the selection of an invalid mode. In cases 1 and 5, the condition ¬PE (V) ¬PE (γ) holds, so that neither the V Command mode nor the γ Command mode is valid (table 3). Because partial effectiveness does not provide a basis for choice, mode selection must be based on priority, selecting the V Command mode when priority is set to SPEED (case 1), and selecting the γ Command mode when priority is set to PATH (case 5). This strategy makes use of control flowing down from higher level, but it leaves cases 2 and 7 unresolved.

TABLE 6. INVERTED VALIDITY TABLE FOR PATH/SPEED COMMAND SUPERMODE

| Condition | | | V Command | γ Command |
|---|---|---|---|---|
| 1. SPEED | ¬PE (V) | ¬PE (γ) | INVALID | INVALID |
| 2. SPEED | ¬PE (V) | PE (γ) | INVALID | INVALID |
| 3. SPEED | PE (V) | ¬PE (γ) | VALID | INVALID |
| 4. SPEED | PE (V) | PE (γ) | VALID | INVALID |
| 5. PATH | ¬PE (V) | ¬PE (γ) | INVALID | INVALID |
| 6. PATH | ¬PE (V) | PE (γ) | INVALID | VALID |
| 7. PATH | PE (V) | ¬PE (γ) | INVALID | INVALID |
| 8. PATH | PE (V) | PE (γ) | INVALID | VALID |

In both cases 2 and 7, the selection could be based on partial effectiveness if the priority were changed; that is, there is logical conflict between priority set at higher level and the relevant information available at the primitive level based on partial effectiveness. In cases 1 and 5, where the condition ¬PE (V) ¬PE (γ) holds and neither mode is valid, it is clear that the dilemma must be resolved at higher level by the invoking entity, because no relevant information is available within the Path/Speed Command supermode. In cases 2 and 7 relevant information is available, but that information is in conflict with the priority specified by the invoking entity. Therefore, because the relevant information available is insufficient to resolve the conflict, the strategy should be the same for cases 2 and 7 as for cases 1 and 5: base mode selection entirely on priority, and annunciate relevant information based on effectiveness violation to the invoking entity for resolution there. This strategy leads to the selection of the V Command mode in case 2, where priority is set to SPEED, and to selection of the γ Command mode in case 7, where the priority is set to PATH. With this system structure, annunciation of the condition ¬PE(V) ¬PE(γ) to the invoking entity constitutes information flowing upward, and the path/speed priority constitutes control flowing downward. The decision determining priority should be made at the lowest level within the control hierarchy at which sufficient information is available.

With the adoption of this policy, which bases mode selection at the primitive level entirely on priority when thrust is saturated, partial effectiveness becomes nonrelevant for the selection strategy, and mode selection based on table 6 when thrust is saturated collapses to the following drastically reduced table:

| Condition | V Command | γ Command |
|---|---|---|
| **¬PP** | TRUE | |
| **PP** | | TRUE |

### Integrated Mode Selection Strategy

When the mode selection strategy for thrust saturation is combined with that for unsaturated thrust discussed previously, the complete mode selection table for the Path/Speed Command supermode is obtained (table 7):

TABLE 7. MODE SELECTION WITHIN PATH/SPEED COMMAND SUPERMODE

<table>
<tr><th>Condition</th><th>γ-V Command</th><th>V Command</th><th>γ Command</th></tr>
<tr><td>¬TS1 ¬TS2</td><td>TRUE</td><td></td><td></td></tr>
<tr><td>TS1 ∪ TS2</td><td></td><td colspan="2">TRUE</td></tr>
<tr><td>¬PP</td><td></td><td>TRUE</td><td></td></tr>
<tr><td>PP</td><td></td><td></td><td>TRUE</td></tr>
<tr><td colspan="4">Definitions</td></tr>
<tr><td colspan="2">¬TS1 ¬TS2: Thrust unsaturated</td><td colspan="2">TS1 ∪ TS2: Thrust saturated</td></tr>
<tr><td colspan="2">¬PP: SPEED priority</td><td colspan="2">PP: PATH priority</td></tr>
</table>

In this *condition-action mode selection table* (appendix D), each column of the table corresponds to a single mode, and each row corresponds to a logical condition. The conditions that must hold for selection of any mode are designated by the entry TRUE in the body of the table in the column corresponding to that mode. For example, for selection of the V Command mode, the conditions ¬PP and TS1 ∪ TS2 must hold, as indicated by the entry TRUE in the second and third rows of the second column. Blank spaces in the body of the table are not relevant to mode selection (appendix D).

The mode selection strategy specified by table 7 constitutes the solution of the synthesis problem for the Path/Speed Command supermode. An alternative form of specification for this mode selection strategy can be obtained by constructing the equivalent statechart (appendix E), which is discussed next.

## Construction of Statechart

### Mode Structure

To enable the state chart for mode selection to reflect the same partitioning that is evident in the decision table (table 7), it is convenient to combine the V Command mode and the γ Command mode to form a supermode that is termed γ OR V Command. The statechart for mode selection illustrated by figure 15 can then be constructed directly from table 7, as explained next. A simplified discussion of statecharts can be found in appendix E.

### Statechart for Mode Selection

Because the γ-V Command mode is selected according to table 7 whenever the condition ¬TS 1 ¬TS2 holds, whatever the previous state, a transition arrow labeled with the condition

¬TS1 ¬TS2 must lead to the γ-V Command mode from the γ OR V Command supermode, and a second arrow labeled ¬TS1 ¬TS2 must lead to the γ-V Command mode from the initial state outside the Path/Speed Command supermode boundary (fig. 15). Similarly, a transition arrow labeled with the condition TS1 ∪ TS2 must lead to the γ OR V Command supermode from the γ-V Command mode. Within the γ OR V Command supermode, a transition arrow labeled with the condition PP must lead from the V Command mode to the γ Command mode, an arrow labeled with the condition ¬PP must lead from the γ Command mode to the V Command mode, and similarly for initialization.

## Statechart for Setting Target Thrust

The strategy for setting target thrust (table 5) can be represented as a concurrent state machine within the Path/Speed Command supermode (fig. 15) by following a similar construction method. It can be seen that the case in which the condition ¬PF1 ¬TS1 ¬TS2 ¬P ¬Q holds, for which the target thrust is not updated (table 5), is represented in the statechart by transition arrows that return to the previous state.

## Other Statecharts

The diagram of figure 15 also shows that the strategy for target screening and for evaluation of the conditions ¬PE (V), ¬PE (γ), and NE can also be represented as concurrent state machines within the Path/Speed Command supermode. Their internal details, which are omitted from figure 15, are specified for target screening by table 4, and for effectiveness evaluation by the definitions of ¬PE (V), ¬PE (y), and NE discussed previously.

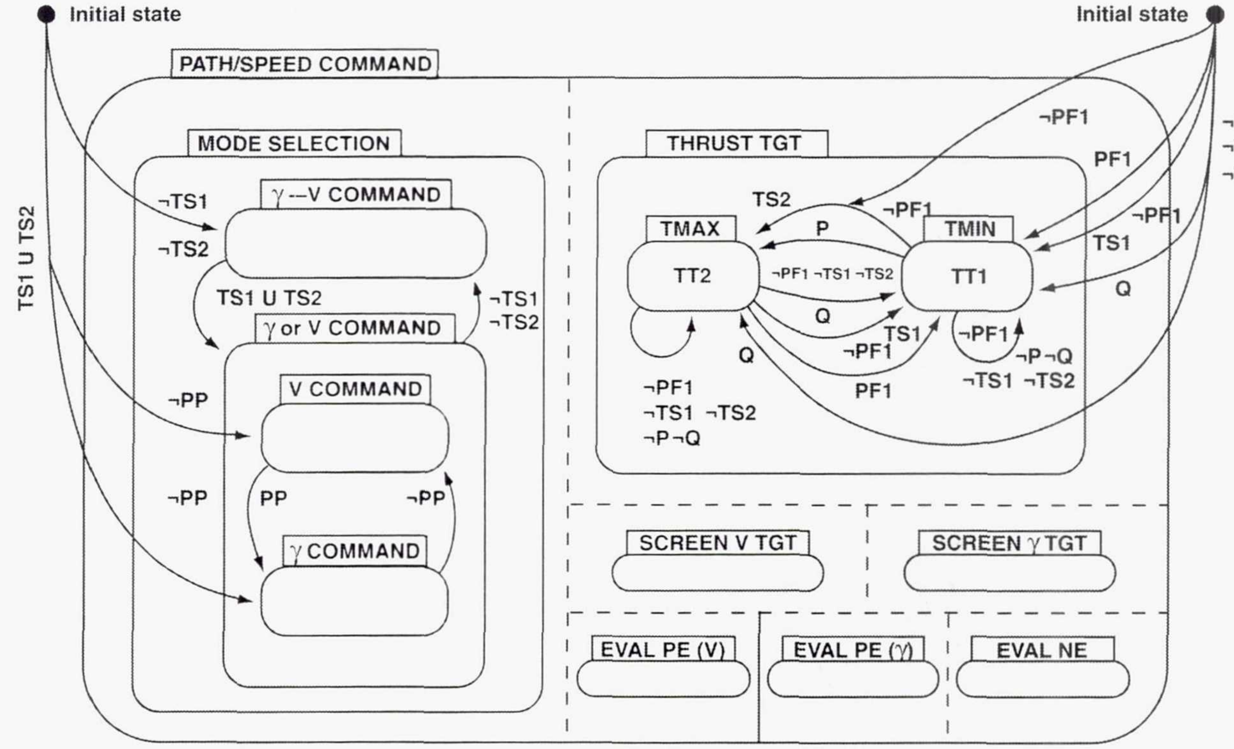


Figure 15. Statechart for Path/Speed Command supermode.

### Statechart for Path/Speed Command Supermode

Figure 15 shows that the statecharts for all the concurrent state machines just described are combined to form the statechart for the Path/Speed Command supermode. The resulting state machine could, of course, also be specified in final form by table 7. Indeed, such decision tables may be preferred to their equivalent statecharts for some purposes, such as providing a reference document that can conveniently be reproduced.

However, the statechart of figure 15 provides a compact representation that is graphically suggestive of the dynamical function of the state machine, and may be preferred to equivalent decision tables for tutorial purposes after the interpretation of statecharts becomes familiar. Statecharts also facilitate verification of the essential logical properties of consistency and completeness.

### Verification of Consistency and Completeness

Mode selection would be *logically inconsistent* if two or more transition arrows leading away from any mode were labeled with the same condition. In that case, mode selection would be ambiguous and would fail to be mutually exclusive, because two or more modes could be selected at the same time.

Mode selection would be *logically incomplete* if any combination of conditions could occur for which no transition is specified. In the statechart diagrams, clutter can be reduced by adopting the convention that, in the absence of any defined transition away from each mode, the system remains in that mode by default. If this convention is adopted, completeness can no longer be assessed directly by inspection of the diagram, and must therefore be verified independently.

For a decision table, consistency would be violated if any two columns of the table contained the same pattern of truth conditions, and completeness would be violated if the disjunction of truth conditions involving each tested quantity failed to hold tautologically (appendix D). For example, in table 7 there are two conditions involving TS1 and TS2. Because the condition TS1 $\cup$ TS2 is logically equivalent to the negation of the condition $\neg$TS1 $\neg$TS2 (Theorem 11, appendix D), their junction (TS1 $\cup$ TS2) $\cup$ ($\neg$TS1 $\neg$TS2) is tautologically true (Theorem 4, appendix D). Therefore, table 7 is logically complete with respect to the conditions TS1 and TS2. Evaluation of complex decision tables for consistency and completeness is accomplished in practice by means of utility programs developed for that purpose.

### Implementation

A more subtle point regarding statechart representation arises during implementation. Conceptually, the concurrency specified by the statechart of figure 15 could be realized by 7 processors (one for each concurrent machine) operating in parallel, each executing a re-entrant program that would repeat continuously. Such a system cannot be implemented exactly by a single sequential processor, because such a processor can execute only one instruction at a time. Therefore, certain temporal approximations must be made during implementation for sequential machines like those installed in current transport aircraft. In consequence, the specification provided by the statechart of figure 15 must be regarded as ambiguous or incomplete for such implementations.

These issues are discussed in detail in appendix E, in which certain changes to standard statechart semantics are proposed that enable exact implementation within a sequential processor, eliminating the need for introduction of temporal approximations during implementation. Such temporal approximations could be a source of design error that could void the behavioral guarantees on which formal validation depends. The same difficulties would, of course, be encountered during implementation from equivalent decision tables. However, based on experience with implementation of similar systems for state estimation and flight control in experimental aircraft for which timing could be critical, the authors are of the opinion that statecharts offer unique clarity for understanding and solving such problems.

## System Properties for Path/Speed Command Supermode

### Geometrical Regions for Mode Selection

For the system specified by the statechart of figure 15, selection of the γ-V Command mode or the γ OR V Command supermode is determined entirely by the thrust saturation conditions

$$\mathrm{TS1} \equiv (\gamma_{\text{POT REF}} \le \gamma_{\text{POT MIN}}) \qquad \mathrm{TS2} \equiv (\gamma_{\text{POT REF}} \ge \gamma_{\text{POT MAX}})$$

which, as already noted, are equivalent to the conditions

$$\mathrm{TS1} \equiv (\gamma \le \gamma_{\text{SPEED MIN}}) \qquad \mathrm{TS2} \equiv (\gamma \ge \gamma_{\text{SPEED MAX}})$$

These conditions are physical, and correspond to well-defined geometrical regions within the (V, γ) plane, as discussed previously (fig. 12). Therefore, each region can be identified by the corresponding mode, as illustrated by figure 16. It can be seen that the γ-V Command mode is selected when the (V, γ) operating point lies within the central region within which thrust saturation is absent, and the γ OR V Command supermode is selected elsewhere. (Within the γ OR V Command supermode, the V Command mode is selected when priority is set to SPEED, and the γ Command mode is selected when priority is set to PATH; these selections are not illustrated by figure 16.)

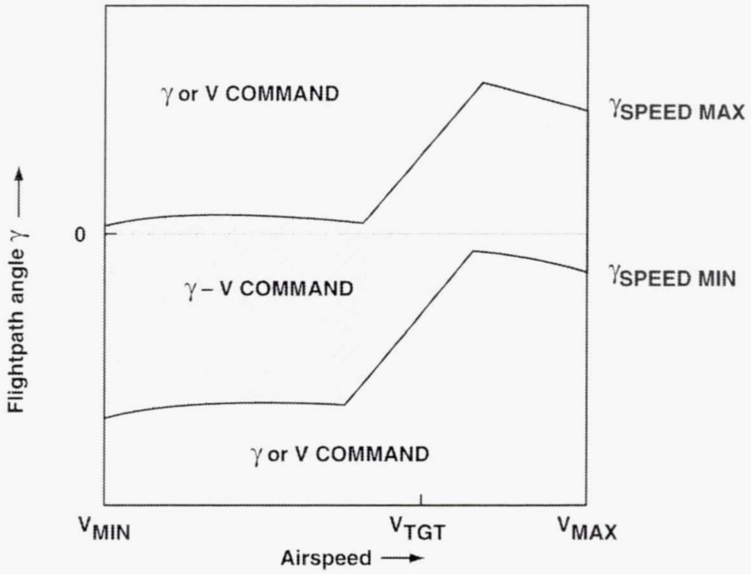


Figure 16. Geometric regions for mode selection within Path/Speed Command supermode.

### Safety Properties

During operation in the Path/Speed Command supermode, the safety properties discussed previously are enforced by screening the targets at the system input, as specified by table 4, and by the envelope protection modes previously described. Therefore, the system behavior is constrained to respect the safety envelope boundaries illustrated by figure 13. Invocation of any envelope protection mode, which should occur only under abnormal conditions, is required to be annunciated by a suitable warning.

### Effectiveness Properties

During operation in the Path/Speed Command supermode, the effectiveness of system behavior is summarized as follows:

**$\gamma$-V Command mode**– During operation in the $\gamma$-V Command mode, capture of both the flightpath target and the airspeed target are assured by the stability properties of the path and speed regulators within the flight control system (figs. 11(c) and 11(e)). Therefore, as defined previously, the $\gamma$-V Command mode is completely effective.

**V Command mode**– During operation in the V Command mode, capture of the airspeed target $V_{TGT}$ is assured by the stability property of the speed regulator within the flight control system (fig. 11(c)).

The flightpath angle $\gamma_{SPEED}$ is captured, which provides an acceptable approximation to the flightpath target $\gamma_{TGT}$ provided that partial effectiveness (that is, PE (V)) holds for the V Command mode.

**$\gamma$ Command mode**– During operation in the $\gamma$ Command mode, capture of the flightpath target $\gamma_{TGT}$ is assured by the stability property of the path regulator within the flight control system (fig. 11(e)).

The airspeed target $V_{TGT}$ is not captured, but an acceptable point of stable airspeed equilibrium is reached provided that partial effectiveness (PE ($\gamma$)) holds for the $\gamma$ Command mode.

Longitudinal acceleration has the correct sign leading to capture of the airspeed target $V_{TGT}$ provided that the normal effectiveness properties NE1 and NE2 hold. Violation of NE1 results in the annunciation "More drag," and violation of NE2 results in the annunciation "More thrust."

**Required annunciations**– All effectiveness violations must be annunciated to the human crew, and also to the invoking entity. Assurance that all such violations will be identified and annunciated as required is provided by logical completeness based on the geometrical regions discussed previously.

## Concluding Remarks

The simplicity of the final mode selection strategy for the Path/Speed Command supermode, which is specified by table 7, stands in sharp contrast to the complexity of the table that would result from brute-force inversion of table 3, which as noted previously would contain 1024 rows.

It should be noted that the Path/Speed Command supermode does not constitute a complete automaton, because the essential path/speed priority must be determined by the invoking entity (that is, a second-level supermode), accounting for all relevant information available at higher level. The means by which that determination is accomplished is discussed in the next section, which describes a supermode on the second level of the mode hierarchy.

## SYNTHESIS OF ALTITUDE COMMAND SUPERMODE

In order to extend the synthesis method to higher level and to develop a system that constitutes a complete automaton, this section discusses the synthesis of a second-level supermode to be termed Altitude Command. Its development makes use of the same methodology employed for synthesis of the Path/Speed Command supermode already described (see the section "Overview of Design Method," steps 1–15). The Path/Speed Command supermode, which lies on the lowest level of the three-level mode hierarchy, accepts flightpath angle and airspeed targets, as explained previously. The Altitude Command supermode, which lies on the second level of the mode hierarchy, makes use of the first-level Path/Speed Command supermode as an internal element, but generalizes its function to enable the system to accept altitude and airspeed targets. These targets can be entered manually by the crew, providing operation functionally similar to that of the Flight Level Change supermode in the autoflight systems of current transport aircraft.

Alternatively, the Altitude Command supermode can be invoked, and its altitude and airspeed targets can be specified, by a supermode on the third level of the mode hierarchy. By this means, operation similar to that of the vertical planning and guidance modes in current aircraft can be provided, while avoiding functional duplication.

In general, it must be expected that moving upward to a higher level in the mode hierarchy will involve consideration of new supermodes, each of which constitutes a hybrid system whose continuous and discrete elements must both be specified. This is the case for the second-level Altitude Command supermode, just as it is for the first-level Path/Speed Command supermode discussed previously. For the Altitude Command supermode, there is only one continuous element: the function of altitude regulation, which is described next.

### Altitude Regulation

Because the aircraft altitude is governed by a differential equation, the aircraft model described previously (eqs. (1), (2), and (4)) must be augmented by the altitude equation

$$DH/dt = V \sin \gamma$$

where the symbol H denotes height above mean sea level. In the flight control system illustrated by figure 11, the height regulator is shown on the left of the broken line in figure 11(a) that separates the Altitude Command supermode (outer-loop control) from the Path/Speed Command supermode (inner-loop control). Its details are illustrated by figure 11(b).

Figure 11(b) shows that the height regulator combines two different height regulation laws. For small height errors of 300 ft or less, the commanded vertical velocity $(dH/dt)_{CMD}$ is linearly proportional to the height error $\Delta H$, providing exponential convergence to the target altitude (lower row of diagram, figure 11(b)). For larger height errors, a parabolic law (upper row of figure 11(b)) is used to obtain constant normal acceleration (approximately 0.lg) during the altitude capture maneuver, avoiding excessive normal accelerations that would be commanded if the exponential law were used for large height errors.

Because both the vertical velocity and the normal acceleration commanded by the parabolic law are matched to those commanded by the exponential law at the point of transition, the commanded vertical velocity $(dH/dt)_{CMD}$ is specified uniquely and continuously as a function of altitude error. It is then divided by inertial speed to obtain the quantity sin $\gamma_{TGT}$, which is limited to ensure realizability before inverting to obtain the flightpath target $\gamma_{TGT}$ (fig. 11(b)).

## Mode Structure

In order to relate the function of the Altitude Command supermode directly to the pilot's task during operation within the National Airspace System, four mode elements termed Climb, Descend, Altitude Capture, and Altitude Hold are specified as internal elements of the Altitude Command supermode. Each of these four modes contains the Path/Speed Command supermode as an internal element. It may be recalled from the previous discussion that the Path/Speed Command supermode contains the primitive modes $\gamma$-V Command, V Command, and $\gamma$ Command as internal elements. The Altitude Capture, Altitude Hold, Climb, and Descend modes are related to each other and to the height regulator law in the following way.

### Altitude Capture and Altitude Hold Modes

For altitude errors smaller than 300 ft, the height regulator makes use of the exponential law, as already described. This exponential law corresponds to the Altitude Hold mode. For larger altitude errors, the height regulator uses the parabolic law, which corresponds to the Altitude Capture mode.

In either the Altitude Capture or the Altitude Hold mode, priority is set to PATH. When the Path/Speed Command supermode is invoked and the PATH priority is passed to it, the primitive $\gamma$-V Command mode is selected if thrust is unsaturated, as already explained. In that case, both the path and speed targets are captured. If thrust saturates, with PATH priority the primitive $\gamma$ Command mode is selected. In that case, only the path target is captured. Therefore, if thrust saturates, the Altitude Capture mode and the Altitude Hold mode both capture the target flightpath $\gamma_{TGT}$ at the expense of the the speed target. Capture of the altitude target $H_{TGT}$ is then assured by the stability property of the height regulator (fig. 11(b)).

### Climb and Descend Modes

For altitude errors so large in magnitude that the target flightpath angle $\gamma_{TGT}$ commanded by the height regulator (fig. 11(b)) lies outside the aircraft performance envelope (fig. 17), this commanded flightpath target should be approximated by selecting the nearest point on the boundary of the performance envelope at the target airspeed as the best physically realizable approximation for steady climb or descent. If the altitude error $\Delta H$ is positive (that is, if the target altitude lies above the prevailing altitude), such operation corresponds to the Climb mode. If the altitude error $\Delta H$ is negative, such operation corresponds to the Descend mode. For example, if the commanded flightpath target $\gamma_{TGT}$ lies above the upper boundary of the performance envelope (point A of figure 17), then point B on the maximum-thrust contour of figure 17 at the target airspeed $V_{TGT}$ should be selected for operation in the Climb mode.

It should be noted that the angular difference $\Delta\gamma$ between $\gamma_{TGT}$ and $\gamma_{POT\ MAX}$ (fig. 17) is limited to a maximum of 3 degrees by the limit imposed on $\gamma_{TGT}$ within the height regulator (fig. 11(b)). This angular increment is sufficient to ensure that maximum thrust is commanded during the climb, even if moderate errors should enter the onboard performance calculation that estimates $\gamma_{POT\ MAX}$. A similar angular difference between $\gamma_{TGT}$ and $\gamma_{POT\ MIN}$ ensures that idle thrust is commanded during descent. These limits imposed on $\gamma_{TGT}$ also serve to confine the corresponding symbol within the field of view of the primary flight display (fig. 11(a)).

In the Climb mode, priority is set to SPEED. When the Path/Speed Command supermode is invoked and the flightpath and speed targets and the SPEED priority are passed to it, the primitive γ-V Command mode is selected if thrust is unsaturated. If the thrust saturates, with SPEED priority the primitive V Command mode is selected, with target thrust set at maximum. In that case, the flightpath angle $\gamma_{SPEED\ MAX}$ is captured, leading to capture of the speed target $V_{TGT}$. Steady climb at point B (fig. 17) then results.

If the commanded flightpath angle $\gamma_{TGT}$ lies below the lower boundary of the performance envelope (point C of figure 17), operation in the Descend mode is obtained in a similar way. Priority is set to SPEED. Invoking the Path/Speed Command supermode and passing to it the flightpath and speed targets and the SPEED priority results in steady descent at point D (fig. 17).

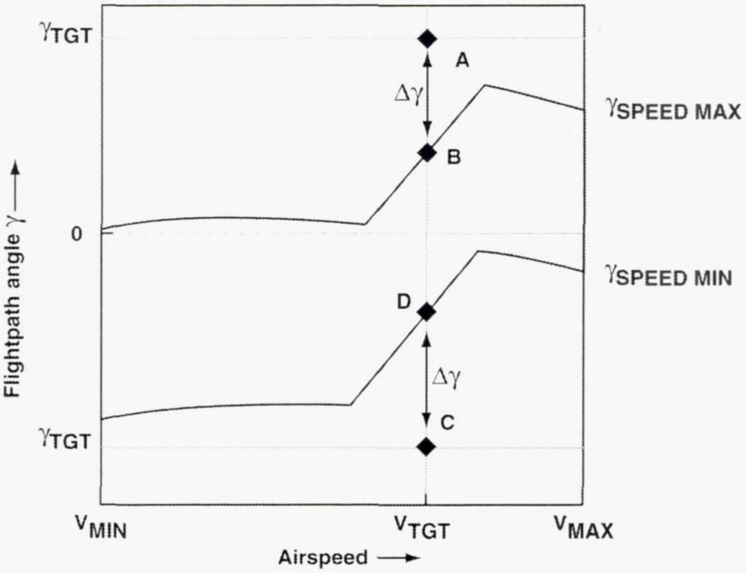


Figure 17. Approximation of flightpath target.

### Path/Speed Priority

It is clear from this discussion that the Climb mode and the Descend mode both require the path/speed priority to be set to SPEED. In contrast, the Altitude Capture mode and the Altitude Hold mode both require that the priority be set to PATH to enable capture of the target altitude. Setting priority is an action that must be associated with mode selection on the second level of the mode hierarchy.

## Functional Specification

Accounting for the mode structure just discussed, the functional specification for the Altitude Command supermode can be summarized as follows. The control flow and information flow within the supermode hierarchy, which are similar for all the second-level supermodes, are illustrated by the block diagram of figure 18. The discussion that follows relates only to the Altitude Command supermode; the other supermodes illustrated by figure 18 are discussed later, and should be ignored for the present.

### Inputs and Outputs

The Altitude Command supermode can be invoked by the human crew via the mode control panel during manual operation (fig. 18), or, alternatively, by a supermode on the third level of the mode hierarchy (not illustrated by figure 18). The Altitude Command supermode accepts from the invoking entity the following two inputs: the altitude target $H_{TGT}$ and the airspeed target $V_{TGT}$. The quantities $V_{MIN}$, $V_{MAX}$, $H_{SAFE}$, and $H_{MAX}$ must also be specified for use in screening the altitude and airspeed targets (fig. 18). The Altitude Command supermode must determine the flightpath target $\gamma_{TGT}$ from the height regulator law (fig. 11(b)). In addition, the Altitude Command supermode must

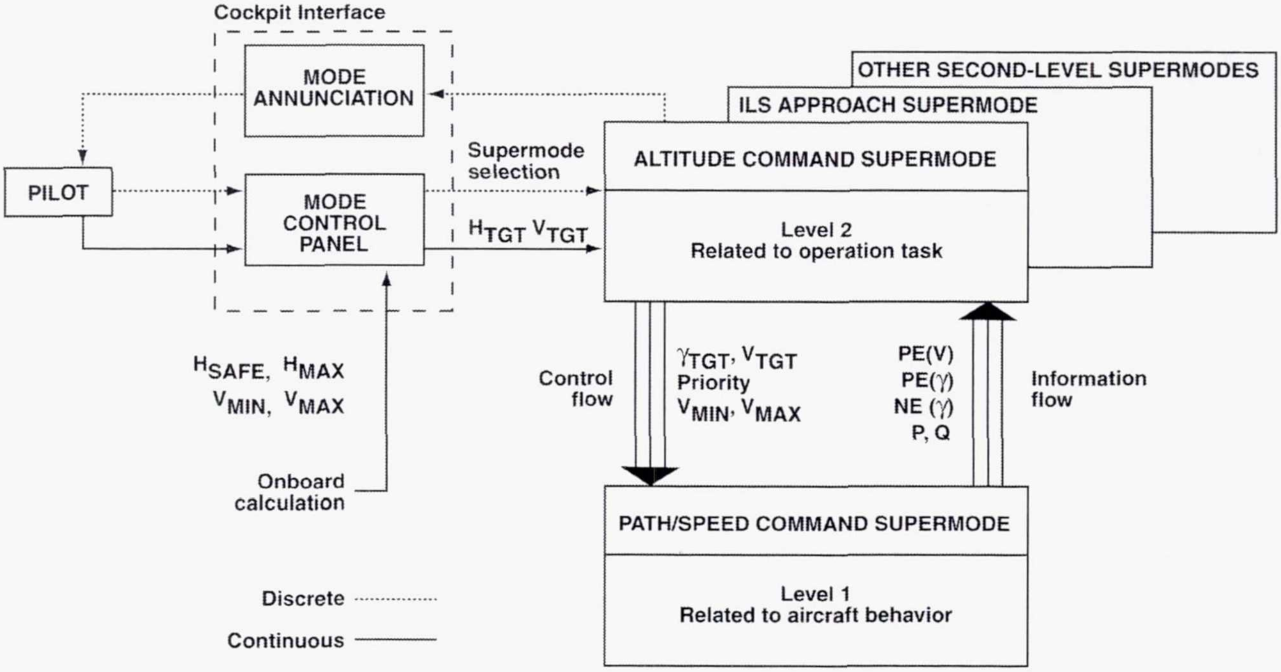


Figure 18. Distributed decision-making: control flow and information flow within supermode hierarchy.

determine the path/speed priority as a Boolean control variable. The Altitude Command supermode invokes the first-level Path/Speed Command supermode, and passes to it the following three outputs: the flightpath target $\gamma_{TGT}$, the airspeed target $V_{TGT}$, and the Boolean path/speed priority (fig. 18).

The Path/Speed Command supermode accepts from the invoking entity (that is, from the Altitude Command supermode) the following three inputs: the flightpath target $\gamma_{TGT}$, the airspeed target $V_{TGT}$, and the Boolean path/speed priority. Based on thrust saturation and on the Boolean path/speed priority, the Path/Speed Command supermode selects one of the three primitive modes γ-V Command, V Command, or γ Command. The selected mode determines the elevator and throttle servo commands from the continuous elements of the flight control system (fig. 11).

In addition, the Path/Speed Command supermode evaluates the partial effectiveness of the γ Command mode and of V Command mode, and the normal effectiveness of the γ Command mode, as previously defined, and returns them to the invoking entity (fig. 18). The Path/Speed Command supermode must also evaluate and return the two conditions P and Q (fig. 18) that will be used by the Altitude Command supermode for mode selection. These Boolean variables were previously defined (eq. (11i) as follows:

$$P \equiv (\gamma_{TGT} \geq \gamma_{SPEED\ MAX}) \qquad Q \equiv (\gamma_{TGT} \leq \gamma_{SPEED\ MIN})$$

### Annunciation requirements

The mode selection status of the primitive γ-V Command mode, γ Command mode, and γ Command mode are also annunciated to the crew, together with the caution and warning messages generated within the Path/Speed Command supermode (fig. 18). All other lower-level details, such as thrust saturation, regulator errors, and the like, are hidden within the Path/Speed Command supermode.

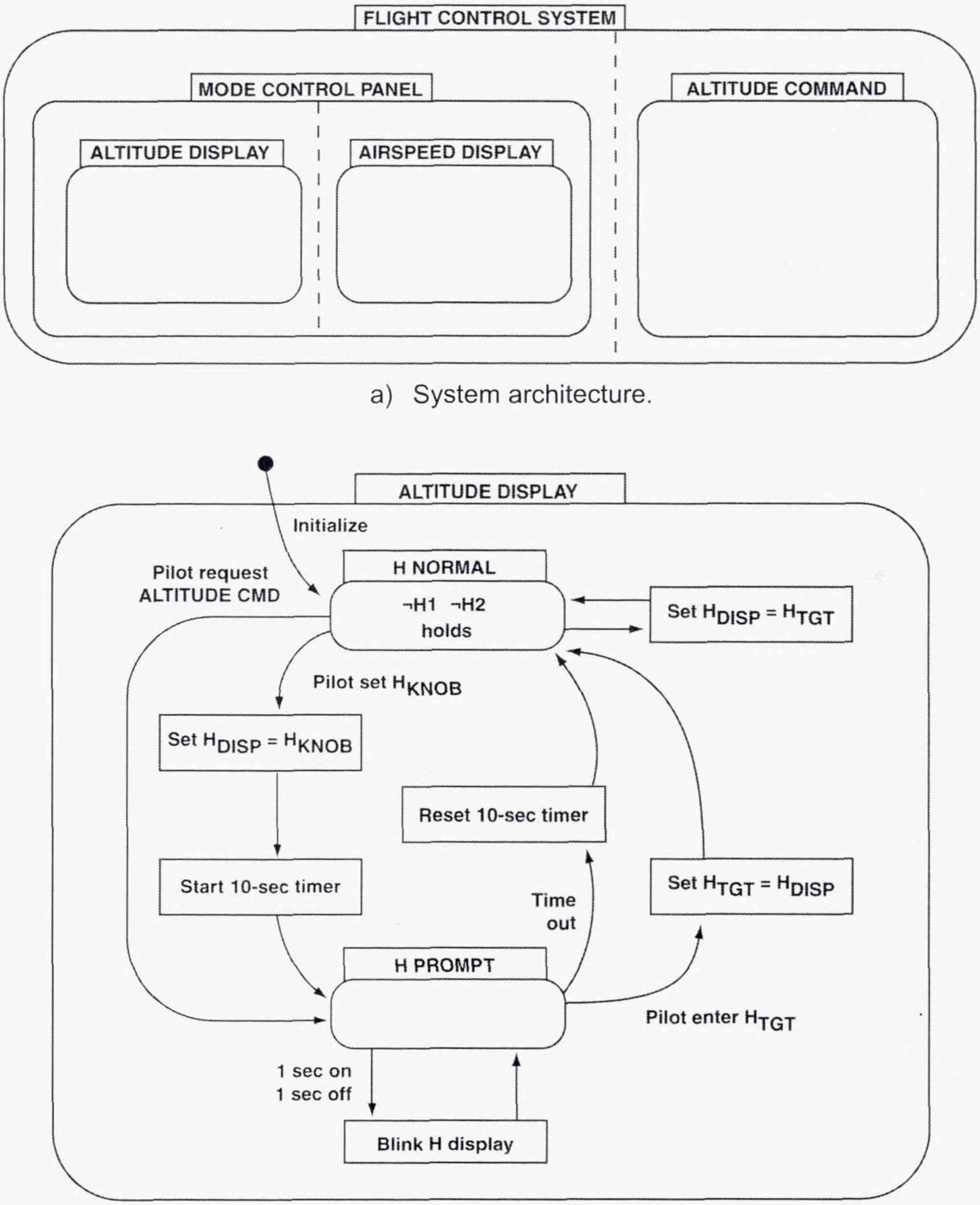
## Mode Control Panel

The Altitude Command supermode is selected manually by means of the Mode Control Panel. A representative implementation for the mode control panel is described briefly in this section, omitting details for clarity. Flight control system architecture relating the Mode Control Panel to the Altitude Command supermode is illustrated in statechart form by figure 19(a), and representative Mode Control Panel functions are illustrated by the statecharts of figures 19(b) and 19(c) and the block diagram of figure 19(d).

This description of the Mode Control Panel is presented in order to illustrate a representative interface with the human crew. Because the focus of the present work is on the Altitude Command supermode, no claims are made for formal validation of the Mode Control Panel functions illustrated by figures 19(b), 19(c), and 19(d). Features of the Mode Control Panel termed display blinking and time-out are described briefly in this section; their purpose is to provide protection against certain human errors to be discussed in detail later.

## Mode Selection

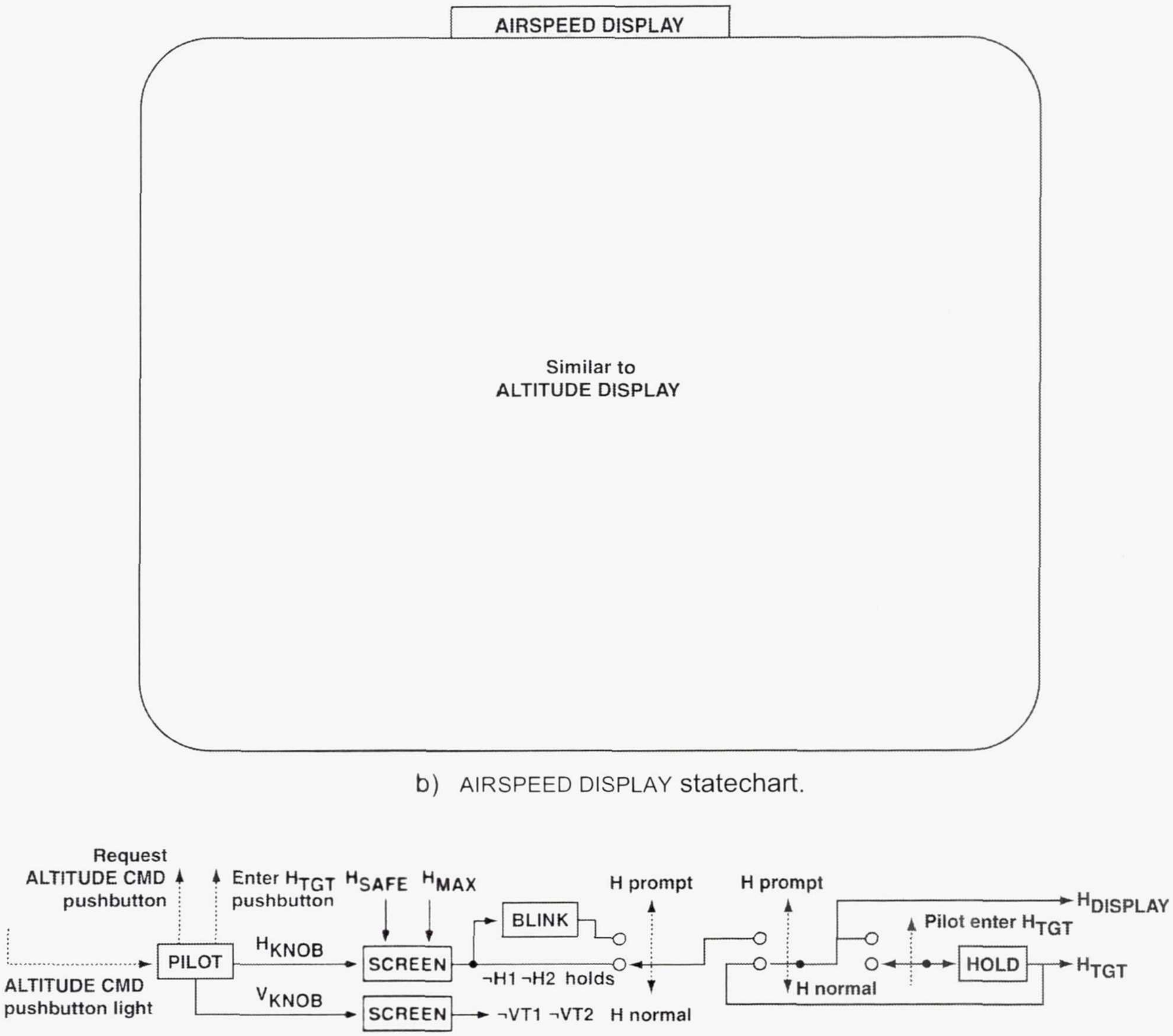
It is assumed that the Altitude Command supermode is one of several second-level supermodes available for manual selection (fig. 18). Supermode selection must be mutually exclusive. (Modes internal to the Altitude Command supermode such as the Climb and Descend modes cannot be selected manually.) The ALTITUDE DISPLAY statechart (fig. 19(b)) and block diagram (fig. 19(d)),



a) System architecture.

b) ALTITUDE DISPLAY statechart.

Figure 19. Mode control panel.



b) AIRSPEED DISPLAY statechart.

d) Block diagram for ALTITUDE DISPLAY (AIRSPEED DISPLAY similar).

Figure 19. Mode control panel (concluded).

show that the crew requests the Altitude Command supermode by pushing a labeled button, which is then magnetically latched. The crew is prompted for entry of altitude and speed targets, which must be completed before the system engages the requested supermode. Prior to engagement, the system continues to operate in the previously engaged supermode, or in manual control. Supermode engagement status is indicated by illumination of its pushbutton (fig. 19(d)). The Altitude Command supermode is de-selected by pushing the illuminated button.

## Altitude and Airspeed Targets

The mode control panel displays the altitude and airspeed targets in display windows. The pilot adjusts the displayed altitude $H_{DISPLAY}$ by turning an adjacent knob (figs. 19(b) and 19(d)). After the altitude is set as desired, it is entered into the system by pressing a button adjacent to the altitude display window, which causes the internally stored altitude target $H_{TGT}$ to be set equal to $H_{DISPLAY}$. If the pilot resets the altitude knob, so that $H_{DISPLAY}$ is no longer equal to $H_{TGT}$, the window blinks to prompt the pilot to enter the new altitude target. If entry is not completed within 10 sec, $H_{DISPLAY}$ again displays the previous $H_{TGT}$. Airspeed targets are entered in a similar manner (figs. 19(c) and 19(d)).

### Mode Annunciation

The engagement status of modes internal to the Altitude Command supermode is annunciated to the crew by a mode display panel. The engagement status of the primitive modes within the Path/Speed Command supermode is also indicated by suitable symbols within the primary flight displays, enabling the crew to anticipate mode transitions as two symbols approach coincidence. This symbology is discussed in detail in the section "Design Recommendations."

## Safety Requirements

In addition to the safety requirements on $V_{TGT}$ and $\gamma_{TGT}$ previously imposed on the Path/Speed Command supermode, further safety requirements on target altitude and airspeed must be added for the Altitude Command supermode, as follows.

### Minimum Altitude

The minimum safe altitude is determined by requirements for safe terrain clearance. This minimum safe altitude $H_{SAFE}$ should be determined from the National Airspace System database as a function of the geographic position of the aircraft. In order to make the Altitude Command supermode suitable for use with area navigation, the restriction $H_{SAFE} \geq 1500$ ft above ground level must be imposed for safety. Operation below 1500 ft above ground level requires more precise navigation, which would be obtained from an approach navigation aid such as an Instrument Landing System (ILS) during operation in a suitable approach mode. Fully automated operation in the Altitude Command supermode can be permitted at minimum altitudes $H_{MIN}$ as low as 400 ft above ground level after takeoff, provided that the target altitude meets the requirement $H_{TGT} \geq H_{SAFE}$.

### Maximum Altitude

The maximum altitude is determined by the reserve thrust required at the target airspeed to avoid small losses of airspeed during coordinated turns, which result from the increase of induced drag (appendix B) owing to the increased lift required in coordinated turning flight (appendix C). Calculations for the generic transport aircraft described in appendix B show that, after accounting for the thrust needed for a coordinated turn with approximately 20 degrees of bank, the thrust required is near that available at the optimal cruising altitude. Therefore, the maximum altitude $H_{MAX}$ is closely approximated by the optimal cruising altitude, which ranges from 28,000 ft to 43,000 ft depending on aircraft wing loading (table B-3, appendix B). The exact value of $H_{MAX}$ under prevailing conditions must be determined from the performance database of the aircraft. The required performance at $H_{MAX}$ is discussed in more detail later.

### Minimum Airspeed

In order to mitigate the adverse safety consequences of invocation of underspeed protection, the minimum airspeed $V_{MIN}$ should be set equal to the minimum-drag airspeed $V_{MIN\ DRAG}$. This mitigation strategy resolves a conflict between safety and effectiveness to be discussed after defining effectiveness for the Altitude Command supermode. In practice, the performance envelope diagram for the generic transport discussed in appendix B (fig. B-7, appendix B) shows negligible performance degradation for airspeeds as low as 20 kt EAS below $V_{MIN\ DRAG}$, owing to the characteristic increase of thrust with decreasing airspeed (fig. B-4). Therefore, the minimum airspeed $V_{MIN}$ for the Altitude Command supermode can be set 20 kt EAS below the minimum-drag speed $V_{MIN\ DRAG}$ (or, in the takeoff configuration with wing flaps extended, 5 kt EAS below the takeoff safety speed $V_2$) without significant performance penalty.

### Maximum Airspeed

The maximum airspeed $V_{MAX}$ is limited by structural considerations at low altitude, and by compressibility (Mach) effects at high altitude (appendix B). For a fixed maximum operating Mach number, the corresponding equivalent airspeed $V_E$ decreases with altitude until the maximum altitude $H_{MAX}$ is reached (eq. (B-7m), appendix B).

### Safety Envelope

The restrictions on altitude and airspeed just discussed combine to limit operation in the Altitude Command supermode to the safety envelope illustrated by figure 20. The limits on prevailing airspeed are enforced by the underspeed and overspeed protection modes discussed previously during development of the Path/Speed Command supermode. The limits $H_{MIN}$ and $H_{SAFE}$ on prevailing altitude can be enforced by the descent path protection mode discussed previously, by making the safe descent flightpath limit $\gamma_{SAFE}$ depend appropriately on $H_{MIN}$ and $H_{SAFE}$ and on inertial speed. Therefore, the limits on prevailing altitude and prevailing airspeed can be enforced by the envelope protection modes already discussed, and impose no new safety requirement on the Altitude Command supermode. On the other hand, the limits on target altitude and target airspeed must be enforced by the target screening algorithms that are included within the Altitude Command supermode, which will be discussed in detail later.

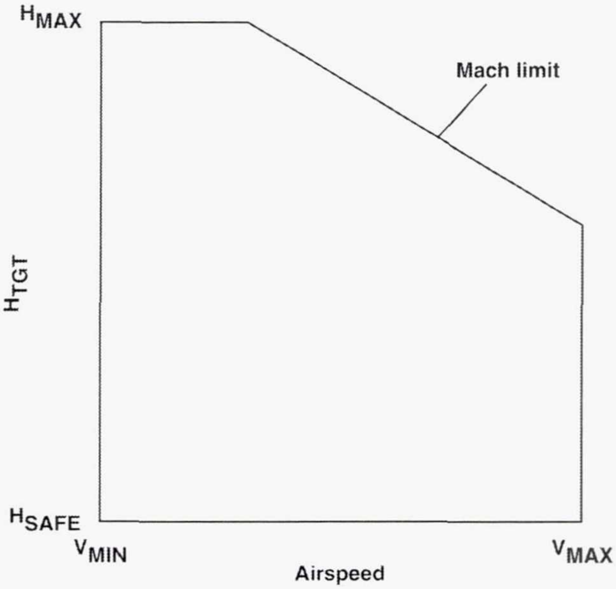


Figure 20. Safety envelope for Altitude Command supermode.

## Effectiveness Requirements

### Effectiveness definitions

The effectiveness properties defined previously for the Path/Speed Command supermode can be generalized as follows for the Altitude Command supermode and its elements, the Altitude Capture, Altitude Hold, Climb, and Descend modes. Complete effectiveness is defined to hold if and only if both the altitude target $H_{TGT}$ and the airspeed target $V_{TGT}$ are captured.

Partial effectiveness is defined to hold for the Altitude Capture and the Altitude Hold modes if and only if the altitude target $H_{TGT}$ is captured, and a point of stable airspeed equilibrium is reached that constitutes an acceptable approximation to $V_{TGT}$. Partial effectiveness is defined to hold for the Climb and Descend modes if and only if the speed target $V_{TGT}$ is captured, and a point of stable path equilibrium is reached that constitutes an acceptable approximation to $\gamma_{TGT}$.

Normal effectiveness is defined to hold if the vertical velocity dH/dt and the normal acceleration both have the correct sign for capture of the altitude target $H_{TGT}$, and the longitudinal acceleration dV/dt has the correct sign for capture of the airspeed target $V_{TGT}$. In practice, the magnitude of the vertical velocity should be required to exceed a nominal threshold value during climb and descent. These stronger conditions are termed climb and descent effectiveness. The property of normal effectiveness is closely related to the smoothness of the altitude and speed capture maneuvers, which are discussed later.

### Altitude Command Supermode

If complete effectiveness holds for the Path/Speed Command supermode as previously defined, which ensures that both the flightpath target $\gamma_{TGT}$ and the airspeed target $V_{TGT}$ are captured, complete effectiveness for the Altitude Command supermode then follows from the stability property of the height regulator. Conditions sufficient to ensure complete effectiveness are discussed later.

### Altitude Capture and Altitude Hold Modes

Because the Altitude Capture and Altitude Hold modes both use the primitive $\gamma$ Command mode during thrust saturation, as already noted, partial effectiveness of the $\gamma$ Command mode is required to ensure partial effectiveness of both the Altitude Capture mode and the Altitude Hold mode. If partial effectiveness holds for the $\gamma$ Command mode, which ensures that the flightpath target $\gamma_{TGT}$ is captured and that an acceptable point of airspeed equilibrium is reached, partial effectiveness of both the Altitude Capture mode and the Altitude Hold mode then follows from the stability property of the height regulator. Denoting partial effectiveness of the Altitude Capture mode and the Altitude Hold mode by PE (CH), this property can be expressed symbolically by the equivalence
$\mathrm{PE\ (CH)} \equiv \mathrm{PE}\ (\gamma)$.

### Climb and Descend Modes

Because the Climb and Descend modes both use the primitive V Command mode during thrust saturation, as already noted, partial effectiveness of the V Command mode is required to ensure partial effectiveness of both the Climb mode and the Descend mode. If partial effectiveness holds for the V Command mode, which ensures that the airspeed target $V_{TGT}$ is captured and that the flight-path angle $\gamma_{SPEED}$ provides an acceptable approximation to the flightpath target $\gamma_{TGT}$, partial effectiveness of both the Climb mode and the Descend mode then follows from the stability property of the speed regulator. Denoting partial effectiveness of the Climb mode and the Descend mode by PE (CD), this property can be expressed symbolically by the equivalence PE (CD) ≡ PE (V).

In addition, climb effectiveness should be required to hold in the Climb mode, and descent effectiveness should be required to hold in the Descend mode, as already noted. The nominal threshold vertical velocity, which is based on the minimum climb performance required at the maximum altitude $H_{MAX}$, is discussed later.

### Annunciation Requirements

Any violation of partial effectiveness of the Altitude Command supermode or any of its internal elements should be annunciated to the human crew and to any other invoking entity. Because partial effectiveness of the Altitude Capture mode and the Altitude Hold mode is required to ensure capture of the target altitude, violation of partial effectiveness PE (CH) could result in an altitude clearance violation.

### Trajectory Prediction and Replanning

Altitude overshoot can be predicted if, at any time during altitude capture, the normal acceleration commanded by the height regulator exceeds the limit on normal acceleration imposed by the normal acceleration limiter (fig. 11(d)). To provide sufficient authority for correction of errors resulting from atmospheric disturbances and the like, the limiter setting $a_{N\ LIM}$ is made equal to twice the maximum normal acceleration expected during altitude capture (fig. 11(d)). If the acceleration commanded by the height regulator, which is equal to the product V $d\gamma_{REF}/dt$ (fig. 11(d)), exceeds this limiter authority setting, then the authority should be increased temporarily (with pilot consent) to avoid altitude overshoot. (For a definition of management by consent, see Billings, 1996, page 104.)

Because the planned trajectory during altitude capture is specified implicitly by flight control system parameters such as limiter authority, such a temporary increase of limiter authority amounts to replanning of the trajectory segment that comprises altitude capture. This replanning illustrates a general design principle that is capable of generalization to higher levels in the mode hierarchy: Any violation of trajectory constraints should be predicted, and should force trajectory replanning so as to satisfy those constraints if possible.

In general, such violations should be annunciated on a case-by-case basis to higher levels within the mode hierarchy (fig. 18) for resolution there (or, ultimately, by the human crew). In some cases, local resolution is possible simply by changing path/speed priority appropriately. In other cases, resolution could require strategic action at the very top of the mode hierarchy such as development of a completely new flight plan, terminated perhaps by emergency landing at an alternate airport. Such extensive replanning is beyond the scope of this report, which focuses on the tactical levels of the mode hierarchy.

However, the distributed decision-making structure developed here lays a foundation for treating such replanning problems by future extension. This concept of distributed decision-making is illustrated by figure 18: Information flows upward to provide a basis for setting policy at higher level, and control flows downward to provide policy for decision-making at lower level.

## Safety Versus Effectiveness

In the trajectory replanning example just discussed, conflict between an external constraint imposed at high level within the mode hierarchy and an internal constraint at lower level was resolved by modifying the lower-level constraint. A related example shows how information available at higher level within the Altitude Command supermode can be used to resolve a conflict between safety and effectiveness by modifying a safety envelope constraint imposed within the lower-level Path/Speed Command supermode.

### Underspeed Protection

Invocation of the Underspeed Protection mode can have adverse consequences for flight safety, even though all a priori safety requirements are satisfied. If terrain clearance becomes critical, it is essential for flight safety to achieve the best climb gradient of which the aircraft is capable. However, this objective is not, in general, achieved by the envelope protection modes previously discussed, which sacrifice effectiveness (that is, performance) in favor of safety requirements imposed a priori.

For example, aircraft performance limitations following engine failure could preclude capture of the flightpath target. Following engine failure, selection of PATH priority could lead to airspeed divergence toward stalling speed, resulting in invocation of underspeed protection. Priority would then be given automatically to speed, and the primitive V Command mode would be selected, as previously discussed. The aircraft would then stabilize at the minimum airspeed $V_{MIN}$ at the flightpath angle determined by maximum available engine-out thrust. The performance envelope of figure 21 shows that this operating point (point A) would result in significantly poorer performance (that is, reduced climb gradient) than that available at the speed for best angle of climb (point B), which lies near the speed for minimum drag (appendix B).

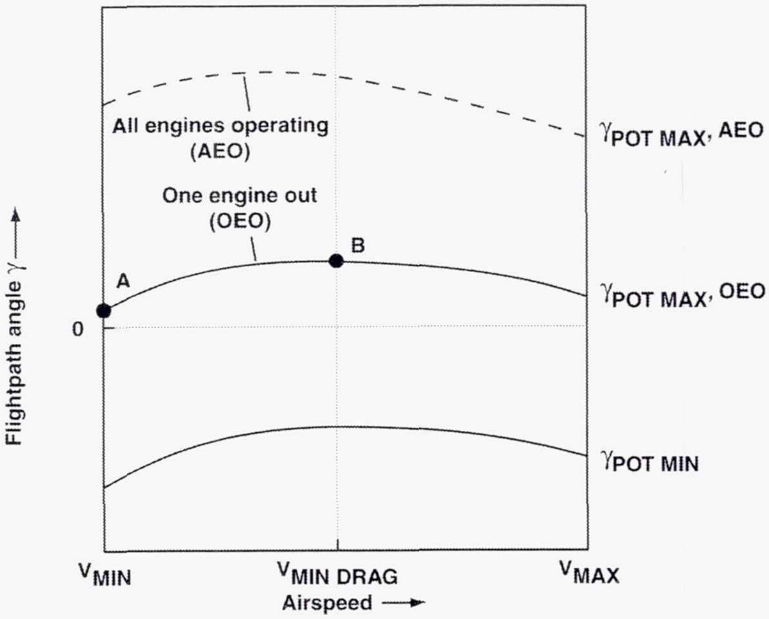


Figure 21. Engine-out performance.

The adverse consequences of this conflict between safety and effectiveness can be mitigated almost completely by imposing an additional restriction on the safe operating envelope limits selected a priori. This additional envelope restriction, which is discussed next, provides improved performance by adjusting the a priori safety envelope limits (fig. 13) so as to match them more closely to the needs of the specific operational situation.

### Mitigation strategy

By redefining the minimum-airspeed safety envelope limit $V_{MIN}$ so as to make it equal to the minimum-drag airspeed $V_{MIN\ DRAG}$ as a default value, the adverse consequences of invocation of underspeed protection can be avoided, and the best possible performance can be assured. In the Underspeed Protection mode, the airspeed is stabilized near the redefined $V_{MIN}$ (that is, near $V_{MIN\ DRAG}$). It can be seen from the performance envelope that the resulting operating point ($V_{MIN\ DRAG}$, $\gamma_{POT\ MAX}$) (that is, point B of figure 21) achieves the maximum available climb angle.

For many operational situations, such as drift-down to lower altitude following engine failure during high-altitude cruising flight, this behavior is optimal. The envelope restriction imposed by setting $V_{MIN}$ equal to $V_{MIN\ DRAG}$ is unimportant during climbing, cruising, or descending flight, because in normal operation the airspeeds scheduled for those flight phases exceed the speed for minimum drag (appendix B). Those are, of course, the flight phases in which the Altitude Command supermode would usually be selected.

However, other situations require operation below the minimum-drag speed, such as climb just after takeoff where obstacle clearance precludes immediate acceleration. Furthermore, some transport aircraft operate below the minimum-drag speed during landing approach. These special cases can be handled by enabling the higher-level entity of the mode hierarchy that invokes the primitive level to adjust the default value of $V_{MIN}$ as necessary (fig. 18). In this way, the numerical value of $V_{MIN}$ can be made to depend on higher-level modes and/or phase of flight in whatever manner is considered desirable, a strategy providing great design flexibility. For safety, it is of course essential that any such redefined envelope limits lie within the a priori safety envelope.

To summarize, basic safety is assured by invocation (if necessary) of envelope protection modes. Restricted envelope limits should depend appropriately on the higher-level supermode engaged and/or the phase of flight. Within those envelope limits, aircraft performance and safety can both be maximized by timely high-level decisions that avoid invocation of envelope protection.

## Mode Validity for Altitude Command Supermode

Mode validity is defined to require that both safety requirements and effectiveness requirements be satisfied. The following validity conditions on the altitude target $H_{TGT}$ and the airspeed target $V_{TGT}$ apply to all the internal elements of the Altitude Command supermode.

### Condition on $H_{TGT}$

As already noted, the altitude target $H_{TGT}$ must lie within the following envelope limits:

$$H_{SAFE} \leq H_{TGT} \leq H_{MAX} \tag{17a}$$

By defining the binary logical conditions H1 ≡ ($H_{TGT} < H_{SAFE}$) and H2 ≡ ($H_{TGT} > H_{MAX}$), condition (17a) can be expressed symbolically as ¬H1 ¬H2.

### Condition on $V_{TGT}$

The validity requirements on the airspeed target $V_{TGT}$ for the Altitude Command superrnode are formally the same as those for the Path/Speed Command supermode:

$$V_{MIN} \leq V_{TGT} \leq V_{MAX} \qquad (17b)$$

By using the binary logical conditions VT1 ≡ ($V_{TGT} < V_{MIN}$) and VT2 ≡ ($V_{TGT} > V_{MAX}$), as defined previously, condition (17b) can be expressed symbolically as ¬VT1 ¬VT2. The numerical value of $V_{MIN}$ is increased during operation in the Altitude Command supermode, as already discussed, in order to mitigate the performance penalty and the associated adverse safety consequences resulting from invocation of underspeed protection.

### Validity of Altitude Capture and Altitude Hold Modes

For validity of both the Altitude Capture and the Altitude Hold modes, partial effectiveness must hold for the $\gamma$ Command mode, as already noted. For validity of the Altitude Hold mode, the condition $|\Delta H| \leq H_0$ must hold for the altitude error $\Delta H$. The altitude deviation threshold $H_0$ is set nominally to 300 ft. Therefore, validity of the Altitude Hold mode assures operation within ATC tolerances.

For validity of the Altitude Capture mode, the condition $|\Delta H| > H_0$ must hold.

### Validity of Climb and Descend Modes

For validity of both the Climb and Descend modes, partial effectiveness must hold for the primitive V Command mode, as already noted. For validity of the Climb mode, the condition $\Delta H \geq 0$ must hold for the altitude error $\Delta H$, and the climb effectiveness condition $dH/dt \geq (dH/dt)_0$ should hold for the vertical velocity dH/dt. Similarly, for validity of the Descend mode the condition $\Delta H < 0$ must hold for the altitude error $\Delta H$, and the descent effectiveness condition $dH/dt \leq -(dH/dt)_0$ should hold for the vertical velocity dH/dt.

If climb effectiveness holds, altitude error is reduced monotonically during climb. Therefore, climb effectiveness ensures that the target altitude will eventually be reached. Similarly, if descent effectiveness holds during descent, the target altitude will eventually be reached. Violation of climb or descent effectiveness should be annunciated to the crew, or to any other higher-level entity invoking the Altitude Command supermode.

It should be noted that violation of climb effectiveness does not imply an immediate threat to continued safe operation in the Climb mode. Therefore, climb effectiveness is not required for validity of the Climb mode, but any violation must result in a cautionary annunciation. Such an annunciation could occur, for example, during manual operation if the pilot failed to follow flight director indications with sufficient accuracy. Similarly, descent effectiveness is not required for validity of the

Descend mode, but any violation must result in a cautionary annunciation. A suitable threshold value $(dH/dt)_0$ for the vertical velocity can be determined as follows.

### Condition on Vertical Velocity

During climb, the vertical velocity decreases with altitude as available thrust decreases (appendix B). At the maximum altitude $H_{MAX}$, a small thrust reserve is necessary to avoid loss of airspeed during turns, as previously noted. In wings-level climbing flight, the same thrust reserve enables a minimal positive vertical velocity to be maintained.

Calculation of this minimum vertical velocity for the generic transport aircraft described in appendix B shows that a nominal vertical velocity of 250 ft/min can be required over the whole altitude range up to $H_{MAX}$:

$$\text{Climb effectiveness:} \quad dH/dt \geq (dH/dt)_0 \tag{18a}$$

where $(dH/dt)_0 = 250$ ft/min.

The threshold vertical velocity for descent can conveniently be chosen as –250 ft/min:

$$\text{Descent effectiveness:} \quad dH/dt \leq -(dH/dt)_0 \tag{18b}$$

where $(Dh/dt)_0 = 250$ ft/min. Climb effectiveness is denoted symbolically as CE, and descent effectiveness as DE.

It is found that, during initial pull-up or push-over before the vertical velocity has stabilized, annunciation based on the conditions (18a) or (18b) could generate nuisance warnings. During initial pull-up or push-over, the normal acceleration limiter (fig. 11(a)) is active, but during steady climb or altitude capture it can be shown that the limiter is inactive. Therefore, nuisance warnings can be suppressed by conjoining conditions (18a) and (18b) with the requirement that the normal acceleration limiter be inactive:

$$|dH/dt| \geq (dH/dt)_0 \quad \text{AND} \quad |V\, d\gamma_{REF}/dt| < |a_{N\,LIM}| \tag{18c}$$

where $(dH/dt)_0 = 250$ ft/min.

### Definition of $H_{MAX}$

At the cruising speed corresponding to $M = 0.825$ (appendix B), the condition ($dH/dt \geq 250$ ft/min) corresponds to the condition ($\gamma_{POT\,MAX} \geq 0.3$ degree) on $\gamma_{POT\,MAX}$. This latter condition will be taken as the defining condition that establishes $H_{MAX}$.

### Summary of Validity Conditions

The validity conditions for the Altitude Command supermode just discussed are summarized by table 8.

TABLE 8. VALIDITY CONDITIONS FOR ALTITUDE COMMAND SUPERMODE

| | Altitude Capture | Altitude Hold | Climb | Descend |
|---|---|---|---|---|
| $H_{TGT}$ | $H_{SAFE} \leq H_{TGT} \leq H_{MAX}$ | | | |
| $V_{TGT}$ | $V_{MIN} \leq V_{TGT} \leq V_{MAX}$ | | | |
| $\Delta H$ | $\lvert\Delta H\rvert > H_0$ | $\lvert\Delta H\rvert \leq H_0$ | $\Delta H \geq 0$ | $\Delta H < 0$ |
| PE (•) | PE ($\gamma$) | | PE (V) | |

## Heuristic Strategy

Validity conditions are not sufficient, in general, to determine mode selection logic. Avoiding the selection of an invalid mode is an obvious requirement, but choosing the most desirable mode selection from several valid possibilities requires a mode selection strategy that goes beyond questions of validity. Such a strategy can be based on the general notion of maximizing system effectiveness. For example, complete effectiveness should be chosen in preference to partial effectiveness, as noted previously during development of the Path/Speed Command supermode.

The mode selection strategy might also have to deal with a situation in which none of the modes are valid. Selection should then favor the choice that minimizes adverse consequences, but determining the least adverse choice requires detailed analysis of each possible flight situation. No general policy, such as sacrificing effectiveness in favor of safety requirements imposed a priori, is certain to minimize adverse consequences under all conditions. After presenting the strategies to be used for screening the altitude and airspeed targets, the mode selection strategy is discussed in detail. Initialization of the Altitude Command and the Path/Speed Command supermodes are discussed after presenting the final mode selection strategy.

### Target Screening

At engagement, the Altitude Command supermode must screen the candidate values of the altitude and airspeed targets to ensure that the target operating point lies within safety limits and within the performance capability of the aircraft. At any time after engagement, the crew (or other invoking entity) must be permitted to change the altitude and airspeed targets as desired, subject only to screening. Continuous screening is necessary to account for changing performance capability, but manually selected targets must not be changed after entry without pilot consent.

Denoting by $H_{KNOB}$ and $V_{KNOB}$ the altitude and airspeed set by the pilot by adjusting the mode control panel knobs (fig. 19), the target screening strategy is summarized by table 9. When the pilot enters $H_{DISP}$, $H_{TGT}$ is set equal to $H_{DISP}$ (fig. 19(b)). Similarly, when the pilot enters $V_{DISP}$, $V_{TGT}$ is set equal to $V_{DISP}$ (fig. 19(c)).

### Partitioning of Validity Table

Because the target screening strategy ensures that the validity conditions on $H_{TGT}$ and $V_{TGT}$ (table 8) always hold, these conditions can be dropped from the table (Theorem 13, appendix D). Other useful simplifications of the validity table can be obtained by means of two heuristic revisions to the mode structure, which are described next.

TABLE 9. TARGET SCREENING STRATEGY FOR ALTITUDE COMMAND SUPERMODE

| | | |
|---|---|---|
| **$H_{KNOB} < H_{SAFE}$**<br>Set $H_{DISP} = H_{SAFE}$ | **$H_{KNOB} > H_{MAX}$**<br>Set $H_{DISP} = H_{MAX}$ | **$H_{SAFE} \leq H_{KNOB} \leq H_{MAX}$**<br>Set $H_{DISP} = H_{KNOB}$ |
| **$V_{KNOB} < V_{MIN}$**<br>Set $V_{DISP} = V_{MIN}$ | **$V_{KNOB} > V_{MAX}$**<br>Set $V_{DISP} = V_{MAX}$ | **$V_{MIN} \leq V_{KNOB} \leq V_{MAX}$**<br>Set $V_{DISP} = V_{KNOB}$ |

**Altitude Capture/Hold supermode**– Because the height regulator laws specify the target flightpath angle $\gamma_{TGT}$ uniquely and continuously as a function of altitude error, as already noted, it is convenient to combine the Altitude Capture and Altitude Hold elements into a single Altitude Capture/Hold supermode. Within this supermode, transitions between the Altitude Capture and Altitude Hold states are controlled by the magnitude of the measured height error (fig. 11(b). Because within the Altitude Capture/Hold supermode one or the other of the conditions on $|\Delta H|$ must always hold, both of the conditions can be dropped from the revised table.

To verify this formally, note that within the Altitude Capture/Hold supermode the conditions $(|\Delta H| \leq H_0)$ and the condition $(|\Delta H| > H_0)$ form a logical disjunction that holds tautologically, because the disjunction $[(|\Delta H| \leq H_0)$ OR $(|\Delta H| > H_0)] \equiv$ TRUE. Therefore, these conditions on $|\Delta H|$ can be dropped from the revised validity table for the Altitude Capture/Hold supermode.

**Climb/Descend supermode**– In a similar manner, combining the Climb mode and the Descend mode into a single Climb/Descend supermode enables the conditions on $\Delta H$ to be dropped from the revised validity table for the Climb/Descend supermode. It is clear that the Climb and Descend modes are functionally similar, differing only in the sign of the altitude error $\Delta H$, which determines the sign of the flightpath target $\gamma_{TGT}$ according to the height regulator law. Therefore, it is convenient to combine the Climb and Descend elements into a single Climb/Descend supermode. Within this supermode, transitions between the Climb and Descend states are controlled by the measured height error $\Delta H$, which must be nonnegative for the Climb mode and negative for the Descend mode. Within the Climb/Descend supermode the disjunction $[(\Delta H \leq 0)$ OR $(\Delta H > 0)]$ holds tautologically; therefore, the conditions on $\Delta H$ can be dropped from the revised validity table for the Climb/Descend supermode.

**Mode selection strategy**– Within the Altitude Capture/Hold supermode and the Climb/Descend supermode, the strategy for selecting the Altitude Capture, Altitude Hold, Climb, and Descend modes and for setting priority can be summarized by the condition-action mode selection process (appendix D) shown in table 10:

TABLE 10. MODE SELECTION WITHIN ALTITUDE CAPTURE/HOLD SUPERMODE AND CLIMB/DESCEND SUPERMODE

<table>
<tr><th>Condition</th><th>Altitude Capture</th><th>Altitude Hold</th><th>Climb</th><th>Descend</th></tr>
<tr><td>ALTITUDE CAPT/HOLD selected</td><td colspan="2">TRUE</td><td colspan="2"></td></tr>
<tr><td>CLIMB/ DESCEND selected</td><td colspan="2"></td><td colspan="2">TRUE</td></tr>
<tr><td>$| \Delta H | \leq H_0$</td><td></td><td>TRUE</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>$| \Delta H | > H_0$</td><td>TRUE</td><td></td></tr>
<tr><td>$\Delta H \geq 0$</td><td colspan="2" rowspan="2"></td><td>TRUE</td><td></td></tr>
<tr><td>$\Delta H < 0$</td><td></td><td>TRUE</td></tr>
<tr><td rowspan="2">Actions</td><td colspan="2">Set PATH priority</td><td colspan="2">Set SPEED priority</td></tr>
<tr><td colspan="4">Select PATH/SPEED COMMAND supermode<br>Enter table 7</td></tr>
</table>

### Inversion of Reduced Validity Table

After removing from table 8 all the conditions just discussed, only the partial effectiveness conditions PE (V) and PE (γ) remain, forming the following reduced validity table:

| | ALTITUDE CAPT/HOLD | CLIMB/DESCENT |
|---|---|---|
| PE (•) | PE (γ) | PE (V) |

This reduced validity table can be inverted by inspection, with the resulting four cases enumerated as shown in table 11.

TABLE 11. INVERSION OF REDUCED VALIDITY TABLE FOR ALTITUDE COMMAND SUPERMODE

| Condition | ALTITUDE CAPT/HOLD | CLIMB/DESCEND |
|---|---|---|
| 1. ¬PE (γ) ¬PE (V) | INVALID | INVALID |
| 2. ¬PE (γ) PE (V) | INVALID | VALID |
| 3. PE (γ) ¬PE (V) | VALID | INVALID |
| 4. PE (γ) PE (V) | VALID | VALID |

Table 11 shows that cases 1, 2, and 3 are abnormal, because each involves invalidity of one or both superrnodes. Mode selection strategy is discussed first for these abnormal conditions.

### Mode Selection Strategy for Abnormal Conditions

If partial effectiveness is violated for the primitive $\gamma$ Command mode but holds for the primitive V Command mode (case 2), then only the Climb/Descend supermode is valid. It should therefore be selected.

If partial effectiveness is violated for the primitive V Command mode but holds for the primitive $\gamma$ Command mode (case 3), then only the Altitude Capture/Hold supermode is valid. It should therefore be selected.

If partial effectiveness is violated for both the primitive $\gamma$ Command and V Command modes (case 1), then neither the Altitude CapturelHold supermode nor the Climb/Descend supermode is valid. The Performance Degradation Theorem (Theorem 1, appendix F) shows that the condition $\neg$PE ($\gamma$) $\neg$PE (V) can hold only if $\gamma_{\mathrm{POT\ MAX}} < 0$; that is, if aircraft performance becomes so severely degraded that available thrust is insufficient for level flight. Detailed analysis of this situation, which could occur following engine failure at high altitude, shows that selection of the Climb/Descend mode is the only appropriate choice. Therefore, in case 1 the Climb/Descend supermode should be selected.

This completes the formulation of the supermode selection strategy for abnormal conditions, for which the strategy is based on considerations of validity. For normal conditions (case 4), for which both supermodes are valid, the selection strategy can be based on considerations of effectiveness, which are discussed next.

### Mode Selection Strategy for Normal Conditions

When partial effectiveness holds for both of the primitive V Command and $\gamma$ Command modes, table 11 shows that both the Altitude Capture/Hold supermode and the Climb/Descend supermode are valid. In that case (the normal situation), the mode selection strategy can be based on the desire for a smooth altitude capture, which can be achieved by avoiding violations of the normal effectiveness properties previously defined. There are two design issues: first, smoothness of normal acceleration and flightpath; and second, smoothness of longitudinal acceleration and airspeed. Both issues can be clarified by examining the geometric stability regions traversed during altitude capture (fig. 22).

**Flightpath smoothness**– If the magnitude of the altitude error $\Delta$H is so large that the target flightpath angle $\gamma_{\mathrm{TGT}}$ commanded by the height regulator (fig. 11(b)) lies in a region where thrust is saturated (that is, the condition P OR Q holds, where P and Q are defined by equation (11i)), the Climb/Descend mode should be selected, as already noted. The system then captures the flightpath $\gamma_{\mathrm{SPEED}}$. If capture of the target airspeed is also complete, the operating point (V, $\gamma$) coincides with the point ($V_{\mathrm{TGT}}$, $\gamma_{\mathrm{SPEED}}$). If airspeed capture is not yet complete, the operating point lies elsewhere (fig. 22) on the $\gamma_{\mathrm{SPEED\ MAX}}$ boundary during climb or on the $\gamma_{\mathrm{SPEED\ MIN}}$ boundary during descent, with the longitudinal acceleration (indicated by arrows in the diagram) directed toward the target airspeed $V_{\mathrm{TGT}}$.

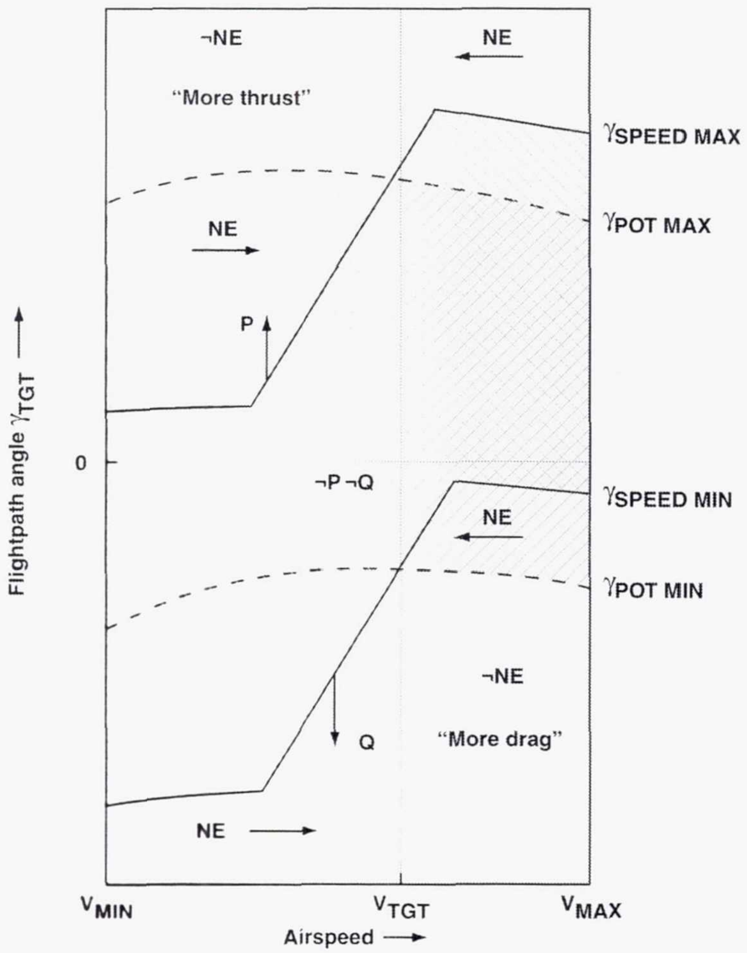


Figure 22. Geometric regions for mode selection within Altitude Command supermode.

As the aircraft approaches the target altitude, the magnitudes of the height error $|\Delta H|$ and of the flightpath target $|\gamma_{TGT}|$ steadily decrease toward zero, so that the condition

$$\neg P \neg Q \equiv (\gamma_{SPEED\ MIN} < \gamma_{TGT} < \gamma_{SPEED\ MAX})$$

must eventually hold (fig. 22). When the Altitude Capture/Hold mode is selected, the system shifts from capture of $\gamma_{SPEED}$ to capture of the target flightpath $\gamma_{TGT}$. Therefore, to ensure smoothness of both normal acceleration and flightpath during the transition, the Altitude Capture/Hold mode should be selected when $\gamma_{TGT}$ coincides with $\gamma_{SPEED}$; that is, as soon as the condition ¬P ¬Q holds. It remains to determine whether this simple mode selection strategy also ensures smoothness of longitudinal acceleration and airspeed.

**Airspeed smoothness**– It will be recalled that the normal effectiveness (NE) property of the primitive $\gamma$ Command mode holds when the longitudinal acceleration has the correct sign leading toward capture of the airspeed target. Violation of normal effectiveness generates cautionary annunciations ("More thrust" or "More drag," as shown by figure 22). Because violation of normal effectiveness leads to an airspeed reversal and possibly to a thrust reversal, such violations should be avoided to ensure a smooth altitude capture, even though normal effectiveness is not required for validity of the Altitude Capture/Hold mode.

During altitude capture, the flightpath target $\gamma_{TGT}$ approaches the V axis monotonically (fig. 22) according to the height regulator law. If $\gamma_{TGT}$ changes continuously, as it does during a normal altitude capture for which the altitude target $H_{TGT}$ remains fixed, the stability property of the path regulator ensures that the actual flightpath angle $\gamma$ tracks the target $\gamma_{TGT}$ closely. Therefore, it may be expected that under normal circumstances, if the condition

$$\neg P \ \neg Q \equiv (\gamma_{\text{SPEED MIN}} < \gamma_{\text{TGT}} < \gamma_{\text{SPEED MAX}})$$

holds, then the condition ($\gamma_{\text{SPEED MIN}} < \gamma < \gamma_{\text{SPEED MAX}}$) also holds, ensuring that thrust is unsaturated (eqs. (11g) and (11h)). The primitive $\gamma$-V Command mode is then selected, for which normal effectiveness always holds. Figure 22 shows that, when the flightpath target enters the central region in which the condition $\neg P \ \neg Q$ holds, it remains in that region as it moves toward the V axis during altitude capture. Therefore, the simple mode selection strategy of selecting the Climb/Descend mode if the condition $P \cup Q$ holds and selecting the Altitude Capture/Hold mode if the condition $\neg P \ \neg Q$ holds ensures that normal effectiveness holds during altitude capture under normal circumstances.

If a small change of the altitude target should be entered by the flight crew during altitude capture, the result would be a discontinuous change in $\gamma_{TGT}$ that could not be tracked closely by the actual flightpath angle $\gamma$. It is then possible that thrust could be saturated and that normal effectiveness could be violated during the short term before path capture is completed, even if the condition $\neg P \ \neg Q$ holds. Under these exceptional circumstances, temporary violation of normal effectiveness can be permitted, provided that partial effectiveness of the $\gamma$ Command mode holds.

**Undesired mode transitions**– Near the boundaries of the central $\neg P \ \neg Q$ region, altitude variations owing to atmospheric disturbances could cause repetitive cycling between the conditions $\neg P \ \neg Q$ and $P \cup Q$, resulting in repetitive mode transitions (“chattering”). Suppression of unwanted mode transitions must be dealt with during implementation, when uncertainties in the measurement and estimation of physical quantities (“noise”) can be accounted for by means of simulation. In order to prevent repetitive mode transitions owing to noise, it will probably be necessary to incorporate hysteresis bands at the boundaries of the geometric regions governing mode transitions. However, hysteresis that prevents repetitive cycling owing to noise might not be sufficient to suppress all undesired mode reversions. Furthermore, the simple mode selection strategy just discussed is more restrictive than necessary to prevent violation of normal effectiveness.

In figure 22, the regions in which normal effectiveness holds for the Altitude Capture/Hold mode are marked with the legend NE. It can be seen that there are regions adjacent to the central $\neg P \ \neg Q$ region (indicated by hatching) in which reversion to the Climb/Descend mode to prevent violation of normal effectiveness is unnecessary, because normal effectiveness would continue to hold unconditionally without reversion. In other regions (indicated by absence of hatching), normal effectiveness holds temporarily for the Altitude Capture/Hold mode, but the airspeed trend (indicated by acceleration/deceleration arrows) would lead to future violation. A mode selection criterion based on prediction of future states would be excessively complex. The situation is further complicated by the fact that the target airspeed must be reset as a function of flight phase (for example, climb, cruise, or descent) according to a schedule of optimal speeds (appendix B). Such changes in target airspeed could cause unnecessary mode reversions by shifting the boundaries of the $\neg P \neg Q$ region horizontally relative to the prevailing operating point.

**Mode latching**– A simple method sometimes used in the past to suppress unwanted mode reversions was to "latch" the mode selection logic; for example, with such latching incorporated, the transition to the Altitude Capture/Hold mode would become a one-way transition. Latching can be regarded as a kind of hysteresis applied to the mode selection logic itself. In contrast to the conventional hysteresis discussed previously, which would blur to a minor extent the regional boundaries treated in this report as mathematically sharp, latching would destroy the relevance of those boundaries for mode selection because, if a latched mode state were entered, it would be maintained without regard to conditions based on geometrical regions.

**Discussion**– Mode latching would be unacceptable for the synthesis method developed in this report. If latching were incorporated, because mode selection would then depend partially on previous mode states, the desirable Markovian property of the selection logic (to be discussed later) would be destroyed, together with the unique relation of the selection logic to the geometrical regions already described. Aircraft behavior would then become dependent not only upon the regions in which the target point ($V_{TGT}$, $\gamma_{TGT}$) and the actual operating point ($V$, $\gamma$) lie, but also upon the sequence in which those regions were traversed. This unwelcome increase in complexity would complicate analysis of system properties and could lead to unpredictable behavior, a situation that should be considered unacceptable for safety-critical systems.

**Future development**– The development of a more sophisticated mode selection strategy that would suppress unnecessary mode reversions without introducing unpredictable system behavior is left as a problem to be addressed in future work. Until simulation based on implementation of the present design has indicated the extent to which unnecessary mode reversions are objectionable, further development seems premature.

### Summary of Supermode Selection Strategy

The complete supermode selection strategy for the Altitude Command supermode is summarized by the following condition-action mode selection table (table 12).

### Altitude Capture Example

To illustrate the sequence of mode selection during performance of a representative task, the same altitude capture task discussed previously is selected. It is assumed that the conditions PE ($\gamma$) and PE (V) hold throughout the altitude capture process; that is, operating conditions are normal.

**Initial conditions**– It is assumed that an aircraft initially in level flight at 15,000 ft with the Altitude Command supermode engaged is cleared to climb to 35,000 ft (Flight Level 350) at an airspeed of 250 kt. In level flight, the condition $\neg P \neg Q$ holds (fig. 22). It then follows from table 12 that, initially, the Altitude Capture/Hold supermode is selected. The mode selection process then continues with table 10.

TABLE 12. SUPERMODE SELECTION WITHIN ALTITUDE COMMAND SUPERMODE

| **Conditions** | **ALTITUDE CAPT/HOLD** | | **CLIMB/DESCEND** | |
|---|---|---|---|---|
| | **NORMAL** | **ABNORMAL** | **NORMAL** | **ABNORMAL** |
| PE ($\gamma$) | TRUE | TRUE | TRUE | |
| $\neg$PE ($\gamma$) | | | | TRUE |
| PE (V) | TRUE | | TRUE | |
| $\neg$PE (V) | | TRUE | | |
| $\neg$P $\neg$Q | TRUE | | | |
| P $\cup$ Q | | | TRUE | |
| **Action** | Enter table 10 | | | |

**Definitions**

$\text{PE}(\gamma) \equiv (\gamma_{\text{TGT}} < \gamma_{\text{POT MAX}})$ OR $(V > V_{\text{MIN DRAG}})$ OR $(\gamma < \gamma_{\text{SPEED MAX}})$

$\text{PE}(V) \equiv \neg[\text{P AND TT1}]$ AND $\neg[\text{Q AND TT2}]$

$\text{P} \equiv (\gamma_{\text{TGT}} \geq \gamma_{\text{SPEED MAX}})$ $\quad$ $\text{Q} \equiv (\gamma_{\text{TGT}} \leq \gamma_{\text{SPEED MIN}})$

$\neg\text{P}\ \neg\text{Q} \equiv (\gamma_{\text{SPEED MIN}} < \gamma_{\text{TGT}} < \gamma_{\text{SPEED MAX}})$

$\text{P} \cup \text{Q} \equiv (\gamma_{\text{TGT}} \geq \gamma_{\text{SPEED MAX}})$ OR $(\gamma_{\text{TGT}} \leq \gamma_{\text{SPEED MIN}})$

It is assumed that the initial altitude error is small, so that the condition $|\Delta H| \leq H_0$ holds. In that case, it follows from table 10 that the Altitude Hold mode is selected, priority is set to PATH, and the first-level Path/Speed Command supermode is invoked. The mode selection process then continues with table 7.

Because by assumption the aircraft is initially in level flight, the thrust is unsaturated, as already noted; that is, the condition $\neg$TS1 $\neg$TS2 holds. It then follows from table 7 that the primitive $\gamma$-V Command mode is selected initially.

**Pull-up**– After verifying the legality and feasibility of the clearance (fig. 1), the flight crew initiates the climb by setting the target altitude to 35,000 ft in the altitude window of the Mode Control Panel, and pressing the ENTER button (fig. 19). Upon entry of the new altitude target, a new flightpath target is calculated as follows. The altitude error becomes (fig. 11(b))

$$\Delta H = 35{,}000 - 15{,}000 = 20{,}000 \text{ ft.}$$

Because the absolute height error exceeds the threshold value $H_0 = 300$ ft, the parabolic altitude capture law is used (fig. 11(b)) to calculate the flightpath target $\gamma_{TGT}$, which is limited to an angle 3 degrees larger than the maximum steady climb angle $\gamma_{POT\ MAX}$ (fig. 11(b)). Therefore, the condition ($\gamma_{TGT} \geq \gamma_{POT\ MAX}$) holds. Because $\gamma_{POT\ MAX} = \gamma_{SPEED\ MAX}$ holds at $V = V_{TGT}$, the condition

$$P \equiv (\gamma_{TGT} \geq \gamma_{SPEED\ MAX})$$

holds on entry of the new altitude target. It then follows from table 12 that the Climb/Descend supermode is selected, and from table 10 that the Climb mode is selected. Priority is set to SPEED, and the Path/Speed Command supermode remains invoked (table 10).

The aircraft pulls up as the $\gamma$-V Command mode attempts to capture $\gamma_{TGT}$, subject to the normal acceleration limit imposed by the acceleration limiter (fig. 11(a)). During pull-up, the increase in flightpath angle is fed forward to command increased thrust, and the airspeed regulator maintains the target airspeed of 250 kt (fig. 11(c)).

**Climb**– When the flightpath angle $\gamma$ reaches $\gamma_{SPEED\ MAX}$ (fig. 22), the thrust saturates (the condition TS2 holds). With SPEED priority, the primitive V Command mode is selected (table 7). The target thrust $\gamma_{POT\ TGT}$ is then set to $\gamma_{POT\ MAX}$ (table 5). The airspeed regulator (fig. 11(c)) maintains the target airspeed of 250 kt during climb.

**Altitude capture**– As the aircraft approaches the target altitude of 35,000 ft, the altitude error is reduced. The target flightpath angle $\gamma_{TGT}$ is likewise reduced, and is no longer limited. When $\gamma_{TGT}$ reaches $\gamma_{SPEED\ MAX}$ and the condition $\neg P \neg Q$ holds (fig. 22), table 12 shows that the Altitude Capture/Hold supermode is selected. Table 10 then shows that, depending on the magnitude of the height error, either the Altitude Capture or the Altitude Hold mode is selected. In either case, priority is set to PATH (table 10), the Path/Speed Command supermode is invoked, and the primitive $\gamma$ Command mode is selected while the thrust remains saturated. It is assumed that the height error initially exceeds the threshold value of 300 ft, so that the Altitude Capture mode is selected. In that case, the altitude capture trajectory is governed by the parabolic law (fig. 11(b)) until the height error has been reduced to 300 ft. When the condition

$$\neg TS2 \equiv (\gamma < \gamma_{SPEED\ MAX})$$

holds, the thrust becomes unsaturated, and the primitive $\gamma$-V Command mode is selected (table 7).

**Altitude hold**– When the height error becomes equal to the threshold value of 300 ft, the Altitude Hold mode is selected (table 10). The aircraft trajectory is then governed by the exponential law (fig. 11(b)), and the height error is nulled. The aircraft stabilizes at the target altitude of 35,000 ft and the target airspeed of 250 kt, completing the example.

### Initialization

For initialization of the Altitude Command supermode, the first task is initialization of the target thrust $\gamma_{POT\ TGT}$, which is carried out within the Path/Speed Command supermode by the strategy for setting target thrust, as discussed previously. Inspection of the strategy for setting target thrust (table 5) shows that this strategy is applicable for initialization except when thrust is unsaturated and the condition $\neg P \neg Q$ holds, in which case the target thrust is not determined.

But in that case, the $\gamma$-V Command mode is selected because thrust is unsaturated (table 7). Because the target thrust $\gamma_{POT\ TGT}$ is nonrelevant for operation in the $\gamma$-V Command mode, it can be initialized arbitrarily to $\gamma_{POT\ MIN}$. Therefore, invoking the Path/Speed Command supermode (with priority assigned arbitrarily) returns the quantities PE ($\gamma$), PE (V), P, and Q, which determine initial mode selection for the Altitude Command supermode according to the usual mode selection strategy (table 12).

A simpler alternative is to initialize the Altitude Command supermode unconditionally to the Climb/ Descend mode. The rationale for this simplified strategy is that, under normal conditions, it corresponds to the supermode selection when $P \cup Q$ holds (table 12), the least restrictive possibility. If the condition PE (V) is violated initially, the mode validity table (table 8) shows that this simplified initialization strategy could result initially in operation that is invalid. Nevertheless, the following analysis shows that this initialization is acceptable. There are two possibilities.

First, if PE ($\gamma$) is also violated, so that $\neg$PE ($\gamma$) $\neg$PE (V) holds, the Climb/Descend mode would be selected according to the mode selection strategy (table 12). Therefore, the simplified initialization strategy is in agreement with the mode selection strategy for this abnormal case.

Second, if PE ($\gamma$) holds, so that PE ($\gamma$) $\neg$PE (V) holds initially, then the Altitude Capture/Hold mode is selected (table 12) after one frame (appendix E) of operation in the Climb/Descend mode. Detailed analysis shows that system behavior during the first frame is the same for operation in the Climb/ Descend mode as it would be in the Altitude Capture/Hold mode, and that during subsequent frames system behavior is governed by the mode selection strategy (table 12), as previously discussed. Therefore, the simplified initialization strategy is acceptable.

**Re-initialization**– If a new altitude target is entered after the Altitude Capture/Hold mode has been selected, so that the condition $P \cup Q$ again holds, the system should revert to the Climb/Descend mode, as already noted. If hysteresis has been applied to the regional boundaries (fig. 22) to suppress repetitive mode transitions owing to noise, as discussed previously, the whole system including the hysteresis bands should be re-initialized, and the Altitude Command supermode should be re-initialized to the Climb/Descend mode.

This re-initialization to the Climb/Descend mode does *not* apply to the entry of a new airspeed target. Such re-initialization must be ruled out because it could allow the aircraft to deviate from the target altitude during capture of the new airspeed target, violating the partial effectiveness of the Altitude Command supermode.

The rationale for this difference in re-initialization strategy is that the altitude target is regarded as an external constraint imposed by ATC that determines the pilot's task. In contrast, the airspeed target can be chosen freely for performance optimization (appendix B). This rationale holds despite the fact that external airspeed restrictions are sometimes imposed by ATC, because in those cases the altitude constraint always remains primary.

## Construction of Statechart

The statechart for supermode selection (fig. 23(a)) can now be constructed directly from table 12. It may be seen that there are two scenarios (columns) for selection of the Altitude Capture/Hold supermode. For normal conditions, the conditions PE ($\gamma$), PE (V), and $\neg$P $\neg$Q must hold. Therefore, transition arrows labeled with those conditions must lead to the Altitude Capture/Hold supermode from the Climb/Descend supermode (fig. 23(a)). For the abnormal condition PE ($\gamma$) $\neg$PE (V), a transition arrow labeled with that condition (shown broken to indicate abnormality) must lead from the Climb/Descend supermode to the Altitude Capture/Hold supermode.

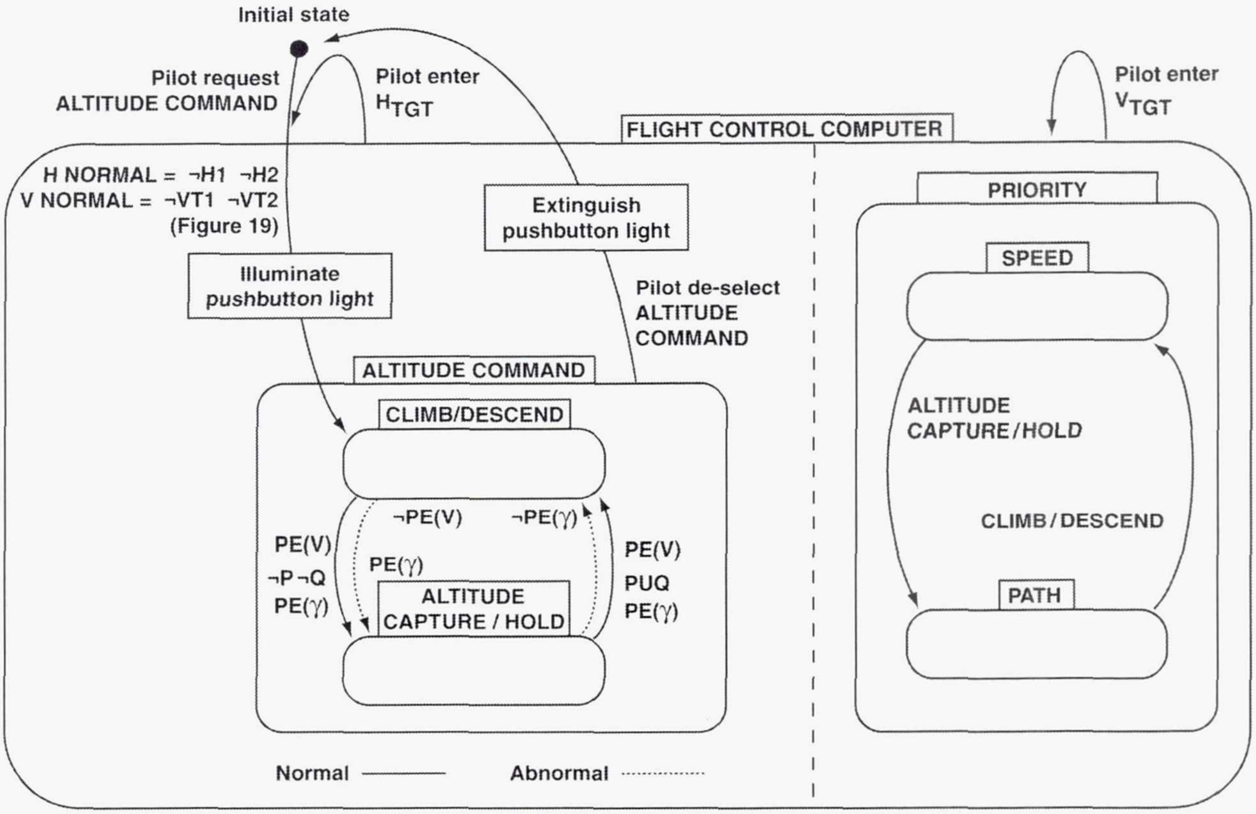
Similarly, table 12 shows that there are two scenarios (columns) for selection of the Climb/Descend supermode. For normal conditions, the conditions PE ($\gamma$), PE (V), and P $\cup$ Q must hold. Therefore, transition arrows labeled with those conditions must lead to the Climb/Descend supermode from the Altitude Capture/Hold supermode. For the abnormal condition $\neg$PE ($\gamma$), a transition arrow labeled with that condition (shown broken to indicate abnormality) must lead from the Altitude Capture/Hold supermode to the Climb/Descend supermode.

By following the construction process just described, the statechart illustrated at the left of figure 23(a) is obtained. The two priority states PATH and SPEED can be represented in statechart form by a two-state machine. The statechart for setting priority illustrated at the right of figure 23(a) can be synthesized directly from table 10.

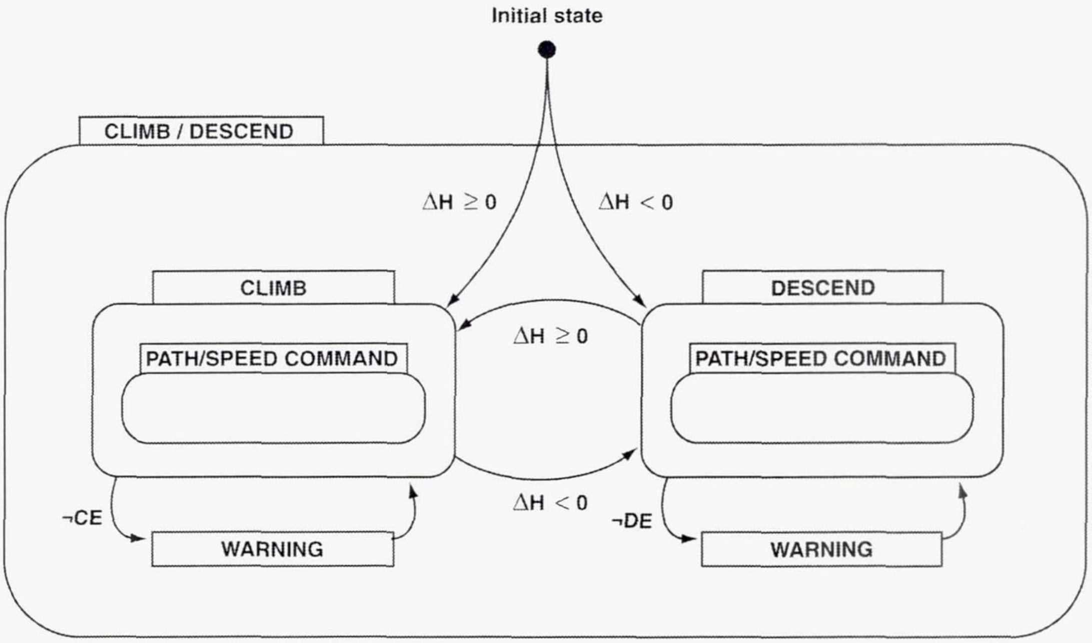
Certain Mode Control Panel actions are indicated schematically in statechart form by figure 23(a), but many details are omitted for clarity. As indicated, new altitude or airspeed targets can be entered at any time after initial engagement. It should be noted that entering a new altitude target $H_{TGT}$ results in re-initializing the Altitude Command supermode to the Climb/Descend mode, as illustrated. However, entering a new airspeed target $V_{TGT}$ has no effect on mode selection, as required by the re-initialization strategy already discussed.

The internal elements of the Climb/Descend supermode can be synthesized directly from table 10, and are illustrated in detail by the statechart of figure 23(b). It can be seen that violation of the climb effectiveness CE or the descent effectiveness DE generate appropriate warnings, which are annunciated to the human crew and also to any other invoking entity.

The internal elements of the Altitude Capture/Hold supermode can also be synthesized directly from table 10. They are illustrated in detail by the statechart of figure 23(c). Figures 23(b) and 23(c) show that the first-level Path/Speed Command supermode (fig. 15) constitutes an internal element of the second-level Climb and Descend modes (fig. 23(b)) and of the second-level Altitude Capture and Altitude Hold modes (fig. 23(c)).

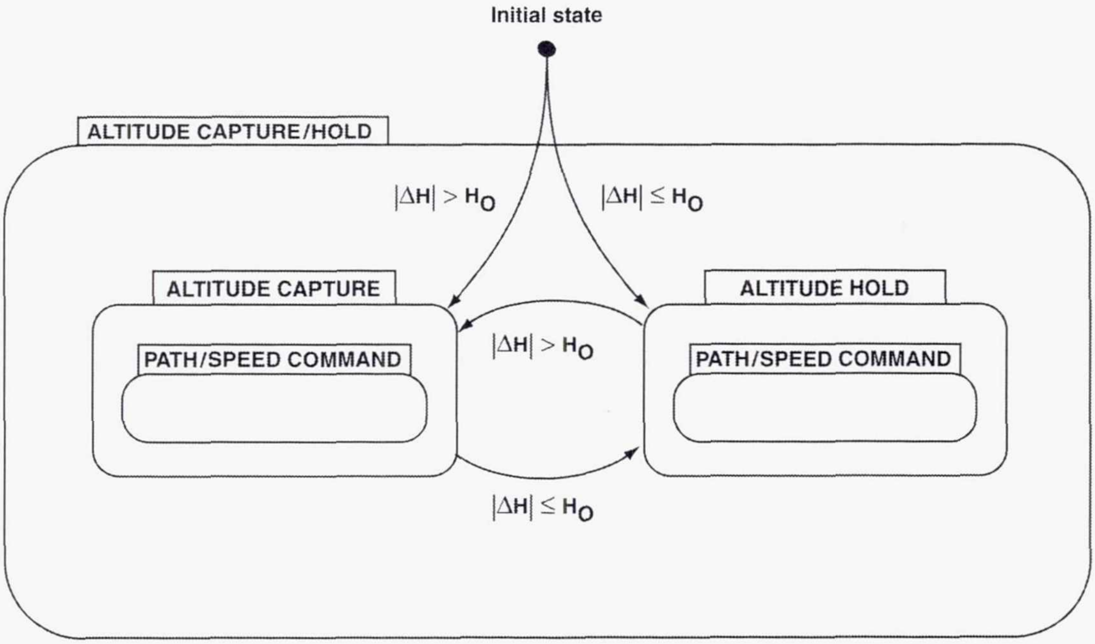


a) Altitude Command and Priority supermodes.



b) Climb/Descend supermode.

Figure 23. Statechart for Altitude Command supermode.



c) Altitude Capture/Hold supermode.

Figure 23. Statechart for Altitude Command supermode (concluded).

## System Properties for Altitude Command Supermode

### Functional Dependence of Mode Selection Logic

Mode selection logic for the Altitude Command supermode is governed by the Boolean variables PE ($\gamma$), PE (V), P, and Q, as already noted. The definitions are repeated for convenience as follows:

$$\neg\text{PE}\ (\gamma) \equiv (\gamma_{\text{TGT}} \geq \gamma_{\text{POT MAX}})\ \text{AND}\ (V \leq V_{\text{MIN DRAG}})\ \text{AND}\ (\gamma \geq \gamma_{\text{SPEED MAX}})$$

$$\neg\text{PE}\ (V) \equiv [P\ \text{AND}\ (\gamma_{\text{POT TGT}} = \gamma_{\text{POT MIN}})]\ \text{OR}\ [Q\ \text{AND}\ (\gamma_{\text{POT TGT}} = \gamma_{\text{POT MAX}})]$$

$$P \equiv (\gamma_{\text{TGT}} \geq \gamma_{\text{SPEED MAX}}) \qquad Q = (\gamma_{\text{TGT}} \leq \gamma_{\text{SPEED MIN}})]$$

From these definitions, it can be seen that their arguments fall into three categories. First, the argument $\gamma_{\text{TGT}}$ is the flightpath target. Second, the arguments $\gamma_{\text{POT MAX}}$, $\gamma_{\text{POT MIN}}$, $V_{\text{MIN DRAG}}$, $\gamma$, and V are physical quantities. Third, the arguments $\gamma_{\text{POT TGT}}$, $\gamma_{\text{SPEED MAX}}$, and $\gamma_{\text{SPEED MIN}}$, are computed quantities whose functional dependence requires further analysis.

Inspection of the strategy for setting the thrust target $\gamma_{\text{POT TGT}}$ (table 5) shows that $\gamma_{\text{POT TGT}}$ depends on the physical quantities $\gamma$, $\gamma_{\text{POT MAX}}$, and $\gamma_{\text{POT MIN}}$, and on the flightpath target $\gamma_{\text{TGT}}$. The definitions (eqs. (8a) and (9d))

$$\gamma_{\text{SPEED MAX}} = \gamma_{\text{POT MAX}} - (1/g)(dV/dt)_{\text{CMD}} \qquad \gamma_{\text{SPEED MIN}} = \gamma_{\text{POT MIN}} - (1/g)(dV/dt)_{\text{CMD}}$$

show that $\gamma_{\text{SPEED MAX}}$ and $\gamma_{\text{SPEED MIN}}$ depend on the physical quantities $\gamma_{\text{POT MAX}}$ and $\gamma_{\text{POT MIN}}$ and on the commanded longitudinal acceleration $(dV/dt)_{\text{CMD}}$, which in turn depends on the airspeed error $\Delta V = V_{\text{TGT}} - V$ and on the acceleration limits, which are physical.

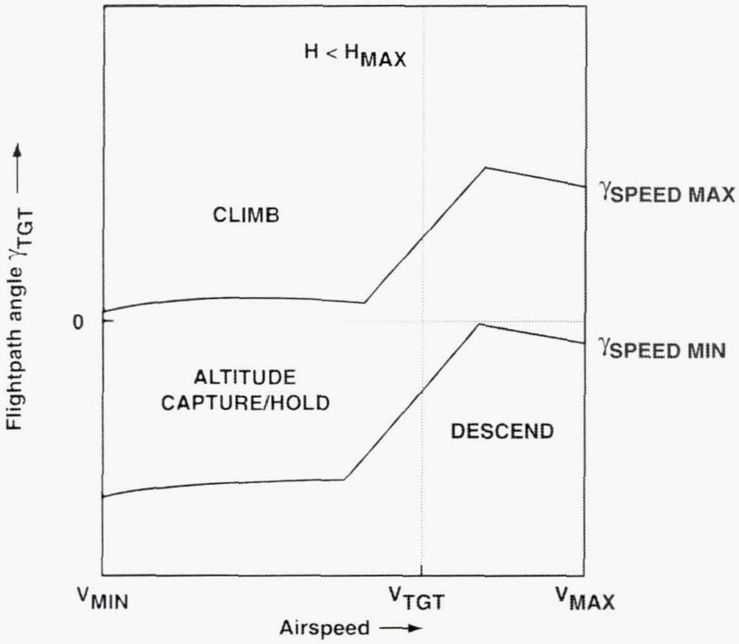
Therefore, it can be concluded that mode selection depends only on the target point ($V_{\text{TGT}}$, $\gamma_{\text{TGT}}$), the operating point ($V$, $\gamma$), and the other physical quantities just mentioned. This functional dependence enables mode selection to be specified geometrically in the following way.

### Geometrical Regions for Mode Selection

The geometrical regions corresponding to selection of the Climb, Descend, and Altitude Capture/Hold modes are illustrated by figures 24(a) and 24(b). It can be seen that, after path capture is complete, mode selection is determined entirely by the geometrical regions in which the target point lies, and is independent of the sequence in which these geometrical regions are traversed.

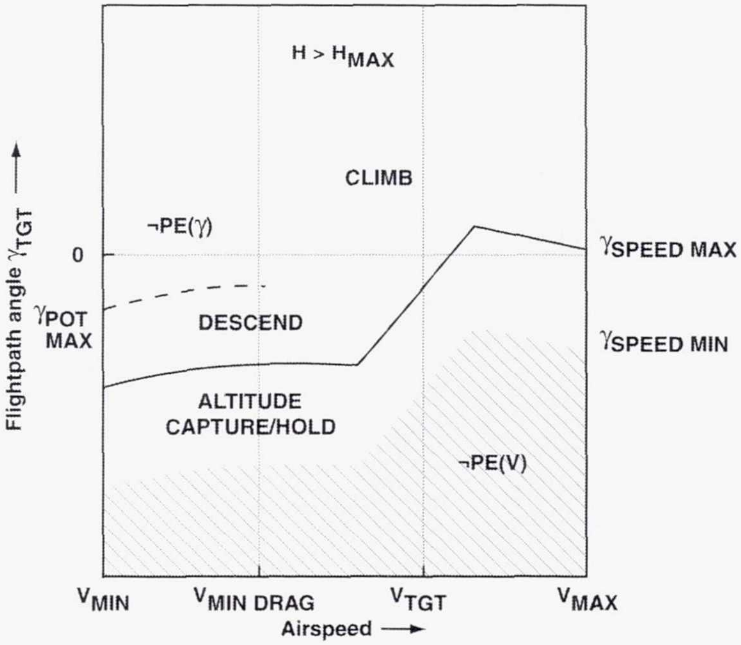
### Markovian Property

This independence of mode selection from mode states previous to the current states is termed the Markovian property of the mode selection strategy (that is, selection of the next set of modes depends only on current modes and on flight conditions). Based on this mathematical property, general theorems can be established that govern system behavior, enabling its safety and effectiveness to be assessed theoretically. Five such theorems are discussed next; these theorems provide a check of system design integrity that is independent of the synthesis process, and enable formal validation of the complete system to be achieved. Proofs of these theorems can be found in appendix F. It should be noted that currently available codes for automated hypothesis testing can provide the basis for a theorem-proving tool; further discussion can be found in appendix F.



a) Normal conditions.

Figure 24. Supermode selection within Altitude Command supermode.



b) Degraded aircraft performance. Thrust target set to maximum thrust ($\gamma_{POT\ TGT} = \gamma_{POT\ MAX}$). $V_{TGT} > V_{MIN\ DRAG}$.

Figure 24. Supermode selection within Altitude Command supermode (concluded).

## Performance Degradation

The first issue to be evaluated relates to the soundness of the design decision to select the Climb/Descend mode in the case where both PE (γ) and PE (V) are violated, so that neither the Climb/Descend mode nor the Altitude Capture/Hold mode is valid. The first theorem shows that this situation can arise only if aircraft performance has become so severely degraded that maximum available thrust is insufficient for level flight:

***Theorem 1 (Performance Degradation Theorem)***

If both PE (γ) and PE (V) are violated, then the condition ($\gamma_{POT\ MAX} < 0$) must hold.

Proof can be found in appendix F.

## Recovery from ¬PE(γ) ¬PE (V)

Severe performance degradation could result from engine failure at high altitude. The aircraft must then descend ("drift down") to a lower altitude at which sufficient thrust is available. Safe recovery from severe performance degradation is demonstrated by the second theorem, which shows that aircraft behavior is near optimal even without intervention by the human crew:

*Theorem 2 (¬PE (γ) ¬PE (V) Recovery Theorem)*

> If both PE (γ) and PE (V) are violated initially, and if the Altitude Command supermode is selected and the target altitude $H_{TGT}$ and the target airspeed $V_{TGT}$ remain fixed, then in the absence of total propulsion failure
>
> **(a)** the condition ($V_{TGT} < V \leq V_{MIN\ DRAG}$) must hold initially;
>
> **(b)** recovery from the condition ¬PE (γ) ¬PE (V) occurs before $V_{TGT}$ is captured;
>
> **(c)** if the conditions ¬PE (γ) PE (V) and TC1 ≡ ($\gamma_{POT\ MAX} < 0$) hold at capture of $V_{TGT}$ and continue to hold thereafter, then the aircraft stabilizes in descent at $V = V_{TGT}$, and the V Command mode remains selected; and
>
> **(d)** if a lower altitude is reached where the condition ($\gamma_{POT\ MAX} = 0$) holds in the long term, then the aircraft stabilizes in level flight at $V = V_{TGT}$.

Proof can be found in appendix F.

Assurance of safe terrain clearance at the final stabilized altitude is provided by appropriate aircraft certification and operation requirements.

It should be noted that the effects of thrust asymmetry on flight control have not been considered, but it is reasonable to suppose that advanced automated control systems in future transport aircraft can incorporate suitable means for handling asymmetric thrust effects, especially those systems with full-authority, fly-by-wire capabilities.

### Recovery from PE(γ) ¬PE (V)

The third theorem shows that violation of PE (V) is self-correcting without intervention by the human crew:

*Theorem 3 (PE (γ) ¬PE (V) Recovery Theorem)*

> If PE (V) is violated initially, but PE (γ) holds, and if the Altitude Command supermode is selected and the target altitude $H_{TGT}$ remains fixed with $H_{TGT} \leq H_{MAX}$, then in the absence of total propulsion failure PE (γ) PE (V) must eventually hold.

Proof can be found in appendix F. The hypothesized condition corresponds to inappropriate setting of the target thrust, so that idle thrust is set when maximum thrust is desired, or the contrary. This condition can be induced by pilot abuses such as entering a lower altitude target during climb or engaging the Altitude Command supermode at an operating point that lies outside the performance envelope of the aircraft. Failures and abuses are discussed in detail in the next section.

### Recovery from ¬PE (γ) PE (V)

The fourth theorem shows that violation of PE (γ) is self-correcting without intervention by the human crew:

***Theorem 4 (¬PE (γ) PE (V) Recovery Theorem)***

> If PE (γ) is violated initially, but PE (V) holds, and if the Altitude Command supermode is selected and the target altitude $H_{TGT}$ remains fixed with $H_{TGT} \leq H_{MAX}$, and if climb effectiveness holds while in the Climb mode, then in the absence of total propulsion failure the condition PE (γ) PE (V) must eventually hold.

Proof can be found in appendix F. The hypothesized condition is classified as abnormal, but holds during steady climb at airspeeds below the speed for minimum drag (for example, just after takeoff). However, because the Climb/Descend supermode is selected during steady climb, PE (γ) is not required for validity. PE (γ) holds when the aircraft approaches the target altitude, provided that the target altitude lies at or below $H_{MAX}$.

The hypothesized condition could also result from engine failure during climb after takeoff, drastically reducing available thrust. However, assurance that the condition ($\gamma_{POT\ MAX} > 0$) continues to hold is provided by the minimum climb gradient required for certification. If engine failure should occur during operation in the Altitude Capture/Hold supermode, for which PE (γ) is required for validity, the system would revert to the Climb/Descend supermode until the aircraft approaches the target altitude more closely, so that the Altitude Capture/Hold supermode again becomes valid. A failure of this kind is discussed in detail in the next section.

### Altitude Command Supermode Effectiveness

The final theorem guarantees the effectiveness of the Altitude Command supermode under general conditions:

***Theorem 5 (Altitude Command Supermode Effectiveness Theorem)***

> If the target altitude $H_{TGT}$ lies at or below $H_{MAX}$ and $H_{TGT}$ remains fixed, and if climb effectiveness holds while in the Climb mode, and if descent effectiveness holds while in the Descend mode, then selection of the Altitude Command supermode ensures that
>
> **(a)** the target altitude $H_{TGT}$ will be captured; and
>
> **(b)** if normal effectiveness of the γ Command mode holds while in the Altitude Capture/Hold mode, then the target airspeed $V_{TGT}$ will also be captured.

Proof can be found in appendix F. It should be noted that violation of any of the effectiveness conditions required by the hypothesis is annunciated to the human crew, and also to any other higher-level entity invoking the Altitude Command supermode. Certainty that such violations will be detected and annunciated as required is provided by the logical completeness properties described previously. Therefore, in the absence of such annunciations, the human crew can have complete confidence in the effectiveness of the Altitude Command supermode.

## Failures and Abuses

Assurance of safe and effective system behavior under all possible conditions is provided by the general theorems just presented. However, it may not yet be clear how the design characterized by these general properties can protect the system against various specific failures and abuses without special-purpose algorithms ("error traps") like those employed in the past to identify specific faults and take remedial action.

This point is crucial for the design of new systems, because the enumeration of specific faults is necessarily incomplete even after extensive operational experience has been accumulated. To illustrate the remedial action provided by general-purpose protective mechanisms, two representative but challenging situations are analyzed in detail.

### Engine Failure After Takeoff

It is assumed that engine failure occurs during climb after takeoff with the Altitude Command supermode engaged and operating in the Altitude Capture/Hold supermode, and that the resulting drastic reduction of available thrust causes the target flightpath angle $\gamma_{TGT}$ commanded by the height regulator to exceed the aircraft performance envelope (fig. 6). It is further assumed that the effects of thrust asymmetry on flight control are dealt with by the continuous elements of the flight control system. Because the condition

$$\neg PE\ (\gamma) \equiv (\gamma_{TGT} \geq \gamma_{POT\ MAX})\ \text{AND}\ (V \leq V_{MIN\ DRAG})\ \text{AND}\ (\gamma \geq \gamma_{SPEED\ MAX})$$

holds, PE ($\gamma$) is violated, causing the Altitude Capture/Hold supermode to become invalid. If the Altitude Capture/Hold supermode were to remain selected and continued operation were attempted, flightpath would be controlled at the expense of a rapid decrease of airspeed, resulting in potentially catastrophic wing stall or other loss of control (fig. 6).

Because the condition $\neg$PE ($\gamma$) PE (V) holds after the engine failure occurs, according to the Altitude Command supermode selection strategy the Climb/Descend supermode is selected because it is the only valid choice, and the system reverts to the Climb mode. In the Climb mode, airspeed is controlled at the expense of flightpath, as expressed by selection of SPEED priority.

After the engine failure occurs, the thrust saturation condition ($\gamma > \gamma_{SPEED\ MAX}$) holds. According to the strategy for setting target thrust, the target thrust $\gamma_{POT\ TGT}$ is set to $\gamma_{POT\ MAX}$, and the primitive V command mode is selected. The system pushes over to capture the flightpath angle $\gamma_{SPEED\ MAX}$, resulting in capture of the target airspeed $V_{TGT}$. Climb then continues in the V Command mode at the flightpath angle $\gamma_{SPEED\ MAX}$, which is the steepest climb angle of which the aircraft is capable at the target airspeed with the maximum thrust available after engine failure (fig. 6). Assurance of safe obstacle clearance at this reduced climb angle is provided by the minimum engine-out climb gradient required for certification. Provided that the target altitude lies at or below the reduced $H_{MAX}$ characterizing the engine-out condition, as the aircraft approaches the target altitude the target flightpath angle commanded by the height regulator is reduced to a value less than $\gamma_{POTMAX}$, the condition PE ($\gamma$) again holds, and PE ($\gamma$) PE (V) holds, as required by the ¬PE ($\gamma$) PE (V) Recovery Theorem. The Altitude Capture/Hold supermode is again selected, and the target altitude is captured.

It should be noted that no identification of engine failure is required to enable the sequence of events just described. To the contrary, the general algorithm for determination of the performance envelope accounts for engine failure, without specific identification of the cause of the reduced performance. It follows that a similar performance reduction resulting from any other cause, such as tailwind shear or failure of landing gear retraction owing to some hydraulic system fault, would result in the same system behavior, and therefore would result in a similar safe recovery.

### Climb Engagement Abuse

It is assumed that, during operation in manual control, the aircraft is pulled up into a steep climb at a flightpath angle exceeding the upper limit of the aircraft performance envelope, and that the Altitude Command supermode is engaged with a target altitude below the altitude prevailing at the instant of engagement. Because the thrust saturation condition ($\gamma > \gamma_{\text{SPEED MAX}}$) holds, according to the strategy for setting target thrust the target thrust $\gamma_{\text{POT TGT}}$ is set to $\gamma_{\text{POT MAX}}$ (table 5). Because the condition Q ≡ ($\gamma_{\text{TGT}} < \gamma_{\text{SPEED MIN}}$) also holds, the partial effectiveness of the V Command mode is violated; that is, the condition

$$\neg\text{PE (V)} \equiv [\text{P AND } (\gamma_{\text{POT TGT}} = \gamma_{\text{POT MIN}})] \text{ OR } [\text{Q AND } (\gamma_{\text{POT TGT}} = \gamma_{\text{POT MAX}})]$$

holds. Therefore, the condition PE ($\gamma$) ¬PE (V) holds.

This abnormal condition results because maximum thrust is required for consistency with the engagement condition, but minimum (idle) thrust is required for consistency with the target condition. Because the condition Q ≡ ($\gamma_{\text{TGT}} < \gamma_{\text{SPEED MIN}}$) holds, if thrust were unsaturated the target thrust $\gamma_{\text{POT TGT}}$ would be set to $\gamma_{\text{POT MIN}}$, but so long as thrust remains saturated at maximum thrust owing to the anomalous engagement condition, according to the thrust setting strategy the target thrust cannot be reset to idle (table 5). If operation in the Climb/Descend mode were attempted, which would be consistent with the target condition, the aircraft would continue to climb with maximum thrust instead of descending with idle thrust as desired.

The Altitude Command supermode selection strategy resolves this problem by selecting the Altitude Capture/Hold supermode because the Climb/Descend supermode is invalid when ¬PE (V) holds (table 12). The system then attempts to capture the target flightpath by pushing over toward $\gamma_{\text{TGT}}$. Since Q ≡ ($\gamma_{\text{TGT}} < \gamma_{\text{SPEED MIN}}$) holds by assumption and the condition ($\gamma_{\text{SPEED MIN}} < 0$) always holds by the property of the longitudinal acceleration limiter, as already noted, the flightpath angle $\gamma$ is driven downward toward zero immediately upon engagement, so that the thrust becomes unsaturated. But when the thrust becomes unsaturated, the target thrust $\gamma_{\text{POT TGT}}$ is reset to $\gamma_{\text{POT MIN}}$ (table 5). Therefore, PE (V) holds, the normal condition PE ($\gamma$) PE (V) holds as required by the PE ($\gamma$) ¬PE (V) Recovery Theorem, and the Climb/Descend supermode becomes valid and is selected. The aircraft then descends in the Descend mode as desired.

The same abnormal condition could be induced after initial engagement if, during climb, the crew entered a target altitude below the instantaneously prevailing altitude, perhaps in response to a request from ATC. A similar abnormal condition could occur during descent, if the crew entered a target altitude higher than the instantaneously prevailing altitude. Whatever the specific details of the abuse leading to the abnormal condition PE (γ) ¬PE (V), assurance of safe recovery is provided by the general PE (γ) ¬PE (V) Recovery Theorem.

### Concluding Remarks

Engineering design of the Altitude Command supermode is now complete. Decision tables and statecharts have been presented that specify the mode selection strategy in forms that, together with semantic conventions described in appendix E, are suitable for exact implementation within single sequential processors like those installed in current transport aircraft. The dynamical behavior of the complete system has been summarized by general theorems that enable formal validation to be achieved.

Furthermore, methodology has been developed for synthesizing mode selection logic directly from design requirements; this methodology can be applied to development of other modes not treated by this report. In the next section, guidelines for simplified development of other second-level and third-level modes are presented briefly.

## OTHER SECOND-LEVEL AND THIRD-LEVEL SUPERMODES

This section presents guidelines for simplified development of other second-level and third-level supermodes based on modifications to the second-level Altitude Command supermode just discussed, and shows how the complete Vehicle Management System could be used in airline service.

### Approach Supermodes

Second-level approach supermodes are needed that are suitable for executing well-defined approach procedures based on ILS facilities, on other ground-based radio aids, or on satellite navigation. These second-level approach supermodes occupy positions within the mode hierarchy similar to that of the Altitude Command supermode, as illustrated by the block diagram of figure 18. If a reliable electronic library of approach procedures (that is, approach plates) were available on board, the flight crew could select the destination procedure by name from a displayed menu; radio frequencies, identifiers, weather minima, the missed approach procedure, and various numerical parameters specific to the selected approach would then be entered automatically into the system.

Tracking of an approach glideslope can be achieved by generalizing the altitude capture and hold function of the Altitude Command supermode (previously described in detail) to define the approach path by means of a dynamically varying target altitude. To avoid an altitude bias error while tracking the glide slope, a commanded flightpath angle equal to the published glide-slope angle must be added to the system within the path regulator (fig. 11(e)). For details, consult a previous publication (Sherry, Youssefi, and Hynes, 1995).

The safety envelope for an approach mode is defined by appropriate limits on lateral and vertical deviations from the intended approach trajectory, and by the published minimum descent altitude. Under normal conditions, approach mode operation should terminate with transition to a flare mode (either manual or automated), or else (under abnormal conditions) with transition to the published missed-approach procedure.

### Go-Around Supermodes

During go-around following a missed approach, initial climb can be achieved in a simple way by specifying targets for climb angle and airspeed, setting priority to SPEED, and invoking the first-level Path/Speed Command supermode. To maximize obstacle clearance, the airspeed target should remain fixed at the final approach airspeed; because no acceleration is then required, all excess thrust is used for climb. The climb angle target can conveniently be set to a fixed angle (perhaps 10 degrees) corresponding to use of maximum thrust at an aircraft weight intermediate between maximum and minimum landing weight. This choice results in use of maximum thrust and maximum available climb angle at heavy weight, and avoids excessive thrust at light weight while ensuring safe obstacle clearance (Lambregts, 1983). After safe obstacle clearance has been achieved, the published altitude at the missed-approach holding fix should be selected as the altitude target, and the second-level Altitude Command supermode should be invoked.

### Departure Supermodes

Second-level departure supermodes are needed that are suitable for executing well-defined departure procedures (Standard Instrument Departures). These second-level departure supermodes can include noise-abatement procedures to be used for departures from specific runways. If a reliable electronic library of such procedures were available on board, the flight crew could select the departure procedure by name from a displayed menu; radio frequencies, identifiers, weather minima, and various numerical parameters specific to the selected procedure would then be entered automatically into the system.

Tracking of the departure flightpath specified for each segment of the departure procedure can be achieved in the same manner as for the Go-Around supermode just described, by specifying targets for climb angle and airspeed, setting priority to SPEED, and invoking the first-level Path/Speed Command supermode. For unrestricted departures, manual operation following takeoff would be continued until the Altitude Command supermode can be selected at 400 ft above ground level.

The safety envelope for a departure mode is defined by appropriate limits on lateral and vertical deviations from the intended departure trajectory. After safe obstacle clearance is achieved, departure mode operation should terminate with transition to the Altitude Command supermode for climb.

### Third-Level Supermodes

At least one third-level Vertical Navigation supermode is needed to enable the aircraft to follow optimized fuel-conservative and time-conservative trajectories during climb and descent (Erzberger, 1982). In general, it must be expected that each third-level supermode would constitute a hybrid system containing both continuous and discrete elements, as pointed out previously. However, in this case such modes can be developed simply by selecting appropriate altitude and airspeed targets from the Reference Flight Path (fig. 1) and invoking the second-level Altitude Command supermode (previously described). By this means, the aircraft can be made to follow any trajectory in the vertical plane (within its performance capability) that is composed by joining straight-line segments in a piecewise-continuous manner, with appropriate smoothing at the joints (control points) between segments. Mathematically, such third-level modes contain only discrete elements, because all hybrid-system aspects of the design are confined to the second level (provided that third-level system dynamical behavior can be approximated as quasi-static, which is assumed).

Nevertheless, if a trajectory with continuous curvature in the vertical plane were generated by varying the altitude target dynamically, the aircraft would follow the curved path with an altitude bias error dependent on path curvature. Elimination of such bias errors would require development of a new path regulator on the third level, instead of making use of the second-level height regulator (fig. 11(b)) already available. The third-level modes would then become hybrid systems, with significant increase of complexity.

It seems likely that curved-path bias errors characterizing the simplified design will be found to be insignificantly small compared with other errors within the modernized ATC system. However, if future ATC system refinement indicates the need, third-level modes that enable an aircraft to follow any feasible path with high accuracy can be developed by applying the synthesis methods described in this report.

## Operational Use of Automated System

Operational use in airline service of the Vehicle Management System described by this report (fig. 1) can be summarized as follows. Takeoff is accomplished manually; it is assumed that all the remaining flight phases are to be automated. At any time, partial automation could be chosen at the pilot's option by selecting augmented manual control. In that case, only the guidance and display functions would be automated; the pilot's manual control task would be to follow the guidance symbology displayed by the primary flight displays.

The initial climb after takeoff would make use of a Departure supermode specialized to the runway in use and selected by name; the specified departure flightpath would include any required noise-abatement procedure.

The climb, cruise, and descent would all make use of a third-level Vertical Navigation supermode, which would select appropriate altitude and airspeed targets for each flight segment from the onboard Reference Flight Path, and would invoke the Altitude Command supermode for execution. Automation of the task of updating the Reference Flight Path with the aid of an air-ground datalink is discussed later, in the section "Future Work."

At the destination airport, the instrument approach would make use of an Approach supermode specialized to the runway in use and selected by name. Landing flare, touchdown, and roll-out would make use of an Autoland supermode similar to those in current use (not treated by this report). For a missed approach, a Go-Around supermode would be used to execute the specified missed-approach procedure.

The flight profile just described shows that, with the exception of takeoff, the whole of a typical airline flight could be carried out by means of only five supermodes selected directly by the flight crew. The simplicity of this mode structure stands in sharp contrast to the complexity and functional duplication (Lambregts, 1983) characterizing the mode structure of the flight management systems installed in current transport aircraft.

# PART III
# CONSEQUENCES

## AIRWORTHINESS, CERTIFICATION, AND COCKPIT INTERFACE DESIGN ISSUES

As stated in the "Introduction," the present work is directed toward three specific objectives: (1) to enable formal validation by developing a practical design procedure for synthesis of hybrid systems that satisfy general safety and effectiveness properties specified a priori; (2) to realize the potential of formal validation for improved operational safety, with particular concern for the design integrity of the complete system, including the human/machine cockpit interface, and to provide a firm technical basis for new certification criteria; and (3) to contribute to the development of more general theoretical methods for synthesizing hybrid systems directly from design requirements.

The first of these three objectives has been achieved. Development of methodology for engineering design of the hybrid system for transport aircraft longitudinal control illustrated by the block diagram of figure 1 is complete. The third objective has been achieved in part by means of this detailed design example (fig. 1); the potential for generalization of the design methodology to other hybrid systems not related to aircraft is discussed later, in the section "Future Work." This section addresses the second objective, by assessing the implications of the design methodology for airworthiness and certification, and for design of the human/machine cockpit interface.

### Airworthiness and Certification Issues

The arguments presented fall into two categories. First, the safety and effectiveness properties imposed on the design of the hybrid system for aircraft control (fig. 1) were not selected arbitrarily. On the contrary, these properties are closely related to design criteria used by the aircraft industry, such as Military Standard 1797 (Anonymous, 1985), and to certification criteria used by regulatory authorities, such as Part 25 of the Federal Air Regulations (Title 14, Code of Federal Regulations, 2001). These criteria are based on more than 40 years of operating experience with both civil and military jet transport aircraft, and specify functional requirements for aerodynamic, structural, and propulsion design. However, for reasons to be discussed, no such broad functional criteria apply to the design of avionic systems.

The authors believe that safety and effectiveness properties similar to those developed in this report can provide a suitable basis for new criteria for design and certification of avionic systems, and that application of such criteria can improve operational safety. This view is supported by analysis of transport aircraft accidents and incidents in which certain actions by the avionic system played a key role, and several such cases are discussed for illustration. These matters are all questions of enginering design, which fall into the first category for discussion.

## Cockpit Interface Design Issues

The second category relates to design of the human/machine cockpit interface. To clarify the human-factors issues involved, an analogy is developed between delegation of a cockpit task to the automated system, and delegation of the same task to a human crewmember. It follows that direct mode selection by the pilot can be regarded as an expression of the pilot's intentions. The discussion shows how the dynamical behavior theorems developed in this report can guarantee achievement of the pilot's intentions under all normal operating conditions; exceptions are required to generate appropriate warnings.

The term trustworthiness is used to characterize this desirable behavior in carrying out an assigned task, and it is argued that today's automated systems often exhibit a lack of trustworthiness that would be recognized as unacceptable if displayed by a human crewmember. Several cases are illustrated, and several heuristic guidelines for system and interface design are proposed. The discussion leads to recommendations that cockpit interface design be integrated into the system design process, and that formal validation be extended to include the cockpit interface.

Before considering the airworthiness and certification issues in detail, two categories of failures involving system and aircraft hardware elements are briefly discussed.

## Hardware Failures

### Sensor and Computer Hardware Validity Requirements

It should be noted that nothing has been said thus far concerning system hardware reliability. Except for a brief discussion in appendix E of certain computational limitations on failure detection, no system errors or failures have been treated, and the synthesis process has been developed under the tacit assumption of ideal system reliability. A complete discussion of these issues lies beyond the scope of this report, and would necessarily be specific to particular systems and aircraft.

However, it is an obvious extension of the concept of mode validity developed previously to add the requirement that for validity of any mode to hold, each of the parameters that enter the computations specified for that mode must be determined with sufficient accuracy. Therefore, sensor or computer hardware failures that violate this requirement must result in mode invalidity. These hardware requirements can be handled within the framework of the synthesis methodology already developed, by conjoining them with other conditions required for mode validity.

### Engine and Aircraft System Failures

In contrast to sensor and computer hardware failures within elements that are internal to the Vehicle Management System, engine failure and other aircraft system failures (within the aircraft hydraulic system, electrical system, and the like) are failures external to the Vehicle Management System to which that system is required to react (fig. 1). Engine failure, which has consequences that are similar for most transport aircraft, has been treated in some detail.

Failures within other aircraft systems have consequences for the Vehicle Management System that are specific to each aircraft type, and cannot be treated in general by means of examples like those presented for engine failure. Nevertheless, when the failure modes and effects are known in detail for a specific aircraft type, aircraft system failures can be handled in a manner similar to that illustrated by the engine failure examples. A simple example based on failure of a single hydraulic system illustrates the process.

### Hydraulic Failure Example

Assume that an aircraft hydraulic system failure renders the yaw damper inoperable while the aircraft is cruising at 35,000 ft, and assume further that this failure makes it necessary to restrict the maximum operating altitude of the aircraft to 25,000 ft. (Otherwise, objectionable motions could result, owing to excitation of the dutch-roll mode at higher altitudes where aerodynamic dutch-roll damping becomes especially weak.) To account for the effect of yaw damper inoperability upon the Vehicle Management System, its detection must cause appropriate restriction of the calculated maximum altitude $H_{MAX}$ of the aircraft (fig. 20). Because by assumption the currently prevailing altitude of the aircraft exceeds $H_{MAX}$, the system is then required (Theorem 5) to generate an appropriate annunciation to the flight crew, prompting the crew to request clearance to a lower altitude.

As pointed out previously, the performance envelope limits of the aircraft must be recalculated continuously by the onboard system to enable target screening. However, after initial system engagement the target screening algorithm is not permitted to change the target altitude value without pilot consent, because that requires a strategic action (obtaining an amended clearance) that is beyond the scope of the tactical system (fig. 1).

As already noted, mode selection depends only on the geometric regions of the performance envelope within which the target point and the operating point lie; mode selection is independent of the sequence in which those regions are traversed (the Markovian property). Therefore, once the aircraft maximum altitude has been restricted appropriately, no other Vehicle Management System modify-cation is required to account for the effect of the assumed hydraulic failure.

## Avionic System Certification

### Background

Historically, the autoflight system (AFS) evolved from the older autopilot. In contrast, the flight management systems (FMSs) in current transport aircraft were developed during the late 1970s. The FMS contains the trajectory synthesis structure corresponding to the third-level supermodes in the system discussed in previous sections of this report, and is much more complex than the AFS. In the first-generation aircraft, weak integration of the FMS with the AFS may have justified their treatment as two separate systems.

The following comments relating to certification issues are based on discussions with the late H. B. (Berk) Greene, FAA certification pilot and system specialist. According to Greene, the regulatory basis for certification of current flight management and flight control systems is that these systems (with the exception of autoland) are regarded as pilot aids that are not to be used under safety-critical conditions (Greene, personal communication, 1994). This long-standing regulatory posture avoids detailed specification of system functionality by placing on the flight crew the burden of determining whether system actions are safe and appropriate, despite mounting human-factors evidence that, faced with the complexity of today's systems, the crew cannot always identify unsafe or inappropriate system actions in time to avoid potentially catastrophic consequences.

### Discussion

It seems to the authors doubtful whether the position that pilot aids are not safety-critical can be maintained for the future, as the level of automation advances and system integration is strengthened. An alternative position that is taken here holds that any violation whatever of the safety and effectiveness properties established by design synthesis should be considered potentially safety-critical.

This alternative view is supported by the analysis of selected transport accidents and incidents that follows. Simple but plausible analysis shows that, although the causes of these accidents and incidents differ greatly in detail, each of them resulted from some violation of the safety and effectiveness properties developed in this report. Furthermore, these problems have not been alleviated by experience, because the most recently designed systems are characterized by violations of safety and effectiveness properties similar to those encountered in previous design generations.

At this point of the discussion readers should refer to appendix G, which summarizes the aspects of four selected accidents and incidents that are considered relevant to issues of system design. No criticism of the design or certification of the systems installed in those aircraft is intended or should be inferred, because the methods employed were based on the state of the art available at the time.

Four aircraft have been selected, with installed systems that represent each avionic design generation since the first digital systems. To relate the systems in the aircraft involved in these accidents and incidents to those synthesized in the previous sections of this report, it should be noted that the AFS in current transport aircraft contains the flight control structure that corresponds to the first-level Path/Speed Command supermode and the second-level Altitude Command supermode discussed previously; the FMS in current transport aircraft corresponds roughly to the third-level supermodes.

### L-1011, Everglades, 1972 (Analog System)

Since the barometric hold mode was engaged directly by the pilot and its purpose was to hold altitude, a reasonable interpretation of effectiveness for that mode would require that barometric altitude be held accurately unless the mode were disengaged intentionally by the pilot. (Bumping the control column is not a suitable means of disengagement because of the possibility of accidental disengagement, and violates the first and second of the seven cockpit interface guidelines to be presented later.) From this viewpoint, it is clear that, because the target altitude was not held, the reversion from the barometric altitude hold mode to the pitch attitude hold mode in the aircraft violated the effectiveness of the barometric altitude hold mode. Furthermore, a crucially important safety property was violated, because (from the present viewpoint) penetration of an appropriate safe descent limit should have caused reversion to an envelope protection mode, with an appropriate warning.

### B-767, San Francisco, Late 1980s (First-Generation Digital System)

Since the supermode engaged by the crew was intended to result in capture of the target altitude, failure to transition to the altitude capture mode violated supermode effectiveness, because the target altitude was not captured. Furthermore, the safety hazard resulting from the altitude overshoot was compounded by the absence of any annunciation.

Detailed analysis shows that the aircraft mode transition logic contained three separate faults. First, no reasonableness test can be logically complete, because only a subset of invalid signals can be detected—logically complete failure detection requires at least duplex redundancy. Second, the numerical criterion used for the reasonableness test was incorrectly related to the performance capability of the aircraft, generating false invalidities. Third, the action taken (that is, skipping the transition test) was logically erroneous, because it permitted continued operation in the climb mode under invalid conditions. The presence of these deficiencies in a certificated system shows the weakness of current avionic certification standards.

### B-737, Early 1990s (Second-Generation Digital System)

Since the ILS approach mode is engaged by the crew in order to track the ILS glide-slope, reversion from the ILS approach mode to the vertical velocity hold mode following failure of the glide-slope receiver in the aircraft violates the effectiveness of the ILS approach mode, because the ILS glide-slope is not tracked.

It should be noted that operation in the vertical velocity hold mode in the aircraft has no exact counterpart in the system described in this report because an "open" descent of this kind is not allowed, but it corresponds roughly to operation in the Altitude Command supermode with the target altitude set to ground level, which could cause the aircraft to fly into the ground. In the system described in this report, such operation would violate two safety properties: first, the validity condition that restricts the target altitude to lie above 1500 ft above ground level; and second, the validity condition that restricts operation in the Altitude Command supermode to measured altitudes above 400 ft above ground level. These safety violations would generate appropriate warnings.

To summarize, glide-slope receiver failure violates the effectiveness of the approach mode in the aircraft without generating an appropriate warning, and subsequent operation in the invalid vertical velocity hold mode following reversion is potentially catastrophic.

### A-330, Toulouse, June 1994 (Third-Generation Digital System)

Since the supermode engaged by the crew was intended to result in capture of the target altitude, it is clear that continued operation in the altitude capture mode after the engine cut violated supermode effectiveness, because the target altitude was not captured. The discussion of the dynamical behavior of transport aircraft presented previously shows that in this situation (path target exceeding maximum steady climb angle), the airspeed must decay until either an underspeed protection mode is invoked or else some flight envelope boundary is violated, unless path control is given up temporarily in favor of airspeed control. In the case of the A-330 accident, the airspeed penetrated the minimum-control-speed boundary for asymmetric thrust (fig. 6), resulting in catastrophic loss of control.

As pointed out previously, the availability of envelope protection modes should not be regarded as reducing the designer's obligation to ensure that the normal modes operate safely. Stated less formally, the authors' position is that, since the envelope protection modes are the last line of defense against potentially catastrophic system behavior (like a parachute), they should not also be the first line of defense. Restricting attention to operation of the normal modes, the A-330 accident can be considered to have resulted from flawed mode selection logic that permitted continued operation in the altitude capture mode following a loss of thrust which rendered that operation invalid. It is also true that, in the case of the A-330 accident, envelope protection modes were not effective for accident prevention.

### Summary

Analysis of the transport aircraft accidents and incidents just presented has shown the broad applicability of airworthiness principles based on the safety and effectiveness properties developed in this report. It follows that new certification criteria using these principles might significantly improve operational safety.

However, realization of the potential for improving operational safety depends on effective communication of system annunciations and warnings to the human crew. The discussion turns next to cockpit interface design.

## Cockpit Interface Design

### Design of Annunciation Systems

A major challenge for cockpit interface design is to develop interfaces that make all necessary information relative to aircraft and system operation available to the human crew, without overloading them or interfering with higher-level tasks. The design synthesis methodology already developed might aid design of improved interfaces, if interface design were based on the required annunciations emerging from the design synthesis process. Because these annunciations constitute essential safety-critical information, redundant information might be eliminated and information overload reduced.

Furthermore, the assurance of logical completeness provided for the annunciation system by formal validation might enable the crew to limit their attention to higher-level system elements while retaining confidence that, in the absence of annunciations, lower-level details of aircraft and system operation are being properly managed; the concept has been termed management by exception (Billings, 1996). By this means, the mental workload associated with system monitoring and supervision might be reduced. These are issues for interface design that are broader than that of information transfer.

### Delegation of Authority

As defined previously, a mode consists of a set of actions that the machine can take; that is, a physical behavior. From the human-factors viewpoint, the pilot selects a mode in order to obtain its characteristic set of actions, so that mode selection constitutes an expression of the pilot's intentions. At the same time, by selecting a mode, the pilot delegates to the system the authority to perform that set of actions. Thus, by analogy, mode selection can be regarded as a command given to the system, just as if it were given to a human crewmember.

### System Behavior

This viewpoint immediately raises the issue of the dynamical behavior of the automated system following the pilot's mode selection, because it is this subsequent behavior that determines whether the pilot's intentions will be carried out. For example, the actions characteristic of the Altitude Command supermode involve capturing (and subsequently holding) the altitude and airspeed targets selected by the crew, as previously described in detail. Therefore, selection of the Altitude Command supermode indicates to the system that the pilot intends to capture these targets.

Assurance that the system will carry out this task properly (that is, assurance of system effectiveness) is provided by Theorem 5 (Altitude Command Supermode Effectiveness Theorem). Conditions sufficient to ensure that the theorem holds are stated explicitly in the hypothesis, and violation of any of these conditions is required to generate an appropriate warning. Certainty that such violations will be detected and annunciated as required is provided by the logical completeness properties described previously. Therefore, in the absence of such annunciations, the crew can have complete confidence in the effectiveness of the Altitude Command supermode.

The term trustworthiness is used to characterize this desirable behavior in carrying out an assigned task, since this term seems equally applicable whether the task is assigned to the automated system or to a human crewmember. With this definition, the discussion that follows shows that today's automated systems often exhibit a lack of trustworthiness that would be recognized as unacceptable if displayed by a human crewmember. This anthropomorphic viewpoint is an interesting one from which to study the accidents and incidents previously analyzed from the system perspective.

### L-1011, Everglades, 1972 (Analog System)

In the L-1011 accident, the system exhibited a lack of trustworthiness because it failed to hold the assigned altitude after being instructed to do so. (It decided to hold pitch attitude instead.) The resulting safety hazard was compounded by absence of an appropriate annunciation; the mode reversion should have generated a warning, but the annunciation in the aircraft was cautionary in nature and was not recognized by the crew.

### B- 767, San Francisco, Late 1980s (First-Generation Digital System)

In the B-767 incidents, the system exhibited a lack of trustworthiness because it failed to capture the assigned altitude after being instructed to do so. (It decided to wait and see whether the vertical velocity, erroneously considered invalid, would become valid again.)

### B-737, Early 1990s (Second-Generation Digital System)

In the B-737 incidents, the system exhibits a lack of trustworthiness because it cannot track the ILS glide-slope after glide-slope receiver failure, and does not provide an effective warning to the pilot. (It decides to hold vertical velocity instead, and provides only a cautionary annunciation—compare with the L-1011 accident.)

It should be noted that, by selecting the ILS approach mode, the pilot has expressed to the system the intention to make an ILS approach. Safety considerations based on terrain clearance require that such an approach must terminate in one of only two ways: under normal conditions, by transition to landing flare; under abnormal conditions, by transition to go-around. Therefore, reversion to the vertical velocity hold mode is inconsistent with the pilot's intentions, and should not be permitted under certification criteria based on safety and effectiveness properties.

It is true that following glide-slope receiver failure the pilot might decide to continue the approach to the localizer-only minimum altitude, and might use the vertical velocity hold mode for that purpose if "open" descent were allowed during final approach. Nevertheless, the descent velocity required in that case would not generally coincide with that chosen by the system, which holds the vertical velocity that happens to prevail at the instant of failure, perhaps during a downward correction to the glideslope. In that case, continued operation in the vertical velocity hold mode would cause the aircraft to fly into the ground short of the runway.

On the other hand, the reported ceiling and visibility might be below localizer-only minimums, so that the appropriate decision would be to go around. The system behavior here goes beyond a simple lack of trustworthiness, greatly exceeding the authority properly delegated to it by the pilot's selection of the ILS approach mode.

### A-330, Toulouse, June 1994 (Third-Generation Digital System)

In the A-330 accident, the system behavior went beyond a simple lack of trustworthiness, because it blindly commanded the altitude capture trajectory appropriate for normal operation with both engines, even after it became clear that this trajectory was beyond the performance capability of the aircraft with one engine out. The catastrophic loss of airspeed resulted.

The resulting safety hazard was compounded by the absence of mode annunciations from the primary cockpit displays at the most critical time. (The system had decided not to bother the flight crew with mode annunciations at excessively steep nose-high pitch attitudes, in order to avoid overloading the crew with unimportant details at a critical time.) The information removed by de-cluttering the primary display was crucial, and its absence may have fatally delayed the pilot's decision to take over manually. This de-cluttering behavior exceeded the authority properly delegated to the system by the pilot's mode selection.

### Discussion

No doubt anthropomorphic comments like those above seem inappropriate when they are applied to automated systems. Nevertheless, they may perhaps drive home the point that pilot "aids" like those just described cannot reasonably be considered *not* to be safety-critical.

### Summary

In all the cases just analyzed, poor annunciation compromised effective supervision of system actions by the human crew. The analysis suggests strongly that the design integrity of the cockpit interface cannot be separated from the design integrity of the underlying aircraft systems. If those systems are not trustworthy in carrying out the pilot's intentions as expressed by mode selection, improved information transfer at the interface cannot remedy those basic system deficiencies.

## Design Recommendations

In order to realize the potential for improved operational safety made possible by the design synthesis methods developed in this report, design of the cockpit interface must be treated as an integral part of the system design process, and current statements of human-centered automation philosophy (Billings, 1996) must be expanded and combined with synthesis methods to form complete system design and certification criteria on which new airworthiness standards can be based. In the meantime, the rationale just presented leads to several heuristic guidelines for system and interface design.

Because mode selection expresses the pilot's intentions directly, the system should never attempt to infer those intentions from observation of aircraft states such as configuration, thrust setting, control deflection, or the like (although *limit* manual control deflection should probably cause reversion to fully manual control, as it does for many current transport aircraft). For the same reason, the system should not change the pilot's selection by automated mode reversion, unless violation of safety limits invokes envelope protection. In that case, appropriate warning should be annunciated, and the system should return to the pilot's selected mode if it again becomes valid.

This prohibition of automated mode reversion does not apply to automated transitions between lower-level elements of a supermode selected by the pilot, because those transitions are included within the set of characteristic actions to be expected by the pilot when selecting that supermode. For example, suppose the pilot selects the second-level Altitude Command supermode. Then automated transitions among its elements Climb, Descend, Altitude Capture, and Altitude Hold are expected actions for which authority has been delegated by the pilot's supermode selection, although they should of course be annunciated as they occur. On the first level of the mode hierarchy, the Path/Speed Command supermode is automatically invoked to perform the necessary control tasks, and transitions between its elements (the three primitive modes) should also be expected. It seems clear that, when selecting a mode, the pilot needs to know exactly what actions to expect over the entire mode hierarchy invoked by that selection.

### Anticipating Mode Transitions

To assist the flight crew in actively supervising operation of the system instead of reacting passively to its initiatives, it seems desirable to provide symbology integrated with the primary flight displays that enables the crew to anticipate at least the primitive mode transitions. The Bray/Hynes head-up display (Bray, 1980; Hynes, Franklin, Hardy, Martin, and Innis, 1989) may be regarded as presenting to the crew a vertical cross-section through the aircraft performance envelope (fig. 25) at the prevailing speed, so that the fundamental dynamical behavior of the aircraft is made visible. (Readers familiar with the well-known Flight Dynamics head-up display currently in airline service will recognize its similarity to the illustrated concept; this similarity results from the fact that NASA-Ames flight research provided a basis for the commercial display development.) As illustrated by figure 25, the symbology required for mode anticipation fits naturally into this flightpath-centered display concept.

### Head-up Display Symbology

Figure 25 illustrates display indications during the maneuvers involved in the altitude capture example discussed previously in detail. As noted previously, the corresponding transition paths in the $(V, \gamma)$ plane are illustrated by the diagrams of figure 14, which for clarity are repeated at the right side of figure 25 with the same vertical (flightpath) scale as that for the diagrams on the left. (In the aircraft cockpit, head-up display scaling is determined by the requirement for conformality of the display with the outside world view.)

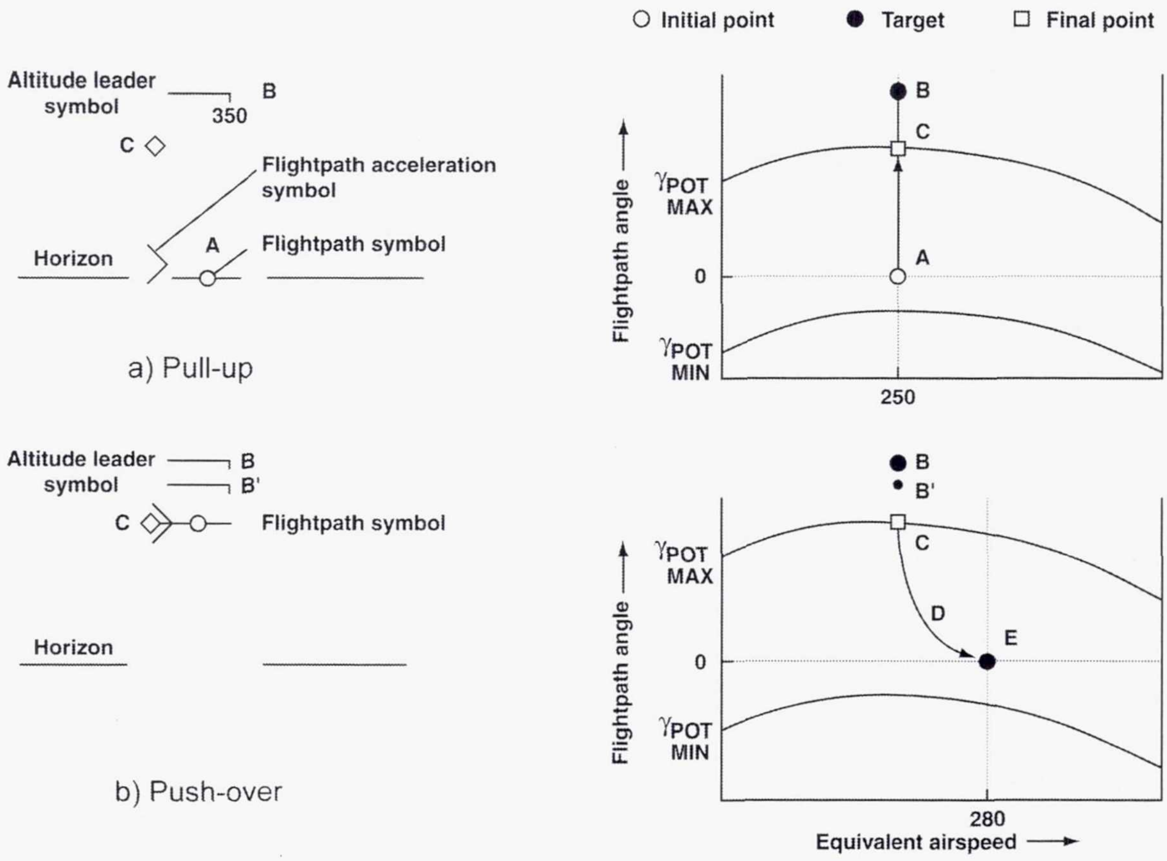


Figure 25. Head-up display symbology.

On the head-up display (left-hand diagrams of figure 25), the flightpath symbol represents the aircraft velocity vector; it lies above the horizon line by an angle equal to the flightpath angle $\gamma$ (fig. 11(a)). The flightpath acceleration symbol (the chevron) at the left of the flightpath symbol displays $\gamma_{POT}$ (fig. 11(a)), which is defined by the expression $\sin \gamma_{POT} = (T - D)/W$ (eq. (1b). The angular difference $\gamma_{POT} - \gamma$ corresponds to the normalized acceleration along the flightpath (eq. (1c). The maximum steady-state climb angle $\gamma_{POT\ MAX}$ available with maximum thrust is indicated by the diamond symbol, and the flightpath angle target $\gamma_{TGT}$ is indicated by the altitude leader symbol (fig. 11(a)), which also displays the altitude target numerically (fig. 25(a)).

## Altitude Capture Example

At the start of the altitude capture example, an aircraft in level flight at 15,000 ft is cleared to climb to 35,000 ft at an airspeed of 250 kt. When the pilot enters the target altitude of 35,000 ft (Flight Level 350) into the mode control panel, the numerals 350 are displayed by the altitude leader symbol (fig. 25(a)). Since the altitude error relative to the target altitude is initially large, the altitude leader symbol is limited to a position (point B, fig. 25(a)) 3 degrees above $\gamma_{POT\ MAX}$ (fig. 11(b)). When the Climb mode is engaged, the diamond symbol appears on the display (point C), indicating $\gamma_{POT\ MAX}$. (If the Descend mode were engaged, the diamond would indicate $\gamma_{POT\ MIN}$.) The diamond is the only symbol added to the head-up display to enable mode anticipation, minimizing display clutter. Since the aircraft is initially in level flight, the flightpath symbol lies on the horizon (point A, fig. 25(a)), and the primitive $\gamma$-V Command mode is selected.

During the initial pull-up at constant airspeed with increasing thrust, the flightpath symbol and the chevron move upward together toward the altitude leader as the $\gamma$-V Command mode attempts to capture $\gamma_{TGT}$ at point B (fig. 25(a)). When the thrust reaches its maximum limit, the chevron coincides with the diamond at point C, and the system transitions to the V Command mode. The pilot can anticipate this primitive mode transition by observing approach of the chevron toward coincidence with the diamond (fig. 25(a)). During the steady climb, the flightpath symbol continues to coincide with the diamond, and the V Command mode remains selected.

As the aircraft approaches the target altitude, the altitude error is reduced (fig. 11(b)), and the altitude leader symbol moves downward (point B') off its limit toward the diamond (point C, fig. 25(b)). When the altitude leader symbol coincides with the diamond (point C), the system transitions to the Altitude Capture mode, and the diamond vanishes. Again the pilot can anticipate this mode transition. Upon transition to the Altitude Capture mode, the airspeed target is reset to 280 kt, as explained previously. During the push-over, the altitude leader symbol and the flightpath symbol move downward together, as the aircraft follows the curved trajectory CDE shown at the right of figure 25(b). When the altitude error is reduced to 300 ft, the altitude leader symbol vanishes to indicate transition to the Altitude Hold mode (fig. 11(b)).

This example shows that, with the symbology described, the pilot can anticipate mode transitions and maintain mode awareness during typical maneuvers by observing the primary flight display, without scanning separate flight mode annunciators. This display was evaluated favorably by NASA and industry test pilots during an exploratory simulation carried out on the NASA-Ames Vertical Motion Simulator (Sherry, Youssefi, and Hynes, 1995). Based on that preliminary evaluation, it seems possible to extend the head-up display concept, including mode anticipation, to conventional head-down primary displays. Further research is needed to verify this conjecture.

### Summary of Cockpit Interface Design Guidelines

The guidelines for cockpit interface design can be summarized as follows:

1. The flight crew should be required to express their intentions clearly by direct mode selection.
2. The automated system should never attempt to infer pilot intention from observation of system states such as aircraft configuration or thrust setting.
3. Automated mode reversion should never over-ride the pilot's mode selection unless an envelope protection mode is invoked for safety.
4. In that case (#3), appropriate warning should be provided, and the system should revert to the pilot's selected mode if that again becomes valid.
5. This limitation on automated mode reversion (#3) does not apply to sub-modes within a pilot-selected supermode, because transitions between such sub-modes are characteristic actions for which authority has been delegated by supermode selection.
6. Symbology integrated with primary flight displays should enable pilots to anticipate primitive mode transitions.
7. Training should familiarize pilots with characteristic actions to be expected for each pilot-selected mode or supermode.

# FUTURE WORK

The elements of the Vehicle Management System treated in detail by this report are indicated by the shaded blocks in the diagram of figure 1, as previously explained, and the elements for which guidelines for simplified development are presented are indicated by cross-hatching. On the right-hand (Tactical) side of the diagram, it can be seen that the continuous elements treated are the guidance function, the outer-loop control, the inner-loop control, and the aircraft itself together with its control surface and throttle servos; the discrete elements treated consist of the mode control logic for the three levels of the mode hierarchy. The navigation function and the corresponding elements for lateral-directional control can be treated by applying the synthesis methodology developed in this report.

It can be seen that, on the right (Tactical) side of the vertical dashed line in figure 1, the two remaining elements are, first, the cockpit interfaces grouped within the display function, and second, the human crew's role in tactical supervision. The first part of the following discussion is concerned with the human-factors research needed to model the cockpit interfaces and the human crew's role in tactical supervision. As previously explained, application of the synthesis method described in this report to the aircraft guidance and control systems results in drastic simplification of mode structure. This simplification can be expected to clarify the problems involved in modeling the crew's role; such modeling could enable complete integration of the human-automation system to be achieved on the tactical level.

On the left (Strategic) side of the diagram (fig. 1), the data-entry task (accomplished manually in current aircraft) is discussed, and the need for automation is pointed out. Brief comments are offered on automation of knowledge-based activities. Finally, the potential for generalizing the synthesis methodology developed in this report to other safety-critical systems such as those for petrochemicals and nuclear power is discussed in some detail.

## Human-Automation Integration

### Tactical Supervision

At a later stage, it may become feasible to extend formal validation to the complete human/machine interface by focusing on the human crew's activities in selecting system modes so as to enable the aircraft to fly a clearance that has already been accepted. Because this selection process is rule-based (after the clearance has been accepted and the desired aircraft trajectory is known), mode selection could, in principle, be fully automated. Alternatively, mode selection can be partially or fully manual, as in current aircraft.

It may be possible to capture the essential aspects of this rule-based pilot behavior by modeling the human pilot's activities in selecting modes and monitoring displays, in effect treating the pilot at the rule-based level as part of the machine. To represent the pilot's behavior realistically, the model must allow mistakes, such as pressing a wrong button, failing to notice an inconspicuous display, skipping a checklist item, or the like. Assessment of the complex system dynamical behavior characterizing modal transitions could then include the contribution of likely pilot errors, enabling formal proof of design integrity to be extended to include the complete human/machine system.

Studies directed toward modeling human supervisory activities would require an interdisciplinary approach coordinating the efforts of human factors research and engineering design groups. Although the results to be expected must remain speculative for the present, the proposed human modeling effort seems worthwhile because pilot behavior at the tactical level is rule-based, and is, furthermore, tightly constrained by the definition of the mode selection/annunciation machine interface (fig. 1).

It should be noted that studies of crew interaction with flight management systems up to this time have necessarily been based on the systems in current aircraft, despite the obstacles that excessive system complexity and flawed design concepts (Lambregts, 1983) have placed in the path toward successful modeling of crew behavior. As mentioned in the "Introduction," system architectures capable of formal validation must make use of a drastically simplified mode structure. Piloted simulation studies of crew errors in the use of such simplified systems known to be free of logical design flaws may be expected to clarify the modeling problem at the tactical level. An example based on the Mode Control Panel specification is presented next to show how human-factors cockpit interface design issues might be identified and resolved by low-level modeling of tactical supervision tasks.

### Mode Control Panel Design Example

The tactical supervisory task selected for analysis is that of manual entry of a desired target altitude into the Mode Control Panel, and of checking the resulting system setup for correctness. It will be shown that even this apparently simple task is vulnerable to human error induced by task interruption, with potentially catastrophic consequences. Furthermore, such errors are a direct result of well-known limitations on the capacity of human short-term memory, and it must be expected that such errors will be committed in flight unless specific precautions are taken to prevent them.

For example, system redundancy could be increased operationally by requiring both pilots to verify the system setup independently whenever the target altitude is changed, but this requirement would increase both task complexity and human workload. Alternatively, by adding to the mode control panel specification the features of display blinking and time-out described in a previous section, the automated system can be made robust against expected operator errors. As mentioned previously, two hypothetical scenarios will show the purpose of these features.

First, assume that the aircraft is in level flight at 10,000 ft with the Altitude Command supermode engaged, and that the pilot wishes to change the target altitude from 10,000 ft to 35,000 ft (Flight Level 350). After setting 35,000 ft into the Mode Control Panel altitude window by turning the altitude knob (fig. 19), the adjacent ENTER button must be pressed to enter the new altitude target into the system. Suppose now that the pilot is interrupted by another task (such as a radio call) before the ENTER button has been pressed. If, owing to a lapse of short-term memory, the pilot forgets that the ENTER button has not been pressed, the setup appears to be correct.

Nevertheless, the aircraft continues to be controlled to the previous internally stored altitude target of 10,000 ft, which differs drastically from the number displayed in the altitude window (35,000 ft), but is not displayed anywhere. If high terrain lies ahead, this situation is potentially catastrophic. The blinking feature added to the specification (figs. 19(b) and 19(d)) provides protection by causing the altitude displayed in the altitude window (35,000 ft in this example) to blink whenever it differs from the internally stored altitude target, alerting the pilot to the fact that the system setup is incomplete. Furthermore, the specification ensures that blinking can cease only when the ENTER button is pressed, which replaces the previous target (10,000 ft) with the new target (35,000 ft), correcting the discrepancy.

A second example based on likely human error illustrates the purpose of the time-out feature. The airspeed target is selected by means of a window and a knob on the Mode Control Panel similar to the altitude window and knob, and adjacent to them (fig. 19(c)). Suppose that the pilot wishes to select a new airspeed target, but turns the altitude knob by mistake. Pilots often use the altitude window as an electronic scratch pad in which to record the clearance altitude. In the absence of a back-up record, the pilot's hypothetical error has destroyed the only onboard record of the clearance altitude. Since reliance on unsupported memory for clearance altitude is potentially catastrophic, the crew has no alternative to an embarrassing query to ATC.

The time-out feature added to the specification (figs. 19(b) and 19(d)) requires that the ENTER button be pressed within 10 sec after the altitude knob is turned; otherwise, the system returns to its previous state, and the previous altitude target is again displayed in the window. In the hypothetical example, the pilot has to wait only a few seconds for time-out to recover the clearance altitude destroyed by mistake. More generally, the time-out feature protects the system against entry of stale data (that is, data no longer operationally relevant) by requiring a connected sequence of actions to take place within a limited time interval.

In both of the examples just discussed, the human errors involved resulted from well-known human limitations that could, in principle, be modeled, and their consequences for system design assessed. It seems likely that research efforts in those directions could enable the need for new features such as blinking and time-out to be identified by a computational synthesis process, instead of relying on experience and ingenuity as at present. The general objective of this research would be detection of usability problems that might otherwise be overlooked until their presence is revealed by incidents or accidents in flight. Furthermore, synthesis methods similar to those described in this report could achieve formal validation of the complete automated system, including the cockpit interfaces, guaranteeing that the addition of features such as display blinking and time-out does not introduce new deficiencies into the design.

### Strategic Planning

On the left (Strategic) side of the block diagram (fig. 1), the planning activities illustrated are in general knowledge-based because they require cause-and-effect understanding of whole disciplines such as meteorology and airline flight operations, as explained previously. However, the rule-based tasks involved in data entry constitute an important exception.

In current aircraft dependent on voice radio communications for transmission of ATC clearances between ground and aircraft, those data-entry tasks must be accomplished manually. As a practical consequence, the flight crew cannot afford the time required for manual data entry in a radar-vectored environment with rapid and unpredictable clearance changes, a situation often encountered during descent into busy terminal areas (Casner, 1994). Under such conditions, the crew is forced to abandon flight-plan updating (accomplished by means of the control display unit (CDU)), and to fall back on second-level (autoflight) operation (accomplished rapidly by means of the mode control panel (MCP)). Efficiency is exchanged for simplicity of operation.

The prospective availability of air-ground communications by two-way datalink will enable elimination of most manual data-entry tasks by means of appropriate automation, allowing automation of the entire Trajectory Synthesis function (left side of figure 1) throughout the flight. (Indeed, with selection of appropriate third-level guidance supermodes, operation of the entire tactical system (right side of figure 1) could then be automated, except the pilot's role supervising execution of the flight plan.) Trajectory synthesis algorithms are well understood (Erzberger, 1982), and form the basis of flight management systems in current aircraft.

In the opinion of the authors, low-level, labor-intensive, error-prone data-entry tasks are inappropriately placed on the human operator in current aircraft. It should be noted that these low-level tasks are sometimes viewed as beneficial because they increase the pilot's involvement in planning. The authors disagree with that view—automation of low-level tasks allows the flight crew to concentrate on the high-level strategic planning and checking tasks (left side of figure 1) that require human judgment. Because the data-entry tasks are rule-based, the necessary automation could in principle be developed by taking the same human modeling approach already discussed for the interfaces involved in tactical supervision.

In contrast, crew modeling at the planning level would involve knowledge-based behavior, and its full scope would extend over the entire strategic flight planning and execution functions. Research directed toward limited-scope decision aiding is currently extending the crew modeling knowledge base (Geddes, 1985). However, within the dynamically varying environment of worldwide flight operations over the 20–25 year service lifetime expected of a new aircraft model, knowledge-based modeling sufficiently comprehensive to provide a basis for formal validation seems too broad in scope to be attempted with present knowledge.

## Generalization to Other Safety-Critical Systems

The question can be raised of the extent to which the design method described in this report could be generalized to enable the solution of hybrid-system control problems in other safety-critical systems, such as the ATC system, or highly automated industrial process-control plants for petrochemicals and nuclear power. No attempt has been made to study this matter systematically, but it seems obvious that most of the steps in the design method discussed previously (fig. 5) could be generalized. One important limitation results from the use of geometrical methods to partition the state space.

### Number of dimensions

Since geometrical methods are most useful when the number of dimensions is small, it is of interest to review briefly the steps by which the six-degree-of-freedom aircraft problem is reduced to two dimensions (appendix B). First, the bilateral symmetry of the aircraft separates its six-degree-of-freedom system of equations into two independent groups, each of three degrees of freedom. Next, the three-degree-of-freedom longitudinal equations are reduced to two degrees of freedom by the assumption that pitch control can generate any desired lift variation without saturation. This assumption is justified for transport aircraft designed to meet applicable flying qualities criteria for manual control, because pitch bandwidth and control power are much higher than required for low-bandwidth automated control of path and speed, especially when modest acceleration limits are imposed for passenger comfort. (High-bandwidth tasks such as automatic landing are not considered in this report.) This pitch-control assumption reduces the longitudinal system from three to two degrees of freedom by eliminating the pitching-moment equation, and enables the aircraft to be treated as a point mass concentrated at its center of gravity. (Simulation of the three-degree-of-freedom system shows response differences that are quite insignificant relative to the response of the two-degree-of-freedom system assumed (appendix B)).

The addition of closed-loop control of path and speed to the basic aircraft could result in additional degrees of freedom associated with differential equations specified by the control laws. However, the nonlinear inverse control employed here cancels aerodynamic and propulsion feedbacks, so that the aircraft in cascade with its inverse is reduced to a simple chain of integrators. In consequence, the specification of primitive modes characterized by the same dynamical behavior as the basic aircraft results in control laws that do not contribute additional degrees of freedom. It should be noted that only this closed-loop behavior is of interest for the design of the hybrid system.

Higher-level modes in the system hierarchy (for example, height control) involve additional degrees of freedom, but separation of the system into hierarchical levels with different time scales (that is, higher-level modes are characterized by less rapid motions) avoids the need to consider all the degrees of freedom at the same time.

All these methods of reducing the number of dimensions can be expected to apply to other systems not related to aircraft. Thus the practical limitation to a small number of dimensions may not prove to be so great an obstacle to generalization of the method as initial impressions might suggest. The same is true with regard to the restriction of desired (target) values of response states so as to lie within the steady-state performance envelope of the aircraft.

### Definition of System Properties

It seems likely that specification of the desired safety and effectiveness properties will be the most difficult element of the design process to generalize. This specification must make a sharp distinction between acceptable versus unacceptable system behavior. On the one hand, from the system design viewpoint the defined properties should be strong and specific, enabling proof of safety and effectiveness theorems that provide assurance of desired system behavior while ruling out behaviors that are undesired or unsafe. If the system can misbehave without generating a warning based on violation of a defined property, then the guarantees provided by formal validation become worthless.

On the other hand, from the human-factors viewpoint the system behavior theorems should hold under weak, general conditions, reducing to a minimum the number of violations that must be detected and annunciated. Unless violations of safety and effectiveness properties are isolated events and the significance of each annunciation is clear, the human operator could be presented with an impossibly demanding task.

Finding an appropriate balance between these conflicting objectives will doubtless require an iterative design process, and the contribution of formal validation to overall design success can be evaluated only a posteriori. It is the authors' hope that the present results may contribute by example to the development of more general theoretical methods for synthesis of hybrid systems. They should also be of practical interest to designers of next-generation transport aircraft avionic systems.

## CONCLUSIONS

Methods have been demonstrated for synthesizing hybrid systems for aircraft longitudinal control directly from design requirements. Systems were developed that satisfy general safety and effectiveness properties, enabling formal validation to be achieved. Three primitive modes were developed for control of path and speed that correspond to the three kinds of dynamical behavior characterizing transport aircraft. Validity conditions for each of these primitive modes were derived directly from the governing differential equations and from safety limits imposed a priori, and were related geometrically to the aircraft performance envelope. The three primitive modes were then combined to form the lowest-level hybrid system. Mode transition logic within this lowest-level supermode was synthesized from the validity conditions, and was simplified by heuristic arguments based on general notions of maximum effectiveness and logical dominance.

In cases where aircraft performance limitations preclude capture of path or speed targets, the exact nature of the limitation is annunciated for resolution at higher levels within the mode hierarchy, or by the human crew, with assurance of logical completeness. Extension to higher-level supermodes was described.

At all times, selection of the next set of modes depends only on currently selected modes and on prevailing flight conditions, and is independent of past history. It was shown that, following engine failure, correct mode selection does not require specific identification of engine failure.

Airworthiness and certification issues were discussed, and analysis of selected transport accidents and incidents led to the suggestion that new airworthiness criteria based on general safety and effectiveness properties might improve operational safety.

These general properties might also provide a basis for design of improved interfaces between aircraft guidance and control systems and the human crew, enabling the crew to limit their attention to higher-level system elements while retaining confidence that lower-level details of aircraft and system operation are being properly managed. By this means, the mental workload associated with system monitoring and supervision might be reduced.

Regarding the pilot's mode selection as a command given to the system enabled analysis of the same accidents and incidents from the human-factors perspective, which suggested that cockpit interface design should be treated as an integral part of the system design process. Several heuristic guidelines for system and interface design were presented.

The need for future work related to human-automation integration was discussed, and it was suggested that modeling the human crew's role in tactical supervision might enable complete integration of the human-automation system to be achieved on the tactical level.

The potential for generalization of the design method described in this report to other safety-critical hybrid systems such as highly automated industrial plants for petrochemicals and nuclear power was discussed briefly.

The present results may contribute by example to the development of more general theoretical methods for synthesis of hybrid systems. They should also be of practical interest to designers of next-generation transport aircraft avionic systems.

# REFERENCES

Anonymous: Military Standard 1797, 1985.

Bartee, Thomas C.: Digital Computer Fundamentals, Sixth ed., McGraw-Hill (New York), 1985.

Billings, Charles E.: Human-centered Aviation Automation: Principles and Guidelines. NASA TM-110381, 1996. See also Billings, Charles E.: Human-centered Aviation Automation: A Concept in Guidelines. NASA TM-103885, 1991

Bray, Richard S.: A Head-Up Display Format for Application to Transport Aircraft Approach and Landing. NASA TM-81199, 1980.

Casner, Stephen M.: Understanding the Determinants of Problem-Solving Behavior in a Complex Environment. Human Factors, vol. 36, no. 4, Dec. 1994.

Dromey, R. G.: Program Derivation, Addison-Wesley (Menlo Park, Calif.), 1989.

Erzberger, Heinz: Automation of On-Board Flightpath Management. Tenth von Karman Memorial Lecture, 24th Israel Annual Conference on Aviation and Astronautics, Tel Aviv and Haifa, Israel, 1982.

Franklin, James A.; Hynes, Charles S.; Hardy, Gordon H.; Martin, James L.; and Innis, Robert C.: Flight Evaluation of Augmented Controls for Approach and Landing of Powered-Lift Aircraft. J. Guidance, Control, and Dynamics, vol. 9, no.5, Sept.–Oct. 1986.

Funabiki, K.; Bando, T.; Tanaka, K.; Hynes, C. S.; and Hardy, G. H.: A Method of Wind Shear Detection for Powered-Lift STOL Aircraft. AIAA-93-3667, 1993.

Geddes, N. D.: Intent Inferencing Using Scripts and Plans. Proc. of the First Annual Aerospace Applications of Artificial Intelligence Conference, Wright-Patterson Air Force Base, Ohio: U. S. Air Force, pp. 160–172, 1985.

Hughes, David; and Dornheim, Michael A.: Accidents Direct Focus on Cockpit Automation. Aviation Week and Space Technology, Jan. 30, 1995, pp. 52–65.

Hynes, Charles S.; Franklin, James A.; Hardy, Gordon H.; Martin, James L.; and Innis, Robert C.: Flight Evaluation of Pursuit Displays for Precision Approach of Powered-Lift Aircraft. J. Guidance, Control, and Dynamics, vol. 12, no. 4, July–Aug., 1989.

Innis, Robert C.; Holzhauser, Curt A.; and Quigley, Hervey C.: Airworthiness Considerations for STOL Aircraft. NASA TN D-5594, 1970.

Lambregts, A. A.: Vertical Flight Path and Speed Control Autopilot Design Using Total Energy Principles. AIAA Paper 83-2239 CP, Oct. 1983.

Leveson, Nancy G.: Safeware: System Safety and Computers. Addison-Wesley (Menlo Park, Calif.), 1995.

Meyer, George; and Cicolani, Luigi: Application of Nonlinear Systems Inverses to Automatic Flight Control Design—Systems Concepts and Flight Evaluations. AGARDograph no. 251, July 1981.

Neumark, S.: Problems of Longitudinal Stability Below Minimum Drag Speed, and Theory of Stability Under Constraint. Aeronautical Research Council of Great Britain R and M, no. 2983, July 1953.

Rasmussen, J.: Outlines of a Hybrid Model of the Process Plant Operator. In Monitoring Behavior and Supervisory Control, T. B. Sheridan and G. Johannsen, eds., Plenum, 1976. See also Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs and symbols and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, no. 13, pp. 257–266.

Rushby, John:. Formal Methods and Their Role in Certification of Critical Syastems. NASA Contractor Report CR–4673, 1995 (also issued as part of the FAA Digital Systems Validation Handbook).

Sherry, Lance; Youssefi, David; and Hynes, Charles S.: A Formalism for the Specification of Operationally Embedded Reactive Avionic Systems. Honeywell Publication C69-5370-001, Honeywell Air Transport System Division (System Engineering Technology), Phoenix, Ariz., 1995.

Sherry, Lance; and McCrobie, Daniel: Behavioral Characteristics of Modem Avionics Software: Implications for Design of the HSCT Cockpit Displays, Pilot Training, and Software Design Tools. Honeywell Publication No. C-69-5370-010, Honeywell, P. O. Box 21111, Phoenix, AZ 85036, 1998.

Taylor, E. S.: Dimensional Analysis for Engineers. Clarendon Press (Oxford), 1974.

Title 14, Code of Federal Regulations. Published by Office of Federal Register, National Archives and Records Administration, U. S. Government Printing Office, Washington, D.C., Jan. 2001.

Wiener, Earl L.: Human Factors of Advanced Technology ("Glass Cockpit") Transport Aircraft. NASA Contractor Report CR–177528, 1989.

## REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*

**1. REPORT DATE** *(DD-MM-YYYY)*
21-12-2007

**2. REPORT TYPE**
Technical Publication

**3. DATES COVERED** *(From - To)*

**4. TITLE AND SUBTITLE**
Synthesis from Design Requirements of a Hybrid System for
Transport Aircraft Longitudinal Control
Volumes I and II

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Charles S. Hynes[1], Gordon H. Hardy[1], and Lance Sherry[2]

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**
411931

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
[1]Ames Research Center, Moffett Field, CA 94035
[2]Honeywell International Inc., Flight Control Systems Design,
Phoenix, AZ

**8. PERFORMING ORGANIZATION REPORT NUMBER**
A-070007

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
National Aeronautics and Space Administration
Washington, D.C. 20546-0001

**10. SPONSORING/MONITOR'S ACRONYM(S)**
NASA

**11. SPONSORING/MONITORING REPORT NUMBER**
NASA/TP–2007-213474

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
Unclassified — Unlimited          Distribution: Standard
Subject Category: 03, 06, 08, 54, 66
Availability: NASA CASI (301) 621-0390

**13. SUPPLEMENTARY NOTES**
Point of Contact: Jeffery A. Schroeder, Ames Research Center, MS 210-4, Moffett Field, CA 94035-0001
(650) 604-4037

**14. ABSTRACT**

Volume I of this report presents a new method for synthesizing hybrid systems directly from design requirements, and applies the method to design of a hybrid system for longitudinal control of transport aircraft. The resulting system satisfies general requirement for safety and effectiveness specified a priori, enabling formal validation to be achieved. Volume II contains seven appendices intended to make the report accessible to readers with backgrounds in human factors, flight dynamics and control, and formal logic.

Major design goals are (1) system design integrity based on proof of correctness at the design level, (2) significant simplification and cost reduction in system development and certification, and (3) improved operational efficiency, with significant alleviation of human-factors problems encountered by pilots in current transport aircraft.

This report provides for the first time a firm technical basis for criteria governing design and certification of avionic systems for transport aircraft. It should be of primary interest to designers of next-generation avionic systems.

**15. SUBJECT TERMS**
Air transportation and safety, Avionics, Aircraft stability and control, Flight controls and autopilots, Man/machine systems, Systems analysis

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Jeffery A. Schroeder |
| Unclassified | Unclassified | Unclassified | Unclassified | 164 | 19b. TELEPHONE *(Include area code)*<br>(650) 604-4037 |