![Universidad Rey Juan Carlos]

Universidad
Rey Juan Carlos

Commercial Aircraft Trajectory Planning based on Multiphase Mixed-Integer Optimal Control

by

Manuel Fernando Soler Arnedo.

5-Year-B.A.E. Universidad Politécnica de Madrid, 2007.
M.A.S. Universidad Politécnica de Madrid, 2011.

THESIS

Submitted in partial fulfillment of the requirements for the degree of European
Doctor of Philosophy in Aerospace Engineering

Supervisors:

Alberto Olivares González.
Ernesto Staffetti Giammaria.

Escuela Técnica Superior de Ingeniería de Telecomunicación
Área de Ingeniería Aeroespacial
Universidad Rey Juan Carlos
Madrid, April 2013.

Commercial Aircraft Trajectory Planning based on Multiphase Mixed-Integer Optimal Control

by

Manuel Fernando Soler Arnedo.

5-Year-B.A.E. Universidad Politécnica de Madrid, 2007.
M.A.S. Universidad Politécnica de Madrid, 2011.

THESIS

Submitted in partial fulfillment of the requirements for the degree of European
Doctor of Philosophy in Aerospace Engineering

Como co-directores de la tesis doctoral titulada: **Commercial Aircraft Trajectory Planning based on Multiphase Mixed-Integer Optimal Control**, presentada por Manuel Fernando Soler Arnedo en el Departamento de Teoría de la Señal y Comunicaciones, hacemos constar que dicha tesis reúne todos los requisitos para poder ser presentada y defendida por el doctorando, por lo que autorizamos su depósito al objeto de que siga los trámites pertinentes para la defensa de la misma.

Co-director: Alberto Olivares González. Co-director: Ernesto Staffetti Giammaria

Escuela Técnica Superior de Ingeniería de Telecomunicación
Área de Ingeniería Aerospacial
Universidad Rey Juan Carlos
Madrid, on April the 9th, 2013.

A mi familia,
con el amor y la gratitud propia
de un niño que sigue creciendo.

Lo que más necesitan, aun los mejores de nuestros buenos estudiantes, es mayor intensidad de vida, mayor actividad para todo, en espíritu y cuerpo: trabajar más, sentir más, pensar más, querer más, jugar más, dormir más, comer más, lavarse más, divertirse más. **Francisco Giner de los Ríos.**

El espíritu de la casa, *como llamaban los residentes al esfuerzo para transmitir la mejor tradición española de educación liberal, quedaba reflejado en una cierta forma de vida construida en torno a la responsabilidad personal, el trabajo, la búsqueda de la excelencia, el culto a la amistad y el ocio creativo, con el fin de que el esfuerzo particular se viera proyectado en la sociedad [...] El* espíritu de la casa *simboliza su modelo de educación integral, basado en la tolerancia, el pluralismo y el diálogo entre distintas disciplinas de las artes y las ciencias, entre diferentes generaciones, y entre tradición y modernidad [...]* **Nota sobe** *El Espíritu de la casa*;

Exposición 100 años de la residencia de estudiantes.
Residencia de estudiantes, Madrid.

La libertad positiva se identifica con la realización plena de las potencialidades del individuo, así como su capacidad para vivir activa y espontaneamente. [...] La victoria de la libertad es solamente posible si la democracia llega a constituir una sociedad en la que el individuo, su desarrollo y felicidad constituyan el fin y el propósito de la cultura; en la que la vida no necesite justificarse por el éxito o por cualquier otra cosa, y en la que el individuo no se vea subordinado ni sea objeto de manipulaciones por parte de ningún otro poder exterior a él mismo, [...]; una sociedad, por fin, en la que la conciencia y los ideales del hombre no resulten de la absorción en el Yo de demandas exteriores y ajenas, sino que sean realmente suyos y expresen propósitos resultantes de la peculiaridad de su Yo.

Erich Fromm,
El miedo a la libertad.

El germanista, que viaja de forma intermitente, cuando y como puede, a lo largo de todo el curso del río que mantiene unido su mundo [...]
Desde la canción de los Nibelungos el Rin y el Danubio se enfrentan y se desafían [...] El Danubio [...] es el río a través del cual se encuentran, se cruzan y se mezclan gentes diversas, en lugar de ser como el Rin un místico guardian de la pureza de la estirpe.

Claudio Magris,
El Danubio.

Al hablar de viaje, por supuesto no tengo en mente una aventura turística. A nuestro entender de reporteros, el viaje significa desafío y esfuerzo, cansancio y sacrificio, cometido difícil y proyecto ambicioso. Cuando recorremos el mundo, sentimos que ocurren cosas importantes, que estamos inmersos en algo de los que somos parte y testigo a la vez, que tenemos una obligación que cumplir y una responsabilidad que asumir. ¿Y de qué somos responsables? Del camino.

Ryszard Kapuscinski.
Encuentro con el otro.

Abstract

The main goal of this dissertation is to develop optimal control techniques for aircraft trajectory planning looking at reduction of fuel consumption, emissions and overfly charges in flight plans. The calculation of a flight plan involves the consideration of multiple factors. They can be classified as either continuous or discrete, and include non-linear aircraft performance, atmospheric conditions, wind conditions, airspace structure, amount of departure fuel, and operational constraints. Moreover, multiple differently characterized flight phases must be considered in flight planning, which typically also involves decision-making processes.

The flight planning problem can be regarded as a trajectory optimization problem. The most natural way to address a trajectory optimization problem is using optimal control techniques. One of the main advantages of using optimal control is that it allows the aircraft continuous non-linear dynamics to be considered. The solution to the problem provides the optimal amount of departure fuel, the optimal four dimensional trajectory (horizontal route and the vertical profile over time), speed, consumption and attitude profiles over time, and the corresponding optimal control inputs of the aircraft.

The multiphase nature of the problem, the non-linear dynamics of the aircraft, and the introduction of integer variables to model decision-making processes lead to the formulation of a multiphase mixed-integer optimal control problem. The duration of the phases is optimized including the switching times as unknowns of the problem, which is modeled using a direct numerical approach. In particular, a collocation method is employed to transcribe the infinite dimensional optimal control problem into a finite dimensional optimization one, which is solved using a mixed integer nonlinear programming solver.

It is shown that the flight planning problem can be effectively tackled using mixed-integer optimal control, considering multiple phases and including decision-making processes. Results show that the efficiency of current flight plans could be substantially improved and that the techniques studied in this thesis have a strong potentiality to be employed in the definition of more efficient flight plans under future operational concepts in air traffic management.

Resumen

El objetivo de esta tesis doctoral es el desarrollo de técnicas de control óptimo para la planificación de trayectorias de aviones comerciales, minimizando el consumo de combustible y los costes de navegación en los planes de vuelo. El cálculo de un plan de vuelo implica la consideración de múltiples factores, los cuales pueden ser considerados de forma general como continuos o discretos. Entre ellos se puede destacar la dinámica no lineal de la aeronave, las condiciones de la atmósfera, el viento, la estructura del espacio aéreo, la cantidad de combustible a despegue, restricciones operativas, etc. Un plan de vuelo viene definido por múltiples fases, cada una caracterizada de manera particular. Además, es habitual tener que modelizar procesos de toma de decisiones. En suma, la obtención de la trayectoria óptima de una aeronave (asociada a un plan de vuelo) en entorno ATM es un problema altamente complejo.

Antecedentes: En la planificación de vuelos, el método más comúnmente utilizado es la descomposición del problema en dos etapas: la optimización de la ruta en 2D primero, y después la optimización de los perfiles de altitud y velocidad sobre la ruta 2D previamente calculada. Esta aproximación se conoce como problema de planificación de rutas. Las líneas aéreas suelen usar este método para calcular sus planes de vuelo diarios. En la optimización de las rutas 2D, la ruta óptima se calcula considerando la estructura del espacio aéreo, esto es, zonas de exclusión aérea, aerovías, waypoints, etc. Algunos de los algoritmos más empleados son: algoritmos de optimización de redes, métodos de programación lineal entera mixta o algoritmos evolutivos. Obtenida la ruta 2D óptima, los perfiles de altitud y la velocidad se calculan usando técnicas heurísticas. La principal ventaja de estos métodos es que son capaces de resolver problemas bastante complejos en tiempos relativamente bajos, siendo mucho más fácil incorporar restricciones específicas (por ejemplo, modelizar procesos de toma de decisiones).

Sin embargo, la solución completa a un problema de planificación de trayectorias es una optimización de la trayectoria en 4D (espacio y tiempo). En este sentido, el problema de optimización de trayectorias puede ser estudiado como un problema de control óptimo en el que el objetivo es encontrar la trayectoria y las entradas de control que guían el estado del avión (considerado como sistema dinámico) entre dos configuraciones, satisfaciendo un conjunto de restricciones sobre el estado y/o las variables de control, mientras se minimiza una cierta función objetivo. En comparación con el método de descomposición antes mencionado, las principales ventajas de utilizar control óptimo son que:

- Permite considerar no linealidades, dando como resultado una dinámica más exacta de la aeronave, de tal forma que la solución proporciona la trayectoria 4D completa, así como la velocidad, el consumo de combustible de la aeronave y los ángulos de actitud (cabeceo, guiñada, y asiento) en función del tiempo.
- Proporciona las leyes de control óptimo de tal manera que la aeronave pueda ser guiada de manera óptima. La ley de control óptimo es útil en el diseño en lazo abierto del autopiloto, dando los valores de referencia para las variables de control.

Objetivos: Para la consecución del objetivo fundamental, esto es, desarrollar técnicas de control óptimo para planificación de trayectorias de aeronaves comerciales, se han alcanzado los siguientes objetivos intermedios:

- Revisión exhaustiva de la literatura sobre teoría de control óptimo y métodos numéricos para la resolución de problemas de control óptimo.
- Construcción de un modelo para abordar el problema de planificación de vuelos. Se contemplan por tanto modelos que incluyan la dinámica no lineal del avión, el viento, las características de la atmósfera, la estructura del espacio aéreo y procesos de toma de decisiones, entre otros.
- Desarrollo de técnicas para la formulación y resolución de problemas de control óptimo que incluyan múltiples fases, eventos discretos y procesos de toma de decisiones.
- Definición de planes de vuelo más eficientes en base a las técnicas anteriores.

Metodología: La naturaleza multifase del problema, la dinámica no lineal de la aeronave, así como la estructura discreta del espacio aéreo, conduce a la formulación de un problema de control óptimo entero mixto multifase. La duración de las fases se ha optimizado incluyendo los tiempos de conmutación entre fases como variables del problema. Se han introducido variables enteras para modelizar los fenómenos discretos. Los procesos de toma de decisiones se han modelizado con variables binarias.

Por lo general, es muy difícil encontrar soluciones analíticas para los problemas de control óptimo. La práctica común es el uso de métodos numéricos. El problema se ha resuelto usando un método numérico directo. En particular, se ha empleado un método de colocación para transcribir el problema de control óptimo de dimensión infinita en un problema de optimización de dimensión finita, que se ha resuelto usando un solver de programación entera mixta no lineal.

Conclusiones: Se muestra que el problema de planificación de vuelos comerciales puede ser tratado de manera eficiente con técnicas de control óptimo entero mixto multifase, dado que permite formular el problema considerando múltiples fases, así como incorporar variables de carácter entero/binario, las cuales permiten modelizar fenómenos discretos y procesos de toma de decisiones. Los resultados muestran que la eficiencia de las trayectorias actuales podría mejorarse substancialmente. Por lo tanto, las técnicas estudiadas en esta tesis doctoral tienen un gran potencial para ser explotadas con el objetivo de definir planes de vuelo más eficientes e, indirectamente, más respetuosos con el medio ambiente.

Acknowledgments

En primer lugar, quiero expresar mi más sincera gratitud hacia mis mentores, compañeros y amigos Dr. Alberto Olivares y Dr. Ernesto Staffetti. Nada de lo realizado durante estos más de cuatro años de concienzudo trabajo y dedicación hubiese sido posible sin su supervisión constante, su dedicación abnegada y su tiempo, sus comentarios y críticas constructivas, siempre buscando la excelencia, y, en definitiva, sin su guía lúcida, sincera y comprometida. De justicia es reconocer no sólo su tutela, si no su coautoría en este trabajo. Gracias, *grazie mille*.

Second, I would like to express my sincere gratitude to Dr. Pierre Bonami for fruitful discussions and tips on the use of BONMIN in the scope of aircraft trajectory optimization. It is also fair to mention that Dr. Bonami was actively involved in composing and reviewing part of the contents in Section 6.1.2.

Third, I'm particularly indebted with Prof. Dr. John Lygeros, who hosted me at the Institut Für Automatic, ETH Zürich, during three months, and with Prof. Dr. Mark Hansen, who hosted me at NEXTOR, UC Berkeley, for five months. Both experiences were vitally and professionally so awe-inspiring that the only thing I can say is: thank you! Also, I would like to mention all those that I have had the pleasure to meet and work with as a visiting scholar in Zürich and San Francisco. Thank you for sharing with me the offices, for nice work and non-work related discussions, for the company in lunches an coffees, for involving me in the departmental life, for sharing your leisure time to make my stay as a foreigner happier, and many other things... From all these people, I would like to enhance the figure of Dr. Maryam Kamgarpour: thank you for the time invested in discussions, meetings, simulations, readings, proof readings, etc. that led to publishing a couple of nice papers, and also for sharing a lot of intercultural experiences. Thank you for your work and your happiness. *Mamnoon*.

No puedo olvidarme de *GMV Aerospace and Defence*, empresa con la que hemos mantenido una estrecha colaboración durante estos más de cuatro años, desarrollando varios proyectos que representan la base inspiradora y han sido de hecho el sustento económico de esta tesis doctoral. En particular, agraceder a Jesús Cegarra y Daniel Zapata su gestión, sus comentarios y sus sugerencias. Gracias a ellos hemos podido acercarnos a la necesaria transferencia de conocimientos entre universidad y empresa.

Contents

Notation

Vectors are not explicitly denoted. Units are denoted in brackets. Superindexes denote phases, whereas subindexes denote discretization samples. We provide a not exhaustive list of terms.

(O, x_w, y_w, z_w)	$=$	wind axis reference frame.
$(\cdot)_w$	$=$	subindex for values expressed in the wind axis reference frame.
(O, x_e, y_e, z_e)	$=$	Earth reference frame.
$(\cdot)_e$	$=$	subindex for values expressed in the Earth reference frame.
$a_{l,max(civ)}$	$=$	maximum longitudinal acceleration for civil flights.
$a_{n,max(civ)}$	$=$	maximum normal acceleration for civil flights.
C_D	$=$	coefficient of drag.
C_{D_0}	$=$	coefficient of parasite drag.
C_{D_i}	$=$	coefficient of induced drag.
C_L	$=$	coefficient of lift.
C_{L_0}	$=$	zero coefficient of lift.
C_{L_α}	$=$	slope of the lift curve.
$C_{L_{max}}$	$=$	maximum coefficient of lift.
$C_{Tc,4}$	$=$	first thrust temperature coefficient.
$C_{V_{min}}$	$=$	minimum speed coefficient.
D	$=$	drag force $= 0.5\rho V^2 S C_D$.
G_t	$=$	temperature gradient on maximum altitude.
G_W	$=$	mass gradient on maximum altitude.
g	$=$	acceleration due to gravity.
h_e	$=$	altitude.
h_{M_0}	$=$	maximum operational height.
h_{max}	$=$	maximum altitude at MTOW under ISA conditions.
h_u	$=$	maximum altitude as a function of the mass of the aircraft.
H_{max}	$=$	threshold altitude for different flap configurations..
L	$=$	lift force $= 0.5\rho V^2 S C_L$.
M	$=$	Mach number.
M_{M0}	$=$	maximum operating Mach number.
m	$=$	mass.

m_{max}	$=$	maximum mass (Maximum Take Off Weigh).
m_{min}	$=$	minimum mass (Operating Empty Weigh).
$\dot{m}$	$=$	fuel flow.
$\dot{m}_{min}$	$=$	minimum fuel flow.
$\hat{q}$	$=$	dynamic pressure, $\hat{q} = \frac{1}{2}\rho V^2$.
R_e	$=$	Earth radius.
S	$=$	reference wing surface area.
T	$=$	thrust.
T_{max}	$=$	maximum thrust.
T_{min}	$=$	minimum thrust.
V	$=$	true airspeed.
V_{CAS}	$=$	calibrated air speed.
V_{M_0}	$=$	maximum operating calibrated airspeed.
V_S	$=$	stall speed.
W	$=$	wind vector, $W = (W_x, W_y, W_z)$.
α	$=$	angle of attack.
χ	$=$	heading angle.
ΔT_{ISA}	$=$	temperature deviation from ISA.
η	$=$	thrust specific fuel flow.
γ	$=$	flight path angle.
λ_e	$=$	longitude.
μ	$=$	bank angle.
$\mu_{max,civ}$	$=$	maximum bank angle for civil flights.
π	$=$	throttle setting.
ρ	$=$	atmospheric density.
ρ_0	$=$	atmospheric density at sea level.
ρ_{11}	$=$	atmospheric density at the tropopause.
θ_e	$=$	latitude.

List of acronyms

ACARE	=	Advisory Council for Aeronautics Research in Europe.
AIP	=	Aeronautical Information Publication.
AIRAC	=	Aeronautical Information Regulation and Control.
AIRE	=	Atlantic Interoperability Initiative to Reduce Emission.
ATC	=	Air Traffic Control.
ASM	=	Air Space Management.
ATFM	=	Air Traffic Flow Management.
ATM	=	Air Traffic Management.
ATS	=	Air Traffic Services.
BADA	=	Base of Aircraft DAta.
CAS	=	Calibrated AirSpeed.
CCD	=	Continuous Climb Departure.
CDA	=	Continuous Descent Approach.
CDM	=	Collaborative Decision Making
CDRs	=	Coded Departure Routes
CGL	=	Chebyshev-Gauss-Lobatto.
DAE	=	Differential-Algebraic Equations.
DOF	=	Degree Of Freedom.
DP	=	Dynamic Programming.
ECP	=	Extended Cutting Plane.
EEC	=	Eurocontrol Experimental Centre.
FAA	=	Federal Aviation Administration.
FIR	=	Flying Information Region.
FSL	=	Forecasts System Laboratory.
FMS	=	Flight Management System.
GCD	=	Great Circle Distance.
GNSS	=	Global Navigation Satellite System.
GRIB	=	GRIdded Binary.
HJB	=	Hamilton-Jacobi-Bellman.
HLGL	=	Hermite-Legendre-Gauss-Lobatto.
IAF	=	Initial Approach Fix.

ICAO   =   International Civil Aviation Organization.
ILS   =   Instrumental Landing System.
IAS   =   Indicated AirSpeed.
IPCC   =   Intergovernmental Panel on Climate Change.
ISA   =   International Standard Atmosphere.
IPOPT   =   Interior Point OPTimizer.
KKT   =   Karush-Kuhn-Tucker.
LG   =   Legendre-Gauss.
LGL   =   Legendre-Gauss-Lobatto.
LGR   =   Legendre-Gauss-Radau.
LW   =   Landing Weight.
NAP   =   Noise Abatement Procedures.
NextGen   =   Next Generation of air transportation system.
Nlp   =   Non-Linear Programming.
Nlp Bb   =   Non-Linear Programming based on Branch and Bound.
NOAA   =   National Oceanic and Atmospheric Administration.
MILP   =   Mixed-Integer Linear Programs.
Minlp   =   Mixed-Integer Non-Linear Programming.
MTOW   =   Maximum TakeOff Weight.
OC   =   Operational Constraints.
ODE   =   Ordinary Differential Equation.
OEW   =   Operating Empty Weight.
PL   =   Payload.
QQ Plot   =   Quantile- Quantile Plot.
RUC   =   Rapid Update Circle.
RF   =   Reserve Fuel.
RWY   =   Runway.
SESAR   =   Single European Sky ATM Research.
SID   =   Standard Instrumental Departure.
SRA   =   Strategic Research Agenda.
STAR   =   Standard Terminal Arrival Route.
TBO   =   Trajectory-Based Operations.
TMA   =   Terminal Maneuvering Area.
TF   =   Trip Fuel.
TPBVP   =   Two Point Boundary Value Problem.
UAV   =   Unmanned Air Vehicle.
UIR   =   Upper Information Region.
USA   =   United States of America.

List of Figures

List of Tables

1

Introduction

1.1   Motivation

A FLIGHT PLAN is an aviation term defined by the International Civil Aviation Organization (ICAO) as: "Specified information provided to air traffic services units, relative to an intended flight or portion of a flight of an aircraft" [1].

A flight plan is prepared on the ground and specified in three different manners: as a document carried by the flight crew, as a digital document to be uploaded into the Flight Management System (FMS), and as a summary plan provided to the Air Transit Services (ATS). It gives information on route, flight levels, speeds, times, and fuel for different flight phases, alternative airports, and other relevant data for the flight, so that the aircraft properly receives support from ATS in order to execute safe operations. Two safety critical aspects must be fulfilled: fuel calculation, to ensure that the aircraft can safely reach the destination, and compliance with Air Traffic Control (ATC) requirements, to minimize the risk of collision.

Flight planning is the process of producing a flight plan to describe a proposed aircraft flight. It requires accurate weather forecasts so that fuel consumption calculation can account for the effects of wind. Furthermore, due to ATC supervision requirements, aircraft flying in controlled airspace must follow predetermined routes.

A route is a description of the path followed by an aircraft when flying between two airports. A complete route often uses several airways. An airway has no physical existence, but can be thought of as a motorway in the sky. On an airway, aircraft fly at different flight levels to avoid collisions and are enforced to fly at specific velocities. On a bi-directional airway, each direction has its own set of flight levels. Each airway starts and finishes at a waypoint, and may contain some intermediate waypoints as well. Airways may cross or join at a waypoint, so an aircraft can change from one airway
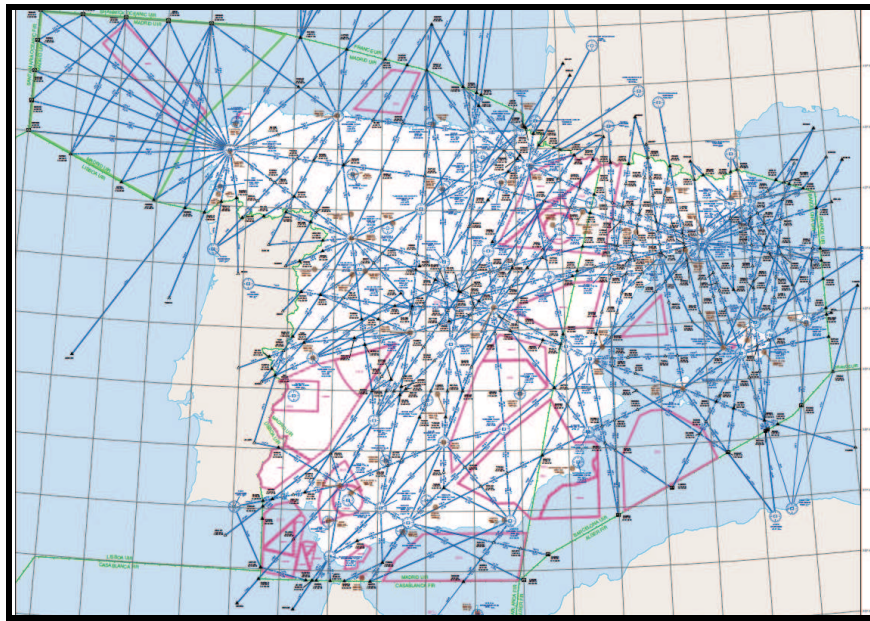
Figure 1.1: En-route upper navigation chart. Source: AIP AENA (http://www.aena.es).

to another at such points. Where there is no suitable airway between two waypoints, ATC may allow a direct waypoint to waypoint routing which does not use an airway. Additionally, there exist special tracks known as ocean tracks, which are used across some oceans. Free flight is also permitted in some areas over the oceans.

The routes of the complete network are referred to as ATS routes. The use of the airspace by ATS routes is restringed in some areas of special use due to military operations, environmental policies, or simply security reasons. ATS routes go across certain regions of airspace known as Flight Information Regions (FIR)[1], in which flight information services and alerting services are provided by ATC. Aviation authorities charge airlines that overfly FIR/UIR regions to cover air traffic control costs. Very different charging schemes are applied including purely traveled distance-based charges, aircraft weight and traveled distance charges, flat rate charges, or communication rate charges [2].

The ATS routes are published in the basic manual for aeronautical information, referred to as Aeronautical Information Publication (AIP). AIP publishes information for en-route and aerodromes in different charts, which are usually updated once a month coinciding with the Aeronautical Information Regulation and Control (AIRAC) cycle. Ocean tracks might change twice a day to take advantage of any favorable wind. In free flight areas the path is defined by the user, and thus finding the optimal path considering the effects of wind is essential.

---

[1]FIR are derived for the lower airspace. In the upper airspace, the regions are called Upper Information Regions (UIR). Both perimeters generally coincide.

Figure 1.2: Instrumental approximation chart. Madrid Barajas RWY 18L.
Source: AIP AENA (http://www.aena.es).

The complete route of an aircraft flying between two airports can be divided in three main parts: origin, en-route, and destination. The en-route phase is defined by a series of waypoints and airways. An en-route upper navigation chart for the Iberian Peninsula is given in Figure 1.1, illustrating the different ATS routes (in blue), the limited areas (in pink), and the UIR regions (in green). However, airports may not be directly connected by airways. Terminal Maneuvering Areas (TMA) are defined to describe a designated area of controlled airspace surrounding an airport (or various close airports) due to high volume of traffic. Operational constraints, arrival procedures, and departure procedures are defined inside an airport TMA.

A flight departing from an airport must follow a Standard Instrument Departure (SID) which defines a pathway from the runway to a waypoint or airway, so that the aircraft can join the en-route sector in a controlled manner. Before landing, an aircraft must follow two different procedures. It must follow first a Standard Terminal

Figure 1.3: SID chart. Madrid Barajas RWY 15R day.
Source: AIP AENA (http://www.aena.es).

Arrival Route (STAR), which defines a pathway from a waypoint or airway to the Initial Approach Fix (IAF). Then, it must proceed from the IAF to the runway following a final approach procedure. Figure 1.2 and Figure 1.3 show a final approach chart and a SID chart in Madrid Barajas TMA, respectively.

An effective flight plan can reduce fuel costs, time-based costs, overflight costs, and lost revenue from payload that can not be carried, simply by choosing efficient routes and altitudes, speed, and the optimal amount of departure fuel.

Optimizing the flight plan to reduce fuel not only saves money, but indirectly reduces the environmental fingerprint of air transportation. An overview on aviation environmental impact is given in [3]. Emissions from aviation account for approximately 3% of total green house effect gases in Europe [4, Section 2.3] and 3.5% of CO_2 man made European emissions [5]. In 1999 the United Nations Intergovernmental Panel on Climate Change (IPCC) forecasted, as aviation was expected to grow to meet increasing demand, that its share of global man made CO_2 emissions would increase to around 3% to 5% in 2050 (in 1999 it was estimated to be 2%) [6].

Reducing thus the aviation environmental fingerprint is peremptory. The Advisory Council for Aeronautics Research in Europe (ACARE) has defined a Strategic Research Agenda (SRA) for aeronautics and air transport in Europe. The SRA has launched the Clean Sky Joint Technology Initiative, which pursues a greener aviation. The objective of Clean Sky is to demonstrate and validate the technological innovations

that are necessary to make major steps towards the environmental goals set by ACARE and to be reached in 2020 when compared to 2000 levels [7] : 50% reduction of CO_2 emissions; 80% reduction of NO_x emissions; 50% reduction of external noise; and a green product life cycle. The reduction in fuel consumption and CO_2 emissions will require contributions from new technologies in aircraft design (engines, airframe materials, and aerodynamics), alternative fuels (bio fuels), improved Air Traffic Management (ATM), and operational efficiency (mission and trajectory management). Indeed, EUROCONTROL estimates that approximately 6% of the burnt fuel and CO_2 emissions related to aviation in Europe is due to inefficiencies in the ATM European system [5, Section 3.5].

In particular, Clean Sky aims at defining new approaches for the optimization of trajectory and mission management by implementing two new concepts referred to as green trajectories and green missions:

- Green trajectories, based on more precise, reliable and predictable tridimensional flight paths, optimized for minimum noise impact and low emission, including agile trajectory management in response to meteorological hazards.

- Green mission from start to finish, with management of new climb, cruise, and descent profiles based on aircraft performances which include noise and allows multi–criteria optimization (noise, emissions, fuel, and time), including weather conditions which could negatively impact the aircraft optimum route resulting in additional fuel consumption.

The improvement of flight planning techniques must be done in coordination with the definition of a new ATM paradigm. ATM, which is responsible for sustainable, efficient, and safe operations in civil aviation, is still a very complex and highly regulated system. Indeed, the need to fit aircraft trajectories to ATM system requirements makes them difficult to be optimized and thus generally suboptimal flight plans are flown. A substantial change in the current ATM paradigm is needed because this system is reaching the limit of its capabilities [8]. Its capacity, efficiency, environmental impact, and flexibility should be improved to accommodate airspace users' requirements [9]. The Next Generation of air transportation system (NextGen), in the United States of America (USA), and the Single European Sky ATM Research (SESAR) Program, in Europe, are facing these challenges aiming at developing a new generation of ATM system.

Looking at Europe, the main goals of SESAR can be summarized as [9]: 3–fold increase the air traffic movements whilst reducing delays; improvement of the safety by a factor of 10; 10% reduction in the effects that aircraft have on the environment; and 50% reduction in the ATM services cost charge to airspace users.

In particular, looking at single flights, SESAR pursues by 2020 [10]: 8–14 [min] gain per flight on average; 300–500 [kg] reduction in fuel consumption per flight on average; 945–1575 [kg] reduction of CO_2 emissions per flight on average. In order to validate this intended savings, SESAR Joint Undertaking collaborates with the US Federal Aviation Administration (FAA) and a large number of international partners in an international program for the reduction of aircraft emissions: AIRE[2] (Atlantic

---

[2]http://www.sesarju.eu/environment/aire

Interoperability Initiative to Reduce Emission). AIRE has been testing more than 9000 trial flights since 2009. A summary of the results achieved in such tests in 2010 and 2011 can be found in [11].

In the future ATM system the trajectory becomes the fundamental element of a new set of operating procedures collectively referred to as Trajectory-Based Operations (TBO) [12, 13]. The ATM system has historically grown along a clearance-based ATC paradigm. TBO will provide the capabilities, decision-support tools, and automation to manage aircraft movement by trajectory. This shift from clearance-based to trajectory-based ATC will enable aircraft to fly negotiated flight paths. In order to understand TBO, two fundamental ideas must be exposed: the concept of the business trajectory, and the idea of trajectory ownership.

The business trajectory is the trajectory that will best meet airline business interests. This business interests may be minimum time for the flight, minimum consumption, minimum cost index, minimum environmental impact, or any other characteristic of the trajectory. Any modification in that trajectory will result in a change in the cost effectiveness of the operation. Thus, the future ATM system should try to keep the overall business trajectory as much intact as possible. The necessary changes to be made on the overall trajectory must consider downstream effects and not only consider a control sector or even a control center. These changes should be proposed by ATC and agreed with the airline, which holds the ownership of the trajectory.

This ownership entails certain rights as well as obligations. The owner will submit the trajectory to ATS authorities in order to be provided with support to perform safe operations. In exchange, the owner accepts the obligation of sharing information, reacting to requests, following clearances, and so forth. In any case, the owner remains the master over the trajectory to approve changes on it. The trajectories (typically into an onboard FMS) will be agreed and updated in real-time.

Derived from the TBO concept of operations to be designed around the notion of business trajectory, the following consequences emerge [12, Chap. 2.4.2.3]:

- Aircraft will fly accurate four dimensional (4D) trajectories (three spatial dimensions and time).

- All necessary data regarding the trajectory will be shared by all concerned actors through system wide information management to facilitate gate to gate collaborative decision-making processes.

- Intervention in the trajectory will happen in full knowledge of the downstream effects and hence it will be possible to pick the option causing the least amount of trajectory distortion.

- The increase in the predictability around the trajectory, together with the provision of decision support tools to assist humans to handle much larger number of trajectories, will enable to increase the capacity into a given airspace volume.

Nevertheless, TBO can only work if the current strict constraints on airspace usage are reduced or eliminated. Indeed, airspace should be seen as an unlimited resource through which the trajectories are allowed to be flown with minimal artificial distortion. Today's view focusing on individual flights and actions to separate them tactically should be replaced by a more strategic view focusing on the trajectories, managing them by accurate planning and reducing tactical intervention to exceptions.

This does not necessary mean an airspace with no constraints and no ATC intervention. Airports, where aircraft must take-off and land, will keep being the bottlenecks of the air navigation system. Even though the predictability of the system will be increased in the future, it will still be subject to certain degree of uncertainty, for instance, due to the atmosphere. Moreover, safety will always be the primary goal in civil aviation. Thus, TBO will work best based upon a constrained paradigm, where ATC will issue constraints to be met and the airline will decide the most economical way to meet them.

Therefore, the underlying motivation of this thesis is to conduct research towards the real implementation of TBO, leading to more flexible, efficient, and environmentally friendly flight planning concepts framed into a constrained based paradigm, considering decision-making processes, reducing airlines operating costs, and allowing green trajectories and green missions to be carried out in the future, which will also result in less emissions to the atmosphere.

1.2 Goals

The main goal of this dissertation is to develop optimal control techniques for optimal aircraft trajectory planning looking at reduction of fuel consumption and overfly costs in flight plans.

To reach this main goal the following intermediate objectives must be achieved:

- Exhaustive revision of the literature on optimal control theory and numerical methods for optimal control.

- Accurate modeling of the elements of the flight planning problem. They include an aircraft performance model, a wind model, a model of the discrete airspace structure, models of operational constraints and operational procedures, models of the different aerodynamic configurations, and finally a model for decision-making processes.

- Definition of techniques to solve optimal control problems that include multiple phases, discrete events, and decision-making processes.

- Based on the devised techniques, design of more efficient flight plans.

- Validation of the approach by solving numerical instances.

1.3   Modeling

As it has been described so far, the calculation of a flight plan is subject to many factors, which makes finding the optimal flight plan a very difficult task. Therefore, modeling the phenomena involved in the calculation of a flight plan in a clever way is key to tackle the problem. Models must be not too complex so that problems are solvable, but yet accurate to represent reality in a reliable way. They include continuous phenomena, discrete phenomena, and decision-making processes. Continuos models include aircraft performance, atmospheric conditions and meteorological conditions. Discrete models include airspace structure, operational constraints and procedures, aerodynamic configurations of the aircraft, and overfly charges.

In this thesis a 3 Degree Of Freedom (DOF) dynamic model that describes the point variable-mass motion of the aircraft has been considered. The Earth has been considered flat, non rotating, and approximate inertial reference frame. The acceleration due to gravity in the atmospheric flight of an aircraft has been considered constant and perpendicular to the surface of Earth. The aircraft has been considered as a conventional jet airplane with fixed engines. The external actions acting on an aircraft can be decomposed into propulsive, aerodynamic, and gravitational. The aircraft has been assumed to have a plane of symmetry, and thus it flies in symmetric flight, i.e., all forces act on the center of gravity and the thrust and the aerodynamic forces lay in the plane of symmetry. This model for aircraft performance has been extensively used in the calculation of trajectory for long-time horizons in which the focus is not on maneuvers (for which more precise assumptions would be needed), but in global performance.

The Eurocontrol Base of Aircraft DAta (BADA) 3.6 [14] has been considered, which provides data for a total of 295 aircraft types, operations, and procedures. BADA also includes models for thrust, consumption, aerodynamics, and performance limitations. In particular, the aerodynamic model considers a parabolic drag polar, and therefore, does not take into account compressibility effects on the aerodynamic behavior of the aircraft. This leads to lower than real drag at high Mach numbers, resulting in higher than real optimum speeds and altitudes. Five different aerodynamic configurations, for take-off, initial climb, cruise, approach, and landing are given. These different aerodynamic configurations which correspond to different flap deflections give rise to different dynamic behaviors of the aircraft. BADA enables the standardization of aircraft model and thus the comparison of achieved results. By proper processing of data, BADA also enables the automation of the flight plan calculation for different types of aircraft.

The model of atmosphere that has been considered is the standard one used in aviation and weather studies, in which the density of air is a function of the altitude given by the profiles established by the International Standard Atmosphere (ISA) in the different layers of the atmosphere.
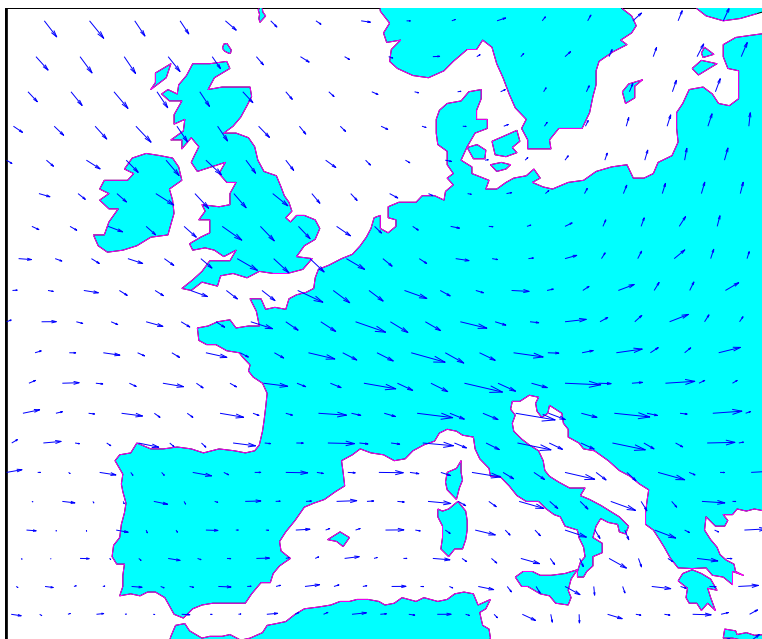
Figure 1.4: Wind field. GRIB data treated with MATLAB software.

Wind has been taken into account due to its important influence in flight planning. As said before, wind forecasts are crucial for the periodic definition of airways in oceans tracks, and obviously for obtaining the optimal path in zones where free flight is permitted. In the future, under TBO concepts, it will become even more crucial. The meteorological predictions have been obtained from the Rapid Update Circle (RUC)[3], a numerical model developed at the National Oceanic and Atmospheric Administration (NOAA) Forecasts System Laboratory (FSL) that incorporates aircraft measurements, balloon soundings, and other sensor data [15, 16]. Forecasts are provided via GRIdded Binary (GRIB) files. GRIB is a format with high compactness and high capacity used by the meteorological institutes of the world to store, share, and manipulate weather data. Data are provided four times a day into a $1° \times 1.25°$ grid and fourteen different barometric altitudes. In particular, GRIB files provide wind forecasts giving the three components (vertical, north and east) of the wind vector at each node of the grid. See Figure 1.4.

In the wind model, we have assumed the wind field to be stationary. Wind in the vertical direction has been assumed to be zero due to its low values. Since the model does not take into account stochasticity, we have considered wind forecasts as deterministic. In order to include wind data into the optimal control problem, an analytical function is needed, which has been obtained via a multiple regression fit of wind discrete data into polynomial functions.

---

[3]http://ruc.noaa.gov/

We turn now to the explanation of the models underlying the discrete phenomena involved in the calculation of a flight plan.

In order to model the airspace structure of ATS routes, we have not considered airways and we have supposed that the aircraft can fly an arbitrary route among waypoints. However considering a complete directed graph structure over the set of waypoints is redundant because aircraft must fly through closer waypoints before reaching farther waypoints. A graph structure which is capable to reflect this simple observation is the multipartite graph structure [17]. Therefore, the airspace structure has been modeled as a complete multipartite graph in which the vertex set of waypoints is partitioned into pairwise independent subsets of waypoints connected if and only if they belong to different (adjacent) partite sets.

The trajectory of an aircraft has inherently a discrete nature due to the presence of multiple flight phases corresponding to different dynamic models. Thus, a discrete dynamics must be considered together with the aircraft continuous dynamics. The discrete dynamics governs the evolution of a set of discrete variables in which the discrete states of the system are stored. The different discrete states give rise to different dynamical subsystems and their corresponding sets of path constraints which govern the continuous motion of the aircraft throughout the phases of the flight. The discrete variables have been included to model the flight phase, the aerodynamic configuration (flap configuration), the dynamic mode (3D, vertical or horizontal motion), the operational procedure (for instance, constant velocity), the atmosphere mode (below or above the troposphere), and the waypoints of the multipartite graph.

The coupling of the discrete structure of a flight plan with the continuous aircraft dynamics results in a hybrid system [18]. In particular the flight of an aircraft has intrinsically the characteristics of a controlled switched dynamical system [19].

Therefore, flights have been modeled as a collection of phases connected by end trigger conditions which make the system "switch" between one phase and the following, resulting thus in a switched system. The end trigger conditions can be of two types: internally forced (autonomous), in which the switching are triggered by capture conditions such as waypoint's coordinates, velocity or altitude; externally forced (controlled), in which the switchings are triggered by an external control input. Each phase is characterized by a determined dynamical subsystem governing the continuos motion of the aircraft and a set of active constraints.

Decision-making processes have been introduced in the model in order to determine the discrete dynamics of the system throughout the different phases. They have been modeled utilizing time-dependent binary control functions and time-independent binary variables.

1.4   Prior work

In commercial aircraft flight planning, the most commonly used approach, called aircraft routing problem [20], is based on a decomposition of the problem into 2D route optimization and profile/speed optimization. Airlines typically rely on this approach to compute their flight plans daily. Some commercial suppliers are Jeppesen JetPlanner[4] and Lufthansa's LIDO[5] flight planning services for airlines.

Given the airspace structure with no–fly zones, airways, waypoints, etc., the approaches to solve the 2D route optimization problem include network optimization based algorithms [21], mixed–integer linear optimization approaches [22, 23], or evolutionary based algorithms [24]. Once the 2D optimal route has been obtained, altitude and velocity profiles can be selected using heuristics. The main advantage of these approaches is that they are able to solve rather complicated, highly constrained problems in which it is easy to include specific restrictions.

However, the most complete strategy to tackle a flight planning problem is to solve a 4D trajectory optimization problem.

Trajectory optimization

The trajectory optimization problem can be studied as an optimal control problem of a dynamic system in which the goal is to find the trajectory and the corresponding control inputs that steer the state of the system between two configurations satisfying a set of constraints on the state and/or control variables while minimizing an objective functional. Compared to the decomposition approaches above mentioned, the fundamental advantage of using optimal control is that allows aircraft dynamics to be considered.   In this manner, without any decomposition, the solution to the problem provides the optimal 4D trajectory, speed profile, consumption profile, and attitude profile.   Moreover, it provides the control inputs that govern the continuous motion of the aircraft.

Typically, it is very difficult to find analytical solutions to optimal control problems, even for the simplest cases.   The common practice is to use numerical methods to get approximated solutions. Numerical approaches to optimal control problems can be mainly classified into indirect [25] and direct methods [26]. Indirect methods are based on the use of necessary optimal conditions [27], whereas the underlying idea in direct methods is to transcribe the infinite dimensional optimal control problem into a finite dimensional optimization one, which can be solved using mathematical programming techniques [28].

Both, direct and indirect methods only ensure local optima.   Even though direct methods present less accuracy than indirect methods [29], its computational efficiency has made them the most used methods for solving real optimal control problems, and in particular for solving aircraft and spacecrafts trajectory optimization problems [26]. Compared with indirect methods, the use of direct methods for solving optimal control

---

[4]http://www.jeppesen.com
[5]http://www.lhsystems.com

problems reduce the size and the complexity of the problem [30]. One of the main disadvantages of using indirect methods has to do with properly setting a suitable initial guess [26, 31]. Moreover, using direct methods is much easier to consider multiple equality and inequality constraints. Furthermore, direct methods allow us to overcome in a natural way one of the classical limitations of optimal control: combining integer and continuous variables.

The existence of different flight phases give rise to a so called multiphase trajectory optimization problem [32], which can be formulated as a multiphase optimal control problem. However, due to the complex nature of commercial aircraft's flight plans, decision variables are needed for better modeling of trajectory optimization problems [33]. Therefore, integer or binary variables are introduced to model decision-making processes in optimal control problems. These problems can be formulated as multiphase Mixed-Integer Optimal Control Problems (Miocp) [34]. The direct collocation methods, such as those based on Hermite-Legrende-Gauss-Lobatto (HLGL) schemes [35, 36] or the so called pseudospectral collocation methods [37, 38], are among the most suitable direct methods to transcribe the infinite dimensional Miocp.

Then, the resulting optimization problem can be solved using Non-Liner Programming (Nlp) [39] and Mixed Integer Non-Linear Programming (Minlp) [40] techniques. A suitable Minlp solver for large-scale, sparse, nonlinear, non-convex optimization problems with a large number of equality and inequality constraints, which include integer variables is Bonmin [41] which implements a Nlp based branch and bound (Nlp Bb) algorithm, and uses Ipopt [42] as Nlp solver. Bonmin and Ipopt are open-source solvers; source codes and binaries are available from Coin-or[6].

We turn now the discussion to the analysis of the prior related work on aircraft trajectory optimization considering both multiple phases and decision-making processes.

In the scope of multiphase optimal control, multiphase (also referred to as multistage) trajectory optimization problems have been solved for many applications in aerospace engineering. Just to name a few, in [43] and [44], an Unmanned Air Vehicle (UAV) flight mission was presented considering the route as a given sequence of waypoints. Multistage problems for space flight trajectory optimization have also been solved. They include multistage launch [45], multistage ascent [46], and multistage orbit transfer [47]. Military applications of multiphase trajectory optimization with radar-range constraints [48] have been also published. Some applications of multiphase optimal control to commercial aircraft are [49], where optimal trajectories with airspace constraints have been discussed, and [32], where the flight of an aircraft was modeled as a collection of phases and procedures and a sequence of minimum-fuel consumption problems was solved.

The fundamental approaches followed for solving multiphase optimal control problems are either based on pseudospectral methods [50, 51] or on concatenating phases [32]. Such approaches have one main drawback: one must add linkage constraints to ensure continuity along time domain, increasing the size and complexity of the problem.

---

[6]http://www.coin-or.org

Few works considering decision-making processes have been presented in the scope of aircraft trajectory optimization: In [33], two examples are solved, an asteroid mission and a refueling mission, using a pseudospectral knotting method to generate a mixed-variable programming problem; in [52, 53], evolutionary algorithms including decision-making processes are derived for space mission planning; a bi-level algorithm was derived in [54] to solve a collision avoidance problem under collaborative decisions. Moreover, to our best knowledge, none application to commercial aircraft trajectory optimization considering decision-making variables within optimal control has been published. Therefore, more efforts in deriving new strategies to solve instances of this problem that include decision-making processes in aircraft trajectory optimization problems are needed.

Operational concepts

We analyze now recent ATM research activity towards greener operational concepts.

In particular, the definition of Continuous Descent Approaches (CDA) has led to important fuel savings and environmental benefits in real scenarios. In contrast to a conventional approach, in the CDA procedure the aircraft stays higher for longer, descending continuously from as higher as possible and avoiding any level segments of flight prior to intercepting the final approach glide path (typically $3°$ with Instrumental Landing Systems (ILS)). A continuous descent requires significantly less engine thrust than prolonged level flight, resulting in less noise, less consumption, and less emissions. Figure 1.5 illustrates a comparison between CDA and conventional approaches.
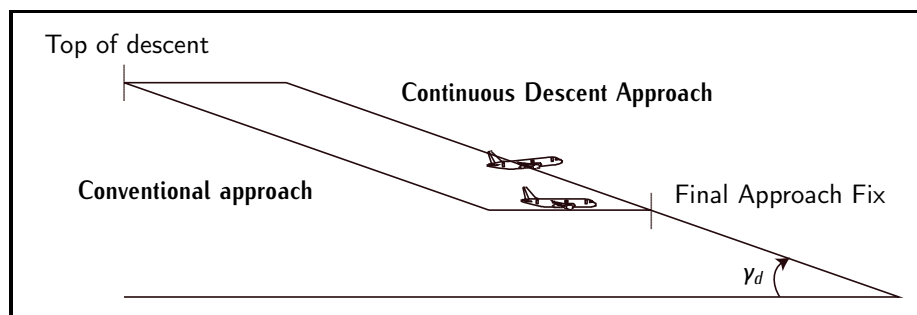
Figure 1.5: Comparison between CDA and a conventional approach.

An optimal control model for advanced CDAs procedures can be found in [55]. CDA simulations in high-density terminal airspace have been presented, for instance, in [56] and [57]. Real flight scenarios have also been studied, for instance, an analysis of tailored arrivals for datalink-enabled CDAs for San Francisco Airport [58], the analysis of the CDA results of project AIRE in Miami and Atlanta airports [59], and the design and flight test evaluation of a CDA procedure for nighttime operation at Louisville International Airport [60].

The simulation and real implementation of Continuous Climb Departures (CCD), also refereed to as green departures, has not received so far the same attention. The idea is analogous to CDA. In contrast to a standardized procedure, a CCD procedure aims at climbing continuously to as higher as possible and avoiding any level segments of flight that will result in excess of noise, fuel consumption, and emissions. So far all approaches to the design of procedures for takeoff and climb have been devoted to individual operations, focusing in defining more efficient, environmentally friendly procedures. Figure 1.6 illustrates a comparison between CCD and conventional departures.
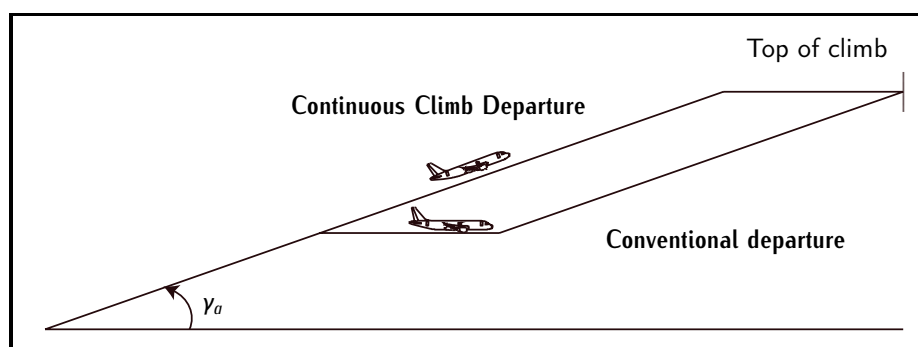
Figure 1.6: Comparison between CCD and a conventional departure.

Extensive work has been done in the field of CCD to design Noise Abatement Procedures (NAP) that minimize noise nuisance around airport during departures. To name a few, in [61, 62] multi-objective optimization strategies for designing departure abatement procedures were presented with application to Girona airport. More detailed background can be found in [63]. In [64] the benefits of aircraft trajectory design to reduce noise and fuel consumption were analyzed. The problem is addressed as an optimal control problem and solved using a Gauss pseudospectral method. In [65] a multi-objective, constrained, nonlinear optimization problem is solved to obtain optimal departure procedures taking into account noise, air quality, and global warming.

Optimal cruise performances have been also extensively studied and discussed. In [66] an aircraft model with mass and speed as state variables is considered, and an indirect method is used to solve the optimal control problem to minimize fuel consumption during cruise phase at constant altitude. The optimal speed decreases as fuel is being burnt during the flight, and savings of 7% when compared to steady (constant speed) profile are reported. Similar conclusions are reported in [67], where the minimum fuel cruise trajectory at constant altitude with fixed arrival time is solved using optimal control, and [68] where the same problem is analyzed including wind effects. Derived from such cruising velocity profiles, new concepts of operation have arisen such as the possibility of absorbing ground delays on air within new airspeed reduction Air Traffic Flow Management (ATFM) concepts [69], resulting in less fuel consumption and emissions. Moreover, research has also been focused on the design of optimal cruising trajectories to mitigate climate impact in persistent contrail formation areas [70].

An observation that arises from the considerations given above is that all the research on optimizing trajectories and operational procedures in ATM has been limited to one portion of the flight. These studies are not exhaustive since the optimal solution of a complete flight could significantly differ if we take into account the influence of all portions of the flight at the same time. Figure 1.7 illustrates a comparison between future and conventional operational concepts for complete flights.
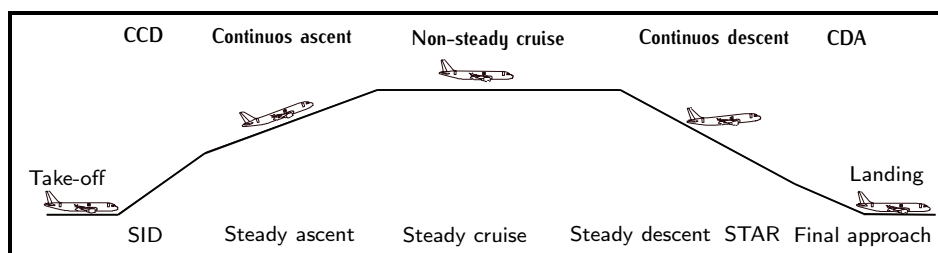
Figure 1.7: Comparison between conventional and future operational concepts

So far, the use of optimal control has been limited to the optimization of trajectories and operational procedures for a single-aircraft, and typically within a portion of the flight (departure and arrival, or cruising performances). In order to cope with the complexity of the ATM system, the study of complete trajectories for a single-aircraft is not enough. A more realistic study should include at least multiple aircraft scenarios, so that potential conflicts and capacity restrictions can be tackled, and uncertainty, so that the obtained trajectories have more chances to be feasible and thus increasing their predictability and the effectiveness of the strategic trajectory planning process. However, this thesis does not consider multiple aircraft scenarios and focuses in deterministic, single-aircraft trajectory optimization.

1.5   Contributions

The original contribution of this dissertation include:

a) A modeling framework for tackling the constraint-based trajectory planning problem using optimal control techniques. The framework is based on switched systems theory and it is capable to reflect continuous dynamics, discrete dynamics, and decision-making processes.

b) A multiphase optimal control approach to aircraft trajectory planning [71]. The multiphase optimal control problem is converted into a NLP problem, first making the unknown switching times part of the state, and then applying a Hermite-Simpson collocation method to convert the dynamic equations of the system into constraints. The resulting NLP problem has been solved using the interior point based nonlinear solver IPOPT. This approach overcomes some of the above mentioned drawbacks in solving multiphase optimal control problems. A minimum fuel application to vertical motion can be found in [72, 73], in which we define and validate what are termed optimized-procedured profiles, proposing a new methodology for flight plan definition. A minimum fuel application to 3D motion can be found in [74, 75]. Moreover, we also present an application to the optimal take-off weight trajectory planning problem. Results show that the efficiency of current flight plans could be substantially improved.

c) A multiphase mixed-integer optimal control approach to trajectory planning [76]. The multiphase mixed integer optimal control problem is converted into a MINLP problem, first making the unknown switching times part of the state, then applying a 5th degree HLGL collocation method to convert the dynamic equations of the system into constraints, and finally including binary variables to model decision-making processes. The resulting MINLP problem has been solved using the NLP BB algorithm implemented in the solver BONMIN. A minimum cost (fuel and overfly charges) application to en-route trajectory planning under wind effects can be found in [76, 77]. The decision-making process arises in determining the optimal sequence of waypoints among a set of alternatives that compose the airspace structure. The approach has also been extended to a collision avoidance problem [78]. It represents a promising approach in a complex environment as ATM since it overcomes a classical optimal control limitation: solving an optimal control problem that combines binary and continuous variables.

Overall, this dissertation aims at defining and validating optimal control based techniques for aircraft trajectory planning towards future operational concepts. Results show that the flight planning problem can be effectively tackled using optimal control based techniques. Therefore, the techniques studied in this thesis have shown a strong potentiality to be further exploited towards defining more efficient flight plans based on TBO.

1.6   Outline of the dissertation

This dissertation is organized as follows:

In Chapter 2 the optimal control problem is presented, stating the problem and discussing the different numerical approaches to finally focus on the so called direct collocation methods, specifically on those based in HLGL schemes.

In Chapter 3 the multiphase mixed-integer optimal control problem is presented. The focus is first on defining briefly what are the characteristics of a switched system. Then, the multiphase optimal control problem is stated, presenting an approach to solve it based on a parameterization of the switching instants. Finally, the multiphase mixed-integer optimal control problem is stated.

In Chapter 4 the models used in this dissertation are presented. They include continuous phenomena, discrete phenomena, and decision-making processes. Continuos models include aircraft performance, atmospheric conditions, and meteorological conditions. Discrete models include airspace structure, operational constraints and procedures, aerodynamic configurations of the aircraft, and overfly charges. Finally, decision-making processes are also modeled.

The multiphase optimal control applications to aircraft trajectory planning are presented in Chapter 5. The optimized-procedured profiles are compared with free-flight and fully-procedured profiles over short and medium range vertical flights. An extension to 3D motion is also presented, solving the optimal take-off weight trajectory planning problem. Numerical results are discussed.

The multiphase mixed-integer optimal control application to en-route trajectory planning is presented in Chapter 6. The Nlp Bb used to solve the problem is presented. Numerical results are discussed.

Finally, in Chapter 7 some conclusions and future lines of research are drawn.

In Appendix A the BADA 3.6 aircraft performance model is reported. Data for an Airbus 320 and Airbus 330 are included.

2

Optimal Control

I N THIS CHAPTER we give general background on the continuous time optimal control problem. We begin with the definition of the problem and the description of the main optimality criteria for its resolution. Optimality conditions can be derived either from Pontryagin's Maximum Principle or from the Hamilton–Jacobi–Bellman equation. However, it is very difficult to solve analytically optimal control problems even for the simplest cases. Numerical methods must be employed. There exist three basic numerical approaches to solve continuos time optimal control problems: dynamic programming methods, indirect methods, and direct methods. The characteristics of these methods will be described and their advantages and disadvantages discussed. Direct methods, in particular direct collocation methods, appear to be the most suitable. Direct collocation methods using Hermite–Legendre–Gauss–Lobatto polynomials habe been used in this thesis for solving continuous time optimal control problems. Thus, a detailed description of these methods will be given.

2.1   Problem definition

Control theory is a discipline that studies the behavior of dynamical systems with control inputs. In general, the aim is to control the state of the dynamical system in some prescribed manner.

Control systems can be classified in two categories: terminal and tracking control. Terminal control refers to the task of guiding a system from a given initial state to a desired final state in a given time, often in the presence of constraints. The objective of tracking control is to maintain the system's state in a close neighborhood of the nominal trajectory.

The control exerted over the system through the control inputs can follow two fundamental strategies: open-loop and closed-loop. The main difference is that in open-loop there is no knowledge of the actual state of the system at a given time, and the control is exerted based upon a model of the system dynamics and an estimate of its state. On the contrary, in a close-loop strategy the actual state of the system is estimated based on measurements and recursively provided to the controller through a feedback loop.

The goal of optimal control theory is to determine the control input that will cause a system to achieve the control objectives, satisfying the constraints, and at the same time optimize some performance criterion.

The trajectory planning problem is in general solved following an open loop terminal control problem. This strategy allows all the constraints acting on the dynamical system, including the dynamic constraints, to be taken into account in such a way that the resulting trajectory is admissible. However this problem has an infinite number of solutions. To eliminate this redundancy optimal control techniques can be used to select only one of them, the trajectory that optimize a given criterion. Once an admissible trajectory or the optimal one has been found, a closed loop tracking control strategy is in general used to follow it.

In this thesis we are interested in commercial aircraft trajectory planning. To solve this problem we use open-loop optimal control techniques. Also, we focus on dynamical systems evolving in continuous time. Then, the general continuous, open-loop optimal control problem studied in this thesis consists in finding the control law and the corresponding trajectory that steers a given dynamic system between an initial and a final state, and satisfies the constraints while minimizing certain optimality criterion.

2.1.1 Preliminary definitions

The dynamical systems studied in this thesis can be modeled by a set of differential state equations and a set of algebraic equations. Such system is called Differential Algebraic Equations (DAE) system and can be defined as follows [34]:

Definition 2.1 (*DAE system*). *A DAE system is a dynamical system of the form:*

$$\dot{x}(t) \;=\; f(t, x(t), u(t), l); \tag{2.1a}$$
$$0 \;=\; g(t, x(t), u(t), l). \tag{2.1b}$$

The set of equations (2.1a) is a set of first-order differential equations and the set of equations (2.1b) is a set of algebraic equations. Variable $t \in [t^I, t^F] \subset \mathbb{R}$ represents time and $l \in \mathbb{R}^{n_l}$ is a vector of parameters. $x(t) : [t^I, t^F] \mapsto \mathbb{R}^{n_x}$ represents the state variables. The state variables $x(t)$ contain both differential and algebraic variables, that is, $x(t) = (y(t), z(t))$. $y(t) : [t^I, t^F] \mapsto \mathbb{R}^{n_y}$ represents the differential states variables, i.e., the state variables with time derivative, and $z(t) : [t^I, t^F] \mapsto \mathbb{R}^{n_z}$ represents the algebraic state variables, i.e., the state variables without time derivative. $u(t) : [t^I, t^F] \mapsto \mathbb{R}^{n_u}$ represents the control functions, also referred to as control inputs, assumed to be measurable. The function $f : [t^I, t^F] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_l} \mapsto \mathbb{R}^{n_x}$ is assumed to be piecewise Lipschitz continuous to ensure existence and uniqueness of a solution. The derivative of the algebraic right hand side function $g : [t^I, t^F] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_l} \mapsto \mathbb{R}^{n_z}$ with respect to z, i.e., $\frac{\partial g}{\partial z} \in \mathbb{R}^{n_z \times n_z}$, is assumed to be regular in order to guarantee System (2.1) is a DAE system of index 1.

We are interested in studying controllable systems, i.e., dynamic systems which have the following property:

Definition 2.2 (*Controllability*). *A dynamical system is said to be controllable if for every vector initial conditions $x^I = x(t^I)$ and every state $x^F \in \mathbb{R}^{n_x}$, there exist a finite time t^F and control function $u(t) \in \mathbb{R}^{n_u}$, $t \in [t^I, t^F]$, such that $x(t^F) = x^F$.*

In particular, we are interested in studying trajectories of such systems. Then, let us define trajectory:

Definition 2.3 (*Trajectory*). *Given the DAE System (2.1), a trajectory is given by*

$$\mathcal{T}(x, u, l) = \{(x(t), u(t), l) \mid t \in [t^I, t^F]\} \tag{2.2}$$

with functions $x : [t^I, t^F] \to \mathbb{R}^{n_x}$, $u : [t^I, t^F] \to \mathbb{R}^{n_u}$ and a parameter vector $l \in \mathbb{R}^{n_l}$ that satisfy set of equations (2.1).

2.1.2   Problem statement

The optimal control problem can be stated as follows:

Problem 2.1 (*Optimal Control Problem*).

$$min\ J(t, x(t), u(t), l) = E(t^F, x(t^F)) +$$

$$\int_{t^I}^{t^F} L(x(t), u(t), l)\,dt;$$

subject to:

$\dot{x}(t) = f(x(t), u(t), l),$ *dynamic equations;*   (OCP)

$0 = g(x(t), u(t), l),$ *algebraic equations;*

$x(t^I) = x^I,$ *initial boundary conditions;*

$\psi(x(t^F)) = 0,$ *terminal boundary conditions;*

$\phi_l \leq \phi[x(t), u(t), p] \leq \phi_u,$ *path constraints.*

The variables t, x(t), u(t), and parameter l are as in Section 2.1.1. Notice that the initial time t^I is fixed and the final time t^F might be fixed or left undetermined. The objective function $J : [t^I, t^F] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_l} \to \mathbb{R}$ is given in Bolza form. It is expressed as the sum of the Mayer term $E(t^F, x(t^F))$ and the Lagrange term $\int_{t^I}^{t^F} L(x(t), u(t), l)\,dt$. Functions $E : [t^I, t^F] \times \mathbb{R}^{n_x} \to \mathbb{R}$ and $L : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_l} \to \mathbb{R}$ are assumed to be twice differentiable. The system is a DAE system in which the right hand side function of the differential set of equations $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_l} \to \mathbb{R}^{n_x}$ is assumed to be piecewise Lipschitz continuous, and the derivative of the algebraic right hand side function $g : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_l} \to \mathbb{R}^{n_z}$ with respect to z is assumed to be regular. $x^I \in \mathbb{R}^{n_x}$ represents the vector of initial conditions given at the initial time t^I and the function $\psi : \mathbb{R}^{n_x} \to \mathbb{R}^{n_q}$ provides the terminal conditions at the final time and it is assumed to be twice differentiable. The system must satisfy algebraic path constraints given by the function $\phi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_l} \to \mathbb{R}^{n_\phi}$ with lower bound $\phi_l \in \mathbb{R}^{n_\phi}$ and upper bound $\phi_u \in \mathbb{R}^{n_\phi}$. Function ϕ is assumed to be twice differentiable.

Let us now briefly explain the different elements of the problem.

The dynamical system is a DAE system as in Equation 2.1 without explicit dependence on time. The system has to be steered from an initial state x^I to a final state throughout the time interval $[t^I, t^F]$. The constraint on the final state is expressed as $\psi(x(t^F)) = 0$. Thus, the boundary conditions can be expressed as:

$$x(t^I) = x^I; \tag{2.3a}$$
$$\psi(x(t^F)) = 0. \tag{2.3b}$$

Moreover, the solution must satisfy algebraic path constraints of the form

$$\phi_l \leq \phi[x(t), u(t), l] \leq \phi_u, \ t \in [t^I, t^F]. \tag{2.4}$$

Notice that such path constraint entail also simple bounds on the state variables

$$x_l \leq x(t) \leq x_u, \ t \in [t^I, t^F],$$

and control variables

$$u_l \leq u(t) \leq u_u, \ t \in [t^I, t^F].$$

A performance index gives a quantitative measure of the performance of the dynamic system over time domain. It can be expressed in terms of final time and value of the state at final time plus the integral of some function of the state, control inputs and time over the domain $t \in [t^I, t^F]$:

$$J(t, x(t), u(t), l) = E(t^F, x(t^F)) + \int_{t^I}^{t^F} L(x(t), u(t), l) dt. \qquad (2.5)$$

Let us now introduce the concepts of admissibility and optimality of trajectories [34]:

Definition 2.4 (*Admissibility*). *A trajectory $\mathcal{T}(x(t), u(t), l)$ is said to be admissible if $x(t)$ is absolutely continuous, $u(t)$ is measurable and essentially bounded, and there exists a 3-tuple $(x(t), u(t), l)$ which satisfies the set of differential–algebraic equations (2.1) and set of Constraints (2.3)-(2.4). A control function $\hat{u}(t)$ is admissible if there exists at least one admissible trajectory $\mathcal{T}(x(t), \hat{u}(t), l)$.*

Definition 2.5 (*Optimallity*). *A trajectory $\mathcal{T}(x^*(t), u^*(t), l^*)$ is said to be globally optimal if it is feasible and*

$$J(t, x^*(t), u^*(t), l^*) \leq J(t, x(t), u(t), l) \qquad (2.6)$$

holds for all admissible trajectories $\mathcal{T}(x(t), u(t), l)$. A trajectory $\mathcal{T}(x^*(t), u^*(t), l^*)$ is said to be locally optimal if it is admissible and there exists a $\delta > 0$ such that equation (2.6) holds for all admissible trajectories within a neighborhood such that

$$||u^*(t) - u(t)|| \leq \delta \ \ \forall t \in [t^I, t^F]; \ ||l^* - l|| \leq \delta. \qquad (2.7)$$

The function $x^*(t)$ is called a (local) optimal state, the vector l^* is called a (local) optimal vector of parameters, and the function $u^*(t)$ is called a (local) optimal control if they are components of a (local) optimal trajectory $\mathcal{T}(x^*(t), u^*(t), l^*)$.

Hence, the optimal control problem (OCP) consists in finding an admissible control $u^*(t)$ such that there exists an admissible trajectory $\mathcal{T}(x^*(t), u^*(t), l^*)$ between the initial state and the final state that minimizes the performance index (2.5). The final time, t^F, may be fixed or left undetermined.

The reader is referred to [25] for a more general overview on different classes of optimal control problems.

2.2   Numerical methods

Typically, optimal control problems are highly nonlinear and it is very difficult to find an analytical solution even for the simplest cases. The common practice is to use numerical methods to obtain solutions.

There are three main approaches to numerically solve continuos time optimal control problems such as problem (OCP):

1. **Dynamic Programming (DP) methods:** The optimality criteria in continuos time is based on the Hamilton–Jacobi–Belman partial differential equation [79].

2. **Indirect methods**: The fundamental characteristic is that they explicitly rely on the necessary conditions of optimality that can be derived from the Pontryagin's Maximum Principle [27]. Bryson and Ho [25] provide a thorough and comprehensive overview of necessary conditions for different types of unconstrained and constrained optimal control problems.

3. **Direct methods:** They can be applied without deriving the necessary condition of optimality. Direct methods are based on a finite dimensional parameterization of the infinite dimensional problem. The finite dimensional problem is typically solved using an optimization method, such as Nlp techniques. Nlp problems can be solved to local optimality relying on the so called Karush–Kuhn–Tucker (KKT) conditions, which give first-order conditions of optimality. These conditions were first derived by Karush in 1939 [80], and some years later, in 1951, independently by Kuhn and Tucker [81].

2.2.1   Dynamic programming methods

The basic idea in using DP is to subdivide the problem to be solved in a number of stages. Each stage is associated with one subproblem and the subproblems are linked together by a recurrence relation. The solution of the whole problem is thus obtained by solving the subproblems using recursive computations. For a more detailed insight in DP and optimal control, the reader is referred to [82].

DP has been extensively applied with success to discrete optimal control problems. Unfortunately, its application is severely restricted in the case of continuous states systems because of the "curse of dimensionality," a term coined by Bellman to describe the problem caused by the exponential increase in the size of the state space.

Therefore, for solving nonlinear, continuous optimal control problems with a large number of variables, e.g., the aircraft trajectory planning problem, DP is clearly not adequate. Other approaches, such as indirect or direct methods, must be used.

2.2.2 Indirect methods

Indirect methods rely on Pontryagin's Maximum Principle [27]. Typically, the optimal control problem is turned into a two point boundary value problem containing the same mathematical information as the original one by means of necessary conditions of optimality. Then, the boundary value problem is discretized by some numerical technique to get a solution. Thus, Indirect methods follow a "first optimize, then discretize" scheme. Numerical techniques for solving this two point boundary value problem can be classified as gradient methods [83], indirect shooting and indirect multiple shooting [84, 85], and indirect collocation [86].

The practical drawbacks of indirect methods are [26, Chap. 4.3], [31]:

- Proper formulations of the necessary conditions of optimality in a numerically suitable way must be derived. Since this formulation is rather complicated, significant knowledge and experience in optimal control is required by the user of an indirect method.

- In order to handle active constraints properly, their switching structure must be guessed.

- Suitable initial guesses of the state variables and, with special relevance, to the adjoint variables must be provided to start the iterative method. State variables have physical meaning, but adjoint variables do not, so that giving a proper initial guess might be hard and a non-proper one usually leads to non-optimal solutions. Even with a reasonable guess for the adjoint variables, the numerical solution of the adjoint equations can be ill conditioned.

- Changes in the problem formulation, e.g., by a modification of the model equations, imply formulating again the optimality conditions of the problem.

- Finally, model functions with low differentiability properties are difficult to tackle with indirect approaches.

Because of these practical difficulties, indirect methods are not suitable to solve highly constrained trajectory planning problems. In fact, rather than indirect approaches, direct methods have been extensively used for solving aerospace trajectory optimization problems in spite of the fact that they present less accuracy than indirect methods [29]. Two comprehensive surveys analyzing direct and indirect methods for trajectory optimization are [30, 87].

2.2.3 Direct methods

The so called direct methods do not use the first-order necessary conditions of the continuous optimal control problem. They convert the infinite dimensional problem into a problem with a finite set of variables, and then solve the finite dimensional problem using optimization methods. Direct methods thus follow a "first discretize, then optimize"

approach. A typical strategy is to convert the infinite problem into a Nlp problem which is solved using mathematical programming techniques [28, 88].

The most important direct numerical methods are direct shooting [89], direct multiple shooting [90] and direct collocation [91]. A good reference on the practical importance of direct methods is [26].

The direct single shooting method has been broadly used because it allows optimal control problems to be easily converted into an Nlp problem with a small number of variables even for very large problems. In single shooting only initial guesses for the control Nlp variables are required. In contrast, it is very sensitive to small perturbations on the initial condition.

The direct multiple shooting method reduces some of the problems that single shooting has. However, the multiple shooting approach increases the size of the problem because additional variables and constraints have to be included. When the problem includes inequality constraints, there is the additional disadvantage that the sequence of unconstrained and constrained arcs has to be specified in advance.

The direct collocation method do not suffer from most of the drawbacks mentioned above, and therefore they are the most suitable for aerospace trajectory optimization problems [26, 30, 87].

A taxonomy of optimal control methods for trajectory optimization is given in Figure 2.1. Notice that this taxonomy is not necessarily exhaustive.
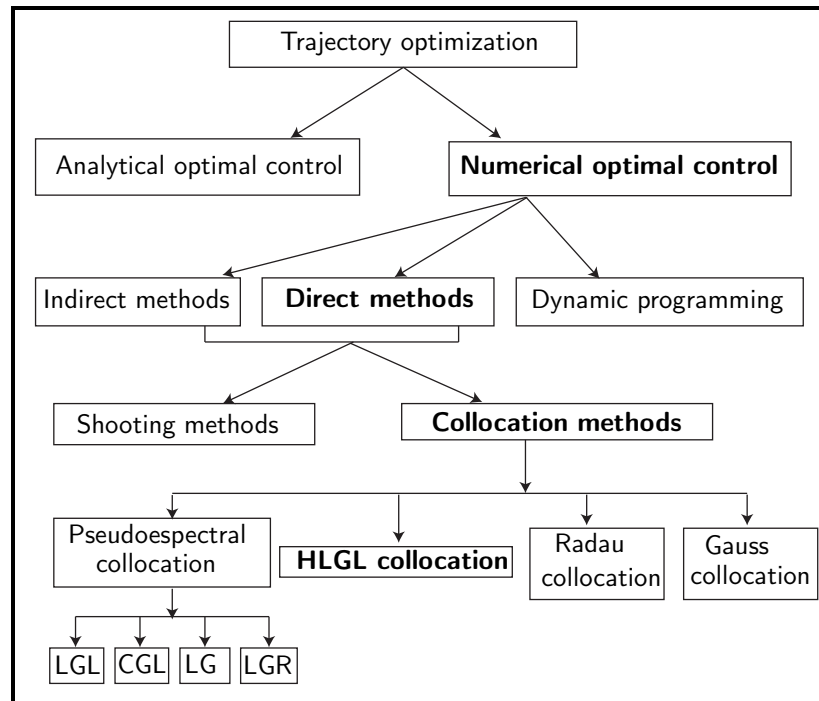
Figure 2.1: Taxonomy of trajectory optimization methods using optimal control.

2.3 Direct collocation methods

Collocation methods enforce the dynamic equations through quadrature rules or interpolation [35, 91]. A suitable interpolating function, or interpolant, is chosen such that it passes through the state values and maintains the state derivatives at the nodes spanning one interval, or subinterval, of time. The interpolant is then evaluated at points between nodes, called collocation points. At each collocation point, a constraint equating the interpolant derivative to the state derivative function is introduced to ensure that the equations of motion are approximately satisfied across the entire interval of time [92].

Collocation methods are characterized by the interpolating function and by the nodes and collocation points they use. One of the simplest methods of collocation is the Hermite–Simpson collocation method [35, 93]. In this method a third-order Hermite interpolating polynomial is used locally within the entire sequence of time subintervals, each solved at the endpoints of a subinterval and collocated at the midpoint. When arranged appropriately, the expression for the collocation constraint corresponds to the Simpson integration rule. A generalization of the method is obtained using the n-th order Hermite interpolating polynomial, and choosing the nodes and collocation points from a set of Legendre-Gauss-Lobatto points defined within the time subintervals. These choices give rise to the Hermite-Legendre-Gauss-Lobatto (HLGL) collocation method [92]. Other collocation methods are based, for instance, on Gauss or Radau collocation schemes [94, 95].

There exist also discretizations for collocation based on pseudospectral methods, which generally use global orthogonal Lagrange polynomial as the interpolants while the nodes are selected as the roots of the derivative of the named polynomial, such as Legendre-Gauss-Lobatto (LGL) (Legendre pseudospectral collocation methods), Chebyshev-Gauss-Lobatto (CGL) (Chebyshev pseudospectral collocation methods), Legendre-Gauss (LG) (Gauss pseudospectral collocation methods), or Legendre-Gauss-Radau (LGR) (Radau pseudospectral collocation methods). Since these methods use global interpolants defined over the entire time interval, the Gauss–Lobatto nodes are clustered near the endpoints.

The reader is referred to [50, 51] and references therein for recent and comprehensive reviews of pseudospectal methods for optimal control.

Needless to say, each method may be more appropriate under certain conditions. The accuracy of such discretizations has been compared in the literature [96]. In this thesis, we will use HLGL collocation methods.

2.3.1   Direct collocation using HLGL polynomials

In this collocation method the resulting polynomial interpolants take the form of a family of modified–Gaussian quadrature rules known as the Gauss–Lobatto rules [35, 36, 93]. In this section we follow [36].

For the sake of simplicity of exposition, let us consider the following unconstrained version of problem (OCP):

$$\min J(t, x(t), u(t)) = E(t^F, x(t^F)) + \int_{t^I}^{t^F} L(x(t), u(t)) dt; \qquad (2.8a)$$

subject to:

$$\dot{x}(t) = f(x(t), u(t)); \qquad (2.8b)$$
$$x(t^I) = x^I; \qquad (2.8c)$$
$$x(t^F) = x^F; \qquad (2.8d)$$

where the initial conditions $x(t^I) = x^I$ are given at the fixed initial time t^I, and the final time t^F is also fixed. The symbols J, E, L, f, t, x, u have the same meaning than in problem (OCP). The vector of parameters l has not been considered.

As Figure 2.2 illustrates, in these methods the time domain is split into a certain number N_d of smaller subintervals:

$$t^I = t_0 < t_1 < \ldots < t_{N_d-1} < t_{N_d} = t^F.$$

In each time subinterval $[t_i, t_{i+1}]$, $x_i = x(t_i)$ and $u_i = u(t_i)$ are the values of the state and controls variables at the grid point t_i, respectively. We can also define $h_i = (t^F - t^I)/N_d$, which is called the integration step size for step i, $i = 0, \ldots, N_d-1$. Thus, the independent variables of the corresponding NLP will be

$$\{x_0, u_0, x_1, u_1, \ldots, x_{N_d}, u_{N_d}\},$$

together with other independent variables according to the integration rule and the control scheme employed. They include, in general, the values of both state and control at the collocation points, e.g., the center point of the subinterval $x_{i,C}, u_{i,C}$ and the controls at the points $u_{i,a}, u_{i,b}$.

Then, the ordinary differential equation (2.8b) is replaced by a finite number of equality constraints called defect equations, that can be written in the most general form as

$$c_i(x_i, x_{i+1}, u_i, u_{i+1}, x_{i,C}, u_{i,a}, u_{i,b}, u_{i,C}) = 0, \; i = 0, \ldots, N_d - 1. \qquad (2.9)$$

Each integration scheme leads to a different formulation of this set of transcribed constraints as explained in the following sections.
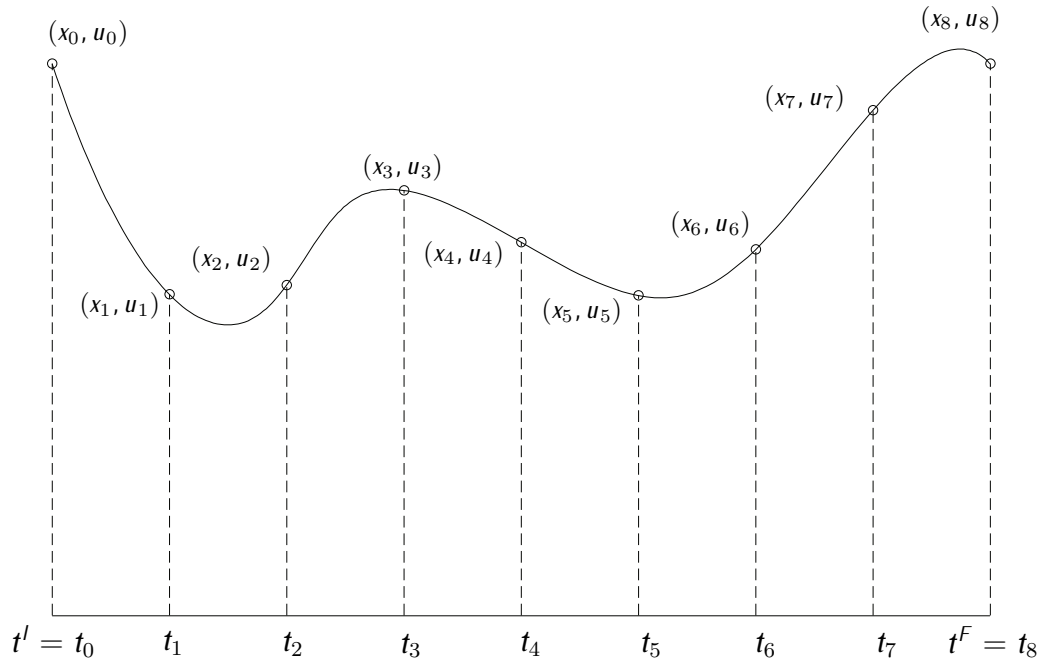
Figure 2.2: Example of time discretization scheme with $N_d = 8$ subintervals.

Basic methods

Consider a simplified form of differential equation (2.8b), $dx/dt = f(t)$.

Basic numerical integration methods to solve this differential equation rely on the trapezoidal rule

$$x(t_{i+1}) - x(t_i) = \int_{t_i}^{t_{i+1}} f(t)dt \approx \frac{h_i}{2}[f(t_i) + f(t_{i+1})], \tag{2.10}$$

with $h_i = (t_{i+1} - t_i)$, where the integrand is approximated with a linear function, and on the Simpson's rule

$$\int_{t_i}^{t_{i+1}} f(t)dt \approx \frac{h_i}{6}[f(t_i) + 4f(t_{i,C}) + f(t_{i+1})], \tag{2.11}$$

in which the integrand is approximated using a quadratic polynomial which depends on the values of the integrand at the endpoints of the subinterval $[t_i, t_{i+1}]$ and at the midpoint $t_{i,C} = (t_{i+1} + t_i)/2$ of this subinterval. These points are called collocation points. Both the trapezoid rule and the Simpson's rule belong to the so called Gauss-Lobatto family of integration rules in which the degree of the integrated polynomial coincides with the number of discrete value of the integrand used to generate the interpolating polynomial. Thus, the trapezoid rule is the second-degree rule and Simpson's rule is the third-degree Gauss-Lobatto integration rule. Higher order Gauss-Lobatto rules can be also used such as the fourth and fifth-degree rules.

An important aspect to be studied is the location of the collocation points when the interpolant is of degree higher than two. This choice depends on the discretization error which will be discussed in the next section.

Error analysis

Truncation errors in numerical integration are of two kinds: local truncation errors and global truncation errors.

The local truncation represents the error in the approximate integration over a subinterval $[t_i, t_{i+1}]$ and it is a function of the subinterval length h_i to an integer power. This power should be as large as possible. Let $p_m(t)$ be a polynomial in time of order m. This polynomial is constructed by choosing the coefficients associated with each term so that the resulting polynomial evaluates the integrand $f(t)$ at selected points in the subinterval $[t_i, t_{i+1}]$. Thus, the integration of $p_m(t)$ is an approximation of the integral of the function $f(t)$. Let the error in the integral of $p_m(t)$ over the subinterval $[t_i, t_{i+1}]$ be defined by

$$\varepsilon(f) = \int_{t_i}^{t_{i+1}} f(t)dt - \int_{t_i}^{t_{i+1}} p_m(t)dt.$$

$\varepsilon(f)$ is called the local truncation error. For instance, the local truncation errors for the trapezoid and Simpson's rules are proportional to h_i^3 and h_i^5, respectively. This means that as h_i is reduced, the local truncation error of the trapezoid rule improves cubically whereas the local truncation error of the Simpson's rule improves with a power 5, that is much faster.

The global truncation error results from using a method to integrate a function over an interval that has been divided into a finite number of subintervals. Thus, the global truncation error results from the accumulation of local truncation errors.

The order of accuracy of the approximation is that power of the step size h_i to which the truncation error is proportional. If the numerical method is employed only over one subinterval, the order of accuracy will be related to the local truncation error. Most commonly, as it is in our case, the numerical method will be employed over a finite number of subintervals, and thus the order of accuracy will be related to the global truncation error.

Because the step size and the number of local truncation error values to be summed are both proportional to the number of subintervals, the global truncation error is of order one less than the local truncation error. Therefore, the order of accuracy for the trapezoid and Simpson's rules are proportional to 2 and 4, respectively.

When analyzing the effectiveness of a numerical method to approximate a solution, the order of accuracy is the most important measure. The greater is the order of accuracy, the greater is the reduction in error if the step size is made smaller. Therefore, as the order of accuracy increases, a specified accuracy may be achieved with larger step sizes.

The relatively high order of accuracy and the simplicity of the Simpson's rule

makes it very attractive to be used in direct transcription methods for optimal control problems. Moreover, the fifth-degree Gauss-Lobatto integration scheme, which has order of accuracy 8, is a good option for large optimal control problems because employing them leads to a consistent reduction of the number of subintervals of the discretization and thus of the number of variables of the problem. However, it is necessary to seek an appropriate trade off between the number of subintervals for the discretization and the integration rule in order to achieve a determined accuracy.

For a deeper insight on the importance of the order of accuracy in HLGL collocation schemes the reader is referred to [97, Chap. 4].

Collocation point determination

In this section the choice of the collocation points will be discussed. The collocation points used to formulate an integration rule must be chosen to increase the order of accuracy of the resulting integration rule to the highest order possible. We consider a family of modified Gaussian integration rules known as the Gauss-Lobatto rules [36].

The collocation points that maximize the power of h_i in the local truncation error are the roots of the corresponding Jacobi polynomials, which are the set of polynomials that are orthogonal on the interval $[-1, 1]$ with respect to the weight function $\mathbb{W} = (1 - \mathbf{s})^\alpha (1 + \mathbf{s})^\beta$. In the Gauss-Lobatto rules, $\alpha = \beta = 1$. A subinterval with endpoints $[t_i, t_{i+1}]$ can be transformed to the interval $[-1, 1]$ using the transformation $\mathbf{s} = 2(t - t_i)/h_i - 1$. The interpolating polynomial can be calculated by interpolating $f(t)$ at the endpoints of the interval $[-1, 1]$ and at the zeros of the corresponding Jacobi polynomial.

Thus, the trapezoid rule is the second-degree rule and Simpson's rule is the third-degree Gauss-Lobatto integration rule.

For the fourth-degree Gauss-Lobatto integration rule the roots of the corresponding Jacobi polynomial (the collocation points) are $\{-\sqrt{\frac{1}{5}}, \sqrt{\frac{1}{5}}\}$, yielding the following approximate integration rule:

$$\int_{t_i}^{t_{i+1}} f(t)dt \approx \frac{h_i}{12}[f(t_i) + 5f(t_{i,C} - \sqrt{\frac{1}{5}}h_i) + 5f(t_{i,C} + \sqrt{\frac{1}{5}}h_i) + f(t_{i+1})], \quad (2.12)$$

which has an order of accuracy of 6.

For the fifth-degree Gauss-Lobatto integration rule the roots of the corresponding Jacobi polynomial (the collocation points) are $\{-\sqrt{\frac{3}{7}}, 0, \sqrt{\frac{3}{7}}\}$, yielding the following approximate integration rule:

$$\int_{t_i}^{t_{i+1}} f(t)dt \approx \frac{h_i}{180}[9f(t_i) + 49f(t_{i,C} - \sqrt{\frac{3}{7}}h_i) + 64f(t_{i,C}) +$$

$$49f(t_{i,C} + \sqrt{\frac{3}{7}}h_i) + 9f(t_{i+1})], \quad (2.13)$$

which has an order of accuracy of 8.

Application to differential equations

Trapezoid defect constraints: Applying the trapezoid rule in equation (2.10) to the differential equation $\frac{dx}{dt} = f(x)$ we obtain the resolution scheme

$$x_{i+1} = x_i + \frac{h_i}{2}(f(x_i) + f(x_{i+1})), i = 0, \ldots, N_d - 1.$$

which give rise to the following constraints

$$c_i^T(x_i, x_{i+1}) = x_i - x_{i+1} + \frac{h_i}{2}(f(x_i) + f(x_{i+1})) = 0, \ i = 0, \ldots, N_d - 1,$$

called the trapezoid system constraints or trapezoid defect constraints.

Hermite–Simpson defect constraints: Consider the differential equation $\frac{dx}{dt} = f(x)$. Simpson's rule in equation (2.11) is formulated considering a quadratic approximation of the integrand, and thus the state as a function of time $x(t)$ must be approximated by a cubic polynomial. Moreover, the polynomial used to interpolate $f(x)$ at the endpoints and center points of the subinterval is obtained as an integration of the above mentioned cubic polynomial. In this case, parameters representing the state at the endpoints x_i and x_{i+1} are used to formulate a constraint. Knowing x_i, x_{i+1}, $f_i = f(x_i)$ and $f_{i+1} = f(x_{i+1})$, a Hermite-cubic polynomial representing the state $x(t)$ between the endpoint times t_i and t_{i+1} can be constructed such as both the values and first derivatives of the interpolant polynomial coincide with the values and first derivatives of function $f(x)$ at the extremes of the subinterval. Figure 2.3 illustrates it. Such polynomial is used to generate an internal collocation point $x_{i,C}$ per subinterval, whose numerical expression is

$$x_{i,C} = \frac{1}{2}(x_i + x_{i+1}) + \frac{h_i}{8}(f(x_i) - f(x_{i+1})), \tag{2.14}$$

where $x_{i,C}$ is a discrete approximation for $x(t)$ at $t_{i,C} = \frac{t_i + t_{i+1}}{2}$ and $i = 0, \ldots, N_d - 1$. The Simpson's system constraint is then formulated using $x_{i,C}$ to evaluate the system equation resulting in a discrete value at center point of the subinterval $f_{i,C} = f(x_{i,C})$. Then, by enforcing $f_{i,C}$ to be equal to the first time derivative of the Hermite–cubic interpolant polynomial at the center point of the subinterval, i.e., $\dot{x}_{i,C} = f_{i,C}$, one defect equation per subinterval is generated:

$$c_i^S(x_i, x_{i+1}) = x_i - x_{i+1} + \frac{h_i}{6}(f(x_i) + 4f(x_{i,C}) + f(x_{i+1})) = 0, \tag{2.15}$$

with $i = 0, \ldots, N_d - 1$. These constraints are known as Hermite–Simpson defect constraints [35, 93].

For the fourth and fifth-degree Gauss-Lobatto integration rules, the process of transforming the integration rule to yield a constraint is analogous to that of Hermite-Simpson defect constraints. It is however more complex, yielding two system constraints per subinterval. For the sake of brevity, we will not go in depth. The reader is referred to [36, 97] for more details.
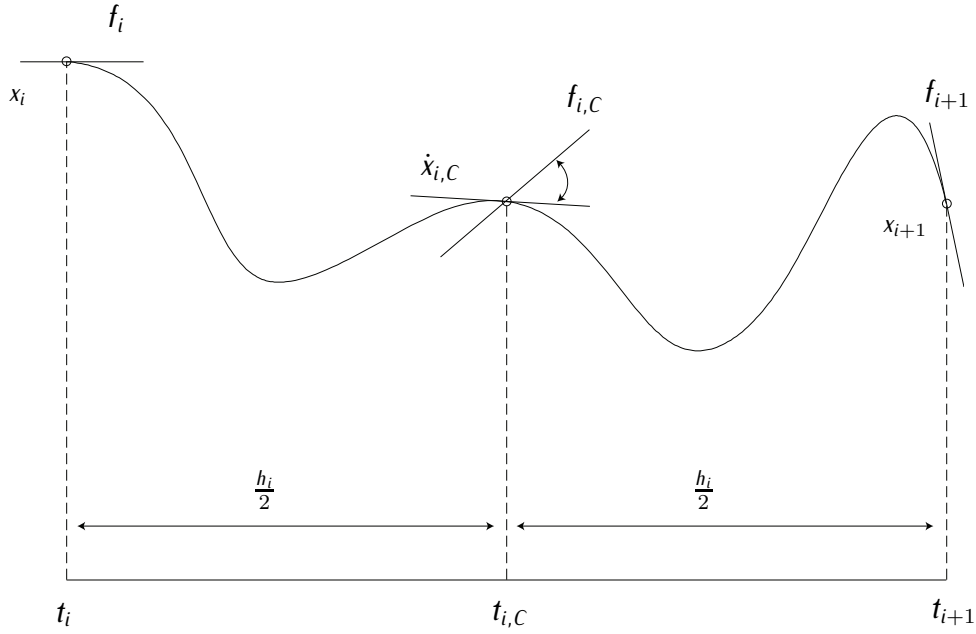
Figure 2.3: Hermite–Simson collocation scheme.

Fourth-degree Gauss–Lobatto defect constraints: Consider the differential equation $\frac{dx}{dt} = f(x)$. The fourth–degree Gauss-Lobatto integration rule is based on the fourth–degree Jacobi polynomial representing the state $x(t)$ between the endpoint times t_i and t_{i+1}. In order to uniquely define the polynomial, five pieces of information are required.

As in the Hermite–Simpson case, four pieces of information are available, namely, x_i, x_{i+1}, $f_i = f(x_i)$, and $f_{i+1} = f(x_{i+1})$. In order to complete the five pieces of information, an additional variable per subinterval is required. This variable is chosen to be the value of the state at the center point $x_{i,C}$.

The fourth–degree polynomial interpolant is used to compute discrete approximations for the state variables at the two collocation points $t_{i,a} = t_{i,C} - \sqrt{\frac{1}{5}}\frac{h_i}{2}$ and $t_{i,b} = t_{i,C} + \sqrt{\frac{1}{5}}\frac{h_i}{2}$, yielding:

$$x_{i,a} = \frac{1}{50}\{(7\sqrt{5}+9)x_i + 32x_{i,C} + (-7\sqrt{5}+9)x_{i+1}+$$
$$h_i[(\sqrt{5}+1)f(x_i) + (\sqrt{5}-1)f(x_{i+1})]\}; \quad (2.16)$$

$$x_{i,b} = \frac{1}{50}\{(-7\sqrt{5}+9)x_i + 32x_{i,C} + (7\sqrt{5}+9)x_{i+1}+$$
$$h_i[(-\sqrt{5}+1)f(x_i) + (-\sqrt{5}-1)f(x_{i+1})]\}; \quad (2.17)$$

with $i = 0, \ldots, N_d - 1$.
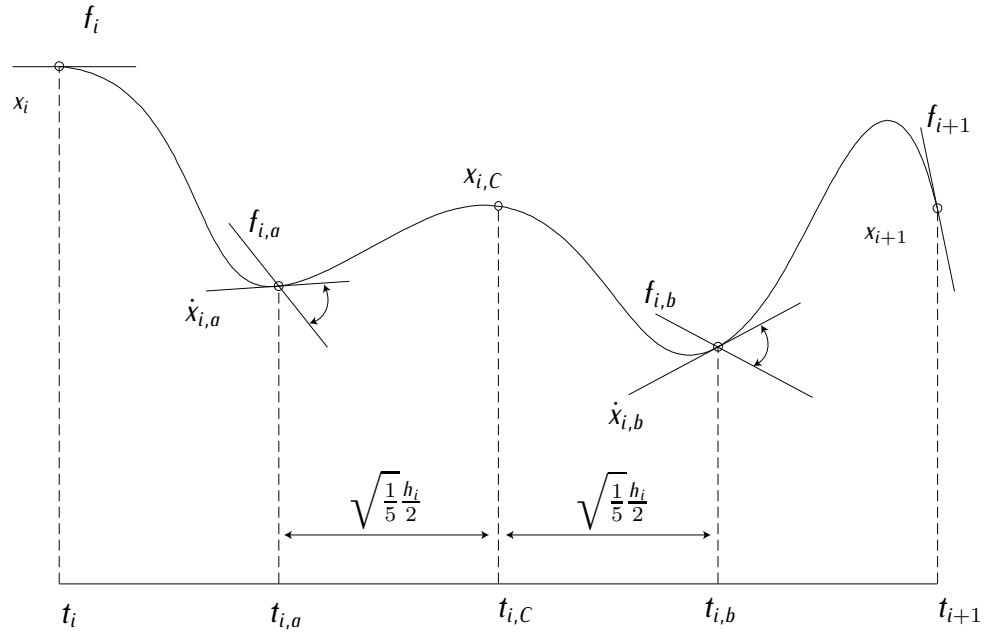
Figure 2.4: Fourth-degree collocation scheme.

As Figure 2.4 illustrates, the fourth-degree Gauss-Lobatto system of constraints is formulated to approximate the differential equation at the internal collocation points $t_{i,a}$ and $t_{i,b}$, resulting in the discrete values $f_{i,a} = f(x_{i,a})$ and $f_{i,b} = f(x_{i,b})$, respectively. Then, $f_{i,a}$ and $f_{i,b}$ are equated with the first derivative of the fourth-degree interpolant polynomial evaluated at the same collocation points, $\dot{x}_{i,a}$ and $\dot{x}_{i,b}$, respectively.

The resulting two defect equations per subinterval are:

$$c_i^{4,a}(x_i, x_{i,C}, x_{i+1}) = \frac{1}{120}\{(32\sqrt{5} + 60)x_i - 72\sqrt{5}x_{i,C} + (32\sqrt{5} - 60)x_{i+1} +$$
$$h_i[(5 + 3\sqrt{5})f(x_i) + 50f(x_{i,a}) + (5 - 3\sqrt{5})f(x_{i+1})]\} = 0; \quad (2.18)$$

$$c_i^{4,b}(x_i, x_{i,C}, x_{i+1}) = \frac{1}{120}\{(-32\sqrt{5}+60)x_i + 72\sqrt{5}x_{i,C} + (-32\sqrt{5}-60)x_{i+1} +$$
$$h_i[(5 - 3\sqrt{5})f(x_i) + 50f(x_{i,b}) + (5 + 3\sqrt{5})f(x_{i+1})]\} = 0; \quad (2.19)$$

with $i = 0, \ldots, N_d - 1$.

Fifth-degree Gauss-Lobatto defect constraints: Consider the differential equation $\frac{dx}{dt} = f(x)$. The fifth-degree Gauss-Lobatto integration rule is based on the fifth-degree Jacobi polynomial representing the state $x(t)$ between the endpoint times t_i and t_{i+1}. In order to uniquely define the polynomial six pieces of information are required. As in the previous case, four pieces of information are available, namely, x_i, x_{i+1}, $f_i = f(x_i)$, and $f_{i+1} = f(x_{i+1})$, and therefore, two additional pieces of information per subinterval are required. One variable is chosen to be the value of the state at the center point of the subinterval (which indeed is one of the three collocation points), $x_{i,C}$. The remaining piece of information results from the evaluation of the function f at the same center point of the subinterval, $f_{i,C} = f(x_{i,C})$.

As Figure 2.5 illustrates, the fifth-degree polynomial interpolant is used to compute discrete approximations for the state variables at the other two collocation points: $t_{i,a} = t_{i,C} - \sqrt{\frac{3}{7}} \frac{h_i}{2}$ and $t_{i,b} = t_{i,C} + \sqrt{\frac{3}{7}} \frac{h_i}{2}$, yielding:

$$
\begin{aligned}
x_{i,a} = \frac{1}{686}\{ & (39\sqrt{21} + 231)x_i + 224x_{i,C} + (-39\sqrt{21} + 231)x_{i+1} + \\
& h_i[(3\sqrt{21} + 21)f(x_i) - 16\sqrt{21}f(x_{i,C}) + (3\sqrt{21} - 21)f(x_{i+1})]\}; \quad (2.20)
\end{aligned}
$$

$$
\begin{aligned}
x_{i,b} = \frac{1}{686}\{ & (-39\sqrt{21} + 231)x_i + 224x_{i,C} + (+39\sqrt{21} + 231)x_{i+1} + \\
& h_i[(-3\sqrt{21} + 21)f(x_i) + 16\sqrt{21}f(x_{i,C}) + (-3\sqrt{21} - 21)f(x_{i+1})]\}; \quad (2.21)
\end{aligned}
$$

with $i = 0, \dots, N_d - 1$.

The fifth-degree Gauss-Lobatto system of constraints is formulated to approximate the differential equation at the two internal collocation points $t_{i,a}$ and $t_{i,b}$, resulting in the discrete values $f_{i,a} = f(x_{i,a})$ and $f_{i,b} = f(x_{i,b})$, respectively. Then, $f_{i,a}$ and $f_{i,b}$ are equated with the first derivative in time of the fifth-degree interpolant polynomial evaluated at the same collocation points, $\dot{x}_{i,a}$ and $\dot{x}_{i,b}$, respectively.

The resulting fifth-degree defect equations per subinterval are:

$$
\begin{aligned}
c_i^{5,a}(x_i, x_{i,C}, x_{i+1}) = \frac{1}{360}\{ & (32\sqrt{21} + 180)x_i - 64\sqrt{21}x_{i,C} + \\
(32\sqrt{21} - 180)x_{i+1} + h_i[(9 + \sqrt{21})f(x_i) & + 98f(x_{i,a}) + 64f(x_{i,C}) + \\
& (9 - \sqrt{21})f(x_{i+1})]\} = 0; \quad (2.22)
\end{aligned}
$$

$$
\begin{aligned}
c_i^{5,b}(x_i, x_{i,C}, x_{i+1}) = \frac{1}{360}\{ & (-32\sqrt{21} + 180)x_i + 64\sqrt{21}x_{i,C} + \\
(-32\sqrt{21} - 180)x_{i+1} + h_i[(9 - \sqrt{21})f(x_i) & + 98f(x_{i,b}) + 64f(x_{i,C}) + \\
& (9 + \sqrt{21})f(x_{i+1})]\} = 0; \quad (2.23)
\end{aligned}
$$

with $i = 0, \dots, N_d - 1$.
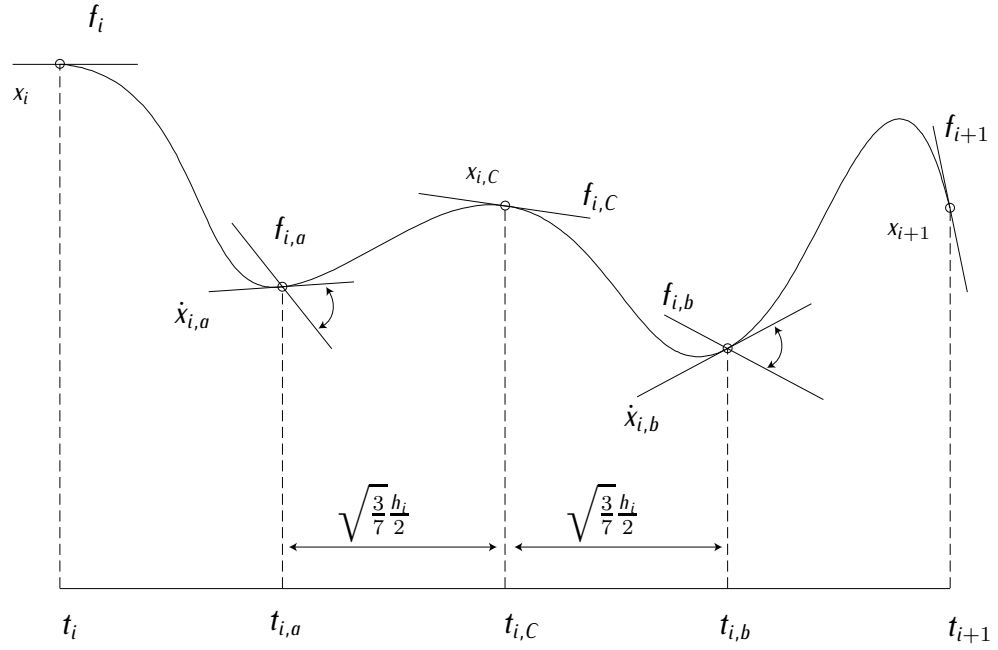
Figure 2.5: Fifth-degree collocation scheme.

Control interpolation schemes: In order to find an optimal control input, the set of differential equations (2.8b) can be solved using any of the above given sets of defect equations.

Within the subinterval $[t_i, t_{i+1}]$, the control variables are also discretized following different interpolation schemes.

In the Hermite–Simpon rule, the control interpolation schemes that have been typically used are three: linear (center control interpolation scheme), i.e., $u_{i,C} = (u_i + u_{i+1})/2$; cubic control interpolation scheme in which a Hermite cubic interpolation is used to determine the control at the center $u_{i,C}$; and free control interpolation scheme, by simply adding $u_{i,C}$ as a new independent variable in each subinterval. In the fourth and fifth-degree, the typical control interpolation scheme is the free one.

See [98, Chap. 3.3] for more details on control interpolation schemes.

For instance, in the case of free control interpolation scheme the fifth-degree defect constraints (2.22–2.23) can be expressed as:

$$c_i^{5,a}(x_i, x_{i,C}, x_{i+1}, u_i, u_{i,a}, u_{i,C}, u_{i+1}) = 0; \qquad (2.24)$$

$$c_i^{5,b}(x_i, x_{i,C}, x_{i+1}, u_i, u_{i,C}, u_{i,b}, u_{i+1}) = 0; \qquad (2.25)$$

$i = 0, \ldots, N_d - 1.$

Nʟᴘ **problem**

In this thesis both Hermite–Simpson and fifth–degree Gauss–Lobatto integration schemes will be used. For the Hermite–Simson integration scheme, a center control interpolation scheme will be employed for the control variables. For the fifth–degree Gauss–Lobatto integration scheme, a free control interpolation scheme will be employed for the control variables. Focusing for instance in the later case, the resulting Nʟᴘ problem can be expressed as follows:

$$\min \; E(x_n) + \int_{t_i}^{t_{i+1}} L(x_i, x_{i,C}, x_{i+1}, u_i, u_{i,a}, u_{i,b}, u_{i,C}, u_{i+1})dt,$$
$$i = 0, \ldots, N_d; \quad (2.26a)$$

subject to:

$$c_i^{5,a}(x_i, x_{i,C}, x_{i+1}, u_i, u_{i,a}, u_{i,C}, u_{i+1}) = 0, \; i = 0, \ldots, N_d - 1; \quad (2.26b)$$

$$c_i^{5,b}(x_i, x_{i,C}, x_{i+1}, u_i, u_{i,C}, u_{i,b}, u_{i+1}) = 0, \; i = 0, \ldots, N_d - 1; \quad (2.26c)$$

$$x_0 = x^I; \quad (2.26d)$$

$$x_{N_d} = x^F; \quad (2.26e)$$

where equation (2.26a) is a Bolza type objective functional, the set of equations (2.26b–2.26c) correspond to the defect equations for the fifth–degree Gauss–Lobatto integration scheme and equations (2.26d–2.26e) correspond to the initial an final constraint, respectively. In this problem, the integral in equation (2.26a) will be approximated according to the integration rule in equation (2.13). Notice that inequality and equality path constraints can be also included in the Nʟᴘ formulation by just expressing them in terms of discretized state and discretized control variables.

Nʟᴘ problems can be solved to local optimality relying in the KKT conditions, which give first-order conditions of optimality. The reader is referred to the following text books [28, 88] for a comprehensive treatment of Nʟᴘ.

It is interesting to point out that it has been demonstrated in [26] that the KKT Nʟᴘ necessary conditions approach the optimal control necessary conditions of optimality as the number of variables grows. Indeed, at the solution of the Nʟᴘ problem, the Lagrange multipliers can be interpreted as discrete approximations to the optimal control adjoint variables [26].

3

Multiphase Mixed-Integer
Optimal Control

I N THIS CHAPTER we introduce the multiphase mixed–integer optimal control problem. In this problem the dynamical system is a switched dynamical system evolving over continuous time. Switched dynamical systems are particular classes of hybrid systems. We begin with a formal description of switched systems. Then, we define the multiphase optimal control problem. In order to tackle such problem, we present an approach based on a parametrization of the switching instants. Finally, we define the multiphase mixed–integer optimal control problem.

3.1   Switched dynamical systems

A switched dynamical system, also referred to as switched system, is a type of hybrid dynamical system. A hybrid system is a dynamical system which exhibits interacting continuous and discrete dynamic behavior. The continuous and discrete dynamics interact at "trigger" times, which are referred to as switching instants. The main characteristic of switched systems is that the continuous state does not exhibit discontinuities at the switching instants, as in the case of general hybrid systems.

A switched system consists of a finite number of dynamical subsystems and a switching sequence. In each subsystem, also referred to as a mode of the system, the dynamics is described by a set of differential–algebraic equations. Such subsystem is also typically subject to a set of constraints. The switching sequence provides the sequence of indexes for subsystems together with the switching times at which the interactions between subsystems (switchings) take place. The switching between different subsystems is governed by certain "rules". These switching rules are termed as discrete dynamics, and they are described by logic constraints. A particular mode at any given time may be triggered by some "external process", such as a controller, computer, or human operator, in which case the switch is said to be controlled. It may also be governed by an "internal process" such as a function of time or state or a combination of both, in which case the switch is said to be autonomous.

3.1.1   Definitions

Let us define the phenomena under which the switchings are triggered:

Definition 3.1 (*Autonomous switching*). *Autonomous switching (also referred to as internal forced switching) is the phenomenon where the dynamical system changes discontinuously, or switches, when the continuous state hits certain regions of the state space. This can be interpreted as switching between different dynamical subsystems due to internal actions.*

Thus, autonomous switching refers to the inherent behavior of the system itself. An example of autonomous switching is the modification in the algebraic equations that govern the evolution of the temperature when the aircraft reaches the tropopause at an altitude of 11000 [m]. In this case, there is a mode change due to different atmospheric conditions that modify the dynamic behavior of the aircraft. See Section 4.1.4 for more information.

Definition 3.2 (*Controlled switching*). *Controlled switching (also referred to as external forced switching) is the phenomenon where the dynamical mode changes in response to a control command, usually with an associated cost. This can be interpreted as switching between different dynamical subsystems due to external actions.*

Consider an aircraft with two different power settings for climbing and cruising, respectively. An example of controlled switching is the action over the proper commands to shift from the climbing to the cruising setting that is made by the pilot commanding the aircraft. This shift from one power setting to another implies a modification in the dynamic equations of the aircraft and can take place within an interval of time, but it is an external agent (the pilot in this case) that triggers the switching. Notice that a controlled switching can be thought as a degree of freedom, so that the switching is to be set in an optimal way.

Remark 3.1. Controlled switchings can sometimes be modeled as autonomous switchings in the problem. An instance of this could be the change between take-off flap configuration and initial climb flap configuration. This is a controlled switch but in strategic in strategic planning, one may preset the switching according to a certain trigger condition, for instance, a given velocity: the switching must be triggered when the aircraft reaches this velocity. See Definition 4.7 for more information.

Let us also formally define a switched system [19]:

Definition 3.3 (*Switched system*). *A switched system is a 3-tuple $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$, where:*

- $\mathcal{G} = (\mathcal{K}, \mathcal{E})$ *is a directed graph indicating the discrete structure of the system. $\mathcal{K} = \{0, \ldots, N-1\}$ is the set of indices for subsystems. The directed edge set $\mathcal{E}$ is a subset of $\mathcal{K} \times \mathcal{K} - \{(k, k) | k \in \mathcal{K}\}$ which contains the valid events, $e \in \mathcal{E}$. For instance, if an event $e = (1, 2), 1, 2 \in \mathcal{K}$, takes place, the systems switches from subsystem 1 to 2.*

- Σ *is the set of dynamical subsystems, so that $\Sigma^k = \{f^k : \mathcal{X}^k \times \mathcal{U}^k \times \mathbb{R}^{n_{lk}} \to \mathbb{R}^{n_{xk}} | k \in \mathcal{K}; \; g^k : \mathcal{X}^k \times \mathcal{U}^k \times \mathbb{R}^{n_{lk}} \to \mathbb{R}^{n_{zk}} | k \in \mathcal{K}\}$, where f^k describes the right-hand side of the differential equation $\dot{x}(t) = f^k(x(t), u(t), l)$ for k^{th} subsystem, g^k describes the algebraic constraints $0 = g^k(x(t), u(t), l)$ for k^{th} subsystem. $\mathcal{X}^k \in \mathbb{R}^{n_{xk}} \subseteq \mathbb{R}^{n_x}$ and $\mathcal{U}^k \in \mathbb{R}^{n_{uk}} \subseteq \mathbb{R}^{n_u}$ are the state and control sets for the kth subsystem, respectively. $x(t) \in \mathbb{R}^{n_x}$ is a n_x–dimensional piecewise state variable and $u(t) \in \mathbb{R}^{n_u}$ is a n_u–dimensional piecewise control input. $l \in \mathbb{R}^{n_l}$ is a vector of parameters.*

- $\mathcal{S} = \mathcal{S}_\mathcal{A} \cup \mathcal{S}_\mathcal{C}$ *provides logic constraints that relate the continuous state and mode switchings. $\mathcal{S}_\mathcal{A}$ corresponds to the set of autonomous (internally forced) switchings, and $\mathcal{S}_\mathcal{C}$ corresponds to the set of controlled (externally forced) switchings. For instance, for an autonomous switch, when the state trajectory intersects a certain set of the state space at subsystem 1, the event $e = (1, 2)$ must be triggered and the system is forced to switch to subsystem 2. For controlled switches, only when the state belongs to a certain set, a transition 1 to 2 is possible. This controlled switch might take place in response to the control law.*

Let us now define the concepts of switching sequence and switching law [19]:

Definition 3.4 (*Switching sequence*). *For a switched system H_{SS}, a switching sequence σ in $t \in [t^I, t^F]$ is defined as*

$$\sigma = [(\tilde{t}^0, e^0), (\tilde{t}^1, e^1), \ldots, (\tilde{t}^{N-1}, e^{N-1})], \tag{3.1}$$

where $0 \leq N - 1 < \infty$, $t^I = \tilde{t}^0 \leq \tilde{t}^1 \cdots \leq \tilde{t}^{N-1} \leq \tilde{t}^N = t^F$ and e^k is the event for superindex $k \in \mathcal{K}$. $\tilde{t}^1, \ldots, \tilde{t}^{N-1}$ are the switchings of the system. The switching sequence σ in $[t^I, t^F]$ is denoted as $\sigma_{[t^I, t^F]}$. In this sequence, the pair $(\tilde{t}^k, e^k)$ indicates that at time $\tilde{t}^k$ the dynamic subsystem of the switched system changes from Σ^{k-1} to Σ^k. As a consequence, in the time subinterval $[\tilde{t}^k, \tilde{t}^{k+1}]$ the system evolution is governed by the dynamic subsystem Σ^k. In the subinterval $[\tilde{t}^{N-1}, t^F]$ the active dynamic subsystem is Σ^{N-1}.

Therefore, a switching sequence consists of the number and value of switching times together with the sequence of active dynamical subsystems and the correspondent sequence of constraints' sets. A switching sequence is given by a switching law.

Definition 3.5 (*Switching law*). *For a switched system H_{SS}, a switching law is defined to be a mapping $s : \mathbb{R}^{n_x} \times \mathbb{R} \to \bigcup_{t^I} \sigma_{[t^I, \infty)}$ which specifies a switching sequence σ for any initial point x^I and any initial time t^I.*

A general reference for hybrid systems is [99]. For more insight in switched systems the reader is referred to [19].

3.2 Multiphase optimal control

Multiphase optimal control problems are simplified cases of optimal control problems of switched dynamical systems in which both the number of switchings and the sequence of events are given. The given sequence of events defines the sequence of active dynamical subsystems with the corresponding sets of constraints. However, in order to obtain the optimal switching sequence of the system, the optimal value of the switching instants is to be determined.

3.2.1 Problem definition

Consider a switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ as in Definition 3.3 and the time interval $t \in [t^I, t^F]$. Consider also a switching sequence $\sigma_{[t^I, t^F]}$ as in Definition 3.4 with a prescribed sequence of active dynamical subsystems $\Sigma = \{\Sigma^0, \Sigma^1, \ldots, \Sigma^{N-1}\}$ and their corresponding sets of constraints.

The multiphase optimal control problem can be stated as follows:

Problem 3.1 (*Multiphase Optimal Control Problem*).

$$min\ J(t, x(t), u(t), l) = \sum_{k=0}^{N-1} \Bigg[E^k(\tilde{t}^{k+1}, x(\tilde{t}^{k+1})) +$$

$$\int_{\tilde{t}^k}^{\tilde{t}^{k+1}} L^k(x(t), u(t), l) dt \Bigg];$$

subject to:

$\dot{x}(t) = f^k(x(t), u(t), l), k = 0, \dots, N-1,$ *multiphase dynamic equations;*

$0 = g^k(x(t), u(t), l), k = 0, \dots, N-1,$ *multiphase algebraic equations;*

$x(t^I) = x^I,$ *initial boundary conditions;*

$\psi(x(t^F)) = 0,$ *final boundary conditions;*

$\phi_{lk} \le \phi^k(x(t), u(t), l) \le \phi_{u^k}, k = 0, \dots, N-1,$

multiphase path constraints;

$0 = \vartheta_{eq}^k(x(\tilde{t}^{k+1}), l), k = 0, \dots, N-2,$

$0 \le \vartheta_{ieq}^k(x(\tilde{t}^{k+1}), l), k = 0, \dots, N-2,$

interior point equality and inequality constraints.

(MOCP)

The variables t, $x(t)$, $u(t)$, k, the switching times $\tilde{t}^1, \dots, \tilde{t}^{N-1}$, and the parameter l are as in Section 3.1. Notice that the initial time t^I is fixed and the final time t^F might be fixed or left undetermined. If the final time is undetermined, it is denoted as $t^F = \tilde{t}^N$. The objective function $J : [t^I, t^F] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_l} \to \mathbb{R}$ is given in multiphase Bolza form. It is expressed as the summatory on superindex $k = 0, \dots, N-1$ of the sum of a Mayer term $E^k(\tilde{t}^{k+1}, x(\tilde{t}^{k+1}))$ and a Lagrange term $\int_{\tilde{t}^k}^{\tilde{t}^{k+1}} L^k(x(t), u(t), l) dt$. Functions $E^k : [\tilde{t}^k, \tilde{t}^{k+1}] \times \mathbb{R}^{n_{x^k}} \to \mathbb{R}$ and $L^k : \mathbb{R}^{n_{x^k}} \times \mathbb{R}^{n_{u^k}} \times \mathbb{R}^{n_{l^k}} \to \mathbb{R}$ are assumed to be twice differentiable. Dimensions $n_{x^k}, n_{u^k}, n_{l^k}$ can be different for each phase. f^k and g^k are as in Definition 3.3. f^k is assumed to be piecewise Lipschitz continuous within the time subinterval $[\tilde{t}^k, \tilde{t}^{k+1}]$, and the derivative of the algebraic right hand side function g^k is assumed to be regular within the time subinterval $[\tilde{t}^k, \tilde{t}^{k+1}]$. $x^I \in \mathbb{R}^{n_x}$ represents the vector of initial conditions given at the initial time t^I. The function $\psi : \mathbb{R}^{n_x} \to \mathbb{R}^{n_q}$ provides the terminal conditions at the final time and it is assumed to be twice differentiable. The system must satisfy multiphase algebraic path constraints within the time subinterval $[\tilde{t}^k, \tilde{t}^{k+1}]$ given by the functions $\phi^k : \mathbb{R}^{n_{x^k}} \times \mathbb{R}^{n_{u^k}} \times \mathbb{R}^{n_{l^k}} \to \mathbb{R}^{n_{\phi^k}}$ with lower bound $\phi_{lk} \in \mathbb{R}^{n_{\phi^k}}$ and upper bound $\phi_{u^k} \in \mathbb{R}^{n_{\phi^k}}$. Function ϕ^k is assumed to be twice differentiable. Dimensions n_{ϕ^k} can be different for each phase. The system must also satisfy equality and inequality interior point constraints given by the functions $\vartheta_{eq}^k \in \mathbb{R}^{n_{eq^k}}$ and $\vartheta_{ieq}^k \in \mathbb{R}^{n_{ieq^k}}$, which are assumed to be twice differentiable. Dimensions n_{eq^k} and n_{ieq^k} can be different for each phase.

Let us now briefly explain the different elements of the problem:

The evolution of the switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ within phase $k \in \mathcal{K}$ is governed by the following differential–algebraic dynamical subsystem:

$$\dot{x} = f^k(x(t), u(t), l), \ t \in [\tilde{t}^k, \tilde{t}^{k+1}]; \tag{3.2a}$$

$$0 = g^k(x(t), u(t), l), \ t \in [\tilde{t}^k, \tilde{t}^{k+1}]; \tag{3.2b}$$

with $k = 0, \ldots, N - 1$.

The switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ has to be steered from an initial state x^I to a final state throughout the time interval $[t^I, t^F]$. The constraints on the final state can be expressed as $\psi(x(t^F)) = 0$. Thus, the boundary conditions are expressed as:

$$x(t^I) = x^I; \tag{3.3a}$$

$$\psi(x(t^F)) = 0. \tag{3.3b}$$

Moreover, in each phase the solution must satisfy algebraic path constraints of the form

$$\phi_{l^k} \leq \phi^k(x(t), u(t), l) \leq \phi_{u^k}, \ t \in [\tilde{t}^k, \tilde{t}^{k+1}]. \tag{3.4}$$

Notice that a particular case of path constraints (3.4) are simple bounds on the piecewise state variables

$$x_{l^k} \leq x(t) \leq x_{u^k}, \ t \in [\tilde{t}^k, \tilde{t}^{k+1}],$$

and piecewise control variables

$$u_{l^k} \leq u(t) \leq u_{u^k}, \ t \in [\tilde{t}^k, \tilde{t}^{k+1}],$$

for $k = 0, \ldots, N - 1$.

The logic constraints $\mathcal{S}$ that relate the continuous state and mode switchings are modeled including interior point equality and inequality constraints:

$$0 = \vartheta_{eq}^k(x(\tilde{t}^{k+1}), l), k = 0, \ldots, N - 2; \tag{3.5a}$$

$$0 \leq \vartheta_{ieq}^k(x(\tilde{t}^{k+1}), l), k = 0, \ldots, N - 2. \tag{3.5b}$$

The following performance index is considered:

$$J(t, x(t), u(t), l) = \sum_{k=0}^{N-1} \left[E^k[\tilde{t}^{k+1}, x(\tilde{t}^{k+1})] + \int_{\tilde{t}^k}^{\tilde{t}^{k+1}} L^k[x(t), u(t), l] dt \right]. \tag{3.6}$$

Definition 2.4 (admissibility) and Definition 2.5 (optimality) can be carried over to problem 3.1.

Hence, the multiphase optimal control problem consists in finding an admissible control u^* and the switching instants, $(\tilde{t}^{1*}, \ldots, \tilde{t}^{N-1*})$, such that the switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ follows an admissible trajectory $\mathcal{T}(x^*, u^*, l^*)$ between the initial and final state that minimizes the performance index (3.6). The final time, $t^F = \tilde{t}^{N*}$, may be fixed or left undetermined.

A typical approach to solve problem (MOCP), as followed for instance in [32], is to solve N sequenced optimal control subproblems. In such approach, linkage constraints are needed to enforce continuity across different phases. The linkage constraints are typically transition conditions of the form

$$(\tilde{t}^{k-}, x(\tilde{t}^{k-}), u(\tilde{t}^{k-})) = (\tilde{t}^{k+}, x(\tilde{t}^{k+}), u(\tilde{t}^{k+})), \ k = 1, \ldots, N-1.$$

Such an approach has one main drawback: one must add linkage constraints to ensure continuity along time domain, increasing the size and complexity of the problem.

3.2.2   Switching times parametrization

We present an approach to tackle the multiphase optimal control problem (MOCP) avoiding the main drawback above mentioned [71]. Problem (MOCP) is converted into a conventional optimal control problem, first making the unknown switching times part of the state and then introducing a new independent variable with respect to which the switching times are fixed [100, 101]. In this reformulated problem there is a linear relation between the new variable and time, but the slope of this linear relation changes on each time interval between two switches. These slopes, which are part of the solution to the optimal control problem, are actually time scaling factors that determine the optimal switching times. Figure 3.1 illustrates it.
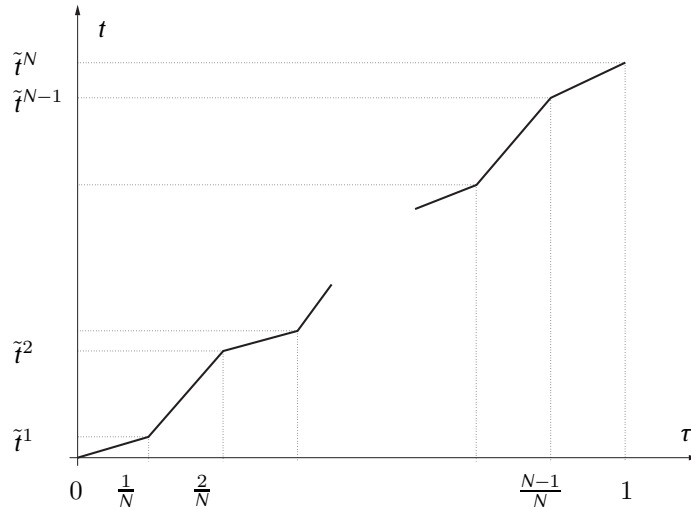

Figure 3.1:   Relation between scaled time, τ, and real (unscaled) time t.

The number of switches, $N - 1$, and the sequence of dynamical subsystems, $\Sigma = \{\Sigma^0, \Sigma^1, \ldots, \Sigma^{N-1}\}$ with their corresponding sets of constraints are considered to be known. They are given by the switching sequence $\sigma_{[t^I, t^F]}$. Without loss of generality, we can assume that $t^I = \tilde{t}^0 = 0$ and $t^F = \tilde{t}^N = 1$. The first step is to introduce the new state variables, $x_{n_x+1}(t), \ldots, x_{n_x+N-1}(t)$, which correspond to the switching times, $\tilde{t}^k$, $k \in \{1, \ldots, N-1\}$ i.e., $x_{n_x+k}(t) = \tilde{t}^k$, with $\dot{x}_{n_x+k}(t) = 0$.

We then introduce the new independent variable, τ. The relation between τ and t changes on each subinterval $[\tilde{t}^k, \tilde{t}^{k+1}]$. We establish piecewise linear correspondence between time, t, and the new independent variable, τ, so that for every chosen fixed point, τ^k, $k = 1, \ldots, N-1$, t equals $\tilde{t}^k$, i.e., $x_{n_x+k}(\tau) = \tilde{t}^k$, with $x'_{n_x+k}(\tau) = 0$, where $(\cdot)'$ denotes the derivative of $(\cdot)$ with respect to the new independent variable, τ. Any monotonically increasing sequence of $N+1$ numbers on interval $[0, 1]$ could be used. We set $\tau^k = k/N, k = 0, \ldots, N$. As a result we obtain the following expression:

$$
t = \begin{cases}
N \cdot x_{n_x+1}(\tau) \cdot \tau, & 0 \leq \tau \leq \frac{1}{N}; \\
\ldots \\
N \cdot (x_{n_x+k+1}(\tau) - x_{n_x+k}(\tau)) \cdot \tau + (k+1) \cdot x_{n_x+k}(\tau) - k \cdot x_{n_x+k+1}(\tau), & \frac{k}{n+1} < \tau \leq \frac{k+1}{N}; \\
\ldots \\
N \cdot (1 - x_{n_x+N-1}(\tau)) \cdot \tau + N \cdot x_{n_x+N-1}(\tau) - (N-1), & \frac{N-1}{N} < \tau \leq 1.
\end{cases}
$$
$$(3.7)$$

By introducing the new independent variable, τ, the set of differential–algebraic equations (3.2a)-(3.2b) that govern the evolution of the switched system on the subinterval $[\tilde{t}^k, \tilde{t}^{k+1}]$ become:

$$
\begin{aligned}
x'(\tau) &= N \cdot (x_{n_x+k+1}(\tau) - x_{n_x+k}(\tau)) \cdot \hat{f}^k(x(\tau), u(\tau), l), \ \tau \in [\tau^k, \tau^{k+1}]; && (3.8a) \\
0 &= \hat{g}^k(x(\tau), u(\tau), l), \ \tau \in [\tau^k, \tau^{k+1}]; && (3.8b)
\end{aligned}
$$

where:

$$
\hat{f}^k(x(\tau), u(\tau), l) = f^k(x(t(\tau)), u(t(\tau)), l), \ \tau \in [\tau^k, \tau^{k+1}]; \tag{3.9a}
$$

$$
\hat{g}^k(x(\tau), u(\tau), l) = g^k(x(t(\tau)), u(t(\tau)), l)), \ \tau \in [\tau^k, \tau^{k+1}]. \tag{3.9b}
$$

Let $\hat{x}$ be the extended state vector

$$
\hat{x}(\tau) = [x_1(\tau), \ldots, x_{n_x}(\tau), x_{n_x+1}(\tau), \ldots, x_{n_x+N-1}(\tau)]^T, \tag{3.10}
$$

where $\hat{x}(\tau) \in \mathbb{R}^{n_x+N-1}$.

The switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ has to be steered from an initial to a final configuration throughout the entire interval $\tau \in [0, 1]$, i.e., initial and final conditions (3.3) become:

$$
\hat{x}(0) = \hat{x}^I; \tag{3.11}
$$

$$
\hat{\psi}(\hat{x}(1)) = 0; \tag{3.12}
$$

where $\hat{x}^I \in \mathbb{R}^{n_x+N-1}$ and the function $\hat{\psi} : \mathbb{R}^{n_x+N-1} \rightarrow \mathbb{R}^{n_q}$.

Remark 3.2. Notice that boundary conditions (3.11–3.12) apply, in all its generality, over the extended state vector. However, since the switching times are to be determined, the boundary conditions do not apply for the last $N - 1$ elements of the extended state vector, i.e., the switching times.

The extended state vector $\hat{x}$ and the control variables are subjected to the following set of inequality path constraints:

$$\hat{\phi}^{k_l} \leq \hat{\phi}^k(\hat{x}(\tau), u(\tau), l) \leq \hat{\phi}^{k_u}, \ \tau \in [\tau^k, \tau^{k+1}], \tag{3.13}$$

where

$$\hat{\phi}^k(\hat{x}(\tau), u(\tau), l) = \phi^k(\hat{x}(t(\tau)), u(t(\tau)), l)), \ \tau \in [\tau^k, \tau^{k+1}].$$

Notice that a particular case of path constraints (3.13) are simple bounds on the extended state variables

$$\hat{x}_{l^k} \leq \hat{x}(\tau) \leq \hat{x}_{u^k}, \ \tau \in [\tau^k, \tau^{k+1}],$$

which include simple bounds on the switching times, and control variables

$$u_{l^k} \leq u(\tau) \leq u_{u^k}, \ \tau \in [\tau^k, \tau^{k+1}].$$

Interior point equality and inequality constraints (3.5) become:

$$0 = \hat{\vartheta}_{eq}^k(\hat{x}(\tau^{k+1}), l), k = 0, \ldots, N - 2; \tag{3.14a}$$

$$0 \leq \hat{\vartheta}_{ieq}^k(\hat{x}(\tau^{k+1}), l), k = 0, \ldots, N - 2. \tag{3.14b}$$

where:

$$\hat{\vartheta}_{eq}^k(\hat{x}(\tau^{k+1}), l) = \vartheta_{eq}^k(x(t(\tau^{k+1})), l)), \ k = 0, \ldots, N - 2;$$

$$\hat{\vartheta}_{ieq}^k(\hat{x}(\tau^{k+1}), l) = \vartheta_{ieq}^k(x(t(\tau^{k+1})), l)), \ k = 0, \ldots, N - 2.$$

Then, let us define on each subinterval $\frac{k}{N} < \tau \leq \frac{k+1}{N}, k = 0, \ldots, N - 1$:

$$\hat{L}^k(\hat{x}(\tau), u(\tau), l) = N \cdot (x_{n_x+k+1}(\tau) - x_{n_x+k}(\tau)) \cdot L^k(x(t(\tau)), u(t(\tau)), l). \tag{3.15}$$

We can rewrite the objective functional (3.6) as

$$\hat{J}(\hat{x}(\tau), u(\tau), l) = E(\hat{x}(1)) + \int_0^{\frac{1}{N}} \hat{L}^0(\hat{x}(\tau), u(\tau), l) d\tau + \cdots$$

$$\cdots + \int_{\frac{N-1}{N}}^1 \hat{L}^n(\hat{x}(\tau), u(\tau), l) d\tau. \tag{3.16}$$

Notice that performance index (3.16) is a Bolza functional as in equation (2.5) with a Mayer term at the end point $\tau = 1$ and a multiphase Lagrange term.

Definition 2.4 (admissibility) and Definition 2.5 (optimality) remain herein valid.

Hence, the problem consists in finding and admissible control $u^*(\tau)$ such that the switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ follows an admissible trajectory $\mathcal{T} = (\hat{x}^*(\tau), u^*(\tau), l^*)$ that minimizes $\hat{J}$ in the extended state space.

The parametrized optimal control problem can be defined as follows:

Problem 3.2 (*Parametrized Optimal Control Problem*).

$$min \, \hat{J}(\hat{x}(\tau), u(\tau), l) = E\left(\hat{x}(1)\right) + \int_0^{\frac{1}{N}} \hat{L}^0(\hat{x}(\tau), u(\tau), l)d\tau + \cdots$$

$$\cdots + \int_{\frac{N-1}{N}}^1 \hat{L}^n(\hat{x}(\tau), u(\tau), l)d\tau;$$

subject to:

$x' = N \cdot (x_{n_x+k+1}(\tau) - x_{n_x+k}(\tau)) \cdot \hat{f}^k(x(\tau), u(\tau), l), k = 0, \ldots, N-1,$
multiphase dynamic equations;

$x'_{n_x+1}(\tau) = \ldots = x'_{n_x+N-1}(\tau) = 0,$ *switching differential constraints;*

$0 = \hat{g}^k(x(\tau), u(\tau), l), k = 0, \ldots, N-1,$ *multiphase algebraic equations;* (POCP)

$\hat{x}(0) = \hat{x}^I,$ *initial boundary conditions;*

$\hat{\psi}(x(1)) = 0,$ *final boundary conditions;*

$\hat{\phi}_{l^k} \leq \hat{\phi}^k(\hat{x}(\tau), u(\tau), l) \leq \hat{\phi}_{u^k}, k = 0, \ldots, N-1,$
multiphase path constraints;

$0 = \hat{\vartheta}^k_{eq}(\hat{x}(\tau^{k+1}), l), k = 0, \ldots, N-2,$

$0 \leq \hat{\vartheta}^k_{ieq}(\hat{x}(\tau^{k+1}), l), k = 0, \ldots, N-2,$

interior point equality and inequality constraints.

In this problem τ, $\hat{x}$, $\hat{x}^I$ and $\hat{\psi}$ are as exposed in this section, while x, u, l and k are as in problem (MOCP). Designators $\hat{f}^k$, $\hat{g}^k$ are equivalent to designators f^k, g^k, respectively, besides introducing the dependence on $t(\tau)$. Designators $\hat{J}$, $\hat{L}^k$, $\hat{\phi}^k$, $\hat{\vartheta}^k_{ieq}$, $\hat{\vartheta}^k_{eq}$ are equivalent to designators J, L^k, E^k, ϕ^k, ϑ^k_{ieq}, ϑ^k_{eq}, respectively, besides introducing the dependence on $t(\tau)$ and the extended state vector $\hat{x}$ instead of x.

Thus, the resulting problem (POCP) is equivalent to problem (MOCP). Problem (POCP) is a conventional optimal control problem and can be solved as shown in Chapter 2. The last $N - 1$ components of the optimal solution of this problem, $\hat{x}^*$, will be the optimal switching times $\tilde{t}^{k*}, k = 1, \ldots, N-1$. If the final time, t^F, is undetermined, an extra element must be added to the extended state vector so that $\hat{x}^T = (x_1, \ldots, x_{n_x}, x_{n_x+1}, \ldots, x_{n_x+N})$ being $x_{n_x+N} = \tilde{t}^N = t^F$.

3.3 Multiphase mixed–integer optimal control

Multiphase mixed–integer optimal control problems are simplified cases of optimal control problems of switched dynamical systems in which the sequence of events for switching is unknown. The optimal switching sequence must be determined and this task is actually a decision-making process. Decision processes are typically modeled using integer or binary variables. See [34, 102, 103].

3.3.1 Problem definition

Preliminary definitions

Let us start with some definitions [34]:

Definition 3.6 (*Integer and binary variables*). *Let $w : [t^I, t^F] \rightarrow \mathbb{R}^{n_w}$ be a measurable function and $v \in \mathbb{R}^{n_v}$ a vector. A time-dependent or time-independent variable, respectively $w_j(t), \ 1 \leq j \leq n_w$ or $v_j, \ 1 \leq j \leq n_v$, is called an integer variable if it is restricted to values in $\mathbb{Z}$. If it is restricted to values in $\{0, 1\}$, it is called a binary variable or, in the case of $w_j(t)$, also a binary control function.*

Definition 3.7 (*Binary control switching*). *Let $w(t)$ be a binary control function vector. If we have a discontinuity in at least one component of $w(t)$ at time $\tilde{t}$, we say that the control function $w(t)$ switches.*

Remark 3.3. Note that if the control function $w(t)$ switches at time $\tilde{t}$, the time point $\tilde{t}$ is the switching time defined in Definition 3.4.

Definition 3.8 (*Feasible switching set*). *The feasible switching set Ψ is the set of time points when a discontinuity in the binary control function vector $w(t)$ may occur. Ψ is either $\Psi_t = \{\tilde{t}^1, \tilde{t}^2, \dots, \tilde{t}^{N-1}\}$, a finite set of possible switching times, or $\Psi_{free} = [t^I, t^F]$, the whole time interval.*

Remark 3.4. Notice that Ψ_t is due to allowance of switchings only at time points from a prefixed given set. In contrast, if $\Psi = \Psi_{free}$, there are no restrictions on the switchings and the controls can switch infinitely often, as $w(t)$ is only assumed to be measurable. An infinite number of switchings is not applicable in practice, therefore a limitation on the number of switchings of the binary control functions must be imposed. This limitation is achieved by a lower limit Ψ_{MIN} on the length of the time subinterval between two consecutive switching times.

Definition 3.9 (*Binary admissibility*). *$w(t)$ is called a binary admissible control function vector on $[t^I, t^F]$, if $w(t) \in \Omega(\Psi)$, where $\Omega(\Psi) = \{w : [t^I, t^F] \rightarrow \{0, 1\}^{n_w}; w(t) \text{ piecewise constant with jumps only at times } \tilde{t}^j \in \Psi \text{ and } \tilde{t}^j - \tilde{t}^{j-1} \geq \Psi_{MIN} \text{ with } j > 1\}.$*

Problem statement

Consider a switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ as in Definition 3.3, and the time interval $t \in [t^I, t^F]$. We can state the multiphase mixed-integer optimal control problem as:

Problem 3.3 (*Multiphase mixed–integer optimal control problem*).

$$min \; J(t, x, u, w, v, l) = \sum_{k=0}^{N-1} \left[E^k[\tilde{t}^{k+1}, x(\tilde{t}^{k+1}), w(\tilde{t}^{k+1}), v] + \right.$$
$$\left. \int_{\tilde{t}^k}^{\tilde{t}^{k+1}} L^k[x(t), u(t), w(t), v, l] dt \right];$$

subject to:

$\dot{x}(t) = f^k(x(t), u(t), w(t), v, l), k = 0, \ldots, N-1,$

multiphase dynamic equations;

$0 = g^k(x(t), u(t), w(t), v, l), k = 0, \ldots, N-1,$

multiphase algebraic equations;   (MMIOCP)

$x(t^I) = x^I,$ *initial boundary conditions;*

$\psi(x(t^F)) = 0,$ *final boundary conditions;*

$\phi_{l^k} \le \phi^k[x(t), u(t), w(t), v, l] \le \phi_{u^k}, k = 0, \ldots, N-1,$

multiphase path constraints;

$0 = \vartheta_{eq}^k[x(\tilde{t}^{k+1}), v, l], k = 0, \ldots, N-2,$

$0 \le \vartheta_{ieq}^k[x(\tilde{t}^{k+1}), v, l], k = 0, \ldots, N-2,$

interior point equality and inequality constraints;

$w(t) \in \Omega(\Psi),$ *binary admissible constraints;*

$v \in \{0, 1\}^{n_v},$ *integer constraints.*

The variables t, $x(t)$, $u(t)$, k, the switching times $\tilde{t}^1, \ldots, \tilde{t}^{N-1}$, the vector x^I, the designator ψ, and the parameter l are as in problem (MOCP). Notice that the initial time t^I is fixed and the final time t^F might be fixed or left undetermined. If the final time is undetermined, it is denoted as $t^F = \tilde{t}^N$. Designators $J, L^k, E^k, f^k, g^k, \phi^k, \vartheta_{ieq}^k, \vartheta_{eq}^k$ are as in problem (MOCP) besides changes of dimension due to the additional integer (binary) variables $v \in \{0, 1\}^{n_v}$ and the binary control functions $w : [t^I, t^F] \to \{0, 1\}^{n_w}$.

Let us now briefly explain the different elements of the problem:

The evolution of the switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ within phase $k \in \mathcal{K}$ is governed by the following differential–algebraic dynamical subsystem:

$$\dot{x} = f^k(x(t), u(t), w(t), v, l), \; t \in [\tilde{t}^k, \tilde{t}^{k+1}]; \quad (3.17a)$$
$$0 = g^k(x(t), u(t), w(t), v, l), \; t \in [\tilde{t}^k, \tilde{t}^{k+1}]; \quad (3.17b)$$

with $k = 0, \ldots, N-1$. The switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ has to be steered from an initial state x^I to a final state throughout the time interval $t \in [t^I, t^F]$. The constraint on the final state is expressed as $\psi(x(t^F)) = 0$. Thus, the boundary conditions can be stated as:

$$x(t^I) = x^I; \tag{3.18a}$$

$$\psi(x(t^F)) = 0. \tag{3.18b}$$

Moreover, the solution must satisfy algebraic path constraints of the form

$$\phi_{l^k} \leq \phi^k[x(t), u(t), w(t), v, l] \leq \phi_{u^k}, \ t \in [\tilde{t}^k, \tilde{t}^{k+1}], k = 0, \ldots, N-1. \tag{3.19}$$

Notice that a particular case of path constraints (3.19) are simple bounds on the piecewise state variables

$$x_{l^k} \leq x(t) \leq x_{u^k}, \ t \in [\tilde{t}^k, \tilde{t}^{k+1}],$$

and piecewise control variables

$$u_{l^k} \leq u(t) \leq u_{u^k}, \ t \in [\tilde{t}^k, \tilde{t}^{k+1}],$$

for $k = 0, \ldots, N-1$.

The logic constraints $\mathcal{S}$ that relate the continuous state and mode switching are modeled including interior point equality and inequality constraints:

$$0 = \vartheta_{eq}^k[x(\tilde{t}^{k+1}), v, l], k = 0, \ldots, N-2; \tag{3.20a}$$

$$0 \leq \vartheta_{ieq}^k[x(\tilde{t}^{k+1}), v, l], k = 0, \ldots, N-2. \tag{3.20b}$$

Binary admissible control functions are considered on the interval $t \in [t^I, t^F]$

$$w(t) \in \Omega(\Psi), \tag{3.21}$$

where $\Psi = \Psi_t = \{\tilde{t}^1, \tilde{t}^2, \ldots, \tilde{t}^{N-1}\}$, prefixing thus the number of switching times, or $\Psi = \Psi_{free}$, in which case the number of switchings is undetermined. Moreover, time independent integer (binary) variables are also included in the problem:

$$v_j \in \{0, 1\}, \ j = 1 \ldots n_v. \tag{3.22}$$

The following performance index is considered:

$$J(t, x, u, w, v, l) = \sum_{k=0}^{N-1} \left[E^k[\tilde{t}^{k+1}, x(\tilde{t}^{k+1}), w(\tilde{t}^{k+1}), v] + \int_{\tilde{t}^k}^{\tilde{t}^{k+1}} L^k(x(t), u(t), w(t), v, l) dt \right]. \tag{3.23}$$

Definition 3.10 (*Feasibility of binary control functions*). *A vector $w(t)$ of binary control functions is said to be feasible if it is binary admissible and there exists an admissible trajectory for problem* (MMIOCP).

Definition 2.3 (trajectory), Definition 2.4 (admissibility), and Definition 2.5 (only in the case of global optimality) can be carried over to problem (MMIOCP). However, the local optimality definition can not be extrapolated to problem (MMIOCP) because it makes no sense to use a δ neighborhood when dealing with binary variables. Instead, a Hamming distance for binary variables can be used. This distance corresponds to $||(\cdot)||_1$ in the space $\{0, 1\}^{n_v}$. See [34, Chap. 1].

Hence, the multiphase mixed–integer optimal control problem consists in finding an admissible control $u^*(t)$, the switching instants, $(\tilde{t}^{1*}, \ldots, \tilde{t}^{N-1*})$, a feasible binary control function $w^*(t)$, and a vector of integer (binary) variables v^* such that the switched system $H_{ss} = [\mathcal{G}, \Sigma, \mathcal{S}]$ follows an admissible trajectory $\mathcal{T}(x^*(t), u^*(t), w^*(t), v^*, l^*)$ between the initial state and the final state that minimizes the performance index (3.23). The final time, $t^F = \tilde{t}^N$, may be fixed or left undetermined.

Remark 3.5. Notice that the set of binary control functions $w^*(t)$ and the vector of integer (binary) variables v^* provide the sequence of events for switching. Thus, the optimal switching sequence σ^* is also obtained as part of the solution to problem (MMIOCP).

An important concept in the numerical resolution of multiphase mixed–integer optimal control problems is the notion of relaxed optimal control problem:

Definition 3.11 (*Relaxed multiphase optimal control problem*). *The relaxation of a multiphase mixed–integer optimal control problem is the multiphase optimal control problem obtained by replacing equation (3.21) and equation (3.22) by*

$$
\begin{aligned}
w(\cdot) &\in \bar{\Omega}(\Psi), \text{ and} \\
v &\in [0, 1]^{n_v}.
\end{aligned}
$$

The relaxed function space $\bar{\Omega}(\Psi)$ is defined as: $\bar{\Omega}(\Psi) = \{w : [t^I, t^F] \to [0, 1]^{n_w}; w(\cdot)$ piecewise constant with jumps only at times $\tilde{t}^k \in \Psi\}$.

4

Problem Modeling

I N THIS CHAPTER we present the models used to solve the aircraft trajectory planning problem. Due to the nature of the problem, continuos phenomena, discrete phenomena, and decision-making processes must be described.

First, we focus on the continuos part of the model. We present the aircraft equations of motion based on a three degrees of freedom model, together with the atmospheric model, the meteorological model, and the aircraft performance model. Additionally, since the continuous motion of the aircraft is typically constrained by performance limitations, path constraints must be taken into account for defining the domain of both states and control variables.

Then, we model the discrete phenomena of a flight. The flight of an aircraft has inherently a discrete nature made of multiple flight phases and multiple dynamic subsystem. Thus, together with the aircraft continuous dynamics, a discrete dynamics must be considered. The discrete dynamics governs the evolution of a set of discrete variables that represent the discrete state of the system. The different discrete states are characterized by different dynamical subsystems and different sets of path constraints which govern the continuous motion of the aircraft through the phases of the flight. Thus, we present the discrete variables that have been included to model the flight phase, the aerodynamic configuration (flap configuration), the dynamic mode (3D, vertical, or horizontal motion), the operational procedure (for instance, constant velocity), the atmosphere mode (below or above the troposphere), and the airspace structure of ATS routes (a set of waypoints).

Finally, we focus on modeling the decision-making processes introducing binary variables.

4.1 Continuous dynamics

A common assumption in aircraft trajectory optimization is to consider a 3–DOF dynamic model that describes the point variable–mass motion of the aircraft over a flat Earth model. For further details on aircraft continuous dynamics see for instance [104].

4.1.1 Reference frames

Definition 4.1 (*Earth reference frame*). *An Earth reference frame $F_e(O_e, x_e, y_e, z_e)$ is a rotating topocentric (measured from the surface of the Earth) system. The origin O_e is any point on the surface of Earth defined by its latitude θ_e and longitude λ_e. Axis z_e points to the center of Earth, x_e lies in the horizontal plane and points to a fixed direction (typically north), and y_e forms a right-handed reference (typically east).*

The system $F_e(O_e, x_e, y_e, z_e)$ is sometimes referred to as navigational system since it is used to represent the trajectory of an aircraft from the departure airport. Typically, in order to express the altitude the axis h_e ($h_e \equiv -z_e$) is used.

Definition 4.2 (*Wind axis frame*). *A wind axes frame $F_w(O_w, x_w, y_w, z_w)$ is linked to the instantaneous aerodynamic velocity of the aircraft. It is a system of axes centered in any point of the symmetry plane (assuming there is one) of the aircraft, typically the center of gravity. Axis x_w points at each instant to the direction of the aerodynamic velocity of the aircraft V. Axis $z_w \equiv -h_w$ lies into the plane of symmetry, perpendicular to x_w and pointing down according to regular aircraft performance. Axis y_b forms a right-handed reference. $r = (x_e, y_e, h_e)$ is the radio-vector representing the position of the aircraft with respect to an Earth reference frame.*
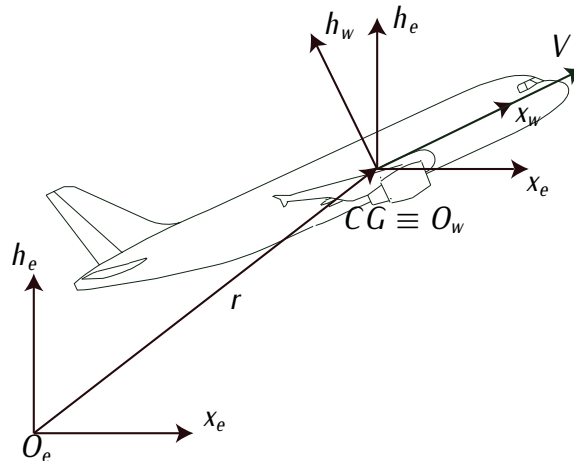

Figure 4.1:   Wind axis reference frame.

4.1.2 Modeling assumptions

Consider also the following hypotheses:

Hypothesis 4.1. Flat Earth model: *The Earth is considered flat, non rotating, and approximate inertial reference frame.*

Hypothesis 4.2. Constant gravity: *The acceleration due to gravity in the atmospheric flight of an aircraft is considered constant ($g = 9.81[m/s^2]$) and perpendicular to the surface of Earth.*

Hypothesis 4.3. Moving atmosphere: *Wind is taken into account. Vertical component is neglected due to its low influence. Only kinematic effects are considered. Dynamic effects of wind are neglected due to their low influence.*

Hypothesis 4.4. 6–DOF model: *The aircraft is considered as a rigid solid with six degrees of freedom, i.e., all dynamic effects associated to elastic deformations, to degrees of freedom of articulated subsystems (flaps, ailerons, etc.), or to the moment of inertia of rotating subsystems (fans, compressors, etc.), are neglected.*

Hypothesis 4.5. Point mass model: *The translational equations are uncoupled from the rotational equations by assuming that the airplane rotational rates are small and that control surface deflections do not affect forces. This leads to consider a 3–DOF dynamic model that describes the point mass motion of the aircraft.*

Hypothesis 4.6. Fixed engines: *The aircraft is assumed to be a conventional jet airplane with fixed engines.*

Hypothesis 4.7. Variable mass: *The aircraft is modeled as variable mass particle.*

Hypothesis 4.8. Forces acting on an aircraft: *We assume that the external forces acting on an aircraft can be generally decomposed into propulsive, aerodynamic, and gravitational forces.*

Hypothesis 4.9. Symmetric flight: *We assume that the aircraft has a plane of symmetry, and that the aircraft flies in symmetric flight, i.e., all forces act on the center of gravity and the thrust and the aerodynamic forces lie on the plane of symmetry.*

Hypothesis 4.10. Small thrust angle of attack: *The thrust angle of attack is assumed to be small.*

4.1.3   Aircraft equations of motion

3D equations of motion

Under Hypothesis 4.1–Hypothesis 4.10, the 3-DOF equations governing the translational 3D motion of an aircraft are:

- three dynamic equations relating forces to translational acceleration;

- three kinematic equations giving the translational position relative to an Earth reference frame; and

- one equation defining the variable-mass characteristics of the airplane over time.

The equations of motion are hence defined by the following Ordinary Differential Equations (ODE) system.

Definition 4.3 (*3-DOF equations of 3D motion*).

$$
\begin{aligned}
&m(t)\dot{V}(t) = T(t) - D(h_e(t), V(t), C_L(t)) - m(t)g\sin\gamma; \\
&m(t)V(t)\dot{\chi}(t)\cos\gamma(t) = L(h_e(t), V(t), m(t))\sin\mu(t); \\
&m(t)V(t)\dot{\gamma}(t) = L(h_e(t), V(t), C_L(t))\cos\mu(t) - m(t)g\cos\gamma(t); \\
&\dot{x}_e(t) = V(t)\cos\gamma(t)\cos\chi(t) + W_x(x_e(t), y_e(t)); \\
&\dot{y}_e(t) = V(t)\cos\gamma(t)\sin\chi(t) + W_y(x_e(t), y_e(t)); \\
&\dot{h}_e(t) = V(t)\sin\gamma(t); \\
&\dot{m}(t) = -T(t)\eta(V(t)).
\end{aligned}
\tag{4.1}
$$

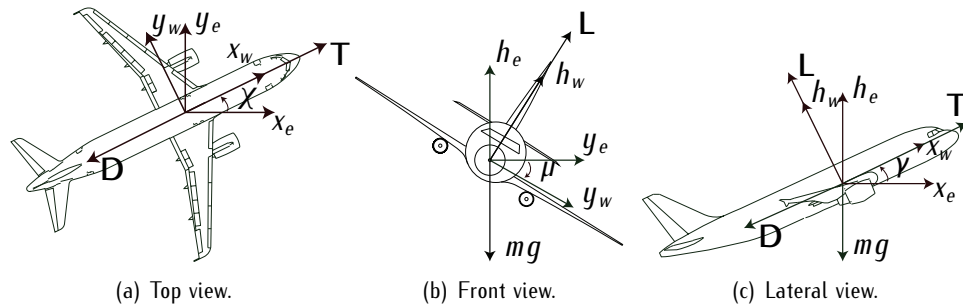(a) Top view.   (b) Front view.   (c) Lateral view.

Figure 4.2: Aircraft forces.

In the above equations:

- The three dynamics equations are expressed in an aircraft based reference frame, the wind axes system $F_w(O, x_w, y_w, z_w)$, usually x_w coincident with the velocity vector.

- The three kinematic equations are expressed in a ground based reference frame, the Earth reference frame $F_e(O_e, x_e, y_e, z_e)$, and are usually referred to as down range (or longitude), cross range (or latitude), and altitude, respectively.

- The vector $r = (x_e, y_e, h_e)$ represents the position of the center of gravity of the aircraft expressed in the Earth reference frame $F_e(O_e, x_e, y_e, z_e)$.

- W_x and W_y denote the components of the wind vector, $W = (W_x, W_y, 0)$, along the axes x_e and y_e of the Earth reference frame $F_e(O_e, x_e, y_e, z_e)$.

- μ, χ and γ are the bank angle, the heading angle, and the flight-path angle, respectively.

- m is the mass of the aircraft, and η is the specific fuel consumption.

- g is the acceleration due to gravity.

- V is the true air speed of the aircraft.

- T is the engines' thrust, i.e., the propulsive force generated by the aircraft's engines. It depends on the altitude h, Mach number M, and throttle π by a relationship $T = T(h, M, \pi)$ that is supposed to be known.

- Lift, $L = C_L S \hat{q}$, and drag, $D = C_D S \hat{q}$, are the components of the aerodynamic force, where C_L is the dimensionless coefficient of lift, C_D is the dimensionless coefficient of drag, $\hat{q} = \frac{1}{2}\rho V^2$ is referred to as dynamic pressure, ρ is the air density, and S is the wet wing surface.

Additional modeling assumptions are the following:

Hypothesis 4.11. Parabolic drag polar: *A parabolic drag polar is assumed, i.e.,* $C_D = C_{D_0} + C_{D_i} C_L^2$.

Hypothesis 4.12. Standard atmosphere model: *The ISA model is used for the atmosphere.*

The set of equations (4.1) contains the following ten variables:

$$(x_e(t), y_e(t), h_e(t), V(t), \gamma(t), \chi(t), \mu(t), m(t), T(t), C_L(t));$$

and seven differential equations. Therefore the system has three DOF, i.e., the system allows three variables which can be used as control variables.

In general the bank angle μ, the engine thrust T, and the lift coefficient C_L are the control variables for the aircraft, $u(t) = (\mu(t), T(t), C_L(t))$. The bank angle is commanded combining rudder and ailerons trims. The thrust is commanded by the engine throttle π. Within standard values, the coefficient of lift can be considered linearly related to the angle of attack according to $C_L = C_{L_0} + C_{L_\alpha} \alpha$. The angle of attack, and thus the coefficient of lift, is commanded by elevator trims.

Remark 4.1. In the calculation of aircraft performance, the lift coefficient will be used as a control variable instead of the angle of attack.

Vertical motion

In the vertical motion of an aircraft, besides Hypothesis 4.1–Hypothesis 4.12, the following additional hypotheses hold:

Hypothesis 4.13. Vertical motion: *The motion of the aircraft is constrained to a vertical plane, i.e., with constant course and thus constant heading angle χ. Without loss of generality, we assume that the heading angle is zero, that is, $\chi = 0$.*

Hypothesis 4.14. Leveled flight: *We suppose that the aircraft performs a leveled wing flight, thus the bank angle is zero, that is, $\mu = 0$.*

Hypothesis 4.15. Vertical actions: *We suppose that there are no actions out of the vertical plane, thus we do not consider wind perpendicular to the plane of motion, that is, $W_y = 0$.*

Then, the 3–DOF equations governing the translational vertical motion of an airplane are:

- two dynamic equations relating forces to translational acceleration;

- two kinematic equations giving the translational position relative to an Earth reference frame; and

- one equation defining the variable–mass characteristics of the airplane over time.

The equations of motion of the aircraft are in this case given by the following ODE system:

Definition 4.4 (*3-DOF equations of vertical motion*).

$$
\begin{aligned}
m(t)\dot{V}(t) &= T(t) - D(h_e(t), V(t), C_L(t)) - m(t)g \sin \gamma(t); \\
m(t)V(t)\dot{\gamma}(t) &= L(h_e(t), V(t), C_L(t)) - m(t)g \cos \gamma(t); \\
\dot{x}_e(t) &= V(t) \cos \gamma(t) \cos \chi(t) + W_x(x_e(t), y_e(t)); \\
\dot{h}_e(t) &= V(t) \sin \gamma(t); \\
\dot{m}(t) &= -T(t)\eta(V(t));
\end{aligned}
\tag{4.2}
$$

where the symbols have the same meaning as in Definition 4.3. In general, the engine thrust T and the lift coefficient C_L are the control variables in a vertical motion, i.e., $u(t) = (T(t), C_L(t))$.

Horizontal motion

In the horizontal motion of an aircraft, besides Hypothesis 4.1–Hypothesis 4.12, the following additional hypotheses hold:

Hypothesis 4.16. Horizontal motion: *The motion of the aircraft is constrained to a horizontal plane, i.e., the altitude is constant and then $\dot{h}_e = 0$.*

Looking at set of equations (4.1) and considering Hypothesis 4.16, it is easy to see that the flight path angle is zero, i.e., $\gamma = 0$.

Then, the 3-DOF equations governing the translational horizontal motion of an airplane are:

- two dynamic equations relating forces to translational acceleration;

- one algebraic equation for the vertical equilibrium;

- two kinematic equations giving the translational position relative to an Earth reference frame; and

- one equation defining the variable-mass characteristics of the airplane over time.

The equations of motion of the aircraft are in this case defined by the following Differential–Algebraic Equations (DAE) system:

Definition 4.5 (*3-DOF equations of horizontal motion*).

$$
\begin{aligned}
m(t)\dot{V}(t) &= T(t) - D(V(t), C_L(t)); \\
m(t)V(t)\dot{\chi}(t) &= L(V(t), C_L(t))\sin\mu(t); \\
0 &= L(V(t), C_L(t))\cos\mu(t) - m(t)g; \\
\dot{x}_e(t) &= V(t)\cos\chi(t) + W_x(x_e(t), y_e(t)); \\
\dot{y}_e(t) &= V(t)\sin\chi(t) + W_y(x_e(t), y_e(t)); \\
\dot{m}(t) &= -T(t)\eta(V(t));
\end{aligned}
\tag{4.3}
$$

where the symbols have the same meaning as in Definition 4.3. In general, the engine thrust T and the bank angle μ are the control variables in the horizontal motion, i.e., $u(t) = (T(t), \mu(t))$.

4.1.4   Performance, atmospheric, and meteorological models

For the complete description of the continuos motion of the aircraft additional models are required, namely, and aircraft performance model, an atmospheric model, and a meteorological model.

Aircraft performance model

BADA 3.6 [14] has been used as aircraft performance model, which provides data for a total of 295 types of aircraft, operations, and procedures. BADA also includes, among others, models for thrust, consumption, aerodynamics, and performance limitations. These models are reported in Appendix A.

Atmospheric model

The standard model of the atmosphere used in aviation and weather studies has been employed. The density of air, ρ, is a function of the altitude, h_e, given by the profiles established by the ISA model in the different layers of the atmosphere. The basic hypotheses of the ISA model are:

Hypothesis 4.17 (*Standard atmosphere*). *ISA complies with the perfect gas equation. In the troposphere the temperature gradient is constant. In the tropopause the temperature is constant.*

For the sake of brevity only the final expressions of the ISA model relating $\rho = f(h_e)$ in the different layers will be reported:

> **Troposphere** $(0 \leq h_e < 11000\ [m])$:
> $\rho/\rho_0 = (1 - 22.558 \cdot 10^{-6} \cdot h_e\ [m])^{4.2559}$;
> **Tropopause** $(h_e = 11000\ [m])$:
> $\rho/\rho_0 = 0.2971;\ and$
> **Inferior part of the stratosphere** $(11000\ [m] \leq h_e \leq 20000\ [m])$:
> $\rho_{11}/\rho_0 = 0.2971 e^{(-157.69 \cdot 10^{-6} \cdot (h_e[m]-11000))}$;

where subindexes "0" and "11" correspond to sea level and tropopause, respectively, and $\rho_0 = 1.225\ [kg/m^3]$.

Meteorological model

The wind velocity can be regarded as the sum of two terms: a deterministic component which represents the meteorological predictions available to the ATC, and a stochastic component which represents the uncertainty in these predictions.

Hypothesis 4.18. Deterministic wind: *The wind is considered as deterministic.*

The meteorological prediction are obtained form the RUC, a numerical model developed at the NOAA forecasts system laboratory [15] and [16]. Forecasts are provided via GRIB files. Data are provided four times a day into a $1° \times 1.25°$ grid and 14 different barometric altitudes.

Hypothesis 4.19. Stationary wind field: *The wind field is assumed to be stationary, i.e., its evolution over time has not been considered.*

In particular, GRIB files provide wind forecasts giving the three components (west, north and vertical) of the wind velocity vector, $W = \{W_x, W_y, W_z\}$, at each node of the grid. GRIB data are given in spherical coordinates, i.e., longitude (λ_e), latitude (θ_e) and altitude (h_e). The spherical coordinates are transformed into plain coordinates by the following transformation:

$$
\begin{aligned}
x_e &= \lambda_e(R_e + h_e)\cos\theta_e; \\
y_e &= \theta_e(R_e + h_e); \\
z_e &= -h_e;
\end{aligned}
\tag{4.4}
$$

where R_e is the radius of Earth.

In order to take into account wind in the optimal control problem, an analytical representation of wind data is needed. There are two main approaches to compute this analytical representation: interpolation and regression. We will give now a very brief overview on both approaches. Most of the introductory books on numerical analysis give detailed information on interpolation. A classical reference is [85]. In regard of regression analysis, we refer the reader, for instance, to [105].

Given a set of points (X_i, Y_i), $i = 1, \ldots, n$, the objective of interpolation is to find a function f within a class of functions that connects them, i.e., such that

$$Y_i = f(X_i).$$

Usually, the interpolation function f is used to approximate a function $Y(X)$ whose values are known only at specified values of X.

In regression analysis, usually the goal is to find a function f within a class of functions that best fits data points (X_i, Y_i), i.e., that which minimizes the following least–square sum:

$$\sum_{i=1}^{n}[Y_i - f(X_i)]^2.$$

Polynomials, are the most common classes of functions used in both approaches. A polynomial of order m, $p_m(X)$, has the form:

$$p_m(X) = \beta_0 + \beta_1 X + \beta_2 X^2 + \cdots + \beta_m X^m,\tag{4.5}$$

where $\beta_0, \ldots \beta_m$ are parameters.

The most commonly used polynomials in polynomial interpolation are cubic splines. Cubic splines are piecewise cubic polynomials with a piecewise continuos first derivative and second derivative.

However, as the number of points to be interpolated grows, the resulting analytical expressions wiggles a lot to reach all of the data points. Moreover, in the case of wind data spline interpolation, the big amount of data results in a great amount of polynomial terms and therefore including them into the optimal control problems is cumbersome and the resolution of the corresponding optimization problem is very difficult.

On the contrary, regression analysis provides rather simple analytical functions for a given set of tabular data, but at a cost of less precision in fitting data, making the inclusion of wind data in the optimal control problem a feasible task. Therefore, in this thesis regression analysis with polynomial functions has been used.

Multiple regression analysis [105, Chap. 8]:   Multiple regression analysis can be regarded as an extension of simple linear regression analysis (which involves only one independent variable) to the situation where two or more independent variables are considered.   The general form of a polynomial regression model for m independent variables is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m + \varepsilon, \tag{4.6}$$

where $\beta_0, \beta_1, \ldots, \beta_m$ are the regression coefficients that need to be estimated. The independent variables $X_1, X_2, \ldots, X_m$ may all be separate basic variables, or some of them may be functions of a few basic variables. For instance, identifying terms in equation (4.5) and equation (4.6): $X_1 = X, X_2 = X^2, \ldots, X_m = X^m$. Y represents an individual observation and ε is the error component reflecting the difference between an individual's observed response Y and the true average response $\mu_{Y|X_1, X_2, \ldots, X_m}$.

The following assumptions must be fulfilled for multiple regression analysis:

Hypothesis 4.20. Existence: *For each specific combination of values of the independent variables $X_1, X_2, \ldots, X_m$, Y is a univariate random variable with a certain probability distribution having finite mean and variance.*

Hypothesis 4.21. Independence: *The Y observations are statistically independent of one another.*

Hypothesis 4.22. Linearity: *The mean value of Y for each specific combination of $X_1, X_2, \ldots, X_m$ is a linear function of $X_1, X_2, \ldots, X_m$. That is,*

$$\mu_{Y|X_1, X_2, \ldots, X_m} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m, \text{ or}$$

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m + \varepsilon.$$

Notice that if some of the independent variables are high-order functions of a few basic independent variables, e.g., $X_2 = X_1^2$ or $X_3 = X_1 X_2$, the expression $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m$ is nonlinear in the basic variables but linear in the parameters $\beta_0, \ldots, \beta_m$.

Hypothesis 4.23. Homoscedasticity: *The variance of Y is the same for any fixed combination of $X_1, X_2, \ldots, X_m$, i.e.,*

$$\sigma^2_{Y|X_1,X_2,\ldots,X_m} = Var(Y|X_1,X_2,\ldots,X_m) = \sigma^2, \quad or \quad \sigma^2_{error|X_1,X_2,\ldots,X_m} = \sigma^2.$$

Hypothesis 4.24. Normality: *For any fixed combination of $X_1, X_2, \ldots, X_m$, the variable Y is normally distributed, i.e.,*

$$Y \sim N(\mu_{Y|X_1,X_2,\ldots,X_m}, \sigma^2), \quad or \quad \varepsilon \sim N(0, \sigma^2).$$

Notice that Hypothesis 4.24 is not necessary for the least-squares fitting of the regression model, but it is required in general for inference making.

There are two basic approaches to estimating the coefficients of a multiple regression equation: the least-squares method and the minimum variance method. In this dissertation, least-squares method is considered.

The least-squares method chooses as the best fitting model the one that minimizes the sum of squares of the distances between the observed responses and those predicted by the fitted model, i.e., it seeks the smaller deviations between observed and predicted values. Thus, the estimated regression equation can be expressed as:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \ldots + \hat{\beta}_m X_m,$$

where $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_m$ correspond to the estimated coefficients and $\hat{Y}$ corresponds to the predicted value. The sum of squares of deviations of n observed Y-values from the corresponding predicted values can be expressed as:

$$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \ldots - \hat{\beta}_m X_{mi})^2. \qquad (4.7)$$

The least-squares solution consists of the values $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_m$, that are called the least-squares estimates, for which equation (4.7) is a minimum. The minimum sum of squares is called the residual sum of squares.

In the case of study in this dissertation, the wind velocity vector W is given in a grid of n points, and we want to estimate the value of the wind velocity vector at every point $r = (x_e, y_e, h_e) \in \mathbb{R}^3$. Notice that Polynomial (4.5) becomes in this case:

$$p_m(r) = \beta_{000} + \beta_{100}x_e + \beta_{010}y_e + \beta_{001}z_e + \beta_{200}x_e^2 + \beta_{110}x_e y_e$$
$$+ \beta_{020}y_e^2 + \beta_{101}x_e z_e + \beta_{011}y_e z_e + \beta_{002}z_e^2 + \cdots + \beta_{00m}z_e^m, \quad (4.8)$$

where $p_m(r) : \mathbb{R}^3 \to \mathbb{R}, r \in \mathbb{R}^3$ and $\beta_{000}, \ldots \beta_{00m}$ are parameters.

In order to estimate the wind velocity vector, a three-dimensional multiple variable regression should be used. However, the following hypotheses are considered:

Hypothesis 4.25. Wind vertical component negligible: *The vertical component of the wind is assumed to be negligible, i.e., $W_z \approx 0$.*

Hypothesis 4.26. Wind components are not correlated: *We assume that the two horizontal components of the wind velocity vector, i.e., W_x, W_y, are non correlated.*

Hypothesis 4.25 is fulfilled in our case since the values of the vertical component of wind are negligible. Hypothesis 4.26 is valid since the linear correlation between W_x and W_y is typically in the intervals $[-0.3, -0.2]$ or $[0.2, 0.3]$. Under these hypotheses, the regression analysis can be carried out over each component independently. Therefore, two independent multiple variable regressions as described previously are used to estimate the east–west and south–north wind components. Hypothesis 4.21–Hypothesis 4.24 are also considered.

In such a form, the wind velocity east–west component can be expressed as:

$$W_x = p_m^x(r) + \varepsilon_x = \beta_{000}^x + \beta_{100}^x x_e + \beta_{010}^x y_e + \beta_{001}^x z_e + \beta_{200}^x x_e^2 + \beta_{110}^x x_e y_e$$
$$+ \beta_{020}^x y_e^2 + \beta_{101}^x x_e z_e + \beta_{011}^x y_e z_e + \beta_{002}^x z_e^2 + \cdots + \beta_{00m}^x z_e^m + \varepsilon_x, \quad (4.9)$$

where the parameters $\beta_{000}^x, \ldots, \beta_{00m}^x$ are referred to as regression coefficients and are to be estimated as explained above.

The wind velocity south–north component can be expressed as:

$$W_y = p_m^y(r) + \varepsilon_y = \beta_{000}^y + \beta_{100}^y x_e + \beta_{010}^y y_e + \beta_{001}^y z_e + \beta_{200}^y x_e^2 + \beta_{110}^y x_e y_e$$
$$+ \beta_{020}^y y_e^2 + \beta_{101}^y x_e z_e + \beta_{011}^y y_e z_e + \beta_{002}^y z_e^2 + \cdots + \beta_{00m}^y z_e^m + \varepsilon_y, \quad (4.10)$$

where the parameters $\beta_{000}^y, \ldots, \beta_{00m}^y$ are referred to as regression coefficients and are to be estimated as explained above.

The last step of regression analysis is to confirm the goodness of fit of the model. Commonly used tests of goodness of fit include the R–squared and the analysis of the patterns of residuals. The goodness of fit will be discussed based upon a practical example in Chapter 6. For more information on regression diagnostics, see for instance [105, Chap. 12].

4.1.5   Performance limitations

The continuous motion of the aircraft is typically constrained by performance limitations that must be taken into consideration to define the domain of both states and control variables. They are as follows:

$$0 \le h_e(t) \le \min[h_{M0}, h_u(m(t))];$$
$$C_{V_{min}} V_{stall}(h_e(t), V(t)) \le V(t) \le V_{Mo};$$
$$M(h_e(t), V(t)) \le M_{M0};$$
$$m_{min} \le m(t) \le m_{max};$$
$$0 \le C_L(t) \le C_{L_{max}};$$

$$T(t) \le T_{max}(h_e(t));$$
$$T_{min}(h_e(t)) \le T(t);$$
$$\mu(t) \le \mu_{max,civ};$$
$$\dot{V}(t) \le a_{l,max(civ)};$$
$$\dot{\gamma}(t) \le \frac{a_{n,max(civ)}}{V(t)}.$$

More information about the performance limitations model, including the definition of the different coefficients, is given in Appendix A. For more details, the reader is referred to [14].

4.2   Discrete dynamics

The flight of an aircraft has inherently a discrete nature made of multiple flight phases and multiple dynamic subsystems. Thus, together with the aircraft continuous dynamics a discrete dynamics must be considered. The discrete dynamics governs the evolution of a set of discrete variables that represent the discrete state of the system. The different discrete states are characterized by different dynamical subsystems and different sets of path constraints which govern the continuous motion of the aircraft through the different phases of the flight.

4.2.1   Flight model

The flight of an aircraft is modeled as a collection of flight phases connected by end trigger conditions which make the system "switch" between phases. Switches can be of two types: internally forced (autonomous switches) or externally forced (controlled switches) [18].

In internally forced switches, the switch is triggered by a capture condition expressed in terms of the components of the state of the system, such as waypoint's coordinates, velocity, or altitude. Externally forced switches are triggered by an external command, for instance by the pilot or by the air traffic controller. Each phase is described by a discrete state and characterized by a determined dynamical subsystem Σ^k governing the continuos motion of the aircraft and a set of path constraints.

Discrete state of the system

The discrete state of the system is characterized by a vector of discrete variables. Each one of these discrete variables takes value within a discrete set as described below:

Ph denotes the flight phase variable, $Ph \in \{0, \dots, N-1\}$;
AC denotes the aerodynamic configuration variable, $AC \in \{TO, IC, CR, AP, LD\}$;
AM denotes the atmosphere mode variable, $AM \in \{Be, Ab\}$;
DM denotes the dynamic mode variable, $DM \in \{3D, HM, VM\}$;
OP denotes the operational procedure variable, $OP \in \{PATH, CAS, CM, T_{max}, T_{min}\}$; and
Wp denotes the waypoint variable, $Wp \in \{p_j : j \in \mathcal{P}\}$.

Definition 4.6 (*Flight phase variable*). *The flight phase variable represents the index of phases constituting the flight:*

$$Ph \in \{0, \dots, N-1\}.$$

Thus, we consider N discrete states for the discrete variable Ph.

Definition 4.7 (*Aerodynamic configuration variable*). *The aerodynamic configuration variable represents the flap configuration of the aircraft, which affects its aerodynamic model. In general, five different aerodynamic configurations can be distinguished: take off (TO), initial climb (IC), cruise (CR), approach (AP), and landing (LD). Thus, $AC \in \{TO, IC, CR, AP, LD\}$.*
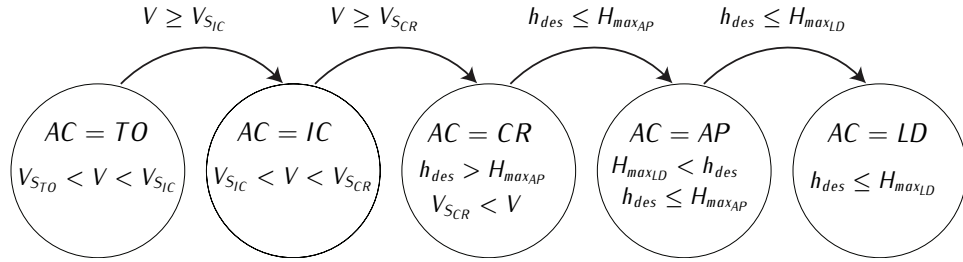

Figure 4.3: Finite state machine for AC.

Thus, we consider five different discrete states for the discrete variable AC. Switching between flap configurations is triggered by two different conditions: stall speed and threshold altitude as described in [14]. Figure 4.3 shows the finite state machine representing the typical evolution of the discrete variable AC during a flight. For instance, during take off $AC = TO$ as long as $V_{S_{TO}} < V \leq V_{S_{IC}}$, and there is a switch between $AC = TO$ and $AC = IC$ when $V \geq V_{S_{IC}}$. Also, during the approach $AC = AP$ as long as $h_{des} \leq H_{max_{AP}}$ and $H_{max_{LD}} < h_{des}$, and there is a switch between $AC = AP$ and $AC = LD$ when $h_{des} \leq H_{max_{LD}}$.

Definition 4.8 (*Dynamic Mode variable*). *The dynamic mode variable represents the ODE/DAE system governing the motion of the aircraft, assuming three-dimensional motion (3D), vertical motion (VM), or horizontal motion (HM). Thus, $DM \in \{3D, VM, HM\}$, where $DM = 3D$ corresponds to set of equations (4.1), $DM = VM$ to set of equations (4.2), and $DM = HM$ to set of equations (4.3).*

Thus, we consider three different discrete states for the discrete variable DM. Figure 4.4 shows the finite state machine representing the typical evolution of the discrete variable DM during a flight. For instance, consider an aircraft whose flight is described by a 3D motion model. In this case $DM = 3D$, and thus $\mu \neq 0$ and $\gamma \neq 0$. There is a switch between $DM = 3D$ and $DM = HM$ when $\gamma = 0$, and there is a switch between $DM = 3D$ and $DM = VM$ when $\mu = 0$.

Definition 4.9 (*Operational procedure variable*). *The operational procedure variable represents the different operational specifications in the different phases of the flight. We assume that $OP \in \{PATH, CAS, CM, T_{max}, T_{min}\}$, which correspond to a constant flight path angle procedure (PATH), a constant calibrated airspeed procedure (CAS), a constant Mach procedure (CM), a maximum thrust procedure (T_{max}), and a minimum thrust procedure (T_{min}), respectively.*
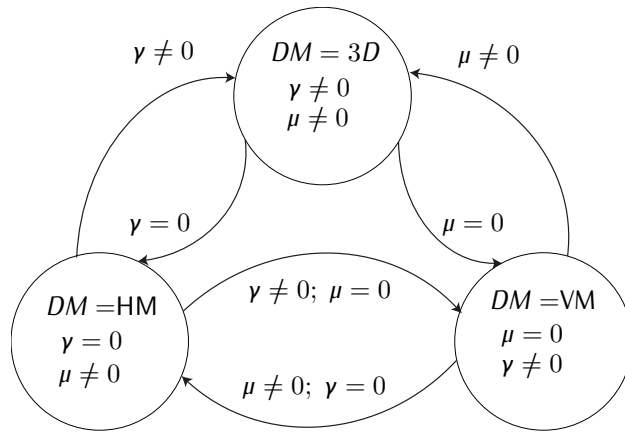
Figure 4.4: Finite state machine for DM.

Thus, we consider five different discrete states for the discrete variable OP.

Remark 4.2. Operational procedure specifications are modeled in the optimal control problem using the algebraic constraints (2.1b).

Definition 4.10 (***Atmosphere mode variable***). *The atmosphere mode variable represents whether the aircraft is below (Be) or above (Ab) the tropopause. Thus, $AM \in \{Be, Ab\}$.*

As described in Section 4.1.4, the ISA model considers a piecewise defined function $T(h)$ with two subfunctions, one applying below the tropopause and the other above the tropopause. Therefore, a discrete variable with two discrete states must be added to reflect this fact. Figure 4.5 shows the finite state machine representing the evolution of the discrete variable AM during a flight.
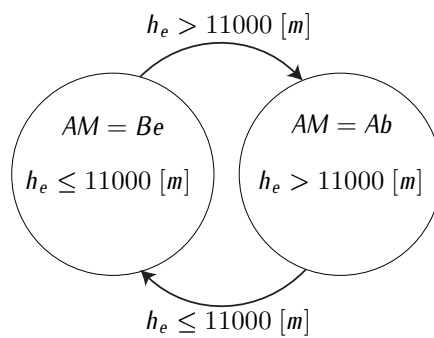
Figure 4.5: Finite state machine for AM.

Definition 4.11 (*Waypoint variable*). *The waypoint variable represents a waypoint within a set $\mathcal{P}$ of waypoints in the airspace:*

$$Wp \in \{p_j : j \in \mathcal{P}\}.$$

This set of waypoints is a subset of the total existing waypoints. It would only include the relevant waypoints for a flight between two given ariports.

4.2.2 Airspace structure

Figure 4.6 shows the waypoints and navigation aids corresponding the the AIRAC cycle published in June 2012. Airways have been omitted for the sake of clarity. A flight plan must be defined specifying a certain number of waypoints which the aircraft is going to fly. The huge number of waypoints in this figure reflects the inherent complexity of defining an efficient flight plan in a structured airspace.
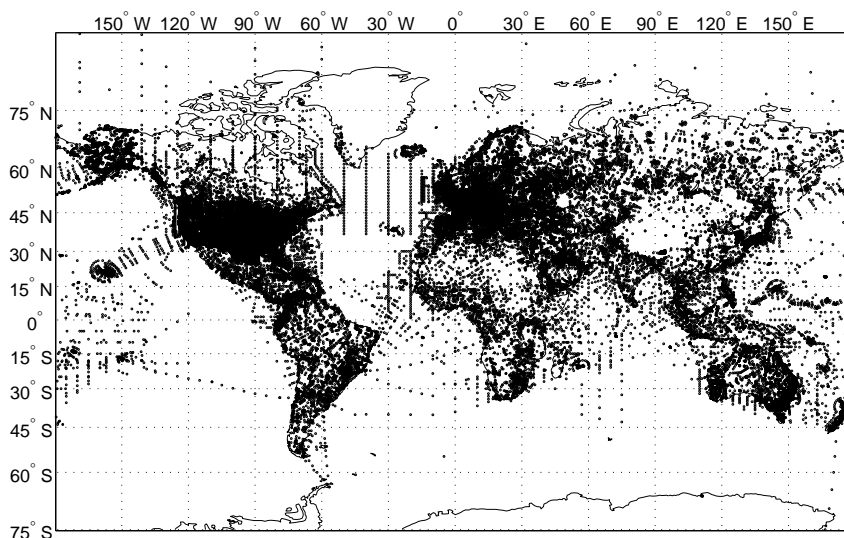
Figure 4.6: Waypoints and navigation aids AIRAC June 2012.

In the model assumed in this dissertation, airways are not considered and it is supposed that the aircraft can fly an arbitrary route among waypoints. However considering a complete directed graph structure over the set of waypoint is redundant because aircraft must fly through closer waypoints before reaching farther waypoints. A graph structure which is capable to reflect this simple observation is the multipartite graph structure.

Thus, the airspace structure is modeled as a complete multipartite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, whose vertex set $\mathcal{V}$ is partitioned into pairwise disjoint subsets which are called partite sets. In this model nodes represent waypoints and arcs represent possible transitions between them. In a complete multipartite graph vertices are adjacent if and only if they belong to different (adjacent) partite sets. The complete multipartite graph considered is composed of a sequence of $N + 1$ partite sets, $\mathcal{V}^0, \mathcal{V}^1, \ldots \mathcal{V}^N$, where $\mathcal{V}^0$ and $\mathcal{V}^N$ contain 1 node each, the initial and final waypoints p^I and p^F, respectively, and $\mathcal{V}^k, k = 1, \ldots, N - 1$, contains n_{v^k} nodes. Let $\mathcal{P} = \{p^{1,1}, \ldots, p^{N-1,n_{v^{N-1}}}\}$ be the collection of waypoints of the partite sets $\mathcal{V}^1, \ldots, \mathcal{V}^{N-1}$. Figure 4.7 illustrates it.
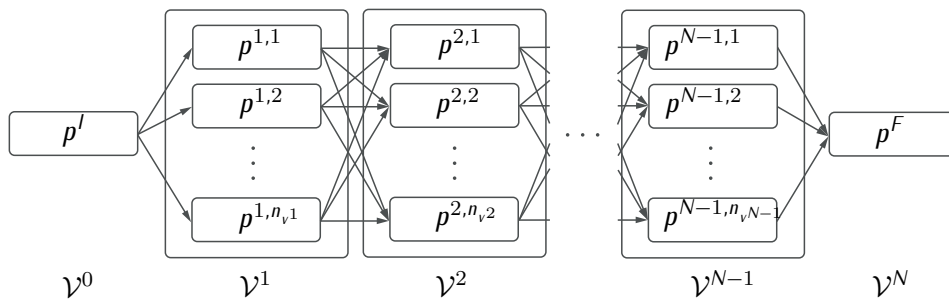
Figure 4.7: Multipartite graph structure.

4.3 Decision-making modeling

Decision-making can be regarded as the processes resulting in the selection of a course of actions or options among several alternatives. Every decision-making process produces a final choice.

Decision-making processes are modeled using the binary admissible functions $w(t)$ and the vector v introduced in Definition 3.6 and Definition 3.7.

Decision-making processes are modeled with time-dependent binary variables or control functions $w(t)$ if and only if the switching set is of the form Ψ_{free} (see Definition 3.8 and Remark 3.4.), in which the system can switch infinitely often. On the contrary, if the switching set is of the form Ψ_t, in which the system switches within a given finite set of switching times, the decision-making process can be modeled using time-independent binary variables v.

In this dissertation the switching set is of the form Ψ_t because, for any given ordered sequence of phases, the switchings can be considered fixed with respect to the scaled time variable τ. Thus, we will use binary variables v to model decision-making processes.

5

Trajectory Planning based on Multiphase Optimal Control

I N THIS CHAPTER we present an application of multiphase optimal control to commercial aircraft trajectory planning: given an aircraft performance model, initial and final conditions, a set of path constraints, and a flight plan model characterized by a collection of phases as defined in Section 4.2.1, we study the problem of finding the control inputs that steer the aircraft from the initial state to the final state and the corresponding 4D trajectory that minimize aircraft's fuel consumption. The sequence of dynamical subsystems with the corresponding sets of constraints, and the end trigger conditions that constitute the flight plan model are supposed to be known. Optimal switching times between phases, including the final time, are to be determined. The problem is formulated as a multiphase optimal control problem. The multiphase optimal control problem is converted into a NLP problem, first making the unknown switching times part of the state as in Section 3.2.2, and then applying a Hermite-Simpson collocation method as in Section 2.3.1 to convert the dynamic equations of the system into constraints. The resulting NLP problem has been solved using the interior point based nonlinear solver IPOPT. Two different cases are studied and discussed. In the first case we present a solution to the vertical trajectory planning problem in the presence of different flight procedure constraints. In the second one we study the tridimensional motion, presenting a solution to the optimal take-off weight trajectory planning problem.

5.1   Multiphase trajectory optimization problem

5.1.1   Problem statement

In this chapter the multiphase optimal control techniques described in Chapter 3 will be applied to the the following flight planning problem: given an aircraft performance model, initial and final conditions, a set of path constraints, and a flight plan model characterized by a collection of phases, we study the problem of finding the control inputs that steer the aircraft from the initial state to the final state and the corresponding 4D trajectory that minimize aircraft's fuel consumption.

Consider a flight planning problem in which the flight plan model is defined by a switching sequence

$$\sigma = [(\tilde{t}^1, e^1), \ldots, (\tilde{t}^{N-1}, e^{N-1})].$$

The sequence of events, $(e^1, \ldots, e^{N-1})$, where $e^1 = (0,1), \ldots, e^{N-1} = (N-2, N-1)$, gives rise to the following sequence of indexes for subsystems $(0, 1, \ldots, N-1)$. This sequence of indexes corresponds to a sequence of phases represented by the phase variable $Ph = k$, $k = 0, 1, \ldots, N-1$. Each phase is characterized by a constituent dynamical subsystem, Σ^k, and the corresponding set of constraints. The switch between phases takes place at switching times $\tilde{t}^1, \cdots, \tilde{t}^{N-1}$, so that:

$$t^I = \tilde{t}^0 \leq \tilde{t}^1 \leq \cdots \leq \tilde{t}^{N-1} \leq \tilde{t}^N = t^F.$$

Thus, for $t \in [\tilde{t}^k, \tilde{t}^{k+1}]$ the system is in phase k, $k = 0, \ldots, N-1$, governed by the dynamical subsystem, Σ^k. A switch between phase k and phase $k+1$ takes place at switching time $\tilde{t}^{k+1}$ due switching conditions $\mathcal{S}^k$, $k \in \{1, \ldots, N-1\}$.

In the problem to be solved the sequence of dynamical subsystems, the corresponding set of constraints, and the switching conditions are known. As a consequence, the sequence of discrete states of the system is a priori known. Therefore, the optimal control to the problem entails finding only the switching instants and the control vector $u(t)$.

Figure 5.1 shows the finite state machine representing the evolution of the discrete trajectory of the flight plan model.

Dynamic and algebraic constraints: The dynamic and algebraic constraints defined in the set of equations (3.2) govern the continuos motion of the aircraft through the air, which have been modeled in Chapter 4. In general, they depend on the aerodynamic configuration of the aircraft, the layer of the atmosphere, the given operational procedure, and the assumptions made on the flying modes. The flying mode of the aircraft can be found in any of the given forms in Definition 4.3, Definition 4.4, and Definition 4.5. The flying mode is governed by the discrete variable DM. The aerodynamic configuration is governed by the discrete variable AC, and the atmospheric layer by the discrete
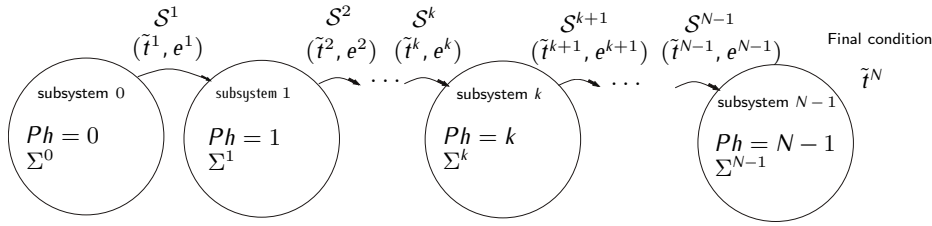
Figure 5.1: Finite state machine for the flight plan model.

variable AM. Moreover, algebraic constraints associated to operational procedures might be defined. The different procedures are governed by the discrete variable OP.

Thus, the dynamical subsystem Σ^k depends on the discrete values taken by the discrete variables AC, DM, AM, and OP, i.e., the number of differential–algebraic equations, the number of state and control variables, and the value of aerodynamic and atmospheric parameters will differ according to the set of discrete values taken by these discrete variables.

Initial and final conditions: The initial condition will be given, in general, by the coordinates at the origin airport, a heading direction corresponding to the take–off runway head, take–off velocity, take–off flight path angle, and take–off mass. This corresponds to a set of values for the vector of state variables at the initial time, $x^I \in \mathbb{R}^{n_x}$. Notice that one or more of these initial conditions might be not specified, such as the initial mass in the optimal take–off weight trajectory planning problem.

In the same manner, the final conditions correspond, in general, to a set of values for the vector of state variables at the final time, $x^F \in \mathbb{R}^{n_x}$. In particular, in order to define a flight plan one needs the coordinates at the destination airport and the heading direction of the landing head of the runway. In the case of the optimal take–off weight trajectory planning problem, the final weight of the aircraft must be specified.

Path constraints: The path constraints of the multiphase optimal control problem given in the set of equations (3.4) refers to physical limits that an aircraft should not exceed. These limitations are described in Section A.2 and are given as inequality constraints in the set of equations (4.11). Notice that some of the limits given in the set of equations (4.11), namely, maximum coefficient of lift, stall speed, and threshold altitude, depend on the aerodynamic configuration, and thus the set of path constraints is a function of the discrete variable AC.

Interior point constraints: Let $\mathcal{S} = \{\mathcal{S}^1, \ldots, \mathcal{S}^{N-1}\}$ be the set of end trigger conditions for switching. In strategic flight planning, end trigger conditions give rise to autonomous or controlled switchings.

Thus, the interior point equality constraints given in the set of equations (3.5a) can be expressed as follows:

$$\mathcal{S}^{k+1} : \; 0 = \vartheta_{eq}^k(x(\tilde{t}^{k+1})), \; k = 0, \dots, N-2. \tag{5.1}$$

Inequality interior point constraints in set of equations (3.5b) might be applied to constraint the switching instants. However, we will not use them in the problem.

Objective function: The objective functional in equation (3.6) becomes:

$$J = \sum_{k=0}^{N-1} \int_{\tilde{t}^k}^{\tilde{t}^{k+1}} \dot{m}^k(t) \, dt, \tag{5.2}$$

where $\dot{m}^k(t)$ is the fuel flow of the aircraft during phase k. This cost functional represents the fuel burnt by the aircraft in flight, which is directly related to the emitted CO_2.

5.1.2   NLP **solution approach**

According to the method described in Section 3.2.2 to incorporate the switching times into the multiphase optimal control problem, the interval $[0, 1]$ must be subdivided into N subintervals $[\tau^k, \tau^{k+1}], k = 0, \dots, N-1$, by means of any monotonically increasing sequence of $N+1$ values. For the sake of simplicity these values can be equally spaced, i.e., we can take $\tau^k = k/N, k = 0, \dots, N$, where $\tau^0 = 0$ and $\tau^N = 1$.

Furthermore, according to the numerical collocation method explained in Section 2.3.1, the subinterval $[\tau^k, \tau^{k+1}], k = 0, \dots, N-1$ is subdivided into N^k subintervals. In this case subintervals having the same amplitude $h^k = (\tau^{k+1}-\tau^k)/N^k$ are considered. They are denoted $[\tau_i^k, \tau_{i+1}^k], i = 0, \dots, N^k - 1$. This discretization gives rise to a subdivision of the interval $[0, 1]$ into $N_d = \sum_{k=0}^{N-1} N^k$ subintervals. The relation between τ_i^k and τ^k, is $\tau_i^k = \tau^k + i \cdot (\tau^{k+1} - \tau^k)/N^k = \tau^k + i \cdot h^k$, $k = 0, \dots, N - 1, i = 0, \dots, N^k$.

Equations (3.7) relate the values of t and τ, so that $\tilde{t}^k$ corresponds to τ^k, $k = 0, \dots, N$ and $\tilde{t}_i^k$ corresponds to τ_i^k, $k = 0, \dots, N-1, i = 0, \dots, N^k$. Moreover, the switching times $\tilde{t}^k$ are related to the values of time $\tilde{t}_i^k$ that correspond to τ_i^k by the relation $\tilde{t}_i^k = \tilde{t}^k + i \cdot (\tilde{t}^{k+1} - \tilde{t}^k)/N^k$, $k = 0, \dots, N-1, i = 0, \dots, N^k$.

Figure 5.2–Figure 5.5 illustrate the discretization of the problem in the two time scales.
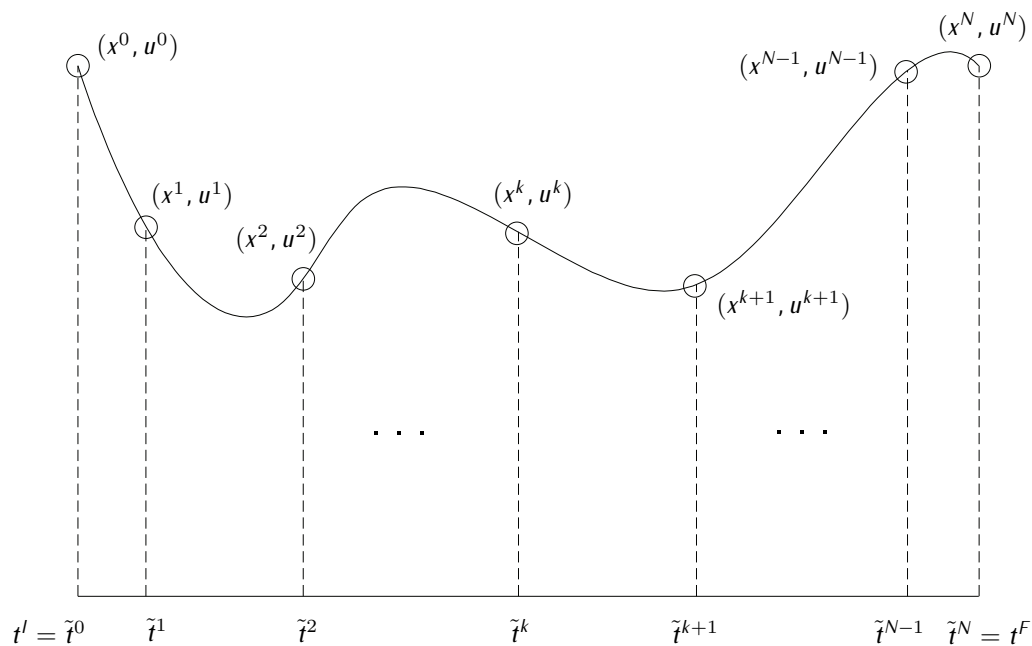
Figure 5.2: Multiphase discretization for time variable t.
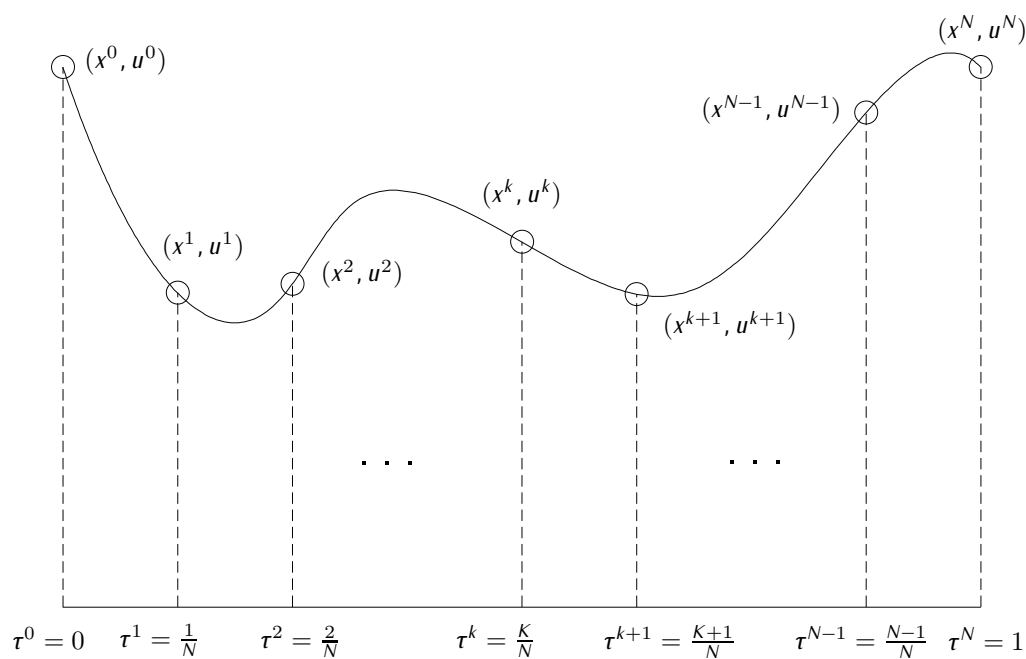
Figure 5.3: Multiphase discretization for time variable τ.
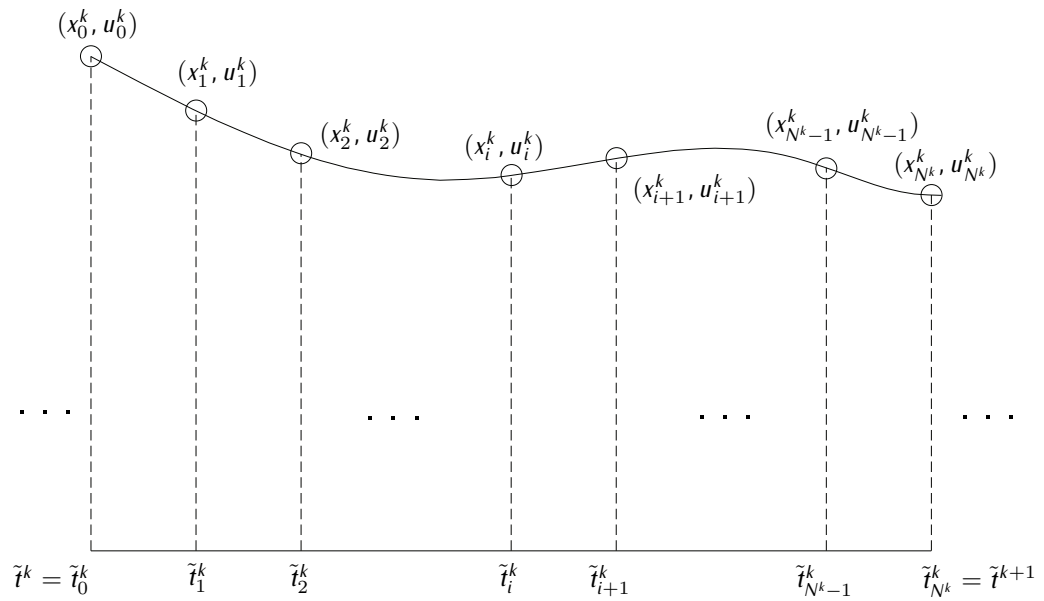
Figure 5.4: Phase discretization for time variable t.

Figure 5.5: Phase discretization for time variable τ.

The Hermite–Simpson collocation scheme described in Section 2.3.1 has been used. In the subinterval $[\tilde{t}_i^k, \tilde{t}_{i+1}^k]$, $k = 0, \ldots, N-1$, $i = 0, \ldots, N^k - 1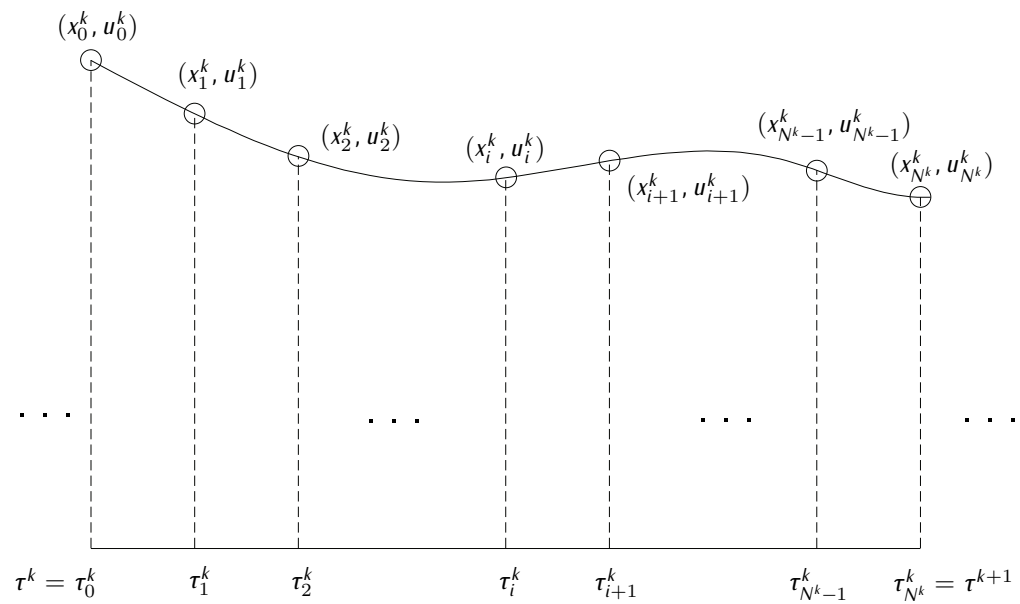$, the collocation point $x_{i,C}^k$ is calculated from the state variables x_i^k and x_{i+1}^k. Figure 2.3 illustrates how the state

variables of the problem are discretized in the subinterval $[\tilde{t}^k_i, \tilde{t}^k_{i+1}]$. A linear control scheme has been used for control variables, that is, $u^k_{i,C} = \frac{1}{2}(u^k_i + u^k_{i+1})$. The discretized control variables represent discrete values for the controls at each discrete time at which the system equation are evaluated, $u^k_i = u(t^k_i)$, $u^k_{i,C} = u(t^k_{i,C})$, $u^k_{i+1} = u(t^k_{i+1})$. For the sake of clarity, let us define $X^k_i = (x^k_i, x^k_{i,C}, x^k_{i+1})$, $U^k_i = (u^k_i, u^k_{i,C}, u^k_{i+1})$, $k = 0, \dots, N-1$, $i = 0, \dots, N^k - 1$.

Thus, the Nlp problem takes the form:

$$\min \; N \cdot \sum_{k=0}^{N-1} \left\{ (\tilde{t}^{k+1} - \tilde{t}^k) \cdot \sum_{i=0}^{N^k-1} \mathcal{Q}(X^k_i, U^k_i) \right\};$$

subject to:

$$\mathcal{A}_S(X^k_i) + N \cdot (\tilde{t}^{k+1} - \tilde{t}^k) \cdot h^k \cdot \mathcal{C}_S(X^k_i, U^k_i) = 0, \quad k = 0, \dots, N-1, i = 0, \dots, N^k - 1; \quad \text{(NLP)}$$

$$0 \leq \phi^k(X^k_i, U^k_i), \quad\quad\quad\quad\quad\quad\quad\quad\quad k = 0, \dots, N-1, i = 0, \dots, N^k - 1;$$

$$0 = \vartheta^k(x^k_0), \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad k = 1, \dots, N-1;$$

$$x^0_0 = x^I, x^{N-1}_{N^k} = x^F;$$

where

$$\mathcal{Q}(X^k_i, U^k_i) = \frac{h^k}{6}(f(x^k_i, u^k_i) + 4f(x^k_{i,C}, u^k_{i,C}) + f(x^k_{i+1}, u^k_{i+1})),$$

and both $\mathcal{A}_S(X^k_i)$ and $\mathcal{C}_S(X^k_i, U^k_i)$ come from the defect equation (2.15) with $k = 0, \dots, N-1$, $i = 0, \dots, N^k - 1$. The unknowns of this problem are

$$(x^k_i, x^k_{i+1}, u^k_i, u^k_{i+1})$$

for $k = 0, \dots, N-1$, $i = 0, \dots, N^k - 1$, together with the switching times $\tilde{t}^k$ for $k = 1, \dots, N-1$. For the sake of clarity, unknowns $\tilde{t}^k$ for $k = 1, \dots, N-1$ have not been renamed as elements of the extended state vector as in Section 3.2.2. Note that $\tilde{t}^N = t^F$ is also a variable of the problem if the final time is unknown.

The multiphase optimal control problem has now been recast as a Nlp problem: minimizing a nonlinear function subject to a number of nonlinear constraints in a space where the variables take values in $\mathbb{R}$. We now turn to the explanation of the solution method to solve problem (NLP).

There exist several off the shelf Nlp solvers to solve such problem. Among the most used ones are for instance Conopt, Knitro, Loqo, Minos, Snopt, and Ipopt. For more information, the reader is referred to the following website: Neos-server solvers.[1]

For the problem considered in this chapter, the Nlp solver Ipopt (Interior Point Optimizer) is the most suitable one because it handles properly large-scale, sparse, non-convex problems, with a large number of equality and inequality constraints. Ipopt is an open source software package for large scale nonlinear optimization problems. It can be used to solve general nonlinear programming problems. Ipopt implements an interior point line search filter method. The mathematical details of Ipopt algorithm can be found in [42]. Source and binary files can be found at Coin-or.[2]

---

[1]http://www.neos-server.org/neos/solvers/index.html
[2]http://www.coin-or.org

5.2   Optimized–procedured vertical trajectory planning

In this section we present the results of a case study in which vertical trajectories are analyzed and discussed. Based on the flight plan model in Section 4.2.1, we define what we term as optimized–procedured flights. In order to evaluate them, short and medium range vertical optimized–procedured flights are compared with free and fully–procedured flights. Aircraft performances, flight procedures, and the resulting consumptions are analyzed and discussed.

5.2.1   Case study

Recall equation (4.2) for the the 3–DOF vertical motion of an aircraft over a flat Earth and neglecting wind effects. The airplane is a conventional jet airplane and BADA 3.6 [14] is used as aircraft performance model (see Appendix A).

Thus, the 3–DOF equations governing the translational vertical motion of an airplane are the following:

$$m(t)\dot{V}(t) = T(t) - D(h(t), V(t), C_L(t)) - m(t)g\sin\gamma(t);$$
$$m(t)V(t)\dot{\gamma}(t) = L(h(t), V(t), C_L(t)) - m(t)g\cos\gamma(t);$$
$$\dot{x}_e(t) = V(t)\cos\gamma(t); \tag{5.3}$$
$$\dot{h}_e(t) = V(t)\sin\gamma(t);$$
$$\dot{m}(t) = -T(t)\eta(V(t));$$

where $x(t) = (V(t), x_e(t), h_e(t), \gamma(t), m(t))$ are the state variables and $u(t) = (T(t), C_L(t))$ are the control inputs.

We consider an Airbus A–320 BADA aircraft model. The different A–320 aerodynamic configurations, governed by the integer variable AC, and the value of the aerodynamic parameters are listed in Table 5.1.

AC	Flap	$C_{L_{max}}$	CD_0	CD_i
TO	1+F	2.43	0.0393	0.0396
IC	1	2.19	0.0242	0.0469
CR	Clean	1.50	0.024	0.0375
AP	2	2.76	0.0456	0.0381
LD	Full	3.09	0.0838	0.0371

Table 5.1: A–320 flap configurations.

The path constraints of the problem are those related with performance limitations (see Section A.2) and are also given in the BADA database manual. For this example, we have used:

$$0 \leq h_e(t) \leq \min[h_{M0}, h_u(m(t)), H_{max}^k(AC)];$$
$$C_{V_{min}} V_S^k(m(t), h_e(t), AC) \leq V(t) \leq V_{Mo};$$
$$M(V(t), h_e(t)) \leq M_{M0}; \tag{5.4}$$
$$m_{min} \leq m(t) \leq m_{max};$$
$$0 \leq C_L(t) \leq C_{L_{max}}^k(AC);$$
$$T_{min}(h_e(t)) \leq T(t) \leq T_{max}(h_e(t));$$

where $h_u = h_{max} + G_t(\Delta T_{ISA} - C_{Tc,4}) + G_W(m_{max} - m)$, and $C_{V_{min}} = 1.3$ (except for TO, where $C_{V_{min}} = 1.2$). The threshold, H_{max}^k, the stall velocity, V_S^k, and the maximum coefficient of lift, $C_{L_{max}}^k$, depend on the flap configuration and, therefore, are governed by the discrete variable AC at each phase, $k = 0, \ldots, N - 1$.

Current ILS set the constant descent path between –2.5° and –3.5°, generally –3°. Within the landing phase, we constraint the flight path angle according to more flexible values of an aircraft's final descent path, i.e.,

$$-6° \leq \gamma_{Landing} \leq -2°.$$

We continue the description of the case study defining the different flights to be compared over short and medium range:

Fully–procedured flights: They reflect current ATM paradigm and have been defined accurately according to typical vertical profiles that are flown nowadays. They can be found in BADA. Modeling flights in such a way enforces the specification of two operational procedures per phase, for instance, to climb with constant calibrated air speed (V_{CAS}) and constant throttle setting, or to perform a steady cruise, i.e., with defined constant Mach and constant altitude.

Optimized–procedured flights: They are based on a relaxation of current procedures by setting, in general, just one procedure per phase, and relaxing some capture conditions for switching, including thus controlled switches. They are defined seeking more flexible and efficient flight planning concepts within a constrained based paradigm.

Free flights: Such profiles are considered as optimal benchmark and represent a comparison baseline. They are useful for a better understanding of optimal performances and a quantitative analysis of current inefficiencies.

Short range

Table 5.2, Table 5.3, and Table 5.4 show, respectively, the fully–procedured, optimized–procedured, and free flight plan models herein used for the numerical experiments. The short range fully–procedured flight has been derived from a real Madrid–Oviedo flight plan following BADA–like flight procedures with airline–defined speed and altitude profile values.

The boundary conditions of the flight are the following: $x_e^I = 0$, $h_e^I = 0$, $V^I = 1.2V_{S_{TO}}$, $\gamma^I = 0.05$ [rad], $m^I = 63070$ [kg]; $x_e^F = 476$ [km], $h_e^F = 0$.

Ph	Name	AC	$\mathcal{S}$	OP	OC*
0	Take Off	TO	$V = 1.3V_S^{IC}$	T_{max}	$V_{CAS} < 250[kt]$
1	Initial Climb	IC	$V = 1.3V_S^{CR}$	T_{max}	$V_{CAS} < 250[kt]$
2	Res. Free Climb	CR	$h_e = 10000[ft]$	T_{max}	$V_{CAS} < 250[kt]$
3	Climb Accel.	CR	$V_{CAS} = 300[kt]$	$h_e = 10000[ft]$, T_{max}	–
4	Climb CAS	CR	CM$= 0.78$	$CAS = 300[kt]$, T_{max}	–
5	Climb Mach	CR	$h_e = FL320$	CM$= 0.78$, T_{max}	–
6	Cruise	CR	–	$h_e = FL320$,CM$= 0.78$	–
7	Descent Mach	CR	$V_{CAS} = 300[kt]$	CM$= 0.78$, T_{min}	–
8	Descent CAS	CR	$h_e = 10000[ft]$	$CAS = 300[kt]$, T_{min}	–
9	Descent Decel.	CR	$V_{CAS} = 250[kt]$	$h_e = 10000[ft]$, T_{min}	–
10	Res. Free Des.	CR	$h_e = 6000[ft]$	T_{min}	$V_{CAS} < 250[kt]$
11	Approach	AP	$h_e = 2000[ft]$	$PATH = -3°$	$V_{CAS} < 250[kt]$
12	Landing	LD	Final cond.	$PATH = -3°$	$V_{CAS} < 250[kt]$

* OC is an acronym that stands for Operational Constraints due to operations near airports.

Table 5.2: Short range: fully–procedured flight plan model.

Ph	Name	AC	$\mathcal{S}$	OP	OC
0	Take Off	TO	$V = 1.3V_S^{IC}$	–	$V_{CAS} < 250[kt]$
1	Initial Climb	IC	$V = 1.3V_S^{CR}$	–	$V_{CAS} < 250[kt]$
2	Res. Free Climb	CR	$h_e = 10000[ft]$	–	$V_{CAS} < 250[kt]$
3	Climb Accel.	CR	–	$h_e = 10000[ft]$	–
4	Climb CAS	CR	–	CAS, T_{max}	–
5	Climb Mach	CR	–	CM, T_{max}	–
6	Cruise	CR	–	HM	–
7	Descent Mach	CR	–	CM	–
8	Descent CAS	CR	$h_e = 10000[ft]$	CAS	–
9	Descent Decel.	CR	–	$h_e = 10000[ft]$	–
10	Res. Free Des.	CR	$h_e = 6000[ft]$	T_{min}	$V_{CAS} < 250[kt]$
11	Approach	AP	$h_e = 2000[ft]$	$PATH$	$V_{CAS} < 250[kt]$
12	Landing	LD	Final cond.	$PATH$	$V_{CAS} < 250[kt]$

Table 5.3: Short range: optimized–procedured flight plan model.

Ph	Name	AC	S	OP	OC
0	Take Off	TO	$V = 1.3V_S^{IC}$	–	$V_{CAS} < 250[kt]$
1	Initial Climb	IC	$V = 1.3V_S^{CR}$	–	$V_{CAS} < 250[kt]$
2	Res. Free Climb	CR	$h_e = 10000[ft]$	–	$V_{CAS} < 250[kt]$
3	Free CL/CR/DS	CR	$h_e = 10000[ft]$	–	
4	Res. Free Descent	CR	$h_e = 6000[ft]$	–	$V_{CAS} < 250[kt]$
5	Approach	AP	$h_e = 2000[ft]$	–	$V_{CAS} < 250[kt]$
6	Landing	LD	Final cond.	–	$V_{CAS} < 250[kt]$

Table 5.4: Short range: free flight plan model.

It can be observed in Table 5.3 that the trigger conditions of the optimized–procedured flight plan model are of two types, either autonomous or controlled switches. If the switch is autonomous, it takes place when the aircraft reaches the capture condition, e.g., when it reaches the prescribed altitude or velocity. On the contrary, switches between phases without capture conditions are controlled switches, and thus, they are given by the control law within the optimal solution. Then, the main differences between fully–procedured and optimized–procedured flight plan models are: less restricted procedures, and the fact that controlled switches are considered. As a consequence, the transition Mach, the cruising altitude, or the constant calibrated speed of descent are not pre–fixed, but are set by the optimal solution, leading the system to an overall minimum fuel consumption.

Medium range

Table 5.5, Table 5.6 and Table 5.7 show, respectively, the fully–procedured, optimized–procedured and free flight flight plan models herein used for the numerical experiments. The medium range fully–procedured flight has been derived from a real Madrid–Berlin flight plan following BADA–like flight procedure with airline–defined speed and altitude profile values [14].

The boundary conditions of the flight are the following: $x_e^I = 0$, $h_e^I = 0$, $V^I = 1.2V_{S_{TO}}$, $\gamma^I = 0.05$ [rad], $m^I = 69415$ [kg]; $x_e^F = 2035$ [km], $h_e^F = 0$.

Ph	Name	AC	S	OP	OC
0	Take Off	TO	$V = 1.3V_S^{IC}$	T_{max}	$V_{CAS} < 250[kt]$
1	Initial Climb	IC	$V = 1.3V_S^{CR}$	T_{max}	$V_{CAS} < 250[kt]$
2	Res. Free Climb	CR	$h_e = 10000[ft]$	T_{max}	$V_{CAS} < 250[kt]$
3	Climb Accel.	CR	$V_{CAS} = 300[kt]$	$h_e = 10000[ft]$, T_{max}	–
4	Climb CAS	CR	CM= 0.78	$CAS = 300[kt]$, T_{max}	–
5	Climb Mach	CR	$h_e = FL360$	CM= 0.78, T_{max}	–
6	Cruise	CR	–	$h_e = FL360$,CM= 0.78	–
7	Descent Mach	CR	$V_{CAS} = 300[kt]$	CM= 0.78, T_{min}	–
8	Descent CAS	CR	$h_e = 10000[ft]$	$CAS = 300[kt]$, T_{min}	–
9	Descent Decel.	CR	$V_{CAS} = 250[kt]$	$h_e = 10000[ft]$, T_{min}	–
10	Res. Free Des.	CR	$h_e = 6000[ft]$	T_{min}	$V_{CAS} < 250[kt]$
11	Approach	AP	$h_e = 2000[ft]$	$PATH = -3°$	$V_{CAS} < 250[kt]$
12	Landing	LD	Final cond.	$PATH = -3°$	$V_{CAS} < 250[kt]$

Table 5.5: Medium range: fully–procedured flight plan model.

Ph	Name	AC	S	OP	OC
0	Take Off	TO	$V = 1.3V_S^{IC}$	–	$V_{CAS} < 250kt$
1	Initial Climb	IC	$V = 1.3V_S^{CR}$	–	$V_{CAS} < 250kt$
2	Res. Free Climb	CR	$h_e = 10000[ft]$	–	$V_{CAS} < 250kt$
3	Climb Accel-	CR	–	$h_e = 10000[ft]$	–
4	Climb CAS	CR	–	CAS, T_{max}	–
5	Climb Mach	CR	–	CM, T_{max}	–
6	Cruise	CR	–	HM	–
7	Descent Mach	CR	–	CM	–
8	Descent CAS	CR	$h_e = 10000[ft]$	CAS	–
9	Descent Decel.	CR	–	$h_e = 10000[ft]$	–
10	Res. Free Descent	CR	$h_e = 6000[ft]$	T_{min}	$V_{CAS} < 250kt$
11	Approach	AP	$h_e = 2000[ft]$	$PATH$	$V_{CAS} < 250kt$
12	Landing	LD	Final cond.	$PATH$	$V_{CAS} < 250kt$

Table 5.6: Medium range: optimized–procedured flight plan model.

Ph	Name	AC	S	OP	OC
0	Take Off	TO	$V = 1.3V_S^{IC}$	–	$V_{CAS} < 250[kt]$
1	Initial Climb	IC	$V = 1.3V_S^{CR}$	–	$V_{CAS} < 250[kt]$
2	Res. Free Climb	CR	$h_e = 10000[ft]$	–	$V_{CAS} < 250[kt]$
3	Free CL/CR/DS	CR	$h_e = 10000[ft]$	–	
4	Res. Free Descent	CR	$h_e = 6000[ft]$	–	$V_{CAS} < 250[kt]$
5	Approach	AP	$h_e = 2000[ft]$	–	$V_{CAS} < 250[kt]$
6	Landing	LD	Final cond.	–	$V_{CAS} < 250[kt]$

Table 5.7: Medium range: free flight plan model.

5.2.2 Numerical results

The computation of the fully–procedured flights have been carried out using a tool combining 3-DOF flight dynamics differential equations with procedure-oriented 3D flight control. More precisely, the aircraft's ODE set (5.3) with the performances based on BADA and ISA atmosphere models is integrated using the set of path constraints (5.4), while controls are properly set so that the aircraft follows the given flight procedures.

Optimized-procedured and free flights are defined, formulated, and solved according to Section 5.1. The resulting sparse NLP problem (NLP) has been solved using IPOPT [42]. IPOPT showed robustness in solving infeasible subproblems in the iterative process even when dealing with infeasible initial guesses. IPOPT showed also robustness when dealing with different initial conditions, showing similar result patterns. Both optimized-procedured and free flights have a discretization grid with $N_d = 650$ ($N^0 = N^1 = \ldots N^{12} = 50$) sample points. In the free flight, for the sake of comparison, the 4th phase is composed by 350 samples. Fully-procedured flights are computed in real times with samples every second. To illustrate computational issues, focusing for instance on the medium range optimized-procedured flight, the resulting NLP problem (NLP) had 4299 variables, 3597 equality constraints, and 4151 inequality constraints. The total computational time on a Mac OS X 2.56 GHz laptop with 4 GB RAM was 1369.86 [s].

Table 5.8 and Table 5.9 show respectively the short and medium range results: they contain the total flight times and times of switching ($\tilde{t}^{k*}[s]$), the consumptions including the accumulate consumption at the end of every phase ($C^k[kg]$), and the constant values that describe the optimized-procedured aircraft performance in the different flight procedures (Value opt).

Ph	$t[s]^1$	$\tilde{t}^{k*}[s]^2$	$\tilde{t}^{k*}[s]^3$	$C[kg]^1$	$C^k[kg]^2$	$C^k[kg]^3$	$Value\ opt^2$
0	–	8.69	8.69	–	18.16	18.16	Free
1	–	25.04	25.06	–	52.94	52.96	Free
2	–	225.3	225.38	–	441.44	441.57	Free
3	–	249.41	–	–	482.71	–	$h_e = 10000$[ft]
4	–	755.409	–	–	1161.36	–	$V_{CAS} = 150.459$[m/s] & T_{max}
5	–	1193.11	–	–	1550.29	–	CM$= 0.7297, T_{max}$
6	–	1243.9	–	–	1554.87	–	$h_e = 10875.4$[m]
7	–	1243.9	–	–	1554.87	–	CM$= 0.6422$
8	–	2416.84	–	–	1681.77	–	$V_{CAS} = 108.406$[m/s]
9	–	2416.84	2368.83	–	1681.77	1661.55	$h_e = 10000$[ft]
10	–	2677.17	2642.83	–	1716.68	1698.26	T_{min}
11	–	2826.22	2853.54	–	1737.66	1728.08	$\gamma = -4.3911°$
12	2592.8	2922.08	2942.26	1967.15	1752.71	1741.09	$\gamma = -4.3911°$

Table 5.8: Vertical motion: short range results. Column for variable Ph corresponds to the optimized-procedured flight. Columns with superindex [1] correspond to the fully-procedured flight, those with superindex [2] to the optimized-procedured flight, and those with superindex [3] to free flight.

Ph	t[s]	$\tilde{t}^{k*}$[s]	$\tilde{t}^{k*}$[s]	C[kg]	C^k[kg]	C^k[kg]	$Value\ opt$
0	–	11.61	11.61	–	24.26	24.26	Free
1	–	33.28	33.29	–	70.25	70.27	Free
2	–	254.66	254.71	–	498.50	498.58	Free
3	–	287.89	–	–	555.43	–	$h_e = 10000$ ft
4	–	796.57	–	–	1258.31	–	$V_{CAS} = 154.817\text{m/s}\ \&\ T_{max}$
5	–	1653.76	–	–	1995.43	–	$M = 0.7249\ \&\ T_{max}$
6	–	7774.07	–	–	6034.37	–	$h_e = 11295.4$ m
7	–	7774.08	–	–	6034.38	–	$M = 0.6772$
8	–	8964.12	–	–	6161.77	–	$V_{CAS} = 111.138$ m/s
9	–	8964.12	8929.65	–	6161.77	6068.77	$h_e = 10000$ ft
10	–	9222.08	9184.45	–	6196.36	6102.95	T_{min}
11	–	9369.81	9388.6	–	6217.16	6131.87	$\gamma = -4.3823°$
12	9403.5	9464.6	9475.87	6529.85	6231.05	6144.66	$\gamma = -4.3823°$

Table 5.9: Vertical motion: medium range results. The interpretation of the columns is the same as in Table 5.8.

Notice that the phases herein named as *Descent Mach* and *Descent Decel.* are eliminated by the optimal solution by setting the corresponding durations to zero. In this manner, the algorithm is capable of eliminating inefficient phases.

Controls, states, and optimal switching instants are represented for short and medium range in Figure 5.6–Figure 5.9. Regarding the state variables, in general, except for the case of γ (Figure 5.7(a) and Figure 5.7(b)), all state variables vary smoothly. γ exhibit high–frequency dynamics at some points near the switchings. This suggests discontinuity around those points. Notice that this behavior is normal since the purpose here is not to capture short duration maneuvers, for which we would need a finer discretization around the switchings. The optimized–procedured flight controls show some high–frequency behavior near the switchings. This is due to not capturing sharp maneuvers near the switchings. Moreover, we can also observe a bang–bang behavior in the phase named *Descent Mach* for both short and medium range. See Figure 5.7(e), Figure 5.7(g), Figure 5.9(e), and Figure 5.9(g). The reason behind this behavior is that the algorithm has neglected this phase resulting in some numerical instability which, however, is not transferred to real time flight controls since the duration of the phase is set to zero. See Figure 5.7(f), Figure 5.7(h), Figure 5.9(f), and Figure 5.9(h). Free flight controls show, on the contrary, a rather smooth behavior.

Results show that the proposed optimized–procedured profiles save 10.9% and 4.6% of fuel for short and medium range, respectively, i.e, 214.44 [kg] and 298.8 [kg], when compared to fully–procedured flights. Furthermore, results also show that free flights achieve 11.5% and 5.9% of fuel savings for short and medium range, respectively, i.e., 226,06 [kg] and 385.19 [kg], when compared to fully–procedured flights. This means that the efficiency of the optimized–procedured flights is very close to the considered optimal benchmark.
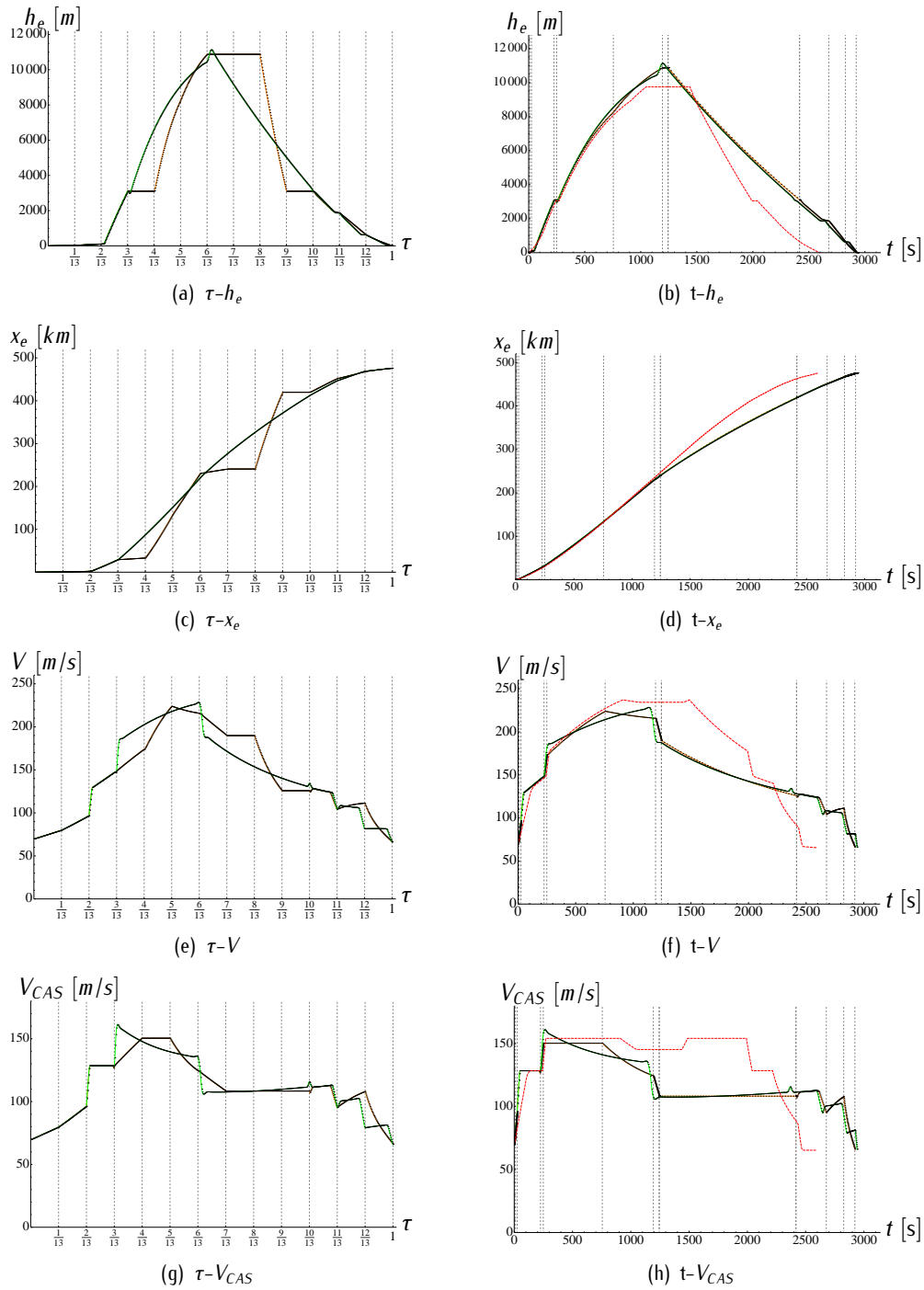
Figure 5.6: Vertical motion: short range altitude, h_e, distance, x_e, True Air Speed, V, and Calibrated AirSpeed, V_{CAS}: dashed red line corresponds to fully-procedured profile; solid dotted-orange line corresponds to optimized-procedure flight (being the dots the computed sample points); solid dotted-green corresponds to free flight (being the dots the computed sample points). Note that the depicted vertical dashed lines correspond to the optimized-procedured flight switching times.
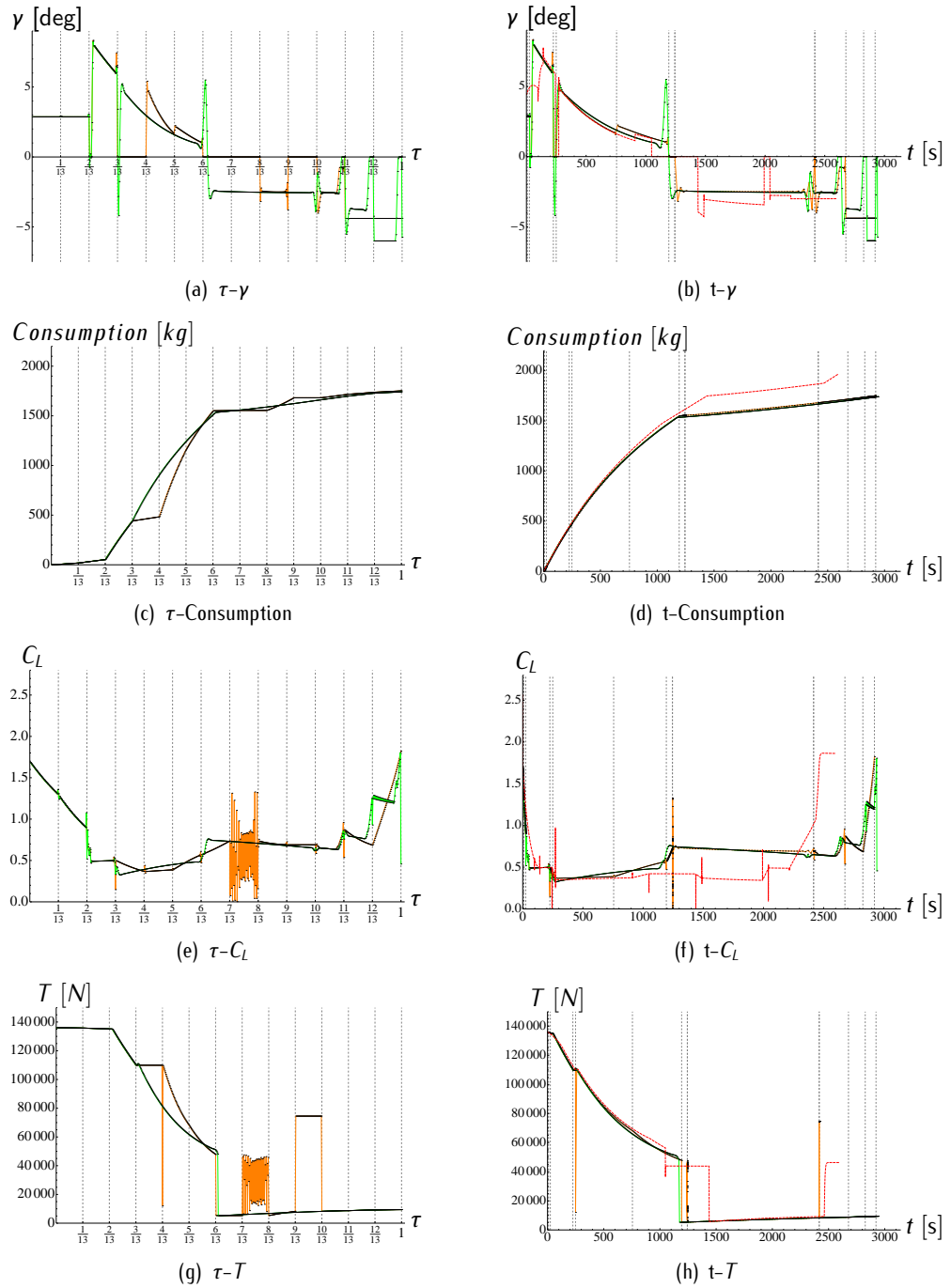
(a) τ–γ

(b) t–γ

(c) τ–Consumption

(d) t–Consumption

(e) τ–C_L

(f) t–C_L

(g) τ–T

(h) t–T

Figure 5.7: Vertical motion: short range consumption, flight path angle, γ, and control laws, T & C_L: dashed red line corresponds to fully-procedured flight; solid dotted-orange line corresponds to optimized-procedure flight (being the dots the computed sample points); solid dotted-green corresponds to free flight (being the dots the computed sample points). Note that the depicted vertical dashed lines correspond to the optimized-procedured flight switching times.
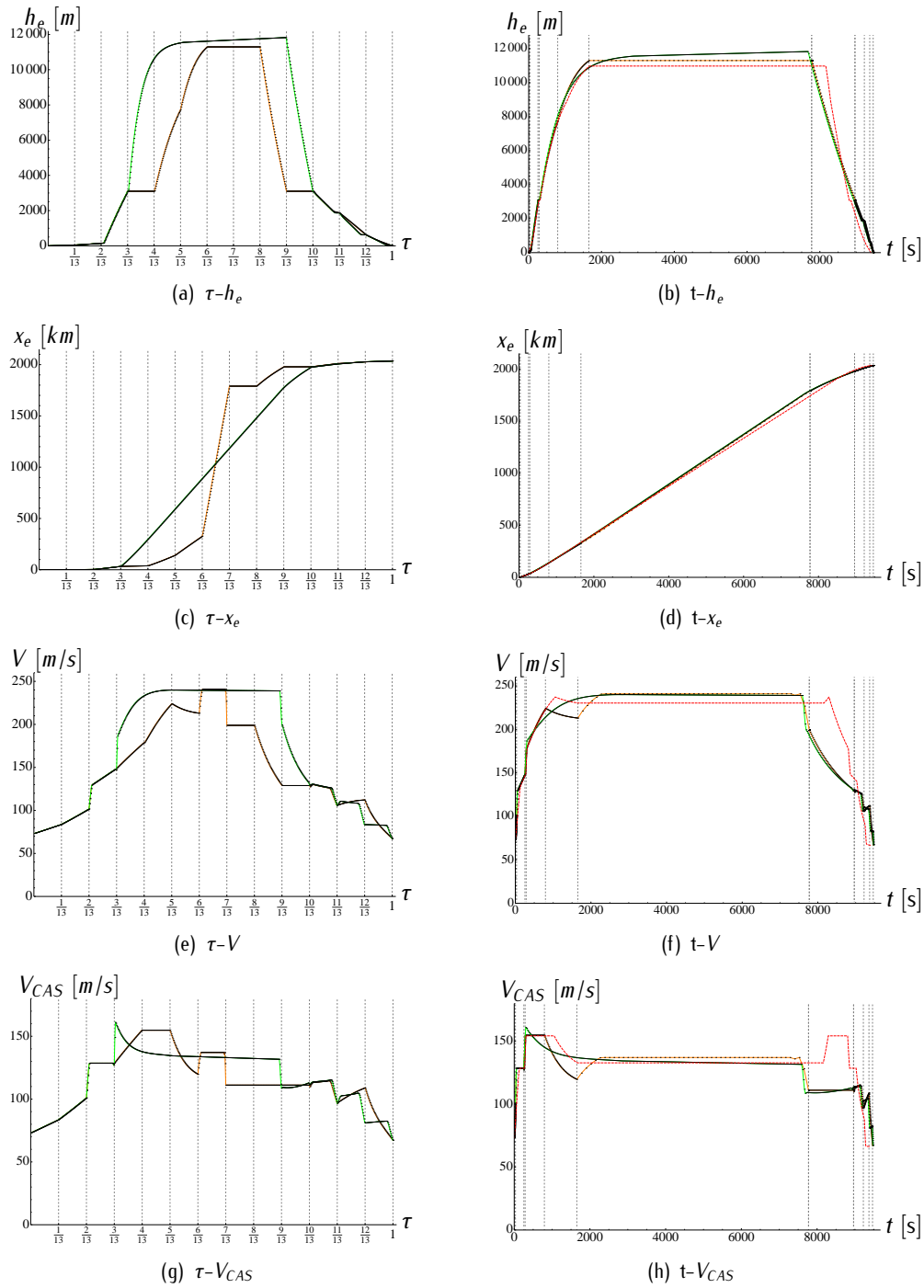
(a) τ–h_e

(b) t–h_e

(c) τ–x_e

(d) t–x_e

(e) τ–V

(f) t–V

(g) τ–V_{CAS}

(h) t–V_{CAS}

Figure 5.8: Vertical motion: medium range altitude, h, distance, x, True Air Speed, V, and Calibrated AirSpeed, V_{CAS}: dashed red line corresponds to fully-procedured flight; solid dotted-orange line corresponds to optimized-procedure flight (being the dots the computed sample points); solid dotted-green corresponds to free flight (being the dots the computed sample points). Note that the depicted vertical dashed lines correspond to the optimized-procedured flight switching times.
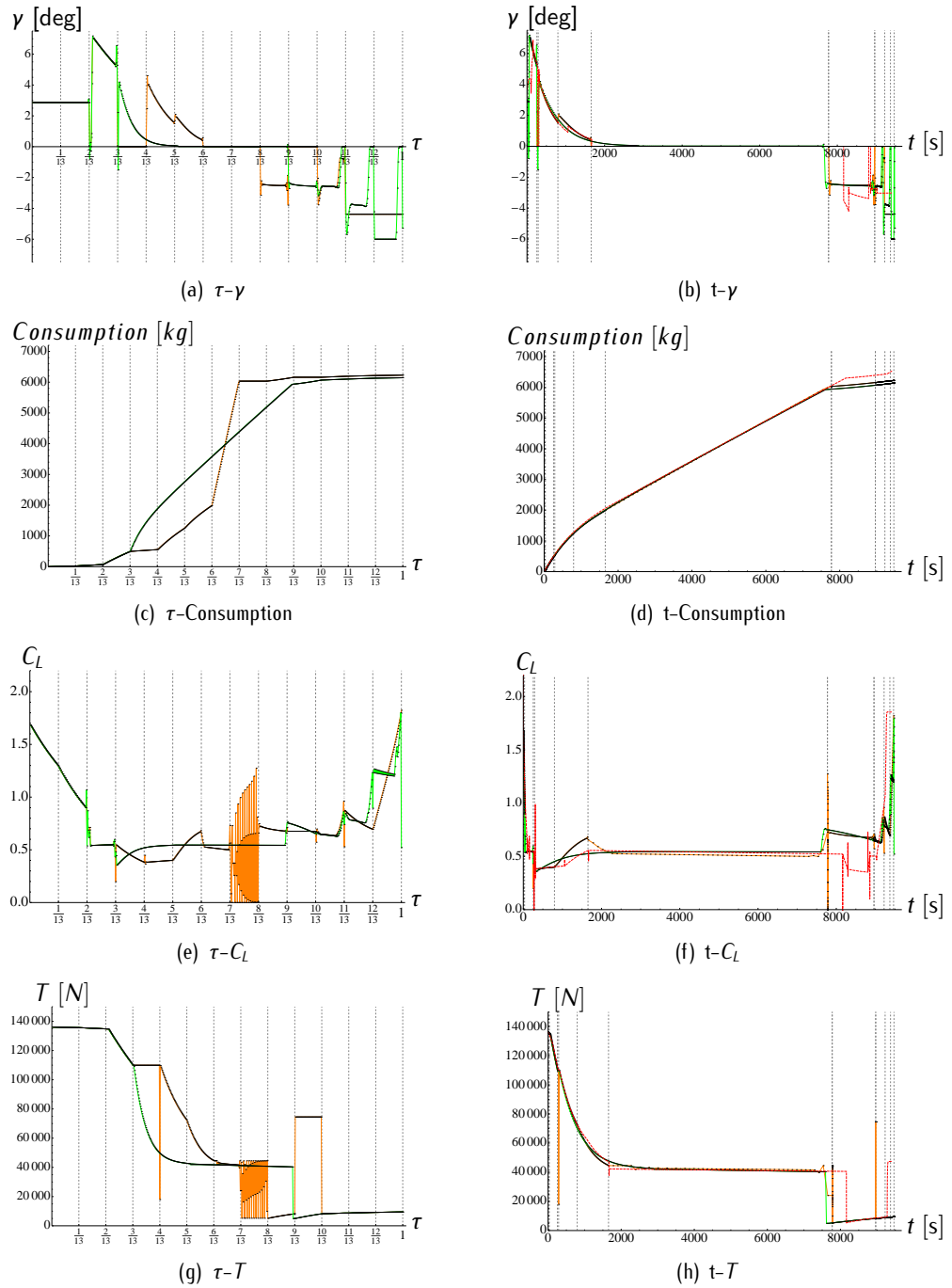
(a) τ–γ

(b) t–γ

(c) τ–Consumption

(d) t–Consumption

(e) τ–C_L

(f) t–C_L

(g) τ–T

(h) t–T

Figure 5.9: Vertical motion: medium range consumption, flight path angle, γ, and control laws, T & C_L: dashed red line corresponds to fully-procedured flight; solid dotted-orange line corresponds to optimized-procedure flight (being the dots the computed sample points); solid dotted-green corresponds to free flight (being the dots the computed sample points). Note that the depicted vertical dashed lines correspond to the optimized-procedured flight switching times.

5.2.3 Discussion on the results

Short range

We first analyze free flight performances. In the optimum solution of the free flight the aircraft seeks to achieve maximum altitude in minimum time, since flying at low altitudes with maximum thrust setting is very fuel consuming. As the aircraft gets higher, it progressively softens its rate of climb until it suddenly performs a sharp climb maneuver to intercept the optimum descent path, thus skipping the cruise phase. With this maneuver, the aircraft consumes the excessive speed with respect to the optimum descent speed, while enabling an anticipated interception of the optimum descent path thanks to the fast altitude gain. Otherwise, the climb would last longer, resulting in greater fuel consumption. The optimum descent path is the result of descending at maximum gradient speed, which allows the aircraft to fly the greatest distance possible at idle thrust, thus minimizing fuel consumption. This optimum speed is the minimum drag speed, also known as base speed. This speed decreases as air density increases, so the aircraft lowers its speed as it descends. The base speed also changes as the aircraft deploys high-lift devices, so the aircraft speed is adjusted by regulating the flight path angle when a change in aerodynamic configuration is to be performed. At the very end of the flight, when the only aim is to land at any flyable speed, the flight path angle is increased to cover the greatest distance possible before touching ground, consuming the excess of speed above minimum speed.

There are two main differences when performing the descent between fully-procedured and free flights. The first one is that fully-procedured flights descend at a speed much higher than the base speed (300kt Vs. 210kt). This permits fully-procedured flights to reduce flight duration at the cost of increasing fuel consumption. The second difference is that fully-procedured flights perform the approach at the ILS $-3°$glide path. When the approach glide path is less inclined than the maximum gradient path (as in this case), the glide path becomes too low for the aircraft to maintain the desired speed at idle thrust. Thus, some extra thrust is required during the approach, resulting in increased fuel consumption.

In the solution of the optimized-procedured flight the aircraft tries always to follow the patterns of free flight optimal performance fulfilling some prefixed procedures. Consequently, it seems that the algorithm aims to minimize climb phase duration since it is very fuel consuming, but at the same time it needs to prevent negative effects on the optimality of the subsequent flight phases. The obtained 292 [kt]/0.73 climb speed is similar to the fully-procedured one in the constant Calibrated AirSpeed (CAS) phase but shows a significantly lower speed for the constant Mach phase because it provides a higher rate of climb. Then, the aircraft performs a short cruise (4 [s]). This is the duration needed to decelerate from the climbing velocity to the descending one. The aircraft intercepts the descending path, performing it at 108 [m/s] (CAS), 211 [kt] approx., considerably different than the fully-procedured 300 [kt]. This is optimal thanks to the minimum drag CAS being almost constant along the descent, consequently neglecting the constant Mach phase. Finally, the constant flight path angle of initial and final approach, $-4.39°$, is more inclined than the fully-procedured one, $-3°$.

Medium range

As in the previous case, medium range optimum free flight seeks to achieve maximum altitude in minimum time. As the aircraft gets higher, it progressively reduces its rate of climb to make a smooth transition to the subsequent pseudo-cruise phase, in which the aircraft asymptotically approaches its operating ceiling. The optimization target during cruise is to maximize the specific range, which is the distance traveled per unit of fuel consumed. As the aircraft mass decreases due to fuel burn, the optimum profile shows an increasing trend in altitude following the also increasing operating ceiling, while speed is conveniently adjusted, typically in a slightly decreasing trend. Such a performance is known as continuous cruise climb. Cruise phase ends when the optimum descent path is intercepted. For the optimum descent path, the principles of free optimal performance explained for short range flights remain valid.

Again, as pointed out for the short-range flights, the optimized-procedured flight tries always to follow the patterns of free flight optimal performance. The obtained 301[kt]/0.72 climb shows a significantly lower speed for the constant Mach phase. Then the aircraft performs the cruise at an altitude that is limited by the operating ceiling at the beginning of cruise (where the ceiling is lower due to the greater mass). Eventually, the aircraft intercepts the descending path, performing it at 111 [m/s] (CAS), 215 [kt] aprox. Finally, the constant flight path angle of initial and final approach, –4.38°, is more inclined than the typical one, –3°. Notice that a closer performance to free flight's continuous cruise climb could have been achieved by defining at least one step climb.

General remarks

Whereas some differences exist in ascent and cruise, the key differences of performance between current and future concept of operations arise in descent phases, where indeed descent velocity and ILS arrival flight path angle exhibit high deviations from what has been shown as optimal benchmark: to descent at base velocity and to perform approach and landing at maximum gradient path.

Focusing on descent, we should consider separately the track going from the top of descent to the initial approach fix, and the track going from that fix to the runway. The first track could be improved without operational problems by just following the profile given by maximum gradient velocity at idle thrust. In lack of that, a constant CAS procedure around the average base velocity could be defined, which as it has been said is not far from the optimal benchmark. The main reason that current flights use a constant CAS up to 300 [kt] is to reduce descent duration. Regarding final and initial approach, it does not seem to be easily achievable to perform the obtained results. Free flight's optimal path shows a very steep path, while performing landing with quasi-level flight. This profile is unsafe because descent path are designed as a trade off between obstacle avoidance handling and not excessive descent rates. Free flight shows rather high descent rate followed by a potentially non-handling obstacles horizontal path. Optimized-procedured flights' descent paths showed, however, higher than nowadays ILS constant path angles, closer to optimal benchmark path, avoiding also potential

obstacles. Such paths would lead to higher descent rates, but lower than free flights' descent paths. The paradigm of Global Navigation Satellite System (GNSS) descent procedures will help defining ad hoc descent paths within safety standards and thus, some of the above obtained fuel savings could be achieved.

It is necessary to point out that BADA aerodynamic model does not take into account compressibility effects on the aerodynamic behavior of the aircraft. This leads to lower than real drag at high mach numbers, resulting in higher than real optimum speeds and altitudes.

5.3 3D optimal take–off weight trajectory planning

We extend the work exposed in the previous section, presenting a solution to the optimal take–off weight trajectory planning problem. We model the three dimensional, spherical motion of the aircraft, introducing new airspace constraints, such as the case of a SID procedure. Aircraft performances, flight procedures, and the resulting fuel consumption are analyzed. Moreover, we discuss the effects of the excess of take–off weight on the optimal 4D trajectory.

5.3.1 Case study

Recall Definition 4.3 for the the 3–DOF, 3D motion of an aircraft over a spherical Earth. We neglect wind effects. The airplane is a conventional jet airplane and BADA 3.6 [14] is used as aircraft performance model.

Thus, the 3–DOF equations governing the translational 3D motion of an airplane are the following:

$$\dot{V}(t)m(t) = T(t) - D(h_e(t), V(t), C_L(t)) - m(t)g\sin\gamma(t);$$
$$\dot{\lambda}_e(t)R_e\cos\theta_e(t) = V(t)\cos\gamma(t)\cos\chi(t);$$
$$\dot{\chi}(t)m(t)V(t)\cos\gamma(t) = L(h_e(t), V(t), C_L(t))\sin\mu(t);$$
$$\dot{\theta}_e(t)R_e = V(t)\cos\gamma(t)\sin\chi(t);$$
$$\dot{\gamma}(t)m(t)V(t) = L(h_e(t), V(t), C_L(t))\cos\mu(t) - m(t)g\cos\gamma(t);$$
$$\dot{h}_e(t) = V(t)\sin\gamma(t);$$
$$\dot{m}(t) = -T(t)\eta(V(t));$$

where $x(t) = (V(t), \lambda(t), \theta(t), h_e(t), \gamma(t), \chi(t), m(t))$ are the state variables and $u(t) = (T(t), C_L(t), \mu(t))$ are the control inputs. The cartesian coordinates x_e, y_e have been transformed into spherical coordinates according to relations (4.4). With abuse of notation, the discrete variable $DM \in \{3D, VM, HM\}$ refers now to spherical dynamics.

We consider an Airbus A–320 BADA aircraft model. The different A–320 aerodynamic configurations, governed by the integer variable AC, and the value of the aerodynamic parameters are listed in Table 5.1.

The path constraints of the problem are those related with performance limitations and can also be found in the BADA database manual. For this experiment, we have used:

$$0 \leq h_e(t) \leq \min[h_{M0}, h_u(m(t)), H_{max^k}(AC)];$$
$$C_{V_{min}} V_S^k(m(t), h_e(t), AC) \leq V(t) \leq V_{Mo};$$
$$M(V(t), h_e(t)) \leq M_{M0};$$
$$m_{min} \leq m(t) \leq m_{max};$$
$$0 \leq C_(t)L \leq C_{L_{max}}^k(AC);$$
$$T_{min}(h_e(t)) \leq T(t) \leq T_{max}(h_e(t));$$
$$\mu(t) \leq \mu_{max,civ};$$
$$\dot{V}(t) \leq a_{l,max(civ)};$$
$$\dot{\gamma}(t) \leq \frac{a_{n,max(civ)}}{V}.$$

The reader is referred to Section A.2 for more details. The threshold, H_{max}^k, the stall velocity, V_S^k, and the maximum coefficient of lift, $C_{L_{max}}^k$, depend on the flap configuration and, therefore, are governed by the discrete variable AC at each phase, $k = 0, \dots, N - 1$.

Within the landing phase, we constraint the flight path angle according to

$$-6^\circ \leq \gamma_{Landing} \leq -2^\circ.$$

The flight plan model used for the numerical experiment is given in Table 5.10. Notice that the aim of such flight plan is at performing an optimized–procedured flight.

Ph	Name	DM	OP	AC	AM	$\mathcal{S}$	OC
0	Take Off	VM	T_{max}	TO	Be	$V = 1.3V_S^{IC}$	$V_{CAS} < 250[kt]$
1	Ini. Climb	VM	T_{max}	IC	Be	$V = 1.3V_S^{CR}$	$V_{CAS} < 250[kt]$
2	MD034	3D	T_{max}	CR	Be	$MD034$	$V_{CAS} < 250[kt]$
3	MD035	3D	–	CR	Be	MD035	$V_{CAS} < 250[kt]$
4	Res. Free Climb	3D	–	CR	Be	$h_e = 10000[ft]$	$V_{CAS} < 250[kt]$
5	RBO	3D	–	CR	Be	RBO	–
6	PINAR	HM	–	CR	Be	PINAR	$h_e \geq 13.000[ft]$
7	Ver. Climb	VM	–	CR	Be	$h_e = 11000[m]$	–
8	Cruise 1	HM	–	CR	Ab	–	–
9	Free Step	3D	–	CR	Ab	–	–
10	Cruise 2	HM	–	CR	Ab	–	–
11	Free Des.	3D	–	CR	Ab	$h_e = 11000[m]$	–
12	Ver. Descent	VM	–	CR	Be	$h_e = 10000[ft]$	–
13	Descent Decel.	HM	–	CR	Be	–	$V_{CAS} < 250[kt]$
14	Res. Free Descent	3D	T_{min}	CR	Be	$h_e = 6000[ft]$	$V_{CAS} < 250[kt]$
15	Approach	VM	$PATH$	AP	Be	$h_e = 2000[ft]$	$V_{CAS} < 250[kt]$
16	Landing	VM	$PATH$	LD	Be	Final cond.	$V_{CAS} < 250[kt]$

Table 5.10: 3D motion: flight plan model.

The boundary conditions of the state variables have been selected according to a realistic flight from Madrid Barajas (LEMD) to Berlin Schoenefeld (EDDB). The departure runway head in Madrid Barajas corresponds to 15L, denoting the initial heading. The arrival runway head in Berlin Schoenefeld corresponds to 25L, denoting the final heading. Notice that, while the departure is modeled to be flown following a SID procedure, the arrival is modeled to be flown as a continuous descent approach from top of descent towards the runway head. Initial velocity and initial flight path angle have been selected according to standard values. In regard of the boundary conditions referring to the mass of the aircraft, notice that the optimal take–off weight trajectory planning problem entails the calculation of the optimal initial mass. Therefore, the initial mass of the aircraft is not specified. On the contrary, the final mass of the aircraft must be given. This can be done assuming we know the minimum required landing weight of the aircraft, which can be expressed as:

$$LW = OEW + PL + RF,$$

where LW stands for Landing Weight, OEW stands for Operating Empty Weight, PL stands for PayLoad, and RF stands for Reserve Fuel. Notice that this means that the aircraft consumes the total amount of Trip Fuel (TP) during the flight.

The boundary conditions are therefore: $\lambda_e^I = -3.56°$, $\theta_e^I = 40.47°$, $h_e^I = 620$ [m], $V^I = 77.40$ [m/s], $\gamma^I = 4.5°$, $\chi^I = -54°$; $\lambda_e^F = 13.52°$, $\theta_e^F = 52.38°$, $h_e^F = 48.67$ [m], $\chi^F = 21°$, $m^F = 58000$ [kg].

The selected SID procedure is Madrid Barajas PINAR1AU shown in Table 5.3.1.
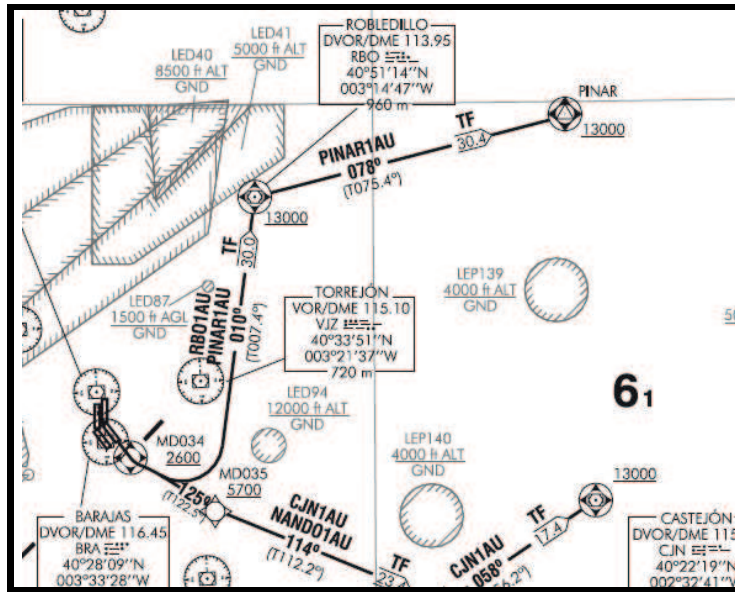
Figure 5.10: SID PINAR.

The definition of the SID can be found in the AIP Spain, published by AENA, where the SID is textually described as:

> To MD034 on heading 144°(Magnetic North) at 2600 [ft] or above, turn left. To MD035 at 5700 [ft] or above, turn left. To RBO at 13000 [ft] or above, turn right. To PINAR at 13000 [ft] or above.

The coordinates of the different waypoints and navaids are:

- MD034 (RNAV waypoint): 40°26'37.3164"N; 003°30'21.2360"W.

- MD035 (RNAV waypoint): 40°21'30.9920"N; 003°19'52.5450"W.

- RBO (VOR/DME): 40°51'14"N; 003°14'47"W.

- PINAR (RNAV waypoint): 40°58'49.0620"N; 002°35'56.9980"W.

5.3.2   Numerical results

The problem is defined, formulated and solved as exposed in Section 5.1. The resulting sparse NLP problem (NLP) has been solved using IPOPT , which showed robustness in solving infeasible subproblems in the iterative process even when dealing with infeasible initial guesses. IPOPT showed also robustness when dealing with different initial conditions, showing similar result patterns. The discretization grid with $N_d = 250$ sample points is as follows: $N^8 = N^{10} = N^{12} = N^{15} = 30$; rest $N^k = 10$. To illustrate computational issues, the resulting NLP problem (NLP) had 2437 variables,

1897 equality constraints, and 1465 inequality constraints. The total computational time on a Mac OS X 2.56 GHz laptop with 4 GB RAM was 1543.93 [s].

Table 5.11 shows the optimal switching times ($\tilde{t}^{k*}$), the accumulated consumption (C^k), and the constant values that describe aircraft performance in the different flight procedures (Value opt.). The optimal initial mass is $m^I = 63307.2$ [kg]. The optimal cruise altitude is $h_e = 12005.9$ [m]. The final descent flight path angle is $\gamma = -4.5$ °. Notice that, as it was already illustrated in the previous example, the algorithm is capable of elimination those phases that are considered inefficient. The phases herein named *Cruise 1* and *Descent Decel.* are eliminated by the algorithm setting the duration to zero.

Phase	$\tilde{t}^{k*}$ [s]	C^k [kg]	Value opt.
1	7.12	13.13	$m^I = 63307.2$ [kg], $\chi = \chi^I$
2	34.16	68.16	$\chi = \chi^I$
3	62.87	126.1	T_{max}
4	196.1	357.0	Free
5	222.8	396.1	Free
6	532.2	673.5	Free
7	873.0	973.8	$h_e = 4030$ [m]
8	1762.36	1903.5	$\chi = 50.2°$
9	1762.36	1903.5	$h_e = 11000$ [m]
10	1803.1	1931.2	Free
11	7380.8	5095.4	$h_e = 12005.9$ [m]
12	7555.7	5110.4	Free
13	8761.04	5240.0	$\chi = 41.4°$
14	8761.04	5240.0	$h_e = 10000$ [ft]
15	9013.95	5273.9.0	T_{min}
16	9157.28	5294.1	$\gamma = -4.5°$
17	9247.31	5307.2	$\gamma = -4.5°$

Table 5.11: Optimal take–off weight trajectory results.

Flight optimal path, optimal control law, and the evolution of the state variables over time are also obtained as part of the solution. We present the optimal path in Figure 5.12, and the SID procedure in Figure 5.11 using Google Earth. Notice that Figure 5.11 represents a zoom–in vision of the compete optimal path given in Figure 5.12.

State variables and control inputs are shown in Figure 5.13–Figure 5.15. Regarding state variables, in general, all of them vary relatively smoothly. γ and V might exhibit high–frequency dynamics at some points near the switchings. As illustrated in the previous example, this suggests discontinuity around those points. Notice that this behavior is normal since the purpose here is not to capture short duration maneuvers, for which we would need again a finer discretization around the switchings. Figure 5.15 shows the behavior of the control inputs. μ evolves rather smoothly, within reasonable

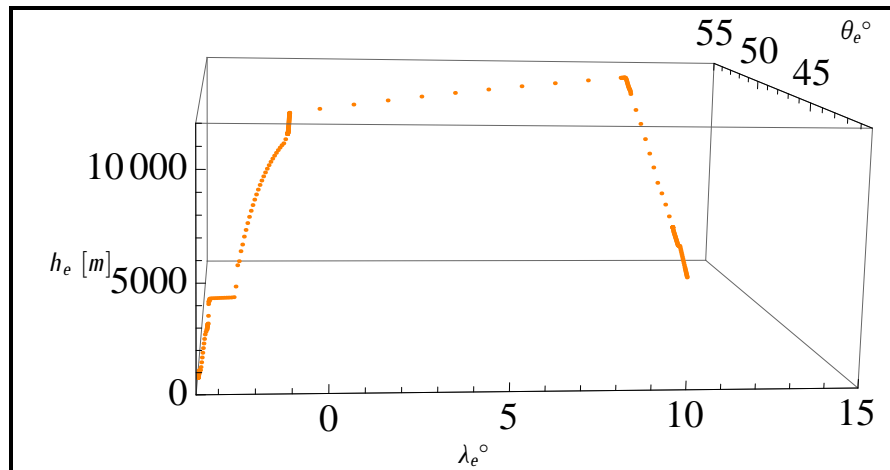Figure 5.11: Flight optimal path: SID.

Figure 5.12: Flight optimal path.

values. C_L shows also reasonable values, even though its evolution can be claimed as *saw-type* most of the time domain and it hits the minimum constraint ($C_L = 0$) at the beginning of the cruise phase. Finally, thrust shows a *bang-bang* behavior during the sixth and seventh phases, both phases defined within the SID procedure. The drastic constrained environment that with find within a SID procedure, together with the fact that the aircraft is allowed to move freely in 3D might have probably caused such instability.
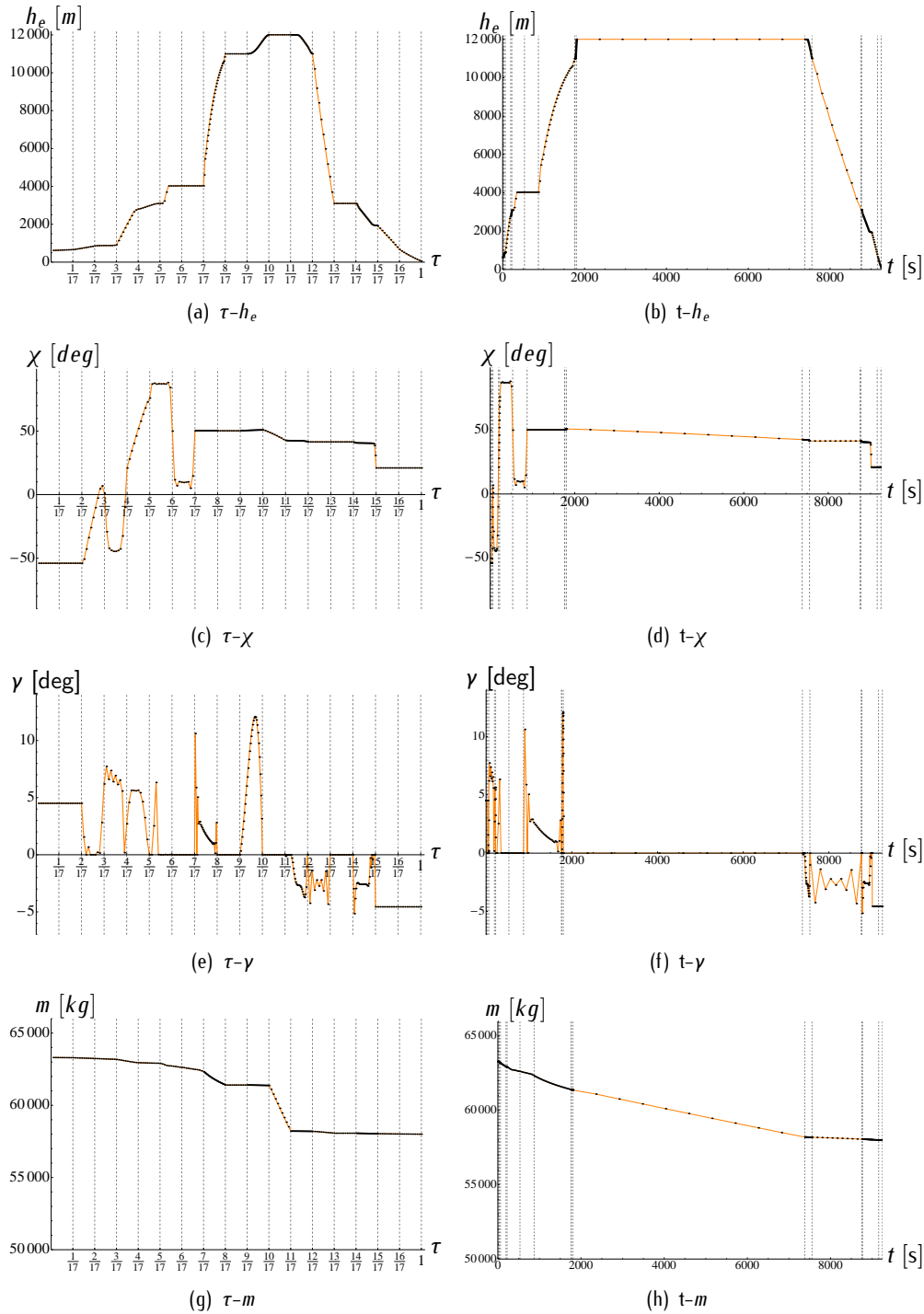
Figure 5.13: State variables with respect to scaled, τ, and unscaled time, t: altitude, heading angle, flight path angle, and consumption: the dots correspond to the computed sample points and the depicted vertical dashed lines correspond to the switching times.
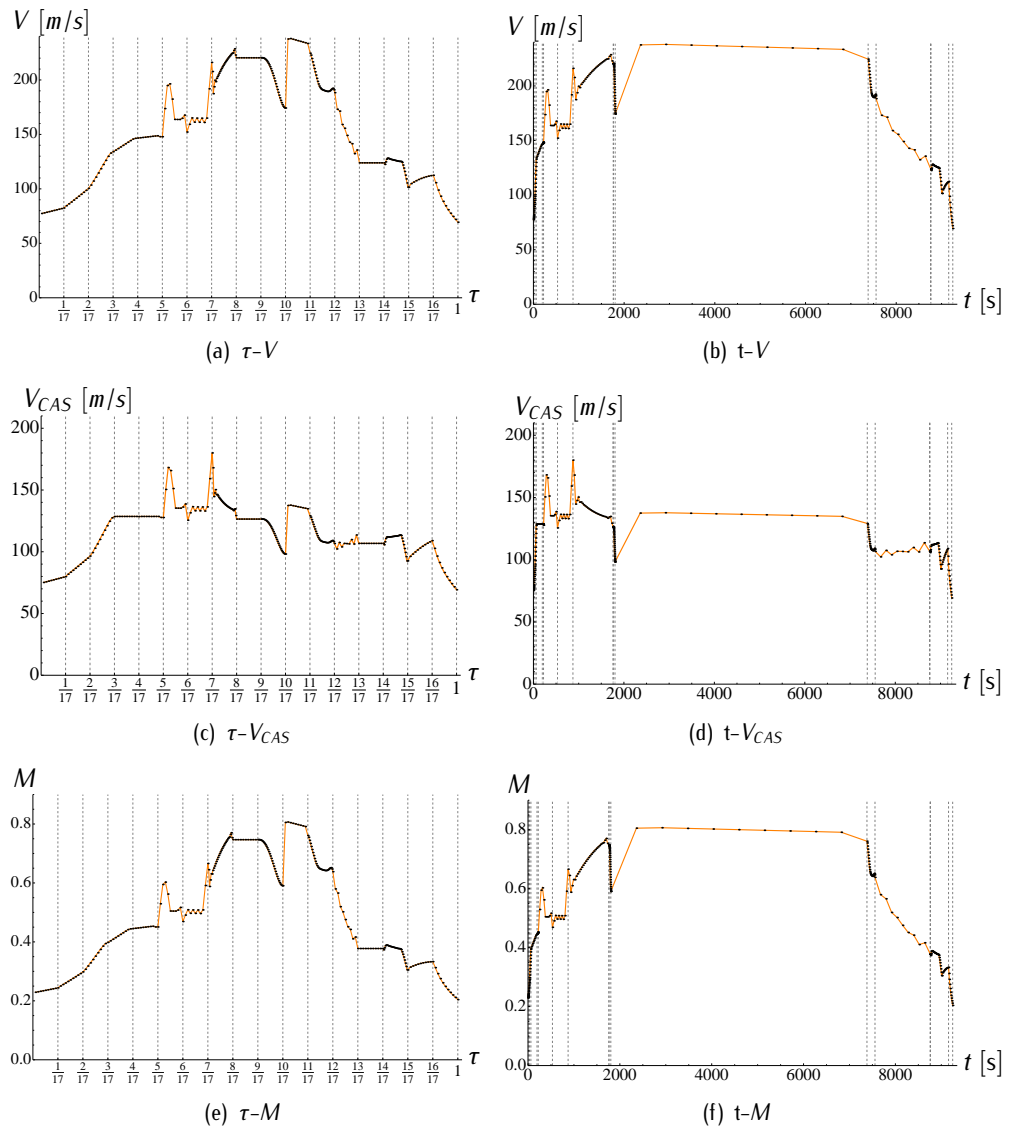
Figure 5.14: True airspeed, calibrated airspeed and Mach number with respect to scaled, τ, and unscaled time, t: the dots correspond to the computed sample points and the depicted vertical dashed lines correspond to the switching times.

Figure 5.15: Control inputs with respect to scaled, τ, and unscaled time, t: the dots correspond to the computed sample points and the depicted vertical dashed lines correspond to the switching times.
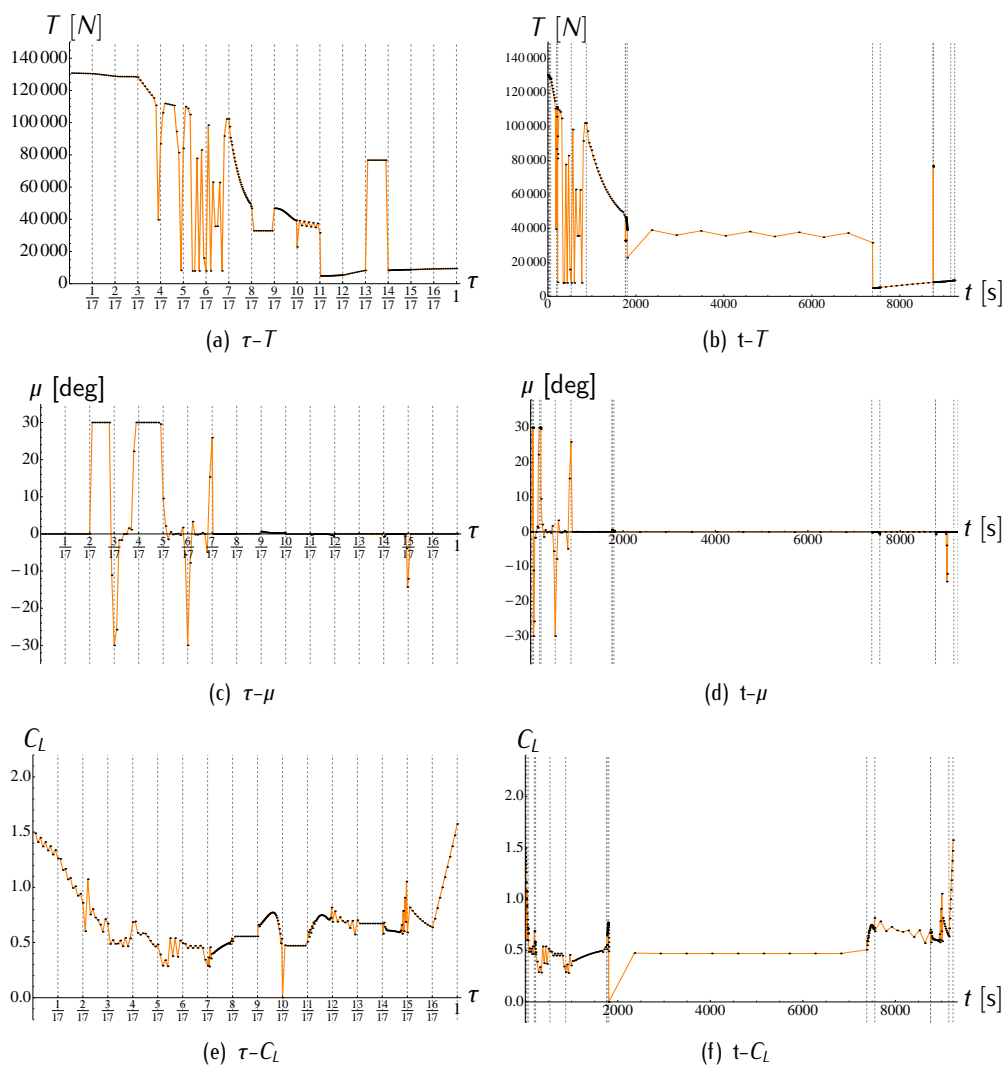
5.3.3   Discussion on the results

We start the discussion analyzing aircraft performances.

In the first part of the flight, the aircraft takes off from the assigned runway head and proceeds on fulfilling the constraints imposed by the SID procedure. After having reached PINAR it starts the ascent phase up to 11000 [m], when the fist part of the cruise starts. The ascent phase is to be flown into a vertical plane, i.e., with constant course, and with no specified procedures in terms of velocity. The velocity profile results in a slightly decreasing CAS profile. After having reached an altitude of 11000 [m], the aircraft is supposed to start a cruise phase at constant altitude and, afterwards, perform a step climb to reach the final cruising altitude. However, cruising at 11000 [m] is considered inefficient and thus neglected. Instead, it immediately starts the step climb towards the optimal cruising altitude. This step climb is freely performed in 3D, where the aircraft increases its rate of climb employing its kinetic energy rapidly gaining altitude, and therefore reducing its Mach number. This maneuver lasts up to the point in which the aircraft is not producing lift anymore, i.e., when $C_L = 0$. At that point, which results in the optimal altitude, the aircraft pitches down to balance itself, acquiring again lift and gaining velocity to start performing the cruise phase.

The cruise phase is performed at constant altitude. Focusing on velocity, first the aircraft needs to speed up as illustrated above. After achieving a stationary value, we observe how the velocity profile slightly decreases as fuel is burnt. This is due to the fact that the optimal cruising speed reduces as the aircraft loses weight due to fuel consumption. At the end of the cruise, it must also adjust the speed to start the descent. Notice that, in case of a finer discretization, we would have observed shorter and smoother velocity changes at the beginning and at the end of the cruise. It is worth to mention that cruising at varying speed profiles is a very efficient strategy.

The aircraft starts then the descent, which is divided in several stages. First, the aircraft performs a 3D free descent down to 11000 [m], from which the aircraft must descent with constant course down to 10000 [ft]. At that point the aircraft is supposed to perform a horizontal deceleration in order to fulfill ATC velocity restrictions. However, this phase is considered inefficient and thus neglected, mainly because the descent is performed at varying velocity and the aircraft is capable of reaching 10000 [ft] below 250 [kt] (CAS). Indeed, the resulting velocity profile follows the patterns of the free flight performances exposed in the previous example. We can observe how the CAS profile is nearly constant below the tropopause. After reaching 10000 [ft], there is a 3D free phase which is employed to orientate the aircraft towards the assigned runway head. Then, the aircraft reaches the Final Approach Fix (FAP), deploys flaps and proceeds to runway. In order to analyze the final descent flight path angle, all exposed in the previous experiment holds. In this case, the value is -4.5°, also more inclined than current ILS procedures. Notice that whole descent is continuously performed, from the top of descent direct to the runway head.

We analyze now how the excess of take–off weight affects aircraft optimal trajectory. For this purpose we conducted a series of computational experiments in which the initial mass is set to a value over the obtained optimal initial mass whereas the final mass is not specified. Initial masses and fuel consumptions are shown in Table 5.12 and Figure 5.16.

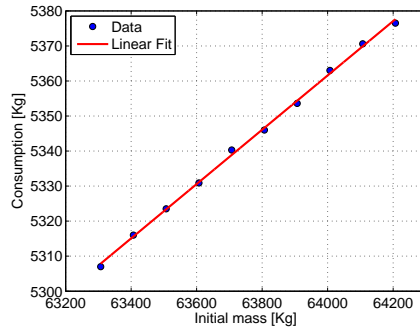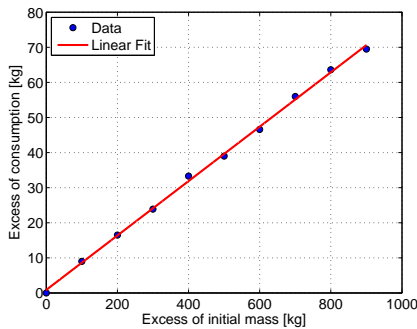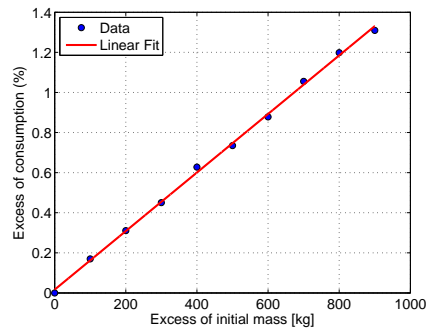| Initial mass [kg]	Consumption [kg]
63307*	5307.2
63407	5316.0
63507	5323.5
63607	5330.9
63707	5340.3
63807	5346.0
63907	5353.6
64007	5363.0
64107	5370.6
64207	5376.5

* Optimal initial mass.

Table 5.12: Consumption comparison.

Figure 5.16: Consumption comparison.

In Figure 5.16 we quantify the effects of the take–off weight excess in terms of extra fuel to be consumed during the trip. Figure 5.17(a) shows it in absolute terms, and Figure 5.17(b) shows it in relative terms, as a percentage of the optimal fuel consumption. The take–off weight excess is linearly related with the excess in fuel consumption. This linear relation can be quantified in 40 [kg] of extra fuel consumed per every extra 500 [kg] of initial take–off mass, i.e., a 0.75% of extra fuel consumption during the flight.

(a) Absolute terms.

(b) Relative terms.

Figure 5.17: Excess of consumption due to excess of take–off weight.

In order to discuss the effects that the initial mass has on performances, we also present altitude and speed profiles in Figure 5.18. For the sake of clarity, we do not explicitly differentiate the different trajectories, but only the optimal one. All of them follow very similar patterns, and this is true for the rest of variables.
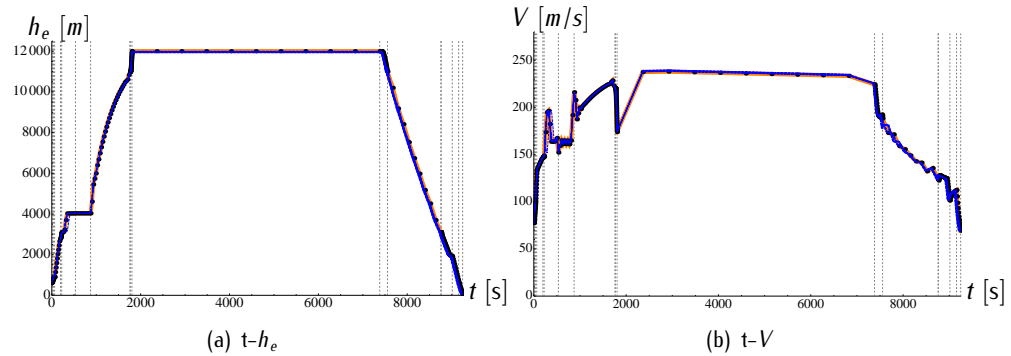
(a) t–h_e

(b) t–V

Figure 5.18: Optimal take-off trajectory: velocity and speed profiles comparison. Optimal take-off weight trajectory is the orange solid–dotted line (black dots correspond to the computed samples) and the rest are all in blue solid lines.

However, if we take a zoom-in look to the cruise phase such as in Figure 5.19, we can observe how the deviation from the optimal take-off weight affects the optimal altitude and the optimal speed profiles during the cruise. The optimal take-off weight leads to an optimal cruising (constant) altitude whose value is higher as the initial mass gets closer to the optimal one. The optimal take-off weight also leads to an optimal cruising (varying) velocity profile whose values are lower as the initial mass gets closer to the optimal one. The higher altitude and lower velocity profiles explain the lower consumption of the optimal take-off weight trajectory.
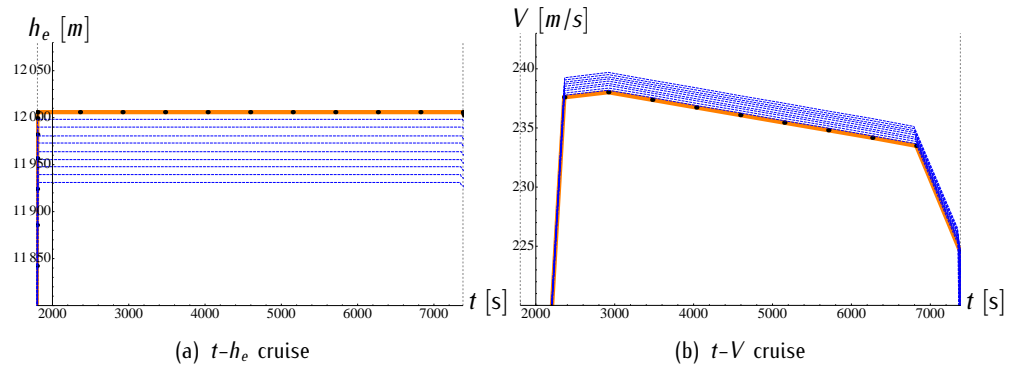

(a) t–h_e cruise

(b) t–V cruise

Figure 5.19: Velocity and speed profile comparison during cruise phase. Optimal take-off weight trajectory is the orange solid-dotted line (black dots correspond to the computed samples) and the rest are all in blue solid lines.

6

Trajectory Planning based on Multiphase Mixed-Integer Optimal Control

I N THIS CHAPTER we present the application of the multiphase mixed–integer optimal control problem to aircraft trajectory planning, considering thus integer and continuous variables. Integer variables model decision-making processes, and continuous variables describe the state of the aircraft which evolves according to differential–algebraic equations. The problem is formulated as a multiphase mixed–integer optimal control problem which is converted into a mixed integer non linear programming problem first making the unknown switching times part of the state as in Section 3.2.2, then applying a fifth–degree Gauss-Lobatto direct collocation method as in Section 2.3.1, and finally introducing binary variables to model the decision-making processes as in Section 4.3. The resulting mixed-integer non linear programming problem has been solved using a nonlinear programming based branch-and-bound algorithm. The approach is applied to the following en-route flight planning problem: given an aircraft point mass model, a wind forecast, an airspace structure, and the flying information regions with their associated overfly costs, we study the problem of finding the control inputs that steer an aircraft from the initial fix to the final fix following a route of waypoints while minimizing the fuel consumption and overfly costs during the flight. The decision-making process arises in determining the optimal sequence of waypoints. The optimal times at which the waypoints are to be overflown are also to be determined. The numerical results are presented and discussed, showing the effectiveness of the approach.

6.1   Multiphase mixed–integer trajectory optimization problem

6.1.1   Problem statement

In this chapter the multiphase mixed–integer optimal control techniques described in Chapter 3 will be applied to the the following en–route flight planning problem: given an aircraft point mass model, a wind forecast, an airspace structure, and the flying information regions with their associated overfly costs, find the control inputs that steer an aircraft from an initial fix to a final fix following a route of waypoints while minimizing the fuel consumption and overfly costs during the flight. The optimal sequence of waypoints is to be determined as well as the optimal overfly times over waypoints. This fight planning problem is a particular case of the general multiphase mixed–integer optimal control problem (MMIOCP) in which the number of switching times is prefixed, i.e., the feasible switching set is $\Psi = \Psi_t = \{\tilde{t}^1, \tilde{t}^2, \ldots, \tilde{t}^{N-1}\}$.

The airspace structure is modeled as a complete multipartite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ as exposed in Section 4.2.2. We assume that the aircraft is constrained to pass through one edge within every partite set of the graph in Figure 4.7, i.e., the aircraft is constrained to fly over the initial waypoint, $N-1$ waypoints, each belonging to a different partite set $\mathcal{V}^1, \ldots \mathcal{V}^{N-1}$, and the final waypoint. Thus, N phases can be identified during the motion of the aircraft.

The selection of the sequence of waypoints results from a decision–making process that is modeled using time–independent binary variables as explained in Section 4.3. In particular, we consider $N-1$ vectors of binary variables $v^k \in \{0, 1\}^{n_{v^k}}, k = 1, \ldots, N-1$, whose components are $v^{k,j}, k = 1, \ldots, N-1, j = 1, \ldots, n_{v^k}$.

Initial and final conditions: The initial and final positions of the aircraft will be given by the coordinates at the initial and final waypoints, $p^I = (\lambda_e^I, \theta_e^I)$ and $p^F = (\lambda_e^F, \theta_e^F)$ respectively. Let $\bar{x}^I = (\lambda_e^I, \theta_e^I), \bar{x}^F = (\lambda_e^F, \theta_e^F)$ be the initial and final horizontal positions or the aircraft. The initial condition x^I will be given by the initial position and a set of feasible initial values for the rest of state variables. The final condition coincides with the final position of the aircraft. The set of conditions (3.18) for initial and final states become:

$$x(t^I) = x^I; \tag{6.1}$$
$$\bar{x}(t^F) = \bar{x}^F. \tag{6.2}$$

Dynamical subsystems: The flight is divided into N phases, and thus there is a dynamical subsystem Σ^k governing the motion of the aircraft within each phase k. The dynamical subsystem Σ^k is characterized by the dynamic and algebraic constraints that were defined in the set of equations (3.17). As already mentioned in Chapter 5, in general, the dynamical subsystem Σ^k is a function of the discrete variables AC, DM, AM, and OP.

We assume the flight is performed in horizontal motion, and thus the differential–algebraic equations correspond to those given in Definition 4.5. As a consequence, the aircraft flies at constant altitude during the entire flight, always in the same atmospheric layer. We also assume that no operational procedure is specified throughout the flight. Moreover, we assume the aircraft flies during the whole flight with cruise configuration. Therefore, the values of the discrete variables AC, DM, AM, and OP are assumed to be known, and thus the dynamical subsystem Σ^k is uniquely determined and does not change during the N phases of the flight.

Path constraints: The path constraints of the general multiphase mixed–integer optimal control problem (MMIOCP) are represented by equation (3.19) and described in Section A.2. As previously mentioned in Chapter 5, the set of path constraints is in general a function of the discrete variable AC. We assume the aircraft flies in cruise configuration, and therefore the value of the discrete variable AC is known. Thus, the corresponding set of path constraints is uniquely determined and does not change during the N phases of the flight.

Interior point constraints: Let $\mathcal{S} = \{\mathcal{S}^1, \ldots, \mathcal{S}^{N-1}\}$ be the set of end trigger conditions. In this problem end trigger conditions correspond only to autonomous switchings. They are expressed by means of capture conditions associated to reaching a waypoint $p \in \mathcal{P}$, which is to be determined.

Each of the binary variables $v^{k,j}$, $k = 1, \ldots, N-1$, $j = 1, \ldots, n_{v^k}$, is associated to a waypoint $p^{k,j}$ of the set $\mathcal{P}$, and $v^{k,j} = 1$ means that the aircraft flies over waypoint j of partite set $\mathcal{V}^k$ at time $\tilde{t}^k$. Let $\bar{X} = \{\bar{x}^{1,1}, \ldots, \bar{x}^{N-1,n_{v^{N-1}}}\}$, with $\bar{x}^{k,j} = (\lambda^{k,j}, \theta^{k,j})$, be the set of waypoint locations. According to Figure 4.7, the interior point equality constraints given in the set of equations (3.20a) can be expressed as follows:

$$\mathcal{S}^k : \ \bar{x}(\tilde{t}^k) = \sum_{j=1}^{n_{v^k}} v^{k,j}\, \bar{x}^{k,j}, k = 1, \ldots, N-1. \tag{6.3}$$

Additional constraints are

$$\sum_{j=1}^{n_{v^k}} v^{k,j} = 1, \ k = 1, \ldots, N-1. \tag{6.4}$$

Condition (6.3) means that, if $v^{k,j} = 1$, it will be $\bar{x}^{k,j}(\tilde{t}^k) = \bar{x}^{k,j}$, that is, the aircraft will overfly waypoint in location $\bar{x}^{k,j}$ at time $\tilde{t}^k$, $k = 1, \ldots, N-1$. Condition (6.4) means that the aircraft must overfly only a single waypoint of partite set $\mathcal{V}^k$ at time $\tilde{t}^k$, $k = 1, \ldots, N-1$. Inequality interior point constraints in equation (3.20b) might be applied to constraint the switching instants. However, we will not use them in the problem.

Objective function: The objective functional in equation (3.23) to be minimized is

$$J = \sum_{k=0}^{N-1} \int_{\tilde{t}^k}^{\tilde{t}^{k+1}} \dot{m}^k(t) \, dt \ + \mathcal{D}, \tag{6.5}$$

where $\dot{m}^k(t)$ is the fuel flow of the aircraft during phase k, and $\mathcal{D}$ is the cost due to overflying charges which is computed taking into account the actual travelled distances in the relevant FIRs/UIRs when flying between couples of waypoints belonging to adjacent partite sets. The calculation of these overflying charges is described in Section 6.2. The expression for $\mathcal{D}$, which is a function of the binary variables $v^{k,j}$, is omitted for the sake of clarity in the exposition.

Notice that the objective functional (6.5) could be seen as a multi–objective functional balancing continuous costs corresponding to fuel consumption and costs corresponding to overflying charges. The natural choiceis to combine these two terms using their monetary cost. In Section 6.2.2, the impact of the costs due to overflying charges will be discussed by varying the weight of these cost in the objective functional.

6.1.2 Minlp **solution approach**

Recall what has already been exposed in Section 5.1.2 about the discretization of the problem for both time scales (see Figure 5.2–Figure 5.5).

The fifth degree Gauss–Lobatto collocation scheme described in Section 2.3.1 has been used. In the subinterval $[\tilde{t}_i^k, \tilde{t}_{i+1}^k]$, $k = 0, \ldots, N-1$, $i = 0, \ldots, N^k - 1$ the collocation points $x_{i,a}^k$ and $x_{i,b}^k$ are calculated from the state variables x_i^k, $x_{i,C}^k$ and x_{i+1}^k. The reader is referred to Figure 2.5 which illustrates how the state variables of the problem are discretized in the subinterval $[\tilde{t}_i^k, \tilde{t}_{i+1}^k]$. A free control scheme has been used for control variables. In this control interpolation scheme the discretized control variables represent discrete values for the controls at each discrete time at which the system equation are evaluated, $u_i^k, u_{i,a}^k, u_{i,C}^k, u_{i,b}^k, u_{i+1}^k$. For the sake of clarity, let us define $X_i^k = (x_i^k, x_{i,C}^k, x_{i+1}^k)$, $U_i^k = (u_i^k, u_{i,a}^k, u_{i,C}^k, u_{i,b}^k, u_{i+1}^k)$, $U_{i,a}^k = (u_i^k, u_{i,a}^k, u_{i,C}^k, u_{i+1}^k)$, and $U_{i,b}^k = (u_i^k, u_{i,C}^k, u_{i,b}^k, u_{i+1}^k)$, $k = 0, \ldots, N-1$, $i = 0, \ldots, N^k - 1$.

Applying the described transformations, the MINLP problem takes the form:

$$\min \; N \cdot \sum_{k=0}^{N-1} \left\{ (\tilde{t}^{k+1} - \tilde{t}^k) \cdot \sum_{i=0}^{N^k-1} \mathcal{Q}(X_i^k, U_i^k) \right\} + \mathcal{D};$$

subject to:

$$\mathcal{A}_{5,a}(X_i^k) + N \cdot (\tilde{t}^{k+1} - \tilde{t}^k) \cdot h^k \cdot \mathcal{C}_{5,a}(X_i^k, U_{i,a}^k) = 0, \quad k = 0, \ldots, N-1, i = 0, \ldots, N^k - 1;$$

$$\mathcal{A}_{5,b}(X_i^k) + N \cdot (\tilde{t}^{k+1} - \tilde{t}^k) \cdot h^k \cdot \mathcal{C}_{5,b}(X_i^k, U_{i,b}^k) = 0, \quad k = 0, \ldots, N-1, i = 0, \ldots, N^k - 1;$$

$$0 \leq \phi^k(X_i^k, U_i^k), \qquad\qquad k = 0, \ldots, N-1, i = 0, \ldots, N^k - 1;$$

$$x_0^0 = x^I, \bar{x}_{N_k}^{N-1} = \bar{x}^F;$$

$$\bar{x}_0^k = \sum_{j=1}^{n_{v^k}} v^{k,j} \cdot \bar{x}^{k,j}, \qquad\qquad k = 1, \ldots, N-1;$$

$$\sum_{j=1}^{n_{v^k}} v^{k,j} = 1, \qquad\qquad k = 1, \ldots, N-1;$$

$$v^{k,j} \in \{0, 1\}, \qquad\qquad k = 1, \ldots, N-1, j = 1, \ldots, n_{v^k};$$

$$\text{(MINLP)}$$

where $\mathcal{Q}(X_i^k, U_i^k)$ is obtained using a fifth–degree integration scheme as follows:

$$\int_{\tilde{t}_i^k}^{\tilde{t}_{i+1}^k} f(t) dt \approx \frac{h^k}{180} [9f(\tilde{t}_i^k) + 49f(\tilde{t}_{i,C}^k - \sqrt{3/7}h^k) +$$
$$64f(\tilde{t}_{i,C}^k) + 49f(\tilde{t}_{i,C}^k + \sqrt{3/7}h^k) + 9f(\tilde{t}_{i+1}^k)],$$

$\mathcal{D}$ is as defined in the previous section, and $\mathcal{A}_{5,a}(X_i^k)$, $\mathcal{C}_{5,a}(X_i^k, U_{i,a}^k)$, $\mathcal{A}_{5,b}(X_i^k)$, and $\mathcal{C}_{5,b}(X_i^k, U_{i,b}^k)$ come from equations (2.22)–(2.23). The unknowns of this problem are

$$(x_i^k, x_{i,C}^k, x_{i+1}^k, u_i^k, u_{i,a}^k, u_{i,C}^k, u_{i,b}^k, u_{i+1}^k)$$

for $k = 0, \ldots, N-1$, $i = 0, \ldots, N^k - 1$, together with $v^{k,j}$ for $k = 1, \ldots, N-1$, $j = 1, \ldots, n_{v^k}$, and the switching times $\tilde{t}^k$ for $k = 1, \ldots, N-1$. For the sake of clarity, unknowns $\tilde{t}^k$ for $k = 1, \ldots, N-1$ have not been renamed as elements of the extended state vector as in Section 3.2.2. Note that $\tilde{t}^N = t^F$ is also a variable of the problem if the final time is not specified.

The multiphase mixed–integer optimal control problem has now been recast as a mixed–integer nonlinear program: minimizing a nonlinear function subject to a number of nonlinear constraints in a space where some of the variables only take $0 - 1$ values while others take value in $\mathbb{R}$. We now turn to the explanation of the numerical method to solve problem (MINLP).

MINLP **resolution**

Fixing all variables $v^{k,j}$, $k = 1, \ldots, N - 1$, $j = 1, \ldots, n_{v^k}$, is equivalent to fixing the sequence of alternatives and if this is done the multiphase MIOCP becomes a conventional multiphase optimal control problem. A simple algorithmic approach could therefore be to enumerate all possible values for $v^{k,j}$, solve the associated multiphase optimal control problems and pick the best solution. Unfortunately, a rapid calculation of the number of problems to solve if one follows this approach shows that it is impractical for more than a handful of possible values for $v^{k,j}$. A common approach to try to address bigger problems is to do an implicit enumeration via a branch–and–bound algorithm [106, 107]. Branch–and–bound is a standard algorithm for integer programming (see for example [108] and references therein). A brief sketch of it in the context of this dissertation is given below, with an emphasis on the particularities that arise in the context of multiphase MIOCP. For a more complete exposition the reader is invited to refer to the references above.

Branch–and–bound is a divide–and–conquer method. The problem is divided by partitioning the set of feasible solutions into smaller and smaller subsets. The conquering is done by computing bounds on the value of the best feasible solution in each subset and discard subsets based on this bound. Branch–and–bound is an exact algorithm when the bound used in the fathoming phase is a valid lower bound. However, the problem of interest here is particular in that systematically obtaining a good lower bound on the value of the multiphase MIOCP is a daunting task.

Indeed, to compute a valid lower bound one has to build a *convex* approximation of the optimization problem where a convex function is minimized over a convex feasible region. Here, there are two sources of non-convexities: the binary variables of the problem and the nonlinear equations used to describe the feasible region. Dealing with binary variables is standard and can simply be done by replacing the set $\{0, 1\}$ with the interval $[0, 1]$, i.e., relaxing them. Dealing with nonlinear and highly non-convex equations is much more involved. Although systematic methods exist to compute convex approximations for such nonlinear equations (for example within the solvers Baron [109, 110] and Couenne [111]) their efficiency is limited. In the experiment herein presented they are not able to yield any good lower bound.

Therefore, the approach does not rely on a true lower bound but rather uses approximate solutions. In that case the procedure is heuristic (i.e. does not return the exact optimal solution). The quality of the final solution depends on the quality of the approximation. To the best of our knowledge, there is no theoretical guarantee on the quality of the approximation. Its practical efficiency is the subject of the computational section of this chapter.

Algorithm 1 NLP BB

---

0. **Initialize.**

   $\Gamma \leftarrow \{(\emptyset, \emptyset)\}$. $\mathcal{B}_\mathcal{U} = \infty$. $v^* \leftarrow$ NONE.

1. **Terminate?**

   Is $\Gamma = \emptyset$? If so, stop and return the sequence described by v^*.

2. **Select.**

   Choose and delete a problem $\mathcal{N}^l = (\mathcal{L}^l, \mathcal{U}^l)$ from Γ.

3. **Evaluate.**

   Solve the RIOCP$(\mathcal{L}^l, \mathcal{U}^l)$. If no solution can be found go to step 1, else let $\mathcal{B}_{\mathcal{L}^l, \mathcal{U}^l}^R$ be its objective function value and $\hat{v}$ be the values for the relaxed binary variables.

4. **Prune.**

   If $\mathcal{B}_{\mathcal{L}^l, \mathcal{U}^l}^R \geq \mathcal{B}_\mathcal{U}$ go to step 1. If $\hat{v} \notin \{0, 1\}^{N-1} \times \{0, 1\}^{n_{v^k}}$ go to step 5, else let $\mathcal{B}_\mathcal{U} \leftarrow \mathcal{B}_{\mathcal{L}, \mathcal{U}}^R$, $v^* \leftarrow \hat{v}$, and delete from Γ all problems with $\mathcal{B}_\mathcal{L}^k \geq \mathcal{B}_\mathcal{U}$. Go to step 1.

5. **Divide.**

   Create two new nodes $\mathcal{N}^{|\Gamma|}$, and $\mathcal{N}^{|\Gamma|+1}$. Choose $\hat{k}$ and $\hat{j}$ such that $\hat{v}^{\hat{k}, \hat{j}} \notin \{0, 1\}$. Let $\mathcal{B}^{|\Gamma|} \leftarrow \mathcal{B}^{|\Gamma|+1} \leftarrow \mathcal{B}_{\mathcal{L}, \mathcal{U}}^R$ and add the problem $\mathcal{N}^{|\Gamma|} = (\mathcal{L}^l \cup v^{\hat{k}, \hat{j}}, \mathcal{U}^l)$ and $\mathcal{N}^{|\Gamma|+1} = (\mathcal{L}^l, \mathcal{U}^l \cup v^{\hat{k}, \hat{j}})$. Go to 1.

Approximations are computed by using the relaxed integer optimal control problem (RIOCP), where the constraints $v^{k,j} \in \{0, 1\}$ are relaxed to $v^{k,j} \in [0, 1]$, for $k = 1, \ldots, N-1$, $j = 1, \ldots, n_{v^k}$. A locally optimal solution to the RIOCP can be computed with a nonlinear programming algorithm, for instance the interior point algorithm implemented by IPOPT. The branch–and–bound framework is then used to find a solution that satisfies also the integrity requirements $v^{k,j} \in \{0, 1\}$, $k = 1, \ldots, N-1$, $j = 1, \ldots, n_{v^k}$. This variant of branch–and–bound is usually called NLP based branch–and–bound or NLP BB for short (for more details, see for example [40] and references therein).

The first step of the branch–and–bound algorithm is to solve the RIOCP. If the solution obtained by solving the RIOCP is integer feasible (all variables $v^{k,j}$ take value 0 or 1) it specifies a sequence of points and the algorithm stops. If no solution to the RIOCP is found the algorithm stops. If an upper bound $\mathcal{B}_\mathcal{U}$ on the value of the optimal solution is known and the value of the solution of the RIOCP is above $\mathcal{B}_\mathcal{U}$, the algorithm also stops (fixing the infeasibility of the solution should increase the objective value of the solution). Otherwise, the algorithm divides the feasible region in two by fixing one of the variables $v^{k,j}$ such that $\hat{v}^{k,j} \notin \{0, 1\}$ to 0 and to 1 successively.

Applying the above steps recursively, yields to a tree Γ of partial assignment for the binary variables. At each node of this tree, a subset $\mathcal{L}$ of the variables v are fixed to 0 and a subset $\mathcal{U}$ are fixed to 1, and a local optimum of the restriction of the Riocp where the variables in $\mathcal{L}$ and $\mathcal{U}$ are fixed is to be sought. This restricted relaxed optimal control program is referred to as Riocp($\mathcal{L},\mathcal{U}$). The value of the upper bound $\mathcal{B}_\mathcal{U}$ is initially $+\infty$ and is updated whenever a new integer feasible solution is found such that the cost is improved.

The pseudo–code of the Nlp Bb is given in Algorithm 1. Several solvers implement this algorithm for example Minlp Bb [112] and Sbb [113]. In this dissertation, the solver Bonmin has been used [41]. Bonmin is an open–source Minlp solver implementing several different algorithms for solving mixed integer nonlinear optimization problems. Source code and binaries of Bonmin are available from Coin–or[1]. Bonmin is called through the Ampl[2] modeling language.

Two critical steps for the practical efficiency of Algorithm 1 which have not been explicated are the selection of the next subproblem to evaluate (step 2), and the choice of the variable to divide the feasible region (step 5). For these two steps standard rules implemented in Bonmin are used. The subproblem selected in step 2 is always the one with lowest $\mathcal{B}^{k,j}$ (best–bound rule). Whereas for choice of the variable $\hat{v}^{k,j}$, a default strategy in Bonmin is used which is a combination of strong–branching and pseudo–costs [40].

6.2   Case study

In this section the results of the application of the method described in the the previous sections to a realistic instance of the aircraft optimization problem will be described. More specifically, the trajectory optimization problem of an A330–301 aircraft performing the en–route part of a flight New York–Rome between the waypoint Yahoo: $p^I = (-69.74°, 41.69°)$, as initial fix, and the waypoint Amtel: $p^F = (11.60°, 43.21°)$, as final fix, is presented. The altitude of the route has been considered constant at 38000 [ft], that is, at flight level 380. The initial conditions of the problem were: $V(t^I) = 235$ [m/s], $\gamma(t^I) = 0°$, $\chi(t^I) = 0°$, $m(t^I) = 174000$ [kg].

The en–route part of the flight has been divided into $N = 9$ phases, also referred to as legs, with an initial and final waypoint, and 8 intermediate waypoints. Thus, 8 partite sets with $n_{v^k} = 5$ waypoints in each set have been considered.

The selection of the number of intervals has been done comparing solutions to the Minlp computed with increasing number of subintervals in each phase until a negligible change in the objective function was observed. Based on this criterion, the fifth–degree Gauss-Lobatto collocation method has been applied with a discretization using a total of 72 subintervals in which for the first 3 phases the number of subintervals is $N^0 = N^1 = N^2 = 12$ and for the other phases $N^3 = N^4 = N^5 = N^6 = N^7 = N^8 = 6$.

---

[1] http://www.coin-or.org

[2] http://www.ampl.com/

Horizontal flight dynamics

The horizontal motion over a spherical Earth including wind effects is considered. Recall Definition 4.5 for the 3–DOF horizontal motion of an aircraft over a spherical Earth including wind effects. The airplane is a conventional jet airplane and BADA 3.6 [14] is used as aircraft performance model (see Appendix A).

The differential–algebraic equations governing the translational horizontal motion of the airplane are the following:

$$
\begin{aligned}
m(t)\dot{V}(t) &= T(t) - D(V(t), C_L(t)); \\
m(t)V(t)\dot{\chi}(t) &= L(V(t), C_L(t)) \sin \mu(t); \\
L(V(t), C_L(t)) \cos \mu(t) &= m(t)g; \\
\dot{\lambda}_e(t)R_e \cos \theta_e(t) &= V(t) \cos \chi(t) + W_x(\lambda_e(t), \theta_e(t)); \\
\dot{\theta}_e(t)R_e &= V(t) \sin \chi(t) + W_y(\lambda_e(t), \theta_e(t)); \\
\dot{m}(t) &= -\eta(V(t))T(t).
\end{aligned}
\tag{6.6}
$$

In general, the engine thrust $T(t)$ and bank angle $\mu(t)$ are the control variables of the aircraft, that is $u(t) = (T(t), \mu(t))$. The thrust is commanded by the engine throttle and the bank angle is commanded combining rudder and ailerons trims. The state vector, x, will be: $x(t) = (\lambda_e(t), \theta_e(t), V(t), \chi(t), m(t))$. The cartesian coordinates x_e, y_e have been transformed into spherical coordinates according to relations (4.4).

The path constraints of the problem are those that define aircraft's flight envelope and restrictions in the control actions. They can be found in BADA database manual [14]. See Appendix A for details. The path constraints are then:

$$
\begin{aligned}
C_{V_{min}}V_S(m(t)) &\leq V(t) \leq V_{M0}; & M(V(t)) &\leq M_{M0}; \\
m_{min} &\leq m(t) \leq m_{max}; & 0 &\leq C_L(t) \leq C_{L_{max}}; \\
T_{min}(h_e(t)) &\leq T(t) \leq T_{max}(h_e(t)); & \mu(t) &\leq \mu_{max,civ}.
\end{aligned}
\tag{6.7}
$$

Waypoints

The waypoints and navaids of the AIRAC cycle published in June 2012 have been considered and a set $\mathcal{P}$ of 8×5 waypoints have been selected from them. For those phases entering or exiting oceanic regions, the waypoints have been selected manually coincident with the FIR/UIR bounds. This was the case of the first, second, and third partite sets. On the contrary, for those phases overflying the intra–European area, the waypoints have been selected randomly according to the following steps. First the trajectory optimization problem has been solved without waypoint constraints, obtaining the free flight trajectory. Then, the subpath from the intersection point between the free flight path and the French FIR/UIR to the final waypoint was considered. Since six phases had to be defined in this subpath, five equidistant points along it have been

$p^{1,1} = (41.11°, -67.00°)$	Dovey	$p^{1,2} = (40.11°, -67.00°)$	Jaboc
$p^{1,3} = (41.78°, -67.00°)$	Vitol	$p^{1,4} = (42.63°, -67.00°)$	Kanni
$p^{1,5} = (43.56°, -67.00°)$	Tusky	$p^{2,1} = (44.93°, -51.00°)$	Vodor
$p^{2,2} = (45.83°, -51.00°)$	Urtak	$p^{2,3} = (46.87°, -51.00°)$	Ronpo
$p^{2,4} = (47.82°, -51.00°)$	Novep	$p^{2,5} = (48.77°, -51.00°)$	Logsu
$p^{3,1} = (46.00°, -8.00°)$	Rivak	$p^{3,2} = (47.00°, -8.00°)$	Lapex
$p^{3,3} = (48.00°, -8.00°)$	Reghi	$p^{3,4} = (49.50°, -8.00°)$	Ratka
$p^{3,5} = (45.00°, -8.00°)$	Begas	$p^{4,1} = (45.93°, -5.22°)$	Erwan
$p^{4,2} = (44.50°, -4.94°)$	Atlen	$p^{4,3} = (44.61°, -5.40°)$	Kolek
$p^{4,4} = (46.32°, -3.69°)$	Novan	$p^{4,5} = (45.08°, -3.86°)$	Pepet
$p^{5,1} = (43.69°, -1.41°)$	Sigos	$p^{5,2} = (45.01°, -0.78°)$	Bmc11
$p^{5,3} = (46.05°, -2.25°)$	Godem	$p^{5,4} = (45.73°, -1.06°)$	Maren
$p^{5,5} = (44.55°, -1.12°)$	Cazaux Ndb	$p^{6,1} = (44.78°, 1.47°)$	Ratra
$p^{6,2} = (43.54°, 1.36°)$	Lfbf 12 Gs	$p^{6,3} = (45.33°, 1.23°)$	Makox
$p^{6,4} = (44.12°, 2.16°)$	Depes	$p^{6,5} = (44.95°, 2.36°)$	Aurillac Ndb
$p^{7,1} = (43.38°, 4.84°)$	Rhone	$p^{7,2} = (44.37°, 5.26°)$	Xirbi
$p^{7,3} = (45.66°, 4.89°)$	Rusit	$p^{7,4} = (45.10°, 5.16°)$	Romam
$p^{7,5} = (46.50°, 4.95°)$	Alura	$p^{8,1} = (42.89°, 8.67°)$	Rapur
$p^{8,2} = (44.59°, 8.66°)$	Testo	$p^{8,3} = (44.04°, 8.03°)$	Lbn32
$p^{8,4} = (45.15°, 7.99°)$	Sirlo	$p^{8,5} = (43.45°, 7.59°)$	Gonto

Table 6.1: Cordinates and designators of the waypoints.

selected. Finally, a random selection of the waypoints has been done using bivariate Gaussian probability density functions centered at the selected points in which the directions of the principal axes of the ellipses that correspond to equidensity contours are parallel to the tangent and normal directions to the subpath at each point. The resulting set $\mathcal{P}$ of waypoints is given in Table 6.1.

Wind Data

The wind forecast of July the 3rd, 2012 has been considered. Focusing on the North Atlantic region at an altitude of 200 [Hpa] ($h_e = 11769$ [m]), the wind forecast vector field is shown in Figure 6.1. Wind vectors (blue arrows) represent the direction and speed of the wind. Longer arrows represent faster winds. There is a region where the phenomena of stronger eastward winds can be identified. This phenomena is referred to as Jet stream, characterized by fast flowing, narrow air currents found in the atmosphere. The main jet streams are located near the tropopause at different latitudes. For aviation, characterizing the jet stream over the North Atlantic ocean is crucial in order to take advantage of favorable winds.
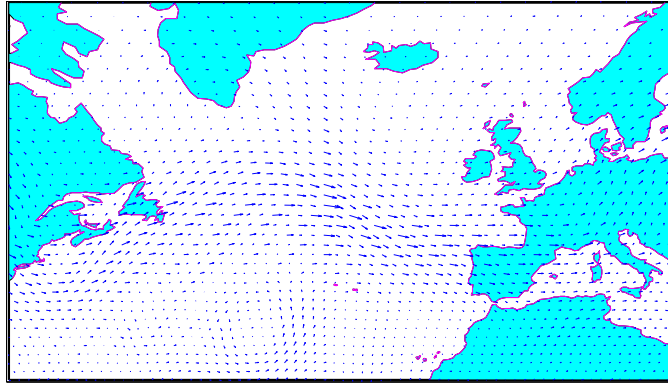

Figure 6.1: Wind vector field July the 3rd, 2012.

As exposed in Section 4.1.4, the NOAA wind forecast tabular data are fitted into analytical functions by means of multivariate regression analysis. A fourth–degree polynomial is fitted to the data. In this way, the east–west component can be expressed as:

$$W_x = \beta_0^x + \beta_{10}^x \lambda_e + \beta_{01}^x \theta_e + \beta_{20}^x \lambda_e^2 + \ldots + \beta_{13}^x \lambda_e \theta_e^3 + \beta_{04}^x \theta_e^4, \quad (6.8)$$

and the north component as:

$$W_y = \beta_0^y + \beta_{10}^y \lambda_e + \beta_{01}^y \theta_e + \beta_{20}^y \lambda_e^2 + \ldots + \beta_{13}^y \lambda_e \theta_e^3 + \beta_{04}^y \theta_e^4, \quad (6.9)$$

where $\beta_0^x, \ldots, \beta_{04}^x$ and $\beta_0^y, \ldots, \beta_{04}^y$ correspond to the coefficients of the regressions applied to the east and north components of the forecast tabular data, respectively. Analytical functions (6.8)–(6.9) are valid within a domain covering Eastern North America, the North Atlantic and Western Europe, i.e., $\lambda_e \in [-70°, 12°]$ and $\theta_e \in [40°, 55°]$. Functions (6.8)–(6.9) can be included in set of equations (6.6). Notice that some of the terms of the fourth–degree polynomials, that have been expressed in their general forms, may be neglected if they are found to be not significant according to their p–values.
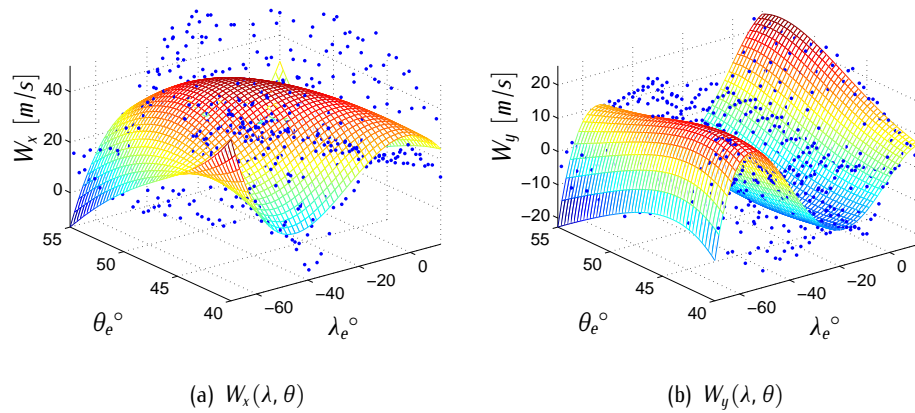
(a) $W_x(\lambda, \theta)$ (b) $W_y(\lambda, \theta)$

Figure 6.2:   North and East component of the wind speed at 200 [HPa] (h=11769 [m]) and the corresponding analytic functions that result from the regression analysis.

Figure 6.2 shows both forecast tabular data (blue dots) and polynomial functions (surfaces) for W_x and W_y at 200 [Hpa] (h=11769 [m]). In order to illustrate the important effects of the jet stream in trajectory optimization, Figure 6.3 shows the great circle distance (minimum distance) path and the free flight path that has been computed including wind effects.
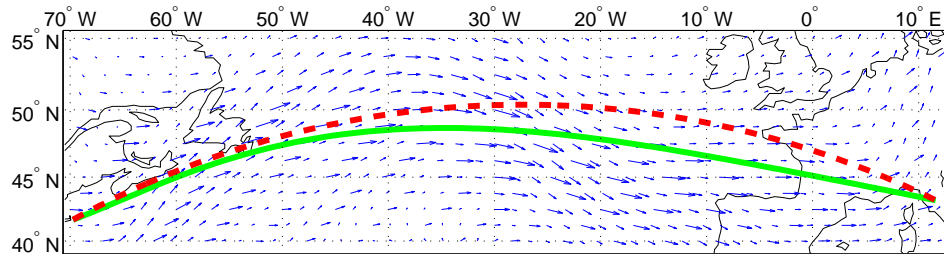

Figure 6.3:   Great circle distance path (red dashed line) and approximated optimal free flight path (solid green line).

The last step of multiple regression analysis is to check the goodness of fit of the model. Commonly used tests of goodness of fit include the multiple R–squared [105, Chap. 10.3] and the analysis of the patterns of residuals [105, Chap. 12.3].

In multiple regression, the multiple R–squared coefficient is a statistical measure of how well a regression approximates the real data points. Multiple R–squared ranges in the interval $[0, 1]$. A value of 1.0 indicates that the regression perfectly fits the data. A value over 0.7 is considered reasonably good, meaning that approximately seventy percent of the variation in the dependent variable can be explained by the least–squares linear combination of independent variables. The multiple R–Squared coefficient was 0.788 for W_x and 0.726 for W_y. This ensures that the predicted values fit reasonably well the data.

The residuals reflect the discrepancy between observed and predicted values that remain after the data have been fitted by the least-squares method. Each residual represents an estimate of the corresponding unobserved error. Hypothesis 4.21– Hypothesis 4.24 assume that these residuals are independent, have a zero mean, the same variance, and follow a normal distribution. The basic strategy underlying residual analysis is to assess the goodness of a regression model according to the behavior of the residuals in fulfilling the above mentioned hypotheses.

The most direct and revealing way to examine a set of residuals is to plot the errors between observed and predicted values against the dependent variables. If the model fitted to the data were correct, the residuals would approximate the random errors that make the relationship between observed and predicted values a statistical relationship fulfilling independence, zero mean, homogeneity in the variance, and normality. Therefore, if the residuals appear to behave randomly, it suggests that the model fits well the data. On the other hand, if non–random behavior is present in the residuals, it is a clear sign that the model fits the data poorly. The residuals are given in Figure 6.4. It can be observed that the errors do not behave completely random, in particular, the errors of the wind north function along the longitude. This suggest that the errors are not normally distributed.
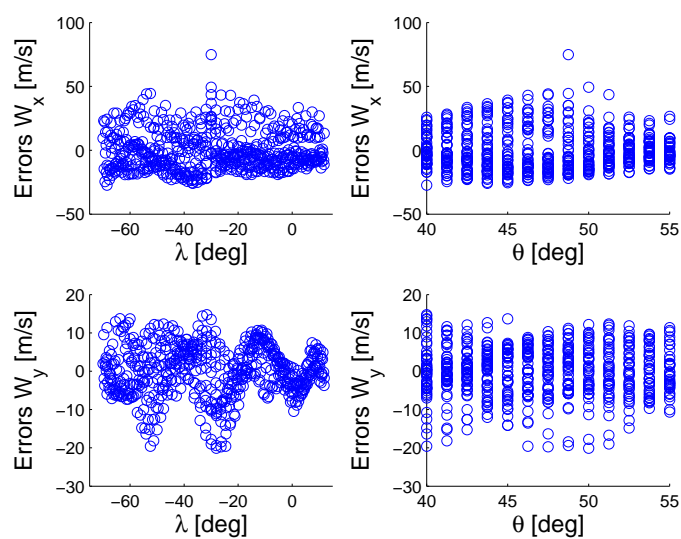

Figure 6.4: Regression residuals.

In order to check whether the errors are normally distributed one can run a Quantile– Quantile Plot (QQ Plot). A QQ–plot is a graphical method for comparing two probability distributions by plotting their quantiles against each other: in this case, the quantiles of the errors versus the quantiles of the standard normal. If data are related linearly, errors are normally distributed. QQ–plots are shown in Figure 6.5. It can be observed how data behave nearly linear in the central range of values, but they deviate within the tails of the the normal distribution. Therefore, we can conclude that errors are not normally distributed.
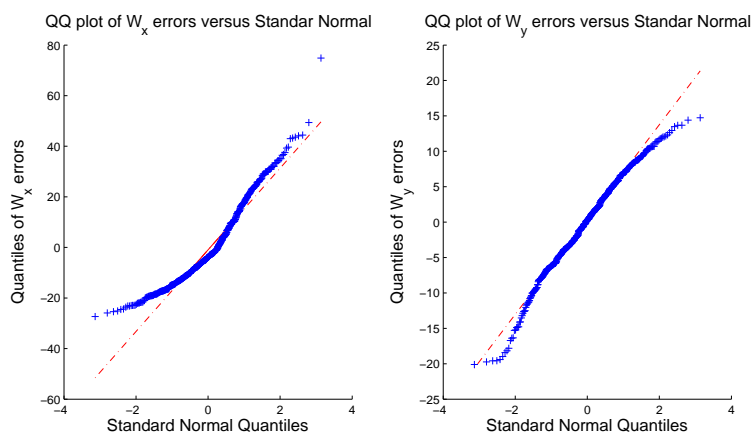
Figure 6.5: QQ-Plots.

A more quantitative criteria for assessing the validity of the normality assumption can be based on standard statistical tests such as the Kolmogorov–Smirnov test. Running such test over the two regressions, it confirms that the residuals are not normally distributed.

Therefore, a first conclusion is that the regression model herein applied is not able to completely capture the tabular data. However, it fits fairly good, being able to capture the fundamental tendencies, such as, the jet stream regions.

As already exposed in Section 4.1.4, regression is much more suitable than interpolation to be used in optimal control problems. Interpolation results in extremely large functions and make the optimization cumbersome. In regard of multiple regression analysis, higher order polynomial have been also tested, resulting in very low improvement in the regression while increasing complexity in the optimization problem. Other regression functions have been tested, for instance, combining polynomial and sinusoidal functions. Although such functions capture rather well some of the peaks, with multiple R-Squared values of about 0.85, the optimization problem does not converge. Thus a fourth-degree polynomial has shown the best trade off between goodness of fit and computational performance.

En-Route Overflying Charges

Figure 6.6 shows the FIR/UIR structure of the North Atlantic airspace. In general, national aviation authorities apply overflying fees for the services they provide. Very different charging schemes are applied including purely traveled distance-based charges, aircraft weight and traveled distance charges, flat rate charges (FR), or communication rate charges (CR) [2].

The charging methodologies in the relevant regions for the flight to be analyzed in the experiment, namely, United States of America (USA), Canada, and Europe, including the North Atlantic oceanic regions, are briefly presented.

In Europe, the standard EUROCONTROL charge formula for en-route services in the
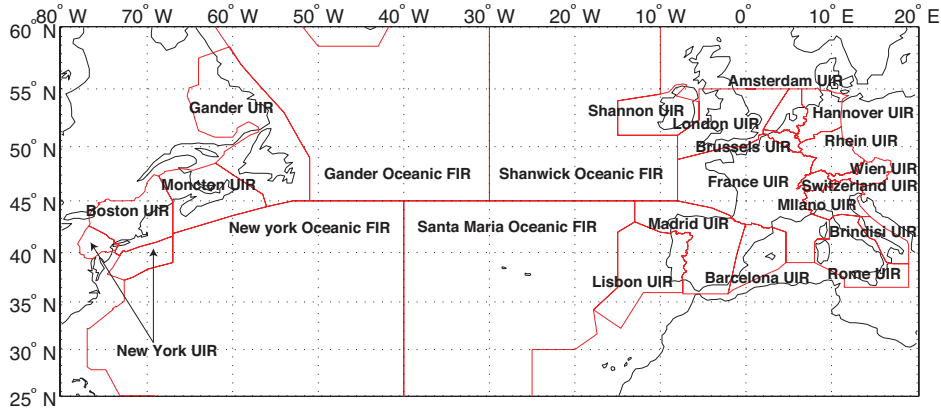
Figure 6.6: Relevant FIR/UIR regions for flights between North America and Europe.

EUROCONTROL members countries is:

$$r_i^{\text{Eur}} = UR_i \frac{GCD_i}{100} \sqrt{\frac{MTOW}{50}},$$

where UR_i is the service unit rate in FIR$_i$ (referring to member country i), GCD_i is the great circle distance in Kilometers [km] traveled in FIR$_i$, and $MTOW$ is the maximum take–off weight in metric tonnes [t] of the aircraft. The unit rates of en–route charges are established by each EUROCONTROL member state and updated every month [114]. In the USA, the Federal Aviation Administration (FAA) only charges overflight fees to operator that fly in the USA controlled airspace, but neither take off nor land in the USA. In the continental airspace, the en–route charges r^{USACon} are \$ 38.44 per 100 nautical miles (measured in Great Circle Distance (GCD)). In the oceanic airspace, the fee r^{USAOc} is \$ 17.22 per 100 nautical miles (in GCD)[3]. Nav Canada applies different fees for its oceanic and continental airspaces. Canadian oceanic charges in Gander Oceanic FIR are based on a flat rate that can be decomposed into navigation fee FR^{GanOc} of C\$ 93.24 and a communication fee CR^{GanOc} of C\$ 22.04 [115]. Canadian continental airspace charges are based on aircraft weight and traveled distance as follows:

$$r^{\text{CanCon}} = UR \cdot GCD \cdot \sqrt{MTOW},$$

where the Unit Rate (UR) is \$ 0.03445, the travelled GDC in [km], and MTOW in [t]. Charges for services provided in the Shanwick Oceanic FIR comprise a flat communication rate CR^{ShOc} of € 45 (charged by Ireland) and a flat navigation fee FR^{ShOc} of £ 65.70 (charged by United Kingdom) [116].

Focusing now in the case study herein analyzed, the overflying cost for a flight from a USA airport to Europe through Canadian continental airspace, Gander Oceanic, and Shanwick Oceanic FIRs can be expressed as

$$r^{\text{CanCon}} + FR^{\text{GanOc}} + CR^{\text{GanOc}} + FR^{\text{ShanOc}} + CR^{\text{ShanOc}} + \sum_i r_i^{\text{Eur}},$$

---

[3]http://www.faa/gov/air_traffic/international_aviation/overflight_fees. Last visited 06–25–2012.

where r_i makes reference to the ith relevant European FIR/UIR. The components of this cost have been defined in Section 6.2. Notice that, since the flight departs from JFK airport, the USA do not apply any navigation fee. The unit rates employed in the European regions are those corresponding to European regions adjusted unit rates applicable to April 2012 flights[4]. For continental Canada and oceanic regions the rates are those given in Section 6.2. All rates have been converted to €. It has been assumed that 1 kg of fuel costs € 1.

Flight plan model

The case we study involves a flight plan with a 9–phases model, which is given in Table 6.2. The values taken by the discrete variables Ph, DM, OP, AC, and AM are preset. On the contrary, the value of the discrete variable Wp is not specified and so are the end trigger conditions $\mathcal{S}^k$, $k = 1, \ldots, N-1$. The waypoint's sequence is to be selected according to the values taken by the binary variables $v^{k,j}$.

Ph	DM	OP	AC	AM	$\mathcal{S}$
0	HM	–	CR	Ab	$\sum_{j=1}^{n_{v1}} v^{1,j} p^{1,j}$
1	HM	–	CR	Ab	$\sum_{j=1}^{n_{v2}} v^{2,j} p^{2,j}$
2	$\mathbb{H}$	–	CR	Ab	$\sum_{j=1}^{n_{v3}} v^{3,j} p^{3,j}$
3	$\mathbb{H}$	–	CR	Ab	$\sum_{j=1}^{n_{v4}} v^{4,j} p^{4,j}$
4	$\mathbb{H}$	–	CR	Ab	$\sum_{j=1}^{n_{v5}} v^{5,j} p^{5,j}$
5	$\mathbb{H}$	–	CR	Ab	$\sum_{j=1}^{n_{v6}} v^{6,j} p^{6,j}$
6	$\mathbb{H}$	–	CR	Ab	$\sum_{j=1}^{n_{v7}} v^{7,j} p^{7,j}$
7	$\mathbb{H}$	–	CR	Ab	$\sum_{j=1}^{n_{v8}} v^{8,j} p^{8,j}$
8	$\mathbb{H}$	–	CR	Ab	p^F

Table 6.2: Horizontal motion: flight plan model.

6.2.1 Results

The computed sequence of waypoints, denoted by the active set of binary variables $v^{k,j}$, is given in Table 6.3. The corresponding route is: YAHOO, DOVEY, VODOR, RIVAK, PEPET, BMC11, RATRA, XIRBI, LBN32, AMTEL.

The approximated optimal path has been depicted in Figure 6.7, where the dots represent the computed discrete samples. The switching and final times of the approximated optimal solution are given in Table 6.4 together with the accumulated consumed fuel at the end of each leg and the overflying costs for each leg. The approximated optimal evolution of both state and control variables within the time domain are represented in Figure 6.8 and Figure 6.9, where the dots represent the computed discrete samples and the vertical lines correspond to the switching times.

---

[4] Santa Maria: € 9.79; United Kingdom France: € 83.23; Spain (continent): € 71.84; France: € 64.63; Italy: € 78.69; Portugal (Lisbon UIR): € 33.06.

$v^{k,j}$	$v^{k,1}$	$v^{k,2}$	$v^{k,3}$	$v^{k,4}$	$v^{k,5}$
$v^{1,j}$	1	0	0	0	0
$v^{2,j}$	1	0	0	0	0
$v^{3,j}$	1	0	0	0	0
$v^{4,j}$	0	0	0	0	1
$v^{5,j}$	0	1	0	0	0
$v^{6,j}$	1	0	0	0	0
$v^{7,j}$	0	1	0	0	0
$v^{8,j}$	0	0	1	0	0

Table 6.3: Switching sequence.

Switching times [s]	Accumulated Consumption [kg]	Overflying Costs [€]
$\tilde{t}^1 = 951.0$	1192.8	0
$\tilde{t}^2 = 6064.3$	7995.1	0
$\tilde{t}^3 = 18407.7$	23150.3	218.2
$\tilde{t}^4 = 19720.1$	24687.4	447.7
$\tilde{t}^5 = 20671.4$	25794.2	326.0
$\tilde{t}^6 = 21388.5$	26626.2	245.9
$\tilde{t}^7 = 22613.5$	28042	413.0
$\tilde{t}^8 = 23533.4$	29103.6	310.5
$\tilde{t}^9 = 24829$	30417.1	399.4

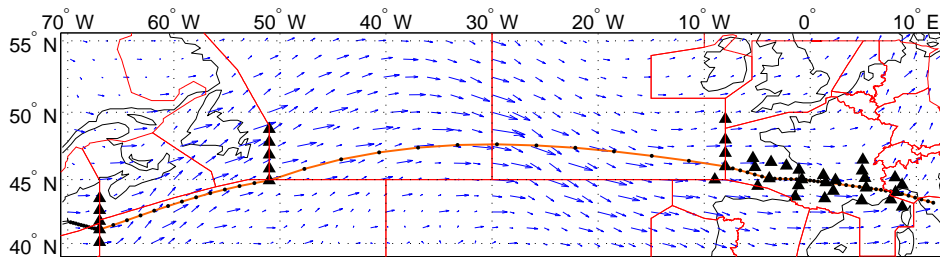Table 6.4: Switching times, fuel consumption, and overflying costs.

Figure 6.7: Approximated optimal path: the dots correspond to the computed samples.
The triangles correspond to the waypoints of set $\mathcal{P}$.

(a) $\tau - V$

(b) $t - V$

(c) $\tau - \chi$

(d) $t - \chi$

(e) $\tau - m$

(f) $t - m$
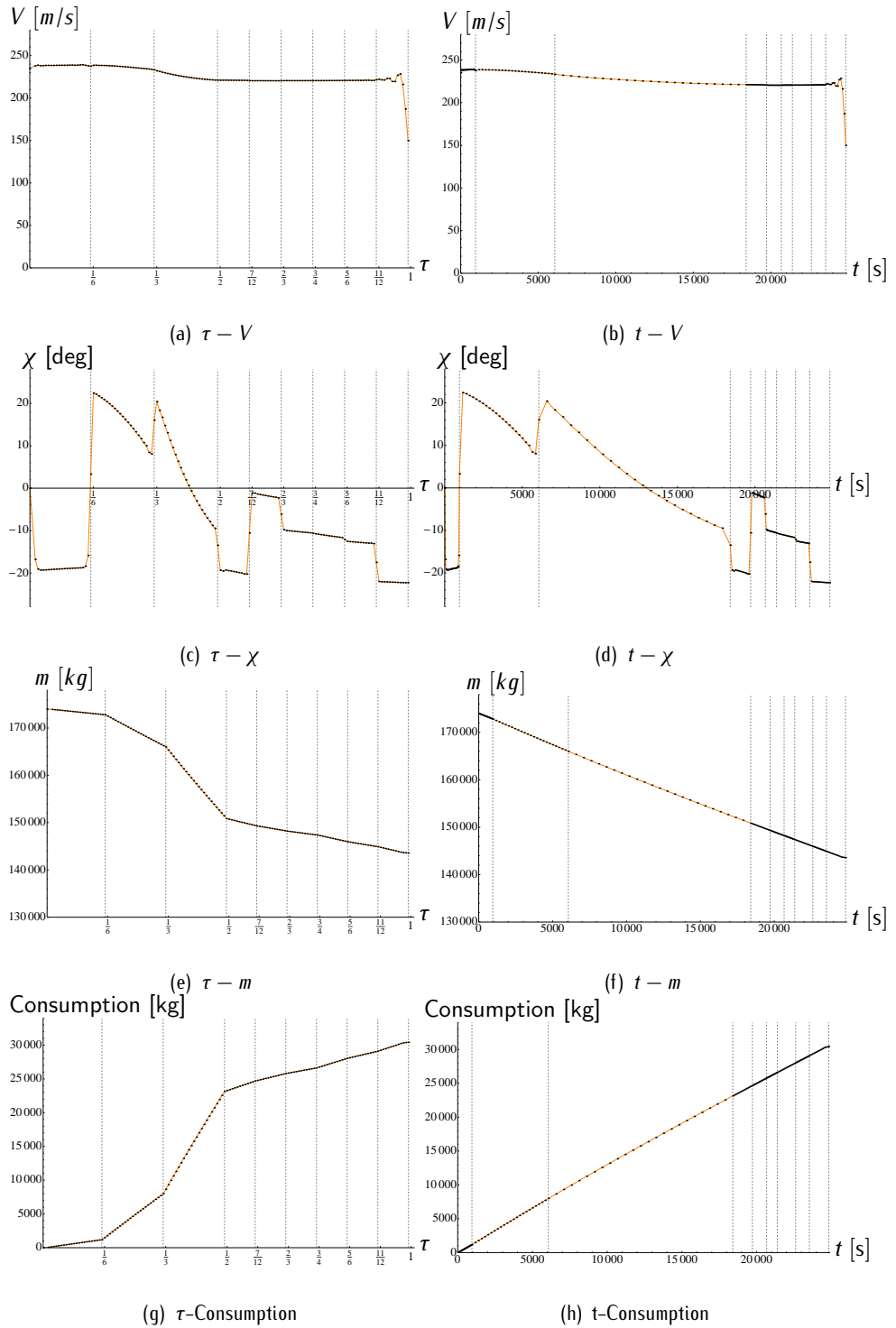
(g) τ-Consumption

(h) t-Consumption

Figure 6.8: State variables V, χ, m; and consumption (dots corresponding to the computed samples). Vertical dashed lines correspond to the switching instants.

(a) $\tau - T$

(b) $t - T$

(c) $\tau - \mu$

(d) $t - \mu$

(e) $\tau - C_L$

(f) $t - C_L$

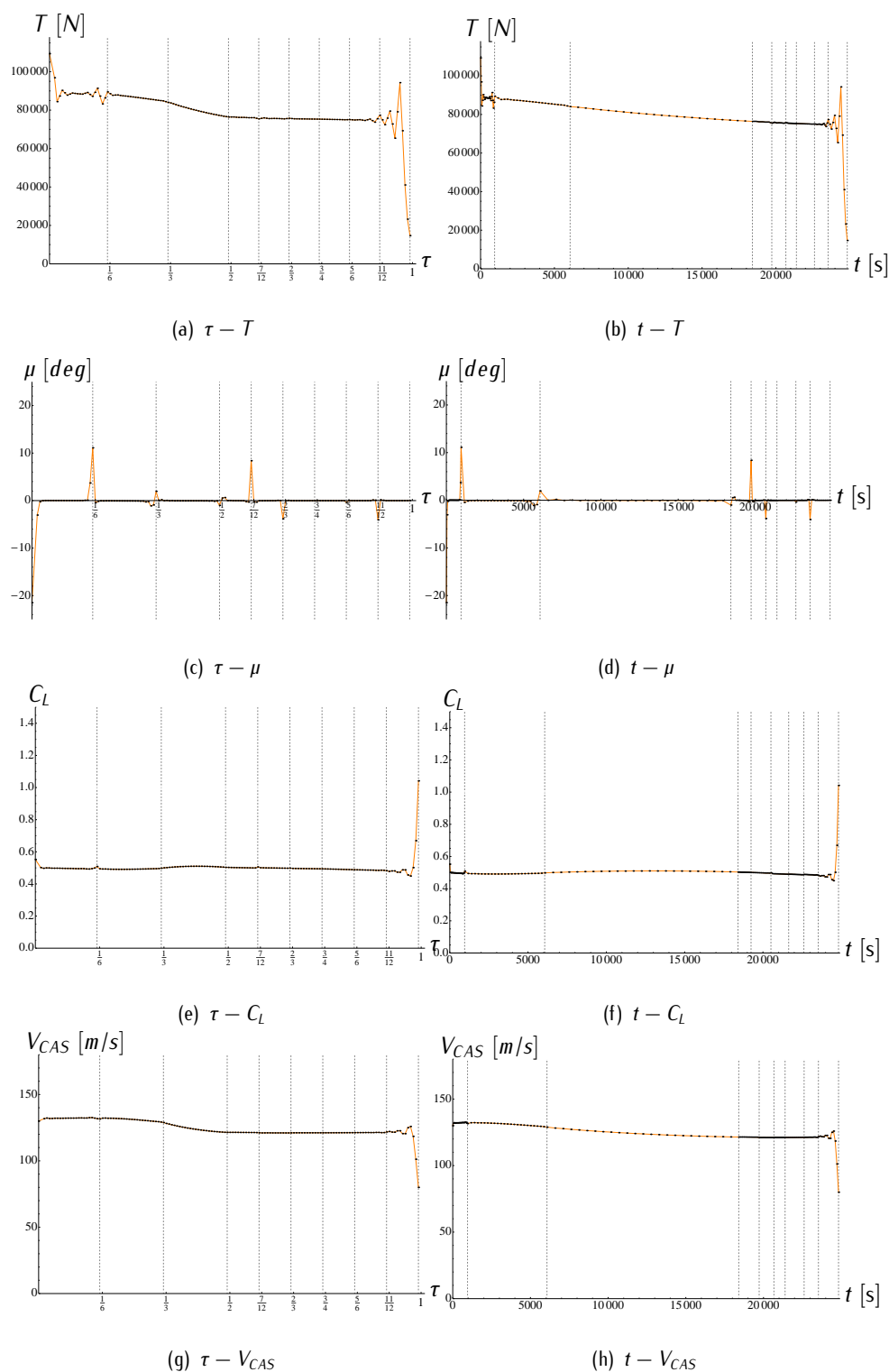(g) $\tau - V_{CAS}$

(h) $t - V_{CAS}$

Figure 6.9: Control variables μ and T; and variables C_L and V_{CAS} (dots corresponding to the computed samples). Vertical dashed lines correspond to the switching instants.

6.2.2   Discussion on the results

The MINLP model used to solve this problem had 1730 variables (40 of them being integer variables) and 1672 constraints. The MINLP solver took 4726 iterations and 30 nodes, with maximum depth 6 on tree, to find a solution whose objective value was € 32778.48.

The velocity of the aircraft of the approximated optimal solution represented in Figure 6.8(b) shows a non-steady behaviour in which it slightly decreases as fuel is burnt. Non steady flight performances are interesting in defining more efficient flight profiles within a future TBO concept of operations. For instance, varying velocity profiles have been shown to save fuel [66] [67], but could also serve to absorb ground delays on air or for defining a more efficient ATFM [69].

The heading angle of the aircraft of the approximated optimal solution represented in Figure 6.8(d) is also non-steady. Instead of flying segments with constant course, the aircraft flies slightly modifying the course by acting on ailerons and rudder so that the bank angle is in general non zero. These maneuvers can be also observed in the evolution of the bank angle reported in Figure 6.9(d).

The trajectory planning problem can be solved on the ground, as part of the strategic planning, or during the tactical planning to redefine the trajectory. In both cases a reduced computational time is a key factor to deliver flight plans or to agree modifications of planned trajectories with ATC. Therefore, a special effort has been spent on reducing the computational time and two different strategies have been combined to achieve this goal.

On the one hand, a 5th degree Gauss-Lobatto integration rule has been used with 72 subintervals. The computation time was 598.69 [s] on a Mac OS X 2.56 [GHz] laptop computer with 4 GB RAM. To give a quantitative measure of the computational time reduction that can be achieved with this integration rule, it is worth mentioning that the resolution of the MINLP problem discretized using a Hermite-Simpson collocation method with 290 subintervals took 2987 [s] on the same computer and the same approximated optimal solution was obtained. This is congruent with the analysis given in [36].

On the other hand, an efficient heuristic has been implemented. As exposed in Section 6.1.2, this heuristic approach has two main steps. First, the integer values are relaxed to the continuous domain $[0, 1]$, and the resulting NLP subproblems are solved to local optimality. Then, using a branch-and-bound framework, a solution that satisfies the integer requirements is sought. In the branching process the branching rules can be seen as heuristics aimed at reducing the size of the search tree, i.e., some regions in which no good integer feasible solution is expected are discarded but there is no theoretical guarantee that supports the choice.

As pointed out before, there exist exact MINLP solvers for problems such as the one presented in this dissertation, although they are typically limited to problems of medium difficulty. To show that exact solvers are not adequate for the problem at hand, the state-of-the-art solver Couenne [111] was tested. It was not able to compute any

feasible trajectory. It is worth stressing out again that the difficulty in solving exactly the aircraft trajectory optimization problem lies not only in its size but also in its highly nonlinear and non–convex nature. To the best of our knowledge, there exist no practical method to compute the optimal solution of this problem. Therefore, assessments on the quality of the solution can only be heuristic. To this purpose, three different tests have been carried out and will be described below.

First, the problem presented in this dissertation has $5^8 = 390625$ feasible solutions for the binary variables. Since it is impractical to solve all the corresponding NLP subproblems and compare the obtained values of the objective function with the solution found by BONMIN, a sample of feasible values for the binary variables have been selected. This sample has been generated based on slight modifications of BONMIN's computed sequence of waypoints, in which one of the BONMIN's computed waypoints is permuted by another waypoint within its partite set, whereas the remaining waypoints are unchanged. This sampling results in 32 integer feasible NLP subproblems, which have been solved using IPOPT and have been compared to BONMIN's computed solution. Results show that BONMIN's computed solution is always the best.

Second, the sensitivity of the MINLP algorithm to costs due to overflying charges has been analyzed. For this purpose, a parameter $\delta \in [0, 1]$ that multiplies the Mayer term in the objective functional (6.5) has been considered. In this way, $\delta = 1$ corresponds to the problem in which overflying charges are considered, whereas $\delta = 0$ corresponds to the problem considering only fuel consumption cost. Table 6.5 shows the performances of the algorithm for different values of the parameter δ. It can be observed that the number of explored nodes decreases as the weight assigned to overflying costs increases, resulting in faster computation and that the selected route is very sensitive to changes of the values of δ. The 7 values of δ reported in Table 6.5 give rise to 4 different routes A, B, C, and D listed in Table 6.6.

δ	Iterations	Nodes	$t_{comp.}[s]$	Obj. func. [€]	Route	Con. [kg]	Overflying Costs [€]
0	21723	401	1835.67	30014.8	D	30014.8	–
0.05	16553	258	1505.15	30390.38	C	30015.7	374.68
0.1	9147	58	953.09	30647.45	B	30410.1	237.35
0.25	7602	49	788.80	31002.87	B	30410.1	592.77
0.5	6095	36	633.11	31595.23	B	30410.1	1185.1
0.75	5873	37	607.07	32187.6	B	30410.1	1777.5
1	**4726**	**30**	**598.69**	**32778.48**	**A**	**30417.1**	**2361.38**

Table 6.5: Sensitivity of the algorithm to changes of the overflying costs.

Route	Waypoints								
A	Yahoo	Dovey	Vodor	Rivak	Pepet	Bmc11	Ratra	Xirbi	Lbn32	Amtel
B	Yahoo	Dovey	Vodor	Rivak	Erwan	Bmc11	Ratra	Xirbi	Lbn32	Amtel
C	Yahoo	Kanni	Ronpo	Lapex	Kolek	Maren	Aurillac Ndb	Xirbi	Lbn32	Amtel
D	Yahoo	Kanni	Novep	Lapex	Kolek	Maren	Aurillac Ndb	Xirbi	Lbn32	Amtel

Table 6.6: Routes A, B, C, and D.

δ	Route A	Route B	Route C	Route D
0	30417.6	30410.5	30016.1	**30014.8**
0.05	30535.6	30528.9	**30390.38**	30690.86
0.75	32188.26	**32187.6**	35629.9	36047.5
1	**32778.48**	32779.96	37501.21	37960.65

Table 6.7: Values of the objective function [€] for different routes.

Another test has been conducted to establish if the algorithm provides the most efficient route for different values of the parameter δ. In Table 6.7 the values of the objective function corresponding to different routes and different values of δ are reported. It can been seen that the algorithm always selected the most efficient route for each value of δ.

Third, the sensitivity of the algorithm to the initial guess of both continuous and discrete variables has been analyzed. To generate the initial guess of the continuous variables of each Nlp subproblem, using information on the physical system is important since a bad initial guess might lead to non-convergence of the Nlp problem. In general, common flight performances are sufficient to create a general initial guess that leads to relatively rapid and robust convergence. Such initial guess for the Nlp subproblems contributes to reduce the overall computational time of the Minlp algorithm since a large amount of Nlp subproblems must be solved in the branch-and-bound algorithm. Moreover, having a good initial guess in terms of convergence is important in the branch-and-bound algorithm since a failure in solving any of the Nlp subproblems might lead to reject a branch of the search tree that contains the optimal solution. It is worth pointing out that every single Nlp subproblem computed in the different tests whose results are reported in Table 6.5 resulted to be feasible and that the interior point strategy implemented in Ipopt to deal with infeasible initial guesses has contributed to this result.

In regard of the initial guesses for the binary variables, as said before, the first step of the branch–and–bound algorithm is to solve the relaxed MINLP problem, which is obtained replacing the integrality constraint $v^k \in \{0, 1\}^{n_{v^k}}$ by $v^k \in [0, 1]^{n_{v^k}}$, $k = 1, \ldots, N-1$. Therefore, the initial guesses for the variables $v^k, k = 1, \ldots, N-1$ are actually vectors of real numbers in the interval $[0, 1]$. Relaxing the integrity of the variables v^q corresponds to request that, at the switching times, the aircraft overflies points that are convex combinations of the waypoints $p^{k,j}$, $k = 1, \ldots, N-1, j = 1, \ldots, n_{v^k}$. In the case in which overfly costs are not considered, the best option is to set the initial values so that its convex combination results in the best achievable solution to the relaxed problem. Using these initial guesses, the results presented in Table 6.5 for $\delta = 0$ have been obtained. Using different initial guesses for the variables $v^k, k = 1, \ldots, N-1$, the algorithm converged to the same solution but with higher computation time.

In the case in which overflying costs are considered, i.e., $\delta \neq 0$, it is not easy to calculate the best achievable solution to the relaxed problem and thus applying the above mentioned strategy to compute the initial guesses for the variables $v^k, k = 1, \ldots, N-1$, is difficult. Therefore, the previously described procedure to compute initial guesses for $\delta = 0$ has been also used for $\delta = 1$ and the algorithm converged with an objective function value of € 32779.13. To check the quality of this strategy, 25 different initial guesses have been generated and tested. The algorithm converged 19 times with a slightly lower value of the objective function € 32778.48, 3 times with € 32779.13, and 3 times with the value € 37497.92 of the objective function. This shows that, if overflying costs are considered, the algorithm is more sensitive to the initial guesses for the binary variables.

7

Conclusions and Future Work

I N THIS DISSERTATION optimal control techniques have been developed to devise more efficient aircraft trajectory planning concepts. Since flight planning involves many factors that can be classified either as continuous or discrete phenomena, together with decision-making processes, a first effort was made in finding a complete modeling framework capable of accurately reflect them. A framework modeling based on multiphase mixed-integer optimal control was proposed to that end. Then, we have concentrated efforts in finding numerical optimal control techniques and algorithms in order to effectively find optimal solutions to the trajectory planning problem. First, a multiphase optimal control approach has been devised in which multiple phases and multiple dynamical subsystems have been considered. Then, decision-making processes have been included in the trajectory planning problem, presenting a realistic multiphase mixed-integer optimal control approach to aircraft trajectory optimization.

The effectiveness of these approaches has been proven solving different single-aircraft, deterministic trajectory optimization problems. The multiphase mixed-integer optimal control approach permits discrete and continuous dynamics to be combined. The decision-making processes have been modeled using binary variables. With these approaches we have been able to find solutions to flight planning problems considering issues such as nonlinear aircraft performances, different flap configurations, different layers of the atmosphere, wind forecasts, amount of departure fuel, multiple operational constraints, and the determination of the route of waypoints to be followed within an airspace structure.

Conclusions

According to the obtained results, we can draw three main conclusions:

a) The multiphase optimal control approach is suitable for effectively define complete, more efficient, single-aircraft trajectories framed into a constrained-based trajectory planning paradigm.

b) The multiphase mixed-integer optimal control approach is suitable for modeling decision making-processes using binary variables and solving fairly realistic flight planning problems. In particular, the combination of integer variables and continuous variables in optimal control problems is viable in a complex environment as it is ATM.

c) The numerical algorithms devised for both approaches are able to find solutions in time frames compatible not only with strategic but also with tactical planning.

Overall, we can conclude that the optimal control techniques studied in this dissertation have a strong potentiality in defining more efficient flight plans towards future operational concepts based on 4D business trajectories.

Open problems and future work

Natural extensions of the applications solved in this thesis are threefold. First, decision-making processes could involve not only interior point constraints as described in Chapter 6, but also the selection of dynamical subsystems as introduced in [78]. Second, decision-making processes could be also included into a three dimensional flight planning problem in which an airspace structure with flight levels and waypoints are considered as introduced in [117]. Third, environmentally friendly trajectories including aircraft CO_2 and NO_x emissions in the objective function and considering areas of persistent contrail formation could be studied. Moreover, more efforts in wind forecast modeling are needed to improve the predictability of the trajectory.

The multiphase mixed-integer optimal control approach presented in this dissertation has also certain limitations that can be considered open problems. First, the (local) optimality of the NLP solution has not been proven in practice. Therefore, efforts in this direction are needed. Second, the MINLP algorithm is only capable of handling a limited number of binary variables to achieve a solution within reasonable computational times. In order to increase the number of binary variables, for instance to model a more realistic airspace graph structure, efforts in finding heuristic approaches capable of solving harder combinatorial problems with similar computational times are needed. Third, the approach is restricted to a fixed number of phases. Future work entails thus extending the approach to problems with variable number of phases.

From an ATM perspective, the deterministic, single-aircraft trajectory optimization approach has also limitations since the flight planning problem entails, among other aspects, considering multiple aircraft and tackling uncertainty. Achievements to these ends would empower the real implementation of 4D business trajectories, leading to increases of capacity, safety, economical efficiency, and reductions in the aviation environmental fingerprint.

Other applications of these methods are, for instance, the design of arrival and departure procedures in high density airports that include decision-making processes.

Finally, it is interesting to point out that the approaches presented in this thesis are not exclusively suitable to commercial aircarft flight planning purposes. Other potential applications include planning UAV missions, such as surveilance misions, or space missions, such as asteroid interception.

A

The Base of Aircraft DAta

THE BASE OF AIRCRAFT DATA (BADA) is a collection of ASCII files which specifies operation performance parameters, airline procedure parameters, and performance summary tables for 295 aircraft types. This information is designed for use in trajectory simulation and prediction algorithms within ATM domain. All files are maintained within a configuration management system at the Eurocontrol Experimental Centre (EEC) in Brétigny-sur-Orge, France. In particular, we expose here BADA Revision 3.6 [14].

This appendix aims to summarize the BADA performance models and BADA performance limitations that have been used in this thesis. For more details and further information on performance models, performance limitation or airline procedure models, the reader is encouraged to see BADA Revision 3.6 [14] and posterior versions.

A.1   BADA performance models

BADA is based in several models for characterizing the 295 aircraft types included in BADA Revision 3.6. We will detail only those models used in this thesis. For an in-depth insight on other useful models, the reader is referred to [14, Chap. 3].

Aircraft type

The aircraft model can be found in the so called *aircraft type block*, in which the following characteristics are given: ICAO aircraft code; number of engines, n_{eng}; engine type; and wake category (heavy, medium or light).

   Four mass BADA operation performance parameters, given in tons, are specified for each aircraft:

- m_{ref}: reference mass;

- m_{max}: maximum mass (MTOW);

- m_{min}: minimum mass (OEW); and

- m_{pyld}: maximum payload mass.

Aircraft operating speeds vary with the aircraft mass. This variation is calculated according to the formula below:

$$V = V_{ref}\sqrt{\frac{m}{m_{ref}}}. \tag{A.1}$$

Notice that all velocities given as parameters in BADA are reference velocities (V_{ref}), and thus, only apply for the reference mass. For any other given mass, the velocity must be calculated using equation (A.1).

Thrust model

The thrust is calculated in Newtons and includes the contribution from all engines. The first thrust to be calculated is the maximum climb thrust because it is used as a reference for the other modes. Under ISA conditions with no temperature deviation the maximum climb thrust is:

$$(T_{max_{climb}})_{ISA} = C_{T_{C1}}\left(1 - \frac{h_e}{C_{T_{C2}}} + C_{T_{C3}}h_e^2\right),$$

where h_e is the altitude above sea level measured in feet, and $C_{T_{C1}}$, $C_{T_{C2}}$, $C_{T_{C3}}$ are BADA operation performance parameters.

The maximum climb thrust is corrected for temperature deviations in the following manner:

$$T_{max_{climb}} = (T_{max_{climb}})_{ISA}(1 - C_{T_{C5}}(\Delta T_{ISA})_{eff}),$$
where
$$(\Delta T_{ISA})_{eff} = \Delta T_{ISA} - C_{T_{C4}}$$
constrained to
$$0 \leq (\Delta T_{ISA})_{eff} C_{T_{C5}} \leq 0.4 \text{ and } C_{T_{C5}} \geq 0,$$

where ΔT_{ISA} measures the deviation from the standard atmosphere in Kelvin degress, and $C_{T_{C4}}$, $C_{T_{C5}}$ are BADA operation performance parameters.

BADA also provides thrust reference performance values for the different phases: For climb: $T = (T_{max_{climb}})$. For cruise, the thrust is derived from the dynamic equations in the case of a cruise at constant speed and constant altitude:

$$T = D + mg \sin \gamma.$$

The descent thrust is set by two coefficients depending on the altitude. If the aircraft is flying above h_{des} (set by BADA as an attribute of the aircraft) then the thrust is:

$$T = C_{des_{high}}(T_{max_{climb}}).$$

If the aircraft is flying below h_{des} (set by BADA as an attribute of the aircraft) then the thrust is:

$$T = C_{des_{low}}(T_{max_{climb}}),$$

which coincides with the thrust when the throttle is in the idle detent position.

Fuel consumption model

The aircraft losses weight while it has the engines running. To model this mass reduction BADA provides a consumption model in which the thrust specific fuel consumption, η, given in $[kg/(s \cdot N)]$, is specified as a function of the true airspeed for jets and truboprop engines. Herein, we will focus on jets. The fuel flow, $\dot{m}$, will be set by the thrust and the specific fuel consumption. The specific fuel consumption and the fuel flow are respectively:

$$\eta = \frac{C_{f1}}{60000}(1 + \frac{3600V}{1852C_{f2}}),$$
$$\dot{m} = -\eta T,$$

where C_{f4}, C_{f2} are BADA operation performance parameters.

The minimum fuel flow, $\dot{m}_{min}$, corresponding to idle thrust or descent conditions is specified in [kg/minute] as a function of altitude above sea level, h [ft]. Cruise fuel flow,

$\dot{m}_{cr}$ is calculated using the thrust specific fuel consumption, η, and a cruise fuel fuel flow factor, C_{fcr}:

$$\dot{m}_{min} = C_{f3}(1 - \tfrac{h_e}{C_{f4}}),$$
$$\dot{m}_{cr} = \eta T C_{fcr},$$

where C_{f3}, C_{f4}, C_{fcr} are BADA operation performance parameters. Notice that $\dot{m}_{min}$ is related to T_{min}.

Aerodynamic model

We give here a brief summary of the aerodynamic characteristics of an aircraft. We consider the aircraft flying in incompressible subsonic regime, so that the effects of Mach number, M, on the aerodynamics are dismissed. We also assume symmetric flight, so that the lateral aerodynamic force, Q, is also necgletable. This means there are two basic aerodynamic forces defined in the wind axes frame, lift, L, and drag, D. If we also assume that L and D do not depend on the deflexion of command surfaces, these aerodynamic forces depend on true airspeed, density of the atmosphere, and geometric parameters of the aircraft. Turning lift and drag non-dimensional by means of the dynamic pressure $(1/2\rho V^2)$ and the wing surface (S), the lift and drag coefficients can be calculated as follows:

$$C_L = \tfrac{2L}{\rho S V^2};$$
$$C_D = \tfrac{2D}{\rho S V^2}.$$

As will be shown in Section A.2.1, C_L can be modeled as a linear function of the angle of attack, α, except for high angles in which linearity is lost and aircraft stalls, reaching the maximum coefficient of lift, $C_{L_{max}}$.

Drag and lift are related by their coefficients in what is know as aircraft polar. C_D is basically parabolic with respect to C_L in incompressible regime, so it is commonly said that the aircraft has a parabolic drag polar, expressed as:

$$C_D = C_{D0} + C_{D_i} C_L^2, \tag{A.2}$$

where C_{D0} is the parasite coefficient, and C_{D_i} is the induced coefficient of the polar. Both are approximately constant in subsonic regime and will be obtained through BADA as operation performance parameters. Note that both C_{D_0} and C_{D_i} vary depending on the flap configuration. BADA provides operation performance parameters for the following five flap configurations: take-off (TO); Initial climbing (IC), cruise (CR), approach (AP), and landing (LD). Notice that BADA uses the notation C_{D2} (instead of C_{D_i}) for the induced coefficient of drag, and thus, the BADA operation performance parameters are

the following:

> $C_{D0,TO}$: parasite drag coefficient (take-off);
>
> $C_{D0,IC}$: parasite drag coefficient (initial climb);
>
> $C_{D0,CR}$: parasite drag coefficient (cruise);
>
> $C_{D0,AP}$: parasite drag coefficient (approach);
>
> $C_{D0,LD}$: parasite drag coefficient (landing);
>
> $C_{D2,TO}$: induced drag coefficient (take-off);
>
> $C_{D2,IC}$: induced drag coefficient (initial climb);
>
> $C_{D2,CR}$: induced drag coefficient (cruise);
>
> $C_{D2,AP}$: induced drag coefficient (approach);
>
> $C_{D2,LD}$: induced drag coefficient (landing);

Given that during landing phase the landing gear must be deployed, the parasite drag is increased. BADA also provides an operation performance parameter to consider this fact:

> $C_{D0,\triangle LDG}$: parasite drag coefficient (landing gear).

Therefore, during landing phase equation (A.2) must be considered as

$$C_D = C_{D0} + C_{D0,\triangle LDG} + C_{D_i} C_L^2,$$

where $C_{D0} + C_{D0,\triangle LDG}$ represents the total parasite drag coefficient.

A.2   BADA performance limitations

Since an aircraft has, among other characteristics, a fixed geometry, a limited strength of the materials used, a maximum thrust for every altitude, the performance is limited. BADA provides a flight envelope model [14, Chap. 3] and global aircraft parameters [14, Chap. 5] to take into account such performance limitations.

A.2.1   Flight envelope model

Maximum speeds

For aerodynamic and structural reasons there is an speed or Mach number that must not be exceeded. If these speeds are exceeded an structural failure may occur, what normally ends up in a major or fatal failure. Furthermore shock waves may appear in some locations around the aircraft what increases the drag and generate vibrations. Thus, the operating maximum speed are:

1. V_{MO}: This speed expressed in IAS[1] may not be deliberately exceeded in any flight phase.

2. M_{MO}: This Mach number may not be deliberately exceeded in any flight phase.

3. V_{FE}: This speed expressed in IAS is established so that it does not exceed the design flaps speed.

4. V_{LO}: This speed expressed in IAS may not exceed the speed at which it's safe to both extend or retract the landing gear.

5. V_{LE}: This speed expressed in IAS may not exceed the speed at which it's safe to fly with the landing gear down and locked.

Minimum speed (Stall)

As there are some maximum speeds there is also a minimum flying speed called stall speed. For aerodynamic reasons, the aircraft, for a certain aerodynamic configuration, can reach only a maximum value of lift coefficient depending on the maximum allowable angle of attack. If the aircraft exceeds the maximum angle of attack, the lift coefficient will drop heavily and the aircraft would be in stall. The relation between the angle of attack and the lift coefficient is shown in Figure A.1.
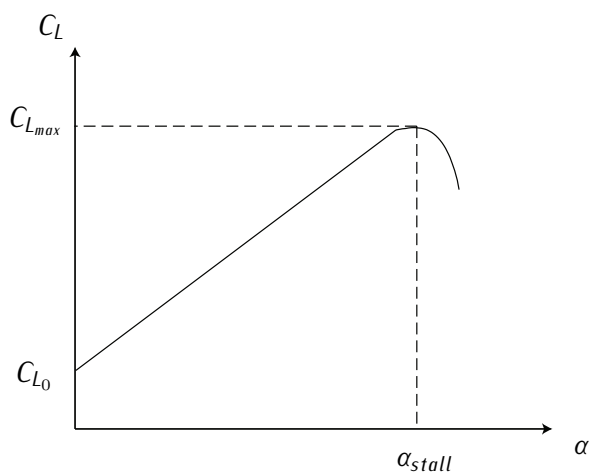

Figure A.1: Lift coefficient curve as a function of angle of attack.

---

[1] V_{IAS} Indicated airspeed (IAS): The direct cockpit instrument reading obtained from the airspeed indicator, uncorrected for variations in atmospheric density, instrument error or installation error. Manufacturers use this airspeed as the basis for determining aircraft performance. Normally the difference between IAS and CAS is very little.

As the lift coefficient expression is $C_L = 2L/\rho S V^2$, and defining the load factor, $n = L/(mg)$, then we can express the stall speed using the maximum lift coefficient as

$$V_S = \sqrt{\frac{2 \cdot n \cdot mg}{\rho S C_{L_{max}}}}. \tag{A.3}$$

It is important to note that the stall speed is a function of the load factor as the maximum lift coefficient is constant for a certain aerodynamic configuration. The load factor is normally equal to 1 but in some situations like in a turn it uses to be greater. Therefore the stall speed depends on how the aircraft is moving. The reference stall speed is when the aircraft is flying straight, with no turns. Below this speed the aircraft can not fly leveled at constant altitude.

The minimum speed for the aircraft is specified by BADA as follows:

$$V_{min} = C_{Vmin,TO} \cdot V_S, \text{ if in take off;} \tag{A.4}$$
$$V_{min} = C_{Vmin} \cdot V_S, \text{ otherwise;} \tag{A.5}$$

where $C_{Vmin,TO}$, C_{Vmin} are BADA global aircraft parameters (see Section A.2.2).

BADA defines stall speeds as aircraft performance parameters, $(V_{stall})_i$ given in knots, for the following five flap configurations: take–off (TO); initial climb (IC); cruise (CR); approach (AP); landing (LD):

$$(V_{stall})_{TO} \leq (V_{stall})_{IC} \leq (V_{stall})_{CR} \leq (V_{stall})_{AP} \leq (V_{stall})_{LD}.$$

Flight envelope

Structural limitations must be considered in order to establish a safe range for certain parameters that should not be exceeded in flight. One major parameter to take into consideration regarding structural stress is the load factor. This parameter has maximum (n_{max}) and minimum (n_{min}) values that correspond to the maximum strength of the aircraft structure. Also maximum and minimum speeds are to be considered. Figure A.2 illustrates it.

The closed area belongs to the flyable zone, where the aircraft structural stress and the speed are within limits. The boundary of stall is calculated using the reference stall velocity exposed in Section A.2.1.

Ceiling

The ceiling is the maximum altitude at which an aircraft can fly. The ceiling is defined in a rectilinear, steady, and symmetrical flight with no wind. At the ceiling the ratio of climb is zero, so the pitch angle of velocity is also zero.
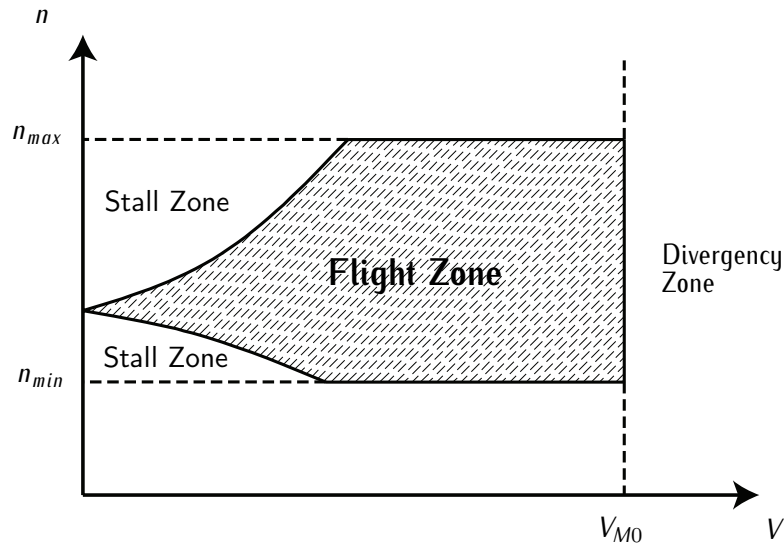
Figure A.2: Sketch of a flight envelope.

Using the dynamic equations: $T = D = \frac{1}{2}\rho S C_D V^2$, and remembering the relations of the drag coefficient with the speed for n=1:

$$C_D = C_{D_0} + C_{D_i}\frac{4(mg)^2}{\rho^2 S^2 V^2},$$

then the necessary thrust to fly under those conditions is:

$$T = \frac{1}{2}\rho S C_{D_0} V^2 + C_{D_i}\frac{2(mg)^2}{\rho S V^2}.$$

Figure A.3 shows the necessary thrust for an airbus A320 with a clean configuration at its reference mass. As can be derived from Figure A.3 a minimum thrust is required. This minimum can be calculated doing the derivative and making it equal to zero, $T_{min} = 2mg\sqrt{C_{D_0}C_{D_i}}$. The altitude at which the maximum available thrust is equal to the minimum one is the ceiling of the aircraft. Using the model described in Section A.1 for the thrust and with an ISA atmosphere, the maximum available thrust is:

$$T_{max} = C_{T_{C1}}(1 - \frac{h}{C_{T_{C2}}} + C_{T_{C3}}h^2) \cdot (1 + C_{T_{C4}}C_{T_{C5}}), \tag{A.6}$$

and after some calculations, herein omitted for the sake of brevity, the operational ceiling (in feet) is:

$$h_{max/act} = \frac{\frac{1}{C_{T_{C2}}} - \sqrt{\frac{1}{C_{T_{C2}}^2} - 4C_{T_{C3}}(1 - \frac{2mg\sqrt{C_{D_0}C_{D_i}}}{C_{T_{C1}}(1+C_{T_{C4}}C_{T_{C5}})})}}{2C_{T_{C3}}}.$$
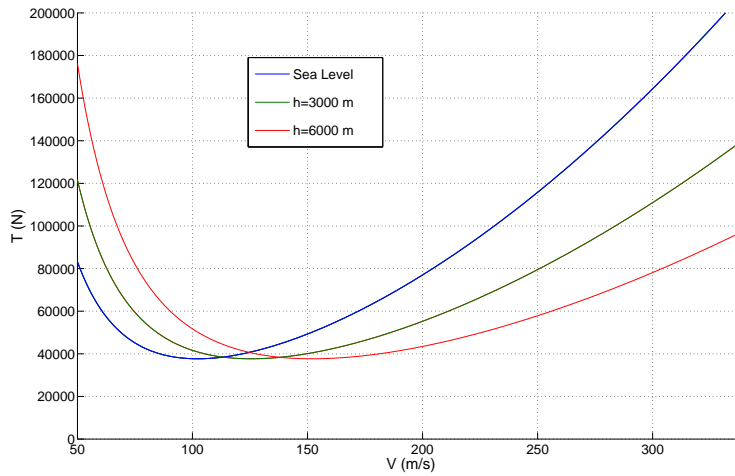
Figure A.3: A–320 thrust required at m_{ref}.

BADA defines the operational ceiling as:

$$h_{max/act} = h_{max} + G_t(\Delta T_{ISA} - C_{T_{C4}}) + G_W(m_{max} - m),$$

where

h_{max} : maximum altitude at MTOW under ISA conditions for maximum mass;

G_W : mass gradient on maximum altitude;

G_t : temperature gradient on maximum altitude;

with:

$G_W \geq 0; G_t \leq 0;$ if $(\Delta T_{ISA} - C_{T_{C4}}) < 0,$ then: $(\Delta T_{ISA} - C_{T_{C4}}) = 0,$

with m is the mass of the aircraft [kg]. Notice that h_{max} is provided by BADA as an operation performance parameter considering a residual 300 f.p.m. Rate Of Climb (ROC). G_w and G_t are also provided by BADA as operation performance parameters.

Finally, BADA defines the maximum altitude as:

$$h_u = MIN[h_{M0}, h_{max/act}],$$

where h_{M0} is the maximum operational height, provided by BADA as an operation performance parameter in feet above see level.

Figure A.4 shows the operational ceiling for an Airbus A320 in an ISA atmosphere. Remember that the ceiling is calculated at the minimum required thrust speed. If the maximum flyable altitude at another speed is to be known, then the equations are to be derived again for a generic speed.
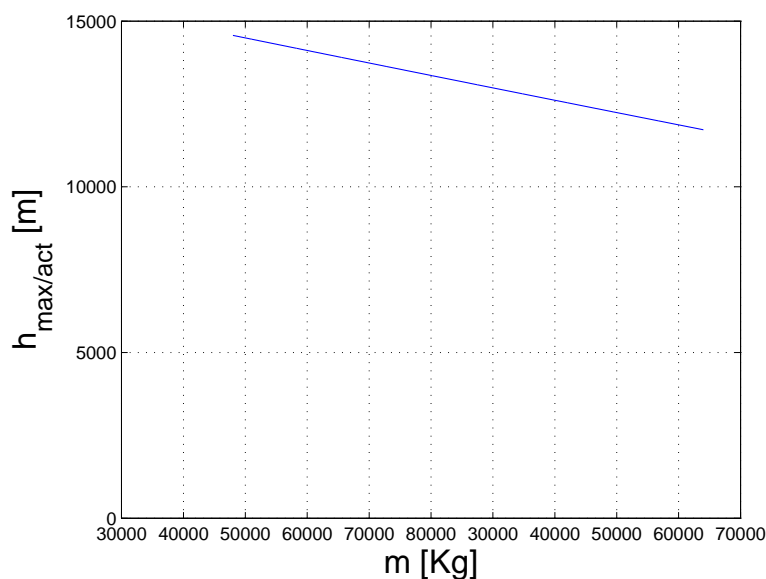
Figure A.4: Operational ceiling for an Airbus A320.

A.2.2   Global aircraft parameters

BADA also defines parameters that do not depend on the aircraft model, but remain valid for all aircrafts. These parameters are defined by BADA as global aircraft parameters. We will detail only those global aircraft parameters used in this thesis. For an in-depth insight on global aircraft parameters, the reader is referred to [14, Chap. 5].

Maximum acceleration

Maximum acceleration parameters are used to limit the increment in longitudinal true airspeed (longitudinal acceleration) and rates of climb or descent (normal acceleration). Among others, BADA defines the following maximum values:

$$a_{l,max\ (civ)} = 2\ [ft/s^2];$$
$$a_{n,max\ (civ)} = 5\ [ft/s^2];$$

where $a_{l,max\ (civ)}$ corresponds to the maximum is the maximum longitudinal acceleration for civil flights, and $a_{n,max\ (civ)}$ corresponds to the maximum normal acceleration for civil flights.

Bank angles

Nominal and maximum bank angles are defined for civil flights. Among others, BADA defines the following maximum values:

$$\mu_{max,civ\ (TO,LD)} = \pm 25°;$$

$$\mu_{max,civ\ (OTHERS)} = \pm 45°;$$

where $\mu_{max,civ\ (TO,LD)}$ corresponds to the maximum bank angles for civil flight during phases of take-off and landing, and $\mu_{max,civ\ (OTHERS)}$ corresponds to the maximum bank angles for civil flight during all other phases.

The limitations on the bank angle leads also to limit the rate of turn ($\dot{\chi}$) as a function of the bank angle: $\dot{\chi} = \frac{g}{V}\tan\mu$.

Configuration altitude threshold

BADA defines threshold altitudes for the following four flap configurations: take-off (TO), initial climb (IC), approach (AP), and landing (LD), so that:

$$H_{max,TO} = 400\ [ft];$$
$$H_{max,IC} = 2000\ [ft];$$
$$H_{max,AP} = 8000\ [ft];$$
$$H_{max,LD} = 3000\ [ft].$$

Notice that the selection of the different flap configurations can be done according to the threshold altitudes above defined, but also considering the different stall speeds.

Minimum speed coefficients

Two minimum speed coefficients are defined by BADA:

$$C_{Vmin,TO} = 1.2;$$
$$C_{Vmin} = 1.3.$$

Notice that such global parameters are to be used in equation (A.4) and equation (A.5).

A.3   A320 performance parameters

```
# Airbus A320-212              C_tdes_low = 0.009437;
# with CFM56_5_A3 engines      C_tdes_high = 0.031014;
# Block Mass (kg)              h_des_ft = 15000.0;
 m_ref = 64000.0;             C_tdes_app = 0.13;
 m_min = 39000.0;             C_tdes_ld = 0.34;
 m_max = 77000.0;             V_des_ref_kn = 310.0;
 m_pyld = 21500.0;            M_des_ref = 0.78;
# Block Aerodynamics           # Block Fuel Consumption
 S = 122.6;                   C_f1 = 0.94;
 CD0_CR = 0.024;              C_f2 = 100000.0;
 CD0_IC = 0.0242;             C_f3 = 8.89;
 CD0_TO = 0.0393;             C_f4 = 81926.0;
 CD0_AP = 0.0456;             C_fcr = 1.06;
 CD0_LD = 0.0838;              # Block Flight Envelope
 CD0_LDG = 0.031200;          h_MO_ft = 34354.0;
 CD2_CR = 0.0375;             h_max_ft = 39000.0;
 CD2_IC = 0.0469;             G_t = -130.0;
 CD2_TO = 0.0396;             G_w = 280.0;
 CD2_AP = 0.0381;              # Block airspeed limitations
 CD2_LD = 0.0371;             V_MO_kn = 350.0;
# Block Engine Thrust          M_MO = 0.82;
 C_tc1 = 136050.0;            (Vstall)_CR = 145.0; # Clean
 C_tc2 = 52238.0;             (Vstall)_IC = 120.0; # 1
 C_tc3 = 2.6637E-11;          (Vstall)_TO = 114.0; # 1+F
 C_tc4 = 10.29;               (Vstall)_AP = 107.0; # 2
 C_tc5 = 0.0058453;           (Vstall)_LD = 101.0; # FULL
```

A.4   A330 performance parameters

```
# Airbus A330-301                    C_tc2 = 54057.0;
# with CF6 80E1 A2 engines           C_tc3 = 1.6042E-11;
# Block Mass (kg)                    C_tc4 = 9.6284;
 m_ref = 174000.0;                   C_tc5 = 0.00804;
 m_min = 126000.0;                   C_tdes_low = 0.026815;
 m_max = 212000.0;                   C_tdes_high = 0.040128;
 m_pyld = 38000.0;                   h_des_ft = 15000.0;
 # Block Flight Envelope             C_tdes_app = 0.18;
  h_MO_ft = 37232.0;                 C_tdes_ld = 0.3;
 h_max_ft = 41000.0;                 V_des_ref_kn = 330.0;
 G_t = -0.224;                       M_des_ref = 0.82;
 G_w = 99.200005;                    # Block Fuel Consumption
# Block Aerodynamics                  C_f1 = 0.9224;
 S = 361.6;                           C_f2 = 1000000.0;
 CD0_CR = 0.0186;                     C_f3 = 21.059;
 CD0_IC = 0.0259;                     C_f4 = 108280.0;
 CD0_TO = 0.0344;                     C_fcr = 0.9184;
 CD0_AP = 0.0555;                    # Block airspeed limitations
 CD0_LD = 0.078;                      V_MO_kn = 330.0;
 CD2_CR = 0.0297;                     M_MO = 0.86;
 CD2_IC = 0.0404;                     (Vstall)_CR = 134.0; # Clean
 CD2_TO = 0.037;                      (Vstall)_IC = 120.0; # 1
 CD2_AP = 0.0325;                     (Vstall)_TO = 112.0; # 1+F
 CD2_LD = 0.0345;                     (Vstall)_AP = 105.0; # 2
# Block Engine Thrust                 (Vstall)_LD = 99.0; # FULL
 C_tc1 = 353620.0;
```

Index

Bibliography

[1] International Civil Aviation Organization (ICAO), "Procedures for air navigation services. Rules of the air and air traffic services," tech. rep., ICAO Doc 4444–RAC/501, 13th Edtion. 1996.
http://www.gcaa.com.gh/extweb/images/stories/ais/icaodoc4444.pdf
[Retrieved 01/08/2012].

[2] M. Gaudet, "Harmonization of aviation user charges in the North Atlantic airspace," Master's thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, 2008.

[3] D. Lee, D. Fahey, P. Forster, P. Newton, R. Wit, L. Lim, B. Owen, and R. Sausen, "Aviation and global climate change in the 21st century," *Atmospheric Environment*, vol. 43, no. 22–23, pp. 3520–3537, 2009.

[4] European Environment Agency (EEA), "Laying the foundations for greener transport. TERM 2011: transport indicators tracking progress towards environmental targets in Europe.," tec. rep., 2011.
http://www.eea.europa.eu/publications [Retrieved 01/09/2012].

[5] Eurocontrol, Performance Review Comission, "Performance review report an assessment of air traffic management in europe during the calendar year 2011: Eurocontrol PRR 2011," tech. rep., Eurocontrol, May 2012.
http://www.eurocontrol.int/prc [Retrieved 01/09/2012].

[6] J. Penner, "Aviation and the global atmosphere: a special report of IPCC working groups I and III in collaboration with the scientific assessment panel to the Montreal protocol on substances that deplete the ozone layer," tech. rep., International Panel of Climate Change (IPCC), 1999.

[7] Advisory Council for Aeronautics Research in Europe (ACARE), "Aeronautics and air transport: beyond vision 2020 (towards 2050)," June 2010. http://www.acare4europe.com [Retrieved 01/05/2011].

[8] SESAR Consortium, "Air transport framework: the current situation, SESAR definition phase milestone deliverable 1," tech. rep., SESAR Consortium, July 2006. http://www.eurocontrol.int/sesar/public [Retrieved 01/09/2009].

[9] SESAR Consortium, "SESAR master plan, SESAR definition phase milestone deliverable 5," tech. rep., SESAR Consortium, April 2008. http://www.eurocontrol.int/sesar/public [Retrieved 01/09/2009].

[10] P. Ky, "Brochure 2 Q4 2009," tech. rep., SESAR Joint Undertaking, http://www.sesarju.eu/news-press/documents-reports [Retrieved 01/02/2010].

[11] European Comission, the FAA and the project AIRE, "Atlantic Interoperability Initiative to Reduce Emission (AIRE). Summary results 2010/2011," tech. rep., SESAR Joint Undertaking, 2012. http://www.sesarju.eu/news-press/documents-reports [Retrieved 01/09/2012].

[12] SESAR Consortium, "The performance target, SESAR definition phase milestone deliverable 2," tech. rep., SESAR Consortium, December 2006. http://www.eurocontrol.int/sesar/public [Retrieved 01/09/2009].

[13] SESAR Consortium, "The ATM target concept, SESAR definition phase milestone deliverable 3," tech. rep., SESAR Consortium, September 2007. http://www.eurocontrol.int/sesar/public [Retrieved 01/09/2009].

[14] A. Nuic, *User Manual for the Base of Aircraft Data (BADA) Revision 3.6.* Eurocontrol Experimental Center, Bretigny, France, 2005. http://www.eurocontrol.int/eec/public [Retrieved 01/09/2009].

[15] S. Benjamin, D. Dévényi, S. Weygandt, K. Brundage, J. Brown, G. Grell, D. Kim, B. Schwartz, T. Smirnova, T. Smith, *et al.*, "An hourly assimilation–forecast cycle: the RUC," *Monthly Weather Review*, vol. 132, pp. 495–518, 2004.

[16] S. Benjamin, K. Brundage, P. Miller, T. Smith, G. Grell, D. Kim, J. Brown, T. Schlatter, and L. Morone, "The rapid update cycle at NMC," in *Preprints, 10th Conference on Numerical Weather Prediction, Portland, Oregon, USA, American Meteorological Society*, pp. 566–568, 1994.

[17] A. Brandstädt, J. Spinrad, and L. Van Bang, *Graph Classes: a Survey.* Monographs on Discrete Mathematics & Applications, Society for Industrial and Applied Mathematics, 1999.

[18] M. S. Branicky, V. S. Borkar, and S. K. Mitter, "A unified framework for hybrid control: model and optimal control theory," *IEEE Transactions on Automatic Control*, vol. 43, no. 1, pp. 31–45, 1998.

[19] X. Xu, *Analysis and design of Switched Systems*. PhD thesis, Department of Electrical Engineering. University of Notre Dame, Notre Dame, Indiana, USA, 2001.

[20] S. Altus, "Flight planning: the forgotten field in airline operations," in *Proceeding of 47th Airline Group of the International Federation of Operation Research Societies Annual Symposium*, Bangkok, Thailand, 2007.

[21] E. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[22] M. Bartholomew-Biggs, S. Parkhurst, and S. Wilson, "Global optimization approaches to an aircraft routing problem," *European Journal of Operational Research*, vol. 146, no. 2, pp. 417–431, 2003.

[23] M. Bartholomew-Biggs, S. Parkhurst, and S. Wilson, "Using DIRECT to solve an aircraft routing problem," *Computational Optimization and Applications*, vol. 21, no. 3, pp. 311–323, 2002.

[24] C. Hewitt and S. Broatch, "A tactical navigation and routing system for low-level flight," tech. rep., General Electric Company-Marconi Avionics, Rochester, Kent, UK, 1995.

[25] A. E. Bryson and Y. Ho, *Applied Optimal Control*. Wiley, 1975.

[26] J. T. Betts, *Practical Methods for Optimal Control and Estimation using Nonlinear Programming*. Advances in Design and Control. Society for Industrial and Applied Mathematics, 2010.

[27] L. Pontryagin, V. Boltyanskii, R. Gamkrelidze, and E. Mishchenko, *The Mathematical Theory of Optimal Processes*. Interscience Publishers, 1962.

[28] J. Nocedal and S. Wright, *Numerical Optimization*. Springer Verlag, 1999.

[29] H. Yan, F. Fahroo, and I. M. Ross, "Accuracy and optimality of direct transcription methods," *Advances in the Astronautical Sciences*, vol. 105, pp. 1613–1630, 2000.

[30] J. T. Betts, "Survey of numerical methods for trajectory optimization," *Journal of Guidance, Control, and Dynamics*, vol. 21, no. 2, pp. 193–207, 1998.

[31] T. Binder, W. Marquardt, L. Blank, W. Dahmen, H. Bock, M. Diehl, J. Schloeder, R. Burlisch, T. Kronseder, and O. Stryk, "Introduction to model based optimization of chemical processes in moving horizons," in *Online Optimization of Large Scale Systems* (J. R. M. Groetschel, S.O. Krumke, ed.), pp. 295–339, Springer Verlag, 2001.

[32] J. T. Betts and E. J. Cramer, "Application of direct transcription to commercial aircraft trajectory optimization," *Journal of Guidance, Control, and Dynamics*, vol. 18, no. 1, pp. 151–159, 1995.

[33] I. Ross and C. D'Souza, "Hybrid optimal control framework for mission planning," *Journal of Guidance, Control, and Dynamics*, vol. 28, no. 4, pp. 686–697, 2005.

[34] S. Sager, *Numerical Methods for Mixed-Integer Optimal Control Problems.* PhD thesis, Interdisciplinary Center for Scientific Computing, Universität Heidelberg, Heidelberg, Germany, 2006.

[35] C. R. Hargraves and S. W. Paris, "Direct trajectory optimization using nonlinear programming and collocation," *Journal of Guidance, Control, and Dynamics*, vol. 10, no. 4, pp. 338–342, 1987.

[36] A. L. Herman and B. A. Conway, "Direct optimization using collocation based on high-order Gauss-Lobatto quadrature rules," *Journal of Guidance, Control, and Dynamics*, vol. 19, no. 3, pp. 592–599, 1996.

[37] F. Fahroo and I. Ross, "Direct trajectory optimization by a Chebyshev pseudospectral method," *Journal of Guidance, Control, and Dynamics*, vol. 25, no. 1, pp. 160–166, 2002.

[38] D. Benson, *A Gauss Pseudospectral Transcription for Optimal Control.* PhD thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, 2004.

[39] M. Bazaraa, H. Sherali, and C. Shetty, *Nonlinear Programming: Theory and Algorithms.* Wilie-Interscience, 2006.

[40] P. Bonami, M. Kilinç, and J. Linderoth, "Algorithms and software for convex mixed integer nonlinear programs," in *Mixed Integer Nonlinear Programming* (J. Lee and S. Leyffer, eds.), vol. 154 of *the IMA Volumes in Mathematics and its Applications*, pp. 1–39, Springer, 2012.

[41] P. Bonami, L. T. Biegler, A. R. Conn, G. Cornuéjols, I. E. Grossmann, C. D. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, and A. Wächter, "An algorithmic framework for convex mixed integer nonlinear programs," *Discrete Optimization*, vol. 5, no. 2, pp. 186–204, 2008.

[42] A. Wächter and L. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, 2006.

[43] T. R. Jorris and R. G. Cobb, "Multiple method 2D trajectory optimization satisfying waypoints and no-fly zone constraints," *Journal of Guidance, Control, and Dynamics*, vol. 31, no. 3, pp. 543–553, 2008.

[44] T. R. Jorris and R. G. Cobb, "Three-dimensional trajectory optimization satisfying waypoints and no-fly constraints," *Journal of Guidance, Control, and Dynamics*, vol. 32, no. 2, pp. 551–572, 2009.

[45] W. Roh and Y. Kim, "Trajectory optimization for a multi-stage launch vehicle using time finite element and direct collocation methods," *Engineering Optimization*, vol. 34, no. 1, pp. 15–32, 2002.

[46] I. Ross and F. Fahroo, "Pseudospectral knotting methods for solving optimal control problems," *Journal of Guidance, Control, and Dynamics*, vol. 27, no. 3, pp. 397–405, 2004.

[47] G. Huntington and A. Rao, "Optimal configuration of spacecraft formations via a Gauss pseudospectral method," *Advances in the Astronautical Sciences*, vol. 120, pp. 33–50, 2005.

[48] M. Norsell, "Multistage trajectory optimization with radar-range constraints," *Journal of Aircraft*, vol. 42, no. 4, pp. 849–857, 2005.

[49] M. Jacobsen and U. Ringertz, "Airspace constraints in aircraft emission trajectory optimization," *Journal of Aircraft*, vol. 47, no. 4, pp. 1256–1265, 2010.

[50] D. Garg, M. Patterson, W. Hager, A. Rao, D. Benson, and G. Huntington, "A unified framework for the numerical solution of optimal control problems using pseudospectral methods," *Automatica*, vol. 46, no. 11, pp. 1843–1851, 2010.

[51] I. Ross and M. Karpenko, "A review of pseudospectral optimal control: from theory to flight," *Annual Reviews in Control*, vol. 36, no. 2, pp. 182–197, 2012.

[52] B. Conway, C. Chilan, and B. Wall, "Evolutionary principles applied to mission planning problems," *Celestial Mechanics and Dynamical Astronomy*, vol. 97, no. 2, pp. 73–86, 2007.

[53] C. Chilan and B. Conway, "Using genetic algorithms for the construction of a space mission automaton," in *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 2316–2323, Trondheim, Norway, 2009.

[54] M. Kamgarpour, M. Soler, C. Tomlin, A. Olivares, and J. Lygeros, "Hybrid optimal control for aircraft trajectory design with a variable sequence of modes," in *Proceedings of the 18th International Federation of Automatic Control World Congress*, vol. 18, pp. 7238–7243, Milano, Italy, 2011.

[55] S. Khardi, L. Abdallah, O. Konovalova, and M. Houacine, "Optimal approach minimizing aircraft noise and fuel consumption," *Acta Acustica United with Acustica*, vol. 96, no. 1, pp. 68–75, 2010.

[56] J. Robinson and M. Kamgarpour, "Benefits of continuous descent operations in high-density terminal airspace under scheduling constraints," in *Procedings of the 10th AIAA Aviation Technology, Integration, and Operations Conference*, pp. 13–15, Fort Worth, Texas, USA, 2010.

[57] Y. Cao, T. Kotegawa, and J. Post, "Evaluation of continuous descent approach as a standard terminal airspace operation," in *Proceedings of the 9th USA/Europe Air Traffic Management Research and Development Seminar*, Berlin, Germany, 2011.

[58] R. Coppenbarger, R. Mead, and D. Sweet, "Field evaluation of the tailored arrivals concept for datalink-enabled continuous descent approach," *Journal of Aircraft*, vol. 46, no. 4, pp. 1200–1209, 2010.

[59] K. Sprong, K. Klein, C. Shiotsuki, J. Arrighi, and S. Liu, "Analysis of AIRE continuous descent arrival operations at Atlanta and Miami," in *Proceedings of the 27th IEEE/AIAA Digital Avionics Systems Conference*, St. Paul, Minnesota, USA, 2008.

[60] J. P. B. Clarke, N. T. Ho, L. Ren, J. A. Brown, K. R. Elmer, K.-O. Tong, and J. K. Wat, "Continuous descent approach: Design and flight test for Louisville International Airport," *Journal of Aircraft*, vol. 41, no. 5, pp. 1054–1066, 2004.

[61] X. Prats, V. Puig, and J. Quevedo, "Equitable aircraft noise-abatement departure procedures," *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 1, pp. 192–203, 2011.

[62] X. Prats, V. Puig, and J. Quevedo, "A multi-objective optimization strategy for designing aircraft noise abatement procedures. Case study at Girona Airport," *Transportation Research Part D: Transport and Environment*, vol. 16, no. 1, pp. 31–41, 2011.

[63] X. Prats, *Contributions to the Optimisation of Aircraft Noise Abatement Procedures*. PhD thesis, Universitat Politècnica de Catalunya, Castelldefels, Barcelona, Spain, 2010.

[64] M. Houacine and S. Khardi, "Gauss pseudospectral method for less noise and fuel consumption from aircraft operations," *Journal of Aircraft*, vol. 47, no. 6, pp. 2152–2159, 2010.

[65] R. Torres, J. Chaptal, C. Bes, and J. Hiriart-Urruty, "Optimal, environmentally friendly departure procedures for civil aircraft," *Journal of Aircraft*, vol. 48, no. 1, pp. 11–22, 2011.

[66] D. M. Pargett and M. D. Ardema, "Flight path optimization at constant altitude," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 4, p. 1197, 2007.

[67] A. Franco, D. Rivas, and A. Valenzuela, "Minimum-fuel cruise at constant altitude with fixed arrival time," *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 1, pp. 280–285, 2010.

[68] A. Franco and D. Rivas, "Minimum-cost cruise at constant altitude of commercial aircraft including wind effects," *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 4, pp. 1253–1260, 2011.

[69] L. Delgado and X. Prats, "En route speed reduction concept for absorbing air traffic flow management delays," *Journal of Aircraft*, vol. 49, no. 1, pp. 214–224, 2012.

[70] M. Kaiser, J. Rosenov, H. Fricke, and M. Schultz, "Tradeoff between optimum altitude and contrail layer to ensure maximum ecological en-route performance using the enhanced trajectory prediction model (ETPM)," in *Proceedings of the 2nd International Conference on Application and Theory of Automation in Command and Control Systems* (E. García, C. Johnson, W. Y. Ochieng, P. Palanque, F. J. Saéz, M. A. Vilaplana, and M. Winckler, eds.), pp. 127–134, London, England, IRIT Press, 2012.

[71] M. Soler, A. Olivares, and E. Staffetti, "Hybrid optimal control approach to commercial aircraft trajectory optimization," *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 3, pp. 985–991, 2010.

[72] M. Soler, D. Zapata, A. Olivares, and E. Staffetti, "Framework for aircraft 4D trajectory planning towards an efficient air traffic management," *Journal of Aircraft*, vol. 49, no. 1, pp. 341–348, 2012.

[73] M. Soler, D. Zapata, A. Olivares, E. Staffetti, and J. Cegarra, "Comparative analysis of commercial aircraft trajectory performance," in *Proceedings of the 2nd International Conference on Engineering and Optimization* (H. Rodrigues, J. Herskovits, *et al.*, eds.), Paper no. 1274, Lisbon, Portugal, 2010.

[74] M. Soler, A. Olivares, and E. Staffetti, "Hybrid optimal control approach to commercial aircrafts 3D multiphase trajectory optimization," in *Proceedings of AIAA Guidance, Navigation, and Control Conference*, Paper no. 2010-8453, Toronto, Canada, 2010.

[75] M. Soler, A. Olivares, E. Staffetti, and J. Cegarra, "Multiphase optimal control applied to 4D business trajectory strategic planning in air traffic management," in *Proceedings of 1st International Conference on Application and Theory of Automation in Command and Control Systems* (E. García, A. Majumdar, P. Palanque, A. Pasquini, F. J. Saéz, and M. Winckler, eds.), pp. 68–78, Barcelona, Spain, IRIT Press, 2011.

[76] P. Bonami, A. Olivares, M.Soler, and E. Staffetti, "Multiphase mixed-integer optimal control approach to aircraft trajectory optimization," *Journal of Guidance, Control, and Dynamics*, 2013. Accepted for publication.

[77] M. Soler, A. Olivares, E. Staffetti, and P. Bonami, "En-route optimal flight planning constrained to pass through waypoints using MINLP," in *Proceedings of the 9th USA/Europe Air Traffic Management Research and Development Seminar*, Paper no. 90, Berlin, Germany, 2011.

[78] M. Soler, M. Kamgarpour, C. Tomlin, and E. Staffetti, "Multiphase mixed-integer optimal control framework to aircraft conflict avoidance," in *Proceedings of the 51st IEEE Conference on Decision and Control*, pp. 1740–1745, Maui, Hawaii, USA, Omnipress, 2012.

[79] R. Bellman, *Dynamic Programming*. Princeton UP, 1957.

[80] W. Karush, *Minima of Functions of Several Variables with Inequalities as Side Constraints*. PhD thesis, Department of Mathematics, University of Chicago, Chicago, Illinois, USA, 1939.

[81] H. Kuhn and A. Tucker, "Nonlinear Programming," in *Proceedings of the 2nd Berkeley Symposium on Mathmatics, Statistics and Probability* (U. of California Press, ed.), pp. 481–492, Berkeley, California, USA, 1951.

[82] D. P. Bertsekas, *Dynamic Programming and Optimal Control, Vol. I, II*. Athena Scientific, 2007.

[83] H. Kelley, "Gradient theory of optimal flight paths," *AIAA Journal*, vol. 30, no. 10, pp. 947–954, 1960.

[84] U. Ascher, R. Mattheij, and R. Russell, *Numerical solution of boundary value Problems for Ordinary Differential Equations*. Society for Industrial and Applied Mathematics, 1995.

[85] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*. Springer Verlag, 2002.

[86] U. Ascher, J. Christiansen, and R. Russell, "A collocation solver for mixed order systems of boundary value problems," *Mathematics of Computation*, vol. 33, no. 146, pp. 659–679, 1979.

[87] O. von Stryk and R. Bulirsch, "Direct and indirect methods for trajectory optimization," *Annals of Operations Research*, vol. 37, no. 1, pp. 357–373, 1992.

[88] P. Gill, W. Murray, and M. Wright, *Practical Optimization*. Academic Press, 1981.

[89] D. Kraft, "On converting optimal control problems into nonlinear programming problems," *Computational Mathematical Programming*, vol. 15, pp. 261–280, 1985.

[90] H. G. Bock and K. J. Plitt, "A multiple shooting algorithm for direct solution of optimal control problems," in *In Proceedings of the 9th International Federation of Automatic Control World Congress*, pp. 242–247, Budapest, Hungary, Pergamon Press, 1984.

[91] T. Tsang, D. Himmelblau, and T. Edgar, "Optimal control via collocation and non-linear programming," *International Journal of Control*, vol. 21, no. 5, pp. 763–768, 1975.

[92] P. Williams, "Hermite-Legendre-Gauss-Lobatto direct transcription methods in trajectory optimization," *Advances in the Astronautical Sciences*, vol. 120, Part I, pp. 465–484, 2005.

[93] O. von Stryk, "Numerical solution of optimal control problems by direct collocation," in *Optimal Control - Calculus of Variations, Optimal Control Theory, and Numerical Methods* (R. Bulirsch, A. Miele, and J. Stoer, eds.), vol. 111 of *International Series of Numerical Mathematics*, pp. 129–143, Birkhäuser, 1993.

[94] C. De Boor and B. Swartz, "Collocation at Gaussian points," *SIAM Journal on Numerical Analysis*, vol. 10, no. 4, pp. 582–606, 1973.

[95] J. Villadsen and W. Stewart, "Solution of boundary-value problems by orthogonal collocation," *Chemical Engineering Science*, vol. 22, no. 11, pp. 1483–1501, 1967.

[96] P. Williams, "A comparison of differentiation and integration based direct transcription methods," in *Proceedings of the AAS/AIAA Space Flight Mechanics Conference, Paper AAS 05-128*, Copper Mountain, Colorado, USA, 2005.

[97] A. Herman, *Improved Collocation Methods with Application to Direct Trajectory Optimization*. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, 1995.

[98] P. Enright, *Optimal Finite-Thrust Spacecraft Trajectories using Direct Transcription and Nonlinear Programming*. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, 1991.

[99] M. Branicky, *Studies in Hybrid Systems: Modeling, Analysis, and Control*. PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, 1995.

[100] X. Xu and P. J. Antsaklis, "Optimal control of switched systems based on parameterization of the switching instants," *IEEE Transactions on Automatic Control*, vol. 49, no. 1, pp. 2–16, 2004.

[101] M. Žefran, *Continuous Methods for Motion Planning*. PhD thesis, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, Pennsylvania, USA, 1996.

[102] S. Sager, "Reformulations and algorithms for the optimization of switching decisions in nonlinear optimal control," *Journal of Process Control*, vol. 19, no. 8, pp. 1238–1247, 2009.

[103] S. Sager, G. Reinelt, and H. Bock, "Direct methods with maximal lower bound for mixed-integer optimal control problems," *Mathematical Programming*, vol. 118, no. 1, pp. 109–149, 2009.

[104] D. G. Hull, *Fundamentals of Aiplane Flight Mechanics*. Springer, 2007.

[105] D. Kleinbaum, L. Kupper, and K. Muller, *Applied Regression Analysis and other Multivariable Methods*. Duxbury Press, 2007.

[106] A. H. Land and A. G. Doig, "An automatic method for solving discrete programming problems," *Econometrica*, vol. 28, pp. 497–520, 1960.

[107] R. J. Dakin, "A tree search algorithm for mixed programming problems," *Computer Journal*, vol. 8, pp. 250–255, 1965.

[108] R. Bixby, M. Fenelon, Z. Gu, E. Rothberg, and R. Wunderling, *The Sharpest Cut*, ch. Mixed-Integer Programming: A Progress Report, pp. 309–325. Series on Optimization, Society for Industrial and Applied Mathematics, 2004.

[109] N. V. Sahinidis, "BARON: A general purpose global optimization software package," *Journal of Global Optimization*, vol. 8, pp. 201–205, 1996.

[110] M. Tawarmalani and N. V. Sahinidis, *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications*. Kluwer Academic Publishers, 2002.

[111] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Waechter, "Branching and bounds tightening techniques for non-convex MINLP," *Optimization Methods and Software*, vol. 24, no. 4, pp. 597–634, 2009.

[112] S. Leyffer, "User manual for MINLP-BB," *University of Dundee, Numerical Analysis Report*, vol. 234, 1999.

[113] M. Bussieck and A. Drud, "SBB: a new solver for mixed integer nonlinear programming," *GAMS Development Corp*, 2001.

[114] EUROCONTROL, "Customer guide to charges version 5.0," February 2011, http://www.eurocontrol.int/crco [Retrieved 01/06/2012].

[115] NAV Canada, "Customer guide to charges," September 2008, http://www.navcanada.ca [Retrieved 01/06/2012].

[116] United Kingdom Civil Aviation Authority, "The civil aviation authority (navigation services charges), specification 2012." Official Record Series 7, March 2012, http://www.caa.co.uk/ors7 [Retrieved 01/06/2012].

[117] A. Olivares, M. Soler, and E. Staffetti, "Multiphase mixed-integer optimal control applied to 4D trajectory planning in air traffic management," in *Proceedings of the 3rd International Conference on Application and Theory of Automation in Command and Control Systems*, Naples, Italy, 2013. Submittted.

Page left intentionally blank

Lo siento, pero no quiero ser emperador. Eso no me va. No quiero gobernar o conquistar a nadie. Me gustaría ayudar a todo el mundo, si fuera posible: a judíos y gentiles; a negros y blancos. Todos queremos ayudarnos mutuamente. Los seres humanos son así. Queremos vivir para la felicidad y no para la miseria ajena. No queremos odiarnos y despreciarnos mutuamente. En este mundo hay sitio para todos. Y la buena tierra es rica y puede proveer a todos.

El camino de la vida puede ser libre y bello; pero hemos perdido ese camino. La avaricia ha envenenado las almas de los hombres, ha levantado en el mundo barricadas de odio, nos ha llevado a la miseria y a la matanza. Hemos aumentado la velocidad. Pero nos hemos encerrado nosotros mismos dentro de ella. La maquinaria, que proporciona abundancia, nos ha dejado en la indigencia. Nuestra ciencia nos ha hecho cínicos; nuestra inteligencia, duros y faltos de sentimientos. Pensamos demasiado y sentimos demasiado poco. Más que maquinaria, necesitamos humanidad. Más que inteligencia, necesitamos amabilidad y cortesía. Sin estas cualidades, la vida será violenta y todo se perderá.

El avión y la radio nos han aproximado más. La verdadera naturaleza de estos adelantos clama por la bondad en el hombre, por la fraternidad universal, por la unidad de todos nosotros. Mi voz está llegando a millones de seres de todo el mundo, a millones de hombres, mujeres y niños desesperados, víctimas de un sistema que tortura y encarcela a personas inocentes. A aquellos que puedan oírme, les digo: No desesperéis.

La desgracia que nos ha caído encima no es más que el paso de la avaricia, la amargura de los hombres, que temen el camino del progreso humano. El odio de los hombres pasará, y los dictadores morirán, y el poder que arrebataron al pueblo volverá al pueblo. Y mientras los hombres mueren, la libertad no perecerá jamás.

¡Soldados! ¡No os entreguéis a esos bestias, que os desprecian, que os esclavizan, que gobiernan vuestras vidas; diciéndoos qué hacer, qué pensar o qué sentir! Que os obligan ha hacer la instrucción, que os mal alimentan, que os tratan como a ganado y os utilizan como carne de cañón. ¡No os entreguéis a esos hombres desnaturalizados, a esos hombres-máquina con inteligencia y corazones de máquina! ¡Vosotros no sois máquinas! ¡Sois hombres! ¡Con el amor de la humanidad en vuestros corazones! ¡No odiéis! ¡Sólo aquellos que no son amados odian, los que no son amados y los desnaturalizados!

¡Soldados! ¡No luchéis por la esclavitud! ¡Luchad por la libertad!

En el capítulo diecisiete de san Lucas está escrito que el reino de Dios se halla dentro del hombre, ¡no de un hombre o de un grupo de hombres, sino de todos los hombres! ¡En vosotros! Vosotros, el pueblo, tenéis el poder, el poder de crear máquinas. ¡El poder de crear felicidad! Vosotros, el pueblo, tenéis el poder de hacer que esta vida sea libre y bella, de hacer de esta vida una maravillosa aventura. Por tanto, en nombre de la democracia, empleemos ese poder, unámonos todos. Lucharemos por un mundo nuevo, por un mundo digno, que dará a los hombres la posibilidad de trabajar, que dará a la juventud un futuro y a los ancianos seguridad.

Prometiéndoos todo esto, las bestias han subido al poder. Pero mienten. No han cumplido esa promesa. ¡Ni la cumplirán! Los dictadores se dan libertad a sí mismos, pero esclavizan al pueblo. Ahora, unámonos para liberar el mundo, para terminar con las barreras nacionales, para terminar con la codicia, con el odio y con la intolerancia. Luchemos por un mundo de la razón, un mundo en el que la ciencia y el progreso lleven la felicidad a todos nosotros. ¡Soldados, en nombre de la democracia, unámonos!

Hannah, ¿puedes oírme? Dondequiera que estés, alza los ojos. ¡Mira, Hannah! ¡Las nubes están desapareciendo! El sol se está abriendo paso a través de ellas. Estamos saliendo de la oscuridad y penetrando en la luz.¡Estamos entrando en un mundo nuevo, un mundo más amable, donde los hombres se elevarán sobre su avaricia, su odio y su brutalidad! ¡Mira, Hannah! ¡Han dado alas al alma del hombre y, por fin, empieza a volar! ¡Vuela hacia el arco iris, hacia la luz de la esperanza! ¡Alza los ojos, Hannah! ¡Alza los ojos!

Alza los ojos, Hanna.
"El Gran Dictador", escrita y dirigida por Charles Chaplin en 1940.