University of Cincinnati

Date: 3/9/2022

I, Alexander L Von Moll, hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in Electrical Engineering.

It is entitled:

Skirmish-Level Tactics via Game-Theoretic Analysis

Student's name: **Alexander L Von Moll**

This work and its defense approved by:

Committee chair: Zachariah Fuchs, Ph.D.

Committee member: David Casbeer, Ph.D.

Committee member: John Gallagher, Ph.D.

Committee member: Meir Pachter, Ph.D.

Committee member: Dieter Vanderelst, Ph.D.

41677

Skirmish-Level Tactics via Game-Theoretic Analysis

Alexander L Von Moll

9 March 2022

Previous Degrees:

B.S. in Aeronautical & Astronautical Engineering, The Ohio State University, 2012

M.S. in Aerospace Engineering, Georgia Institute of Technology, 2016

Degree: Doctor of Philosophy (Electrical Engineering)
Department: Electrical Engineering and Computer Science
College: College of Engineering and Applied Science
Committee:
    Zachariah Fuchs (Chair)
    David Casbeer
    Meir Pachter
    Dieter Vanderelst
    John Gallagher

ABSTRACT

Supremacy in armed conflict comes not merely from superiority in capability or numbers but from how assets are used, down to the maneuvers of individual vehicles and munitions. This document outlines a research plan focused on skirmish-level tactics to militarily relevant scenarios. Skirmish-level refers to both the size of the adversarial engagement – generally one vs. one, two vs. one, and/or one vs. two – as well as the fact that the goal or objective of each team is well-established. The problem areas include pursuit-evasion and target guarding, either of which may be considered as sub-problems within military missions such as air-to-air combat, suppression/defense of ground-based assets, etc. In most cases, the tactics considered are comprised of the control policy of the agents (i.e., their spatial maneuvers), but may also include role assignment (e.g, whether to act as a decoy or striker) as well as discrete decisions (e.g., whether to engage or retreat). Skirmish-level tactics are important because they can provide insight into how to approach larger scale conflicts (many vs. many, many objectives, many decisions). Machine learning approaches such as reinforcement learning and neural networks have been demonstrated to be capable of developing controllers for large teams of agents. However, the performance of these controllers compared to the optimal (or equilibrium) policies is generally unknown. Differential Game Theory provides the means to obtain a rigorous solution to relevant scenarios in the form of saddle-point equilibrium control policies and the min max (or max min) cost / reward in the case of zero-sum games. When the equilibrium control policies can be obtained analytically, they are suitable for onboard / real-time implementation. Some challenges associated with the classical Differential Game Theory approach are explored herein. These challenges arise mainly due to the presence of singularities, which may appear in even the simplest differential games. The utility of skirmish-level solutions is demonstrated in (i) the multiple pursuer, single evader

differential games, (ii) multi-agent turret defense scenarios, and (iii) engage or retreat scenarios. In its culmination, this work contributes differential game and optimal control solutions to novel scenarios, numerical techniques for computing singular surfaces, approximations for computationally-intensive solutions, and techniques for addressing scenarios with multiple stages or outcomes.

A man can only stumble for so long before he either falls or stands up straight.

— Brandon Sanderson, Mistborn Trilogy

ACKNOWLEDGMENTS

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

1v1   one-versus-one

1v2   one-versus-two

2v1   two-versus-one

AC   Apollonius circle

AI   Artificial Intelligence

BVP   Boundary Value Problem

BVR   Beyond Visual Range

DP   Dynamic Programming

DDP   Differential Dynamic Programming

DGT   Differential Game Theory

DS   Dispersal Surface

FIS   Fuzzy Inference System

FNE   Feedback Nash Equilibrium

FSE	Feedback Stackelberg Equilibrium
G2C	Game of Two Cars
GA	Genetic Algorithm
GoE	Game of Engagement
GP	Gaussian Process
GSE	Global Stackelberg Equilibrium
HCDG	Homicidal Chauffeur Differential Game
HJB	Hamilton-Jacobi-Bellman
HJI	Hamilton-Jacobi-Isaacs
iLQR	iterative LQR
IRDG	Isotropic Rocket Differential Game
KDE	Kernel Density Estimator
LOS	Line of Sight
LQR	Linear Quadratic Regulator
MDP	Markov Decision Process
Mv1	many-versus-one
MvN	many-versus-multi
NLP	nonlinear program
NN	Neural Network
OCR	Optimal Constrained Retreat
PDE	partial differential equation
PE	Pure Evasion
PID	proportional-integral-derivative
PMP	Pontryagin's Maximum Principle
PN	Proportional Navigation
POMDP	Partially Observable Markov Decision Process
PP	Pure Pursuit
RL	Reinforcement Learning
SPDG	Suicidal Pedestrian Differential Game

TAD   Target-Attacker-Defender

TADDG   Target-Attacker-Defender Differential Game

TDDG   Turret Defense Differential Game

TEoR   Turret Engage or Retreat

TPBVP   two-point boundary value problem

TRPDG   Turret-Runner-Penetrator Differential Game

The road must be trod, but it will be very hard. And neither strength nor wisdom will carry us far upon it. This quest may be attempted by the weak with as much hope as the strong. Yet it is oft the course of deeds that move the wheels of the world: Small hands do them because they must, while the eyes of the great are elsewhere.

— J.R.R. Tolkien, The Fellowship of the Ring

To Christina, for choosing this quest with me.

Part I

INITIAL CONDITIONS

*Of the deluge of material written on game theory,
most concerns general theorems and results, often
of the highest calibre mathematics but very little of
usable techniques for obtaining practical answers.*

— Rufus Isaacs, Differential Games [130]

1

INTRODUCTION

Strategy in adversarial conflict is broad in scope, ranging from military doctrine to behaviors and adaptations found in natural organisms. Within the myriad of applications (e.g., military, biological, economic, games, sports) one may conceive of strategies or tactics at many levels – from the organizational level (i.e., cooperation among groups of entities) down to the individual level (the decisions of a pilot, an animal, an investor, etc.). Then there is always the challenge of the unpredictable actions and response of the adversary. One may go to great lengths to understand the setting or environment for the conflict, the capability of the adversary, and even the current state of the adversary. However, in the end, the outcome of the conflict generally depends on what the adversary *will do*, both now and in the future.

This work is concerned with strategies and tactics at the skirmish level with an emphasis on militarily relevant scenarios (though, most of the reality of the end application will be abstracted out of the analysis). The term "skirmish" bears no precise mathematical definition but is used throughout to denote a conflict between small teams of agents, including one-versus-one (1v1) conflicts. Most of the analysis is centered on 1v1, two-versus-one (2v1), one-versus-two (1v2), and many-versus-one (Mv1) conflicts. In addition to small team sizes, skirmish is meant to connote that the scenario can be boiled down to a specific objective for each team. This is in contrast with conflicts which are larger

and broader in scope, such as a military mission which may be comprised of many sub-tasks or sub-goals, or even a military campaign comprised of many missions across a wide variety of assets.

The two main problem areas of interest are pursuit-evasion and target guarding. In the former, a Pursuer (or group of Pursuers) seeks to capture an Evader (or group of Evaders). In the latter, a Defender seeks to prevent an Attacker from reaching a target set of states, which could represent a region of interest or position of mobile agents. Note, throughout the remainder of the document, the convention of capitalizing agent names is adhered to. For both of these problem areas, the primary concern is over the spatial maneuvers of the agents. Fig. 1.1 shows the notional relationship between various layers of autonomous (or automatic) control within an agent, e.g., aircraft. As noted in the figure, the research proposed in this document pertains mostly to the top two layers: the Mission Manager and the Outer Loop. For example, the analyses and algorithms proposed herein may prescribe heading reference commands for an aircraft which are based off of, or predicated upon, "higher-level" decisions such as which task the aircraft should complete. It is assumed that a "lower-level" (Inner Loop) controller exists which sends actuator commands to the aircraft in order to track the reference commands provided by the Outer Loop. In turn, the Inner Loop provides vehicle state information (such as position) to the Outer Loop. The Outer Loop provides an estimate of performance for the assigned task which is used by the Mission Manager to assess whether to make an update, e.g., reassign tasks or re-form teams. Possible functions of the Mission Manager, in the context of pursuit-evasion and target guarding, may include assigning a role to the vehicle or assigning tasks, e.g., a particular Evader for a Pursuer to target. Other possible roles are to act as a decoy in order to draw attention away from a teammate, to act as a striker and aim at a target, or a blocker in order to prevent cut off certain paths for the adversary, etc. Additionally, the Mission Manager may be responsible for discrete decisions such as whether an Attacker should engage a Defender or retreat altogether.

Larger, more complex, conflicts may involve larger teams of agents (i.e., many versus many), many different objectives, and many differ-
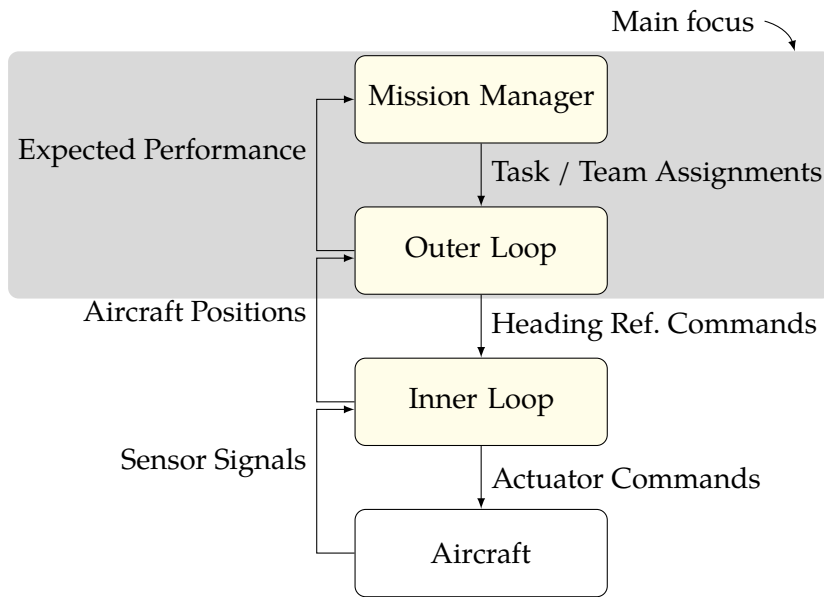
Figure 1.1: Notional block diagram describing the layers of control for, e.g., an autonomous aircraft.

ent decisions. Moreover, the structure of the conflict may change over time. Fig. 1.2 shows an abstract example of such a conflict taking place between 5 mobile agents, comprising the Blue team, and 5 mobile agents with 3 turrets comprising the Red team. As humans, we have a better chance of understanding larger engagements if we can understand and reason about what is optimal in each of the constituent skirmish-level conflicts, which are, by comparison, "atomic".

In the following parts, Differential Game Theory (DGT) is used to examine these skirmish-level engagements, which provides a mechanism to quantitatively analyze agent performance while simultaneously providing insight into the underlying interactions between agents. Solutions to problems formulated within DGT are comprised of the saddle-point equilibrium strategies of each player as well as the Value function which represents the equilibrium outcome of the conflict as a function of the state of the system. These saddle-point solutions are powerful because they give a guarantee to each player that they can do no worse than the Value of the game if they implement the equilibrium strategy, *regardless of the actions of the adversary*.

The proposed research plan may be summarized as: obtain differential game solutions to novel skirmishes, extend existing solutions
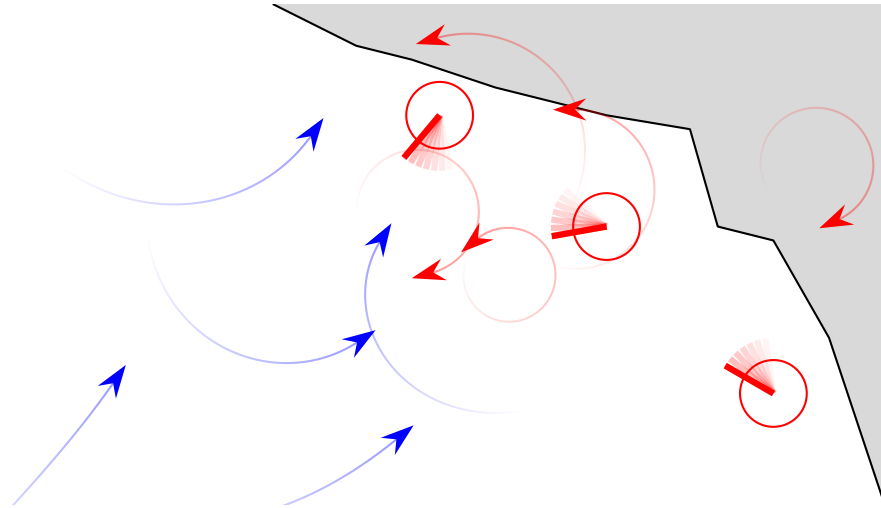
Figure 1.2: Conceptual illustration of a conflict between a Blue team and Red team. The Red team has 3 turrets in addition to its mobile agents. The grey region may represent a target region for the Blue team – they may wish to retreat there, or perhaps enter it and inflict damage or find particular targets. Naturally, one may think of assignments between the various agents before attempting to compute their control; but even this may hinge on the higher mission-level goals of each team. The Red team may, for example, wish to minimize casualties or minimize the number of Blue agents that can enter the grey region.

to more complex skirmishes, develop/compare numerical techniques, and develop approximations. Throughout, there is an emphasis on the practical application of differential game and optimal control solutions – the goal is for an individual agent to be able to compute its optimal/equilibrium control input using onboard computational resources in an amount of time appropriate for the particular task. Thus, the goal is generally to obtain closed-form analytic solutions to these problems. When this is not possible, parametric solutions with a reduced dimensionality are also developed, which can be quickly and efficiently solved using traditional root-finding methods. Lastly, when neither analytic nor parametric solutions are feasible, suitable approximations are sought.

In the following Section, a classical pursuit-evasion scenario involving one Pursuer and one Evader is formulated and solved using DGT to introduce some of the key concepts, such as players, cost functionals, control variables, equilibrium, and Value. Chapter 2 contains a literature survey which provides the necessary context and identifies key themes for the proposed research. Chapter 3 specifies the proposed

plan as a series of research objectives; it concludes with a discussion about the challenges and strengths of the approach taken.

In this section, some DGT concepts are demonstrated upon a classic pursuit-evasion problem. Fig. 1.3 depicts the scenario. The Pursuer, P, wishes to capture the Evader, E, in minimum time, while the latter wishes to delay capture as long as possible. Capture occurs when P is coincident with E. The environment is simply the 2-D plane ($\mathbb{R}^2$) – there are no obstacles or external disturbances (such as wind).
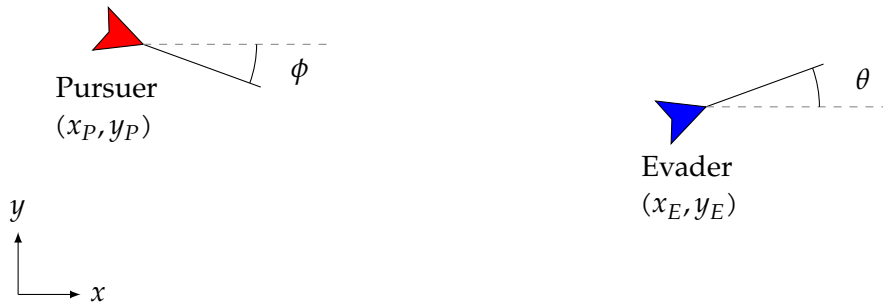
Figure 1.3: Schematic representation of the 1v1 pursuit-evasion problem.

1.1.1 *Problem Formulation*

The state of each player is simply their Cartesian coordinates on the (x, y)-plane, $P \equiv (x_P, y_P)$ and $E \equiv (x_E, y_E)$. The overall state of the system, $\mathbf{x}$, is the concatenation of their coordinates; its evolution through time is given by the following set of ordinary differential equations,

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x}_E \\ \dot{y}_E \\ \dot{x}_P \\ \dot{y}_P \end{bmatrix} = \begin{bmatrix} v_E \cos \theta \\ v_E \sin \theta \\ v_P \cos \phi \\ v_P \sin \phi \end{bmatrix}, \tag{1.1}$$

where v_E and v_P are the speeds of the Evader and Pursuer, respectively. Eq. (1.1) is referred to as the dynamics of the system (although, strictly speaking, this is a kinematic model). Each player controls their instantaneous heading, thus they can "turn on a dime". This motion model

is generally referred to as simple motion (sometimes also as first-order or single integrator). A solution to (1.1) is referred to as a trajectory [1]. For a given state, $\mathbf{x}$, and particular Evader and Pursuer strategies, $\theta(t)$ and $\phi(t)$, respectively, the cost/payoff functional is

$$J\left(\theta(t), \phi(t); \mathbf{x}\right) = \int_0^{t_f} 1 \, d\tau, \tag{1.2}$$

where t_f, the final time, is the time at which capture occurs. The Value function, if it exists, is defined as

$$V(\mathbf{x}) = \max_{\theta(t)} \min_{\phi(t)} J\left(\theta(t), \phi(t); \mathbf{x}\right) = \min_{\phi(t)} \max_{\theta(t)} J\left(\theta(t), \phi(t); \mathbf{x}\right). \tag{1.3}$$

Termination of the game occurs when P captures E – in this case, point capture is specified (i.e., E and P must be coincident). Let the terminal surface be defined as the collection of state variables that satisfy the point capture condition:

$$\mathcal{T} = \{\mathbf{x} \mid x_P - x_E = y_P - y_E = 0\}. \tag{1.4}$$

Then the terminal time is

$$t_f = \min t, \qquad \text{s.t. } \mathbf{x}(t) \in \mathcal{T}. \tag{1.5}$$

The Pursuer and Evader's equilibrium control strategies are denoted by ϕ^* and θ^*, respectively, and obey the saddle-point equilibrium property,

$$J(\theta(t), \phi^*(t); \mathbf{x}) \leq \underbrace{J(\theta^*(t), \phi^*(t); \mathbf{x})}_{V} \leq J(\theta^*(t), \phi(t); \mathbf{x}), \qquad \forall \theta, \phi, \mathbf{x}. \tag{1.6}$$

In other words, neither player has incentive to deviate from the equilibrium strategies.

It is clear that if $v_E \geq v_P$, there exists an Evader heading to guarantee capture *never* occurs. Therefore, let $v_E < v_P$ in order for the Value of the game to be finite. Additionally, assume that the players' initial positions are not coincident. Furthermore, it is assumed that both play-

---

1 Lipschitz continuity of the trajectory is not necessary; clearly, with simple motion, an agent may turn suddenly, producing a sharp corner in the path.

ers have full state and parameter information but do not have access to each others' current control input.

1.1.2 *Necessary Conditions for Equilibrium*

The analysis begins with derivation of the first-order necessary conditions for equilibrium. Note that in many places the functional arguments are omitted for convenience (and when clarity is not hindered)[2]. First, the system's Hamiltonian is formed:

$$
\begin{aligned}
\mathscr{H}\left(\mathbf{x}, \boldsymbol{\lambda}, \theta, \phi\right) &= \boldsymbol{\lambda} \cdot \dot{\mathbf{x}} + 1 \\
&= \lambda_{x_E} v_E \cos \theta + \lambda_{y_E} v_E \sin \theta \\
&\quad + \lambda_{x_P} v_P \cos \phi + \lambda_{y_P} v_P \sin \phi + 1,
\end{aligned}
\tag{1.7}
$$

where $\boldsymbol{\lambda} \equiv \begin{bmatrix} \lambda_{x_E} & \lambda_{y_E} & \lambda_{x_P} & \lambda_{y_P} \end{bmatrix}^\top$ is a vector of adjoint variables. The equilibrium adjoint dynamics must satisfy

$$
\begin{aligned}
\dot{\lambda}_{x_E} &= -\frac{\partial \mathscr{H}}{\partial x_E} = 0 \\
\dot{\lambda}_{x_y} &= -\frac{\partial \mathscr{H}}{\partial y_E} = 0 \\
\dot{\lambda}_{x_P} &= -\frac{\partial \mathscr{H}}{\partial x_P} = 0 \\
\dot{\lambda}_{y_P} &= -\frac{\partial \mathscr{H}}{\partial y_P} = 0.
\end{aligned}
\tag{1.8}
$$

As is typical with simple motion, the equilibrium adjoint dynamics are zero, which implies that the adjoint variables are constant. Next, it is necessary for the Evader and Pursuer controls to maximize and minimize the Hamiltonian, respectively. Fortunately, $\mathscr{H}$ is a separable function of θ and ϕ, and thus these operations can be done independently. To maximize $\mathscr{H}$, E's equilibrium control must be such that the vector $\begin{bmatrix} \cos \theta^* & \sin \theta^* \end{bmatrix}^\top$ must be parallel to $\begin{bmatrix} \lambda_{x_E} & \lambda_{y_E} \end{bmatrix}^\top$, and thus

$$
\cos \theta^*(t) = \frac{\lambda_{x_E}}{\sqrt{\lambda_{x_E}^2 + \lambda_{y_E}^2}}, \qquad \sin \theta^*(t) = \frac{\lambda_{y_E}}{\sqrt{\lambda_{x_E}^2 + \lambda_{y_E}^2}}.
\tag{1.9}
$$

---

2 This practice is adopted from Y. C. Ho *et al.* [125].

The equilibrium control for P, which minimizes $\mathcal{H}$, is obtained similarly:

$$\cos \phi^*(t) = \frac{-\lambda_{x_P}}{\sqrt{\lambda_{x_P}^2 + \lambda_{y_P}^2}}, \qquad \sin \phi^*(t) = \frac{-\lambda_{y_P}}{\sqrt{\lambda_{x_P}^2 + \lambda_{y_P}^2}}. \qquad (1.10)$$

Since the equilibrium adjoint variables are constant (from (1.8)), it must be the case that the equilibrium headings are also be constant, $\theta^*(t) = \theta^*$ and $\phi^*(t) = \phi^*$. Thus, the paths that E and P take from their initial conditions to the point of capture are straight lines.

1.1.3   *Geometric Interpretation*

Now that it has been established that the equilibrium paths are straight lines it is helpful to proceed with a more geometric analysis. For convenience, define $\alpha \equiv \frac{v_E}{v_P}$. Consider all (constant) Evader headings $\theta \in [0, 2\pi]$. There is a corresponding (constant) Pursuer heading, ϕ, which results in capture. For each capture point, the ratio of distances from the initial positions of E and P must be α since the players arrive at the capture point simultaneously. It so happens that the locus of these capture points is a circle: the Apollonius circle (AC) [259, Appendix 2].
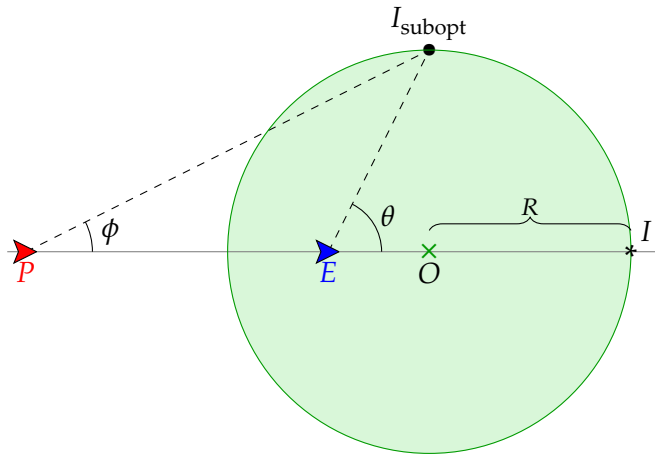
Figure 1.4: AC associated with a pursuit-evasion scenario involving a faster Pursuer and point capture (adapted from [245]).

Fig. 1.4 shows the AC for this scenario. The AC center and radius are

$$\overline{EO} = \frac{\alpha^2}{1-\alpha^2}\overline{PE}, \qquad R = \frac{\alpha}{1-\alpha^2}\overline{PE}, \tag{1.11}$$

where, e.g., $\overline{PE}$ is the Euclidean distance between the points P and E. Without loss of generality, Fig. 1.4 shows the Line of Sight (LOS) aligned with the x-axis.

Now, the problem of obtaining the equilibrium headings may be distilled to finding the point on the AC which gives E the maximum capture time. Equivalently, this is the point on the circle which is farthest from E. It is obvious, then, that the equilibrium aim point must be the point I in Fig. 1.4, which is the point on the AC which is antipodal from P. Then the associated equilibrium headings are aligned with the LOS $\overrightarrow{PE}$. This strategy is referred to as Pure Pursuit (PP) for the Pursuer and Pure Evasion (PE) for the Evader. In general, the strategy may be written as

$$\cos\theta^*(\mathbf{x}) = \cos\phi^*(\mathbf{x}) = \frac{x_E - x_P}{\overline{PE}}, \qquad \sin\theta^*(\mathbf{x}) = \sin\phi^*(\mathbf{x}) = \frac{y_E - y_P}{\overline{PE}} \tag{1.12}$$

If both P and E take a straight-line path to the point I, the associated capture time is $\frac{\overline{PE}}{v_P - v_E}$, which gives a candidate for the Value function:

$$V(\mathbf{x}) = \frac{\sqrt{(x_E - x_P)^2 + (y_E - y_P)^2}}{v_P - v_E}. \tag{1.13}$$

1.1.4 *Sufficient Conditions*

In order to verify that the candidate Value function (1.3) indeed gives the saddle-point equilibrium value of the differential game it is sufficient to show that V is continuous and continuously differentiable[3] (i.e., $V \in C^1$) and that it satisfies the associated Hamilton-Jacobi-Isaacs

---

3 Except along singular surfaces.

(HJI) partial differential equation (PDE) [130]. The HJI for this differential game is

$$\max_\theta \min_\phi \{\nabla_\mathbf{x} V \cdot \dot{\mathbf{x}} + 1\} = \min_\phi \max_\theta \{\nabla_\mathbf{x} V \cdot \dot{\mathbf{x}} + 1\} = 0 \qquad (1.14)$$

The min max terms are accounted for by substituting the equilibrium controls (1.12) into (1.1):

$$\nabla_\mathbf{x} V \cdot \dot{\mathbf{x}}^* + 1 = 0$$

$$\begin{bmatrix} \frac{x_E - x_P}{(v_P - v_E)\overline{PE}} \\ \frac{y_E - y_P}{(v_P - v_E)\overline{PE}} \\ \frac{-(x_E - x_P)}{(v_P - v_E)\overline{PE}} \\ \frac{-(y_E - y_P)}{(v_P - v_E)\overline{PE}} \end{bmatrix} \cdot \begin{bmatrix} \frac{v_E(x_E - x_P)}{\overline{PE}} \\ \frac{v_E(y_E - y_P)}{\overline{PE}} \\ \frac{v_P(x_E - x_P)}{\overline{PE}} \\ \frac{v_P(y_E - y_P)}{\overline{PE}} \end{bmatrix} + 1 = 0$$

$$\frac{-(v_P - v_E)\left((x_E - x_P)^2 + (y_E - y_P)^2\right)}{(v_P - v_E)\overline{PE}^2} + 1 = 0$$

$$-1 + 1 = 0.$$

Additionally, it is clear that V and $\nabla_\mathbf{x} V$ exist as long as $v_P \neq v_E$ and $\overline{PE} \neq 0$ – both of which were listed as assumptions earlier. Thus the candidate Value function is continuous and continuously differentiable and satisfies the HJI which ensures that it is, indeed, the solution to the game.

1.1.5 *Pitfalls*

It is tempting to think that the PP strategy guarantees capture on or inside the AC. However, the saddle-point property of the Value function applies only to the particular cost/payoff functional, which was the capture *time*. At this point, there is no guarantee of *location* of capture under any Evader strategy. In particular, consider a class of Evader strategies that involve picking a constant heading. Fig. 1.5 shows the result when P implements the equilibrium strategy (for the game of min max capture time) of Pure Pursuit. The orange curve is the locus of capture locations, and the players' trajectories are shown for a partic-

ular θ. The Pursuer's trajectory happens to have a closed-form analytic solution known as a pursuit curve (c.f. [220]). Of course, the capture time is still less than the Value of the game, but this discussion is included to highlight the need to carefully consider which policies offer which guarantees.

1.1.6 *A Second Pursuer*

Now suppose a second Pursuer is added to the scenario; let the Pursuers be denoted P_1 and P_2. This new Pursuer "cooperates" with the first in the sense that P_1 and P_2 are controlled by one "player" whose control input consists of $\mathbf{u}_P \equiv \begin{bmatrix} \phi_1 & \phi_2 \end{bmatrix}^\top$. Rufus Isaacs introduced this problem as the "Two Cutters and Fugitive Ship" in his seminal book on differential games [130]. Fig. 1.6 shows Isaacs' proposed solution for the case of simultaneous capture by both Pursuers. For simultaneous capture, the equilibrium strategies may be summarized as: all three agents aim at the point I which is the intersection of P_1 and P_2's ACs that is furthest from E.

Despite the clean geometric interpretation of the solution, its verification (via the process demonstrated in Section 1.1.4) is nontrivial,
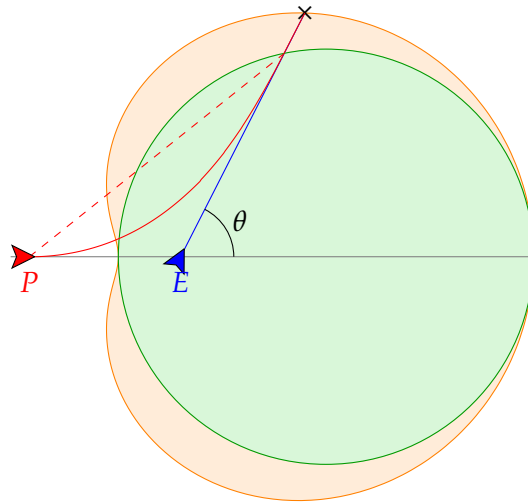
Figure 1.5: Pure Pursuit versus straight-line Evader paths showing "leakage" outside of the initial AC. The solid red curve is longer than the dashed red curve corresponding to the Pursuer's path if it knew $\theta(t)$, however it is still shorter than if E were to implement Pure Evasion.
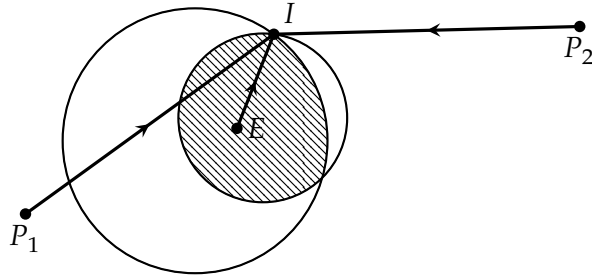
Figure 1.6: Simultaneous capture solution for Isaacs' Two Cutters and Fugitive Ship problem, adapted from [130].

and was only recently completed [107]. Part of the difficulty of the verification is due to the possibility for the game to degenerate into solo capture by either P_1 or P_2 in particular regions of the state space. Then there is the Dispersal Surface (DS), a singularity for which the gradient of V does not exist. Along the DS, the equilibrium headings are not uniquely defined; either AC intersection is equally optimal because they are equidistant from E. This singularity occurs when E lies on the line segment joining the two Pursuers.

Although the DS is generally considered to be relatively benign (compared to the other types of singularities which may appear in differential games, c.f., [19, Fig. 11]), this particular DS has prompted some focused attention recently. For example, in [186], it was shown that, in the vicinity of the DS, Isaacs' proposed strategy can cause the Pursuers to effectively *slow down* if the Evader stands still. Ultimately, the Evader is able to achieve the Value of the game (i.e., the min max capture time) without even moving. The problem lies in the fact that, without knowing E's instantaneous control action, the P's cannot know which AC intersection to aim at. A resolution for this particular dilemma has been proposed in [165], however, it should be clear at this point that, compared to the 1v1 game where there was only one type of termination and no singularities, the 2v1 game is vastly more complex.

Sometimes it seems as though each new step towards AI, rather than producing something which everyone agrees is real intelligence, merely reveals what real intelligence is not.

— Douglas Hofstadter, Gödel, Escher, Bach:
An Eternal Golden Braid

2

LITERATURE SURVEY

2.1 INTRODUCTION

This literature survey is intended to provide an overview of research works related to the overall topic of skirmish-level tactics from a control science and theory perspective. Along the way, the connections to existing military (vehicle-level) tactics [211], military technology plans [237], and military-inspired scenarios [67] will become obvious. The survey is not intended to be exhaustive; its purpose is to provide context for the proposed research questions and preliminary results presented herein.

The remainder of the survey is organized as follows. Section 2.2 provides a brief description of control science and various approaches to designing a controller. Section 2.3 contains a discussion of optimal control, from its theoretical underpinnings, to numerical methods and applications to higher-level problems such as path planning. Section 2.4 constitutes the bulk of the survey, covering DGT from its history, to applications, solution approaches, singularities, and open research questions. Finally, Section 2.5 examines several works which are particularly relevant to the proposed research plan, some of which are the bases for proposed extensions, while others are focused on related scenarios or apply similar methodologies.

2.2   ENGINEERING APPROACHES TO CONTROLLER DESIGN

An old joke that most Aerospace Engineering students have heard before goes something like this:

> What did the pilot say over the intercom to the passengers after losing control of the airplane?
>
> "Everyone please move to the left-hand side of the plane!"

In an abstract sense, *control* refers to the design of an input signal to a system in order to achieve a desired outcome based on the measured outputs of the system. The desired outcome may be to stabilize an unstable system, reject disturbances from the environment or unmodeled dynamics, alter the dynamic response of a system, or even to reach some desired state. Design of a controller (or control system), then, entails the mapping of observed outputs to an appropriate input corresponding to the desired outcome. As mentioned in [151], stability is always a chief concern or criteria in designing a controller. Traditional control design approaches are based on analysis in the frequency domain and thus hinged upon the so-called gain and phase margins (i.e., Bode plot) which, together, imply the stability and responsiveness of the controlled system. Underneath the gain and phase margin concepts lie transfer functions, which describe input/output relationships on signals in the frequency domain. For a system to be stable, the poles of the system's transfer function must have negative real parts (i.e., lie on the left-hand side of the real-complex plane – hence the joke, above).

A simple example of a traditional control system is that of a reference governor. Its purpose is to maintain the system at some given (fixed) reference; it can be thought of as rejecting noise or disturbances from the environment. Thus there is some notion of an error signal which represents a deviation from the reference, which is to be driven to zero by the controller. A typical approach to designing a reference governor begins by linearizing the (assumed to be known) system dynamics about the desired reference. Then various combinations of the error signal, its derivative, and its integral may be used to stabilize the system.

This is, of course, the famous, and ubiquitous, proportional-integral-derivative (PID) controller (c.f. [54, 174]), which has even been applied to the control of quadrotor control [29]. The meaning of stability, in this context, means that the error signal tends to zero (rather than exploding), as long as the system does not stray too far from the reference. This process may be repeated for many different reference points across the state space and the control gains may be *scheduled* so as to be able to transition from one reference point to another. Many real-world systems operate in this fashion today [149]. Indeed, approaches similar to this have even been employed to scenarios relevant to this research proposal, such as in [208]. There, a multi-robot cooperative control strategy is designed to entrap a set of prey, and the nonlinear dynamics are linearized about an operating point for each hunter.

Frequency domain analysis, however, is replaced with linear algebra in what is referred to as "modern" control. One is concerned with *states* rather than signals, and the stability of a system is governed by eigenvalues of the state transition matrix, which maps the current states to state derivatives. Two significant branches of (linear) modern control theory are worth mentioning at this point: optimal control (e.g., the Linear Quadratic Regulator (LQR) [124]) and robust control (e.g., $\mathcal{H}$-∞ control [275]). They represent two important design philosophies: best possible performance, and guaranteed level of performance in spite of uncertainty, respectively. It will become clear, later, that a saddle-point equilibrium strategy within a game is both *optimal* (in some sense) while also being *robust* to any action of the opponent. Note that the LQR formulation has been applied extensively to differential games and optimal control problems which are relevant to the current study. For example, [208] applies LQR to obtain the pursuers' control in an entrapment problem, and [147] applies LQR to the Target-Attacker-Defender (TAD) problem.

There are also methods for control design which do not rely on linearization. In nonlinear control, the designer typically proposes a candidate controller and Lyapunov function, $\mathcal{V}$. The Lyapunov function represents some notion of absolute error or energy of the system. If it can be shown that $\mathcal{V}(t = t_0) > 0$ and $\dot{\mathcal{V}}(t) < 0$ for all t, it is gen-

erally sufficient for saying that the system will be stabilized around some equilibrium. In the context of the present study, this method can be used, e.g., to obtain a controller which can guarantee a pursuer can capture an evader, regardless of the evader's strategy. This type of controller is *robust* to the evader's inputs, but may not necessarily be *optimal* w.r.t. any performance measure. A similar concept was employed to guarantee capture of an evader in a bounded environment in [128].

As the field of Artificial Intelligence (AI) has grown and matured in recent decades, many AI concepts have made their way into control systems, control design, and applications, including (but not limited to) Neural Network (NN), Genetic Algorithm (GA), Fuzzy Inference System (FIS), Reinforcement Learning (RL), and other data-based approaches. Reference [149] gives an overview of the ways in which AI can and have been applied to and integrated within control systems. Neural networks have, for example, been used to generate inputs for a flapping flight controller [61], adaptive flight control [173], and a reactive driving controller for a small robot [78]. Genetic algorithms have been applied to the control design process, e.g., in [78] to design neural network weights, in [48] as a pre-processing step for other control design approaches, and in [9] to design an open-loop control input to maneuver behind a turning opponent. Fuzzy inference systems have also been applied to so-called tail-chase scenarios, e.g., in [10], as well as to combat scenarios with teams of vehicles [67]. Reinforcement learning is a process by which an agent, or control system, is rewarded for taking "good" actions and penalized for taking "bad" actions; the quantity of literature on this subject is vast. Most notably, reinforcement learning has been applied to multiple-pursuer, single-evader games (a topic of interest in the present study) in, e.g., [28], and more famously in DARPA's AlphaDogfight Trials wherein agents competed in 1v1 simulated F-16 combat [119, 194]. The application of RL specifically to optimal control and differential games has been covered generally in [257] and in works by Schwartz, e.g., [14, 205]. Finally, one of the central themes in AI and machine learning is the balance between *exploration* and *exploitation*. This tradeoff is ever-present in data-driven control approaches based on Gaussian Process (GP) [56, 57, 137, 189,

243]. There, a GP is used to model the uncertain dynamics – the model more closely matches reality in the vicinity of points which have been previously visited. Despite all of the work that has been done, the judicious application of machine learning and AI to control remains a rich area for further development and future research [135].

Already, the reader may sense some divergence from the original notion of *control* wherein a central concern is the stability of a system. At this point, it is important to introduce the notion of *guidance*, whose definition is included, below:

> "Guidance is about the determination of the maneuvering commands to steer the vehicle to fly a trajectory that satisfies the specified terminal/targeting condition as well as other pertinent constraints, and, if required, optimizes a defined performance."
>
> —Ping Lu, 2021 [151]

Thus, the present study is more closely related to guidance in the sense that the chief concern is, e.g., the heading of a vehicle or turning direction of an agent on a 2-D plane. In practice, this information may be used as a reference for a controller at a lower level of abstraction to follow (e.g., the rudder controller may move the rudder to orient the vehicle along the prescribed heading). The term *guidance* has also long been used in connection with missiles and interception [114, 175, 220]. Obviously, missile guidance is a military relevant application, but it also has a strong relation to pursuit-evasion, an important area within differential games [129]. Note that one of the most widely used missile guidance laws, Proportional Navigation (PN), is based upon nullifying the LOS rate to the target, thereby placing it upon a collision course (c.f., [113]). One can easily spot the relationship to classical control concepts – the LOS rate, in this case, is the error signal and the control is in proportion to the error (i.e., the 'P' part of PID).

2.3   OPTIMAL CONTROL

This section is focused on optimal control in a broader sense than the linear (LQR) approach mentioned earlier. Perhaps one of the earliest rigorous formulations of an optimal control problem can be found in Goddard's rocket example [115]. Therein, the objective was to maximize the terminal altitude of a rocket with a limited fuel weight. From a modeling perspective, the example is relatively low fidelity: there are three states, altitude, velocity, and mass, with simple dynamics. The biggest complication comes from the introduction of the nonlinear drag force term, which is a function of altitude (due to varying air density throughout the atmosphere), as well as velocity. Another slight challenge comes from the fact that the control input is bounded (a case which is not easily accounted for in the traditional LQR formulation). Despite its relative simplicity, the solution of the problem is nontrivial, involving a sequence of non-singular and singular arcs.

The main impetus for the inclusion of this discussion in the present study is simply the fact that any adversarial scenario (i.e., game) can be turned into an optimal control problem by assuming a particular control strategy of the opponent. Thus the theory behind optimal control and the methods of solving these problems is relevant to, e.g., skirmish-level tactics wherein the opponent's strategy can safely be assumed. One example is the scenario of evasion from a missile. There are a couple strategies which are particularly prevalent both in the literature and on fielded systems: PP and PN [220]. Therefore, for certain applications, it may be reasonable to assume the pursuer is employing one or other of these strategies.

A second reason for including this discussion on optimal control is that many of the analytical tools carry over to the differential game realm, sometimes with some slight modifications (e.g., the relation between the Hamilton-Jacobi-Bellman (HJB) and the HJI equations, which, interestingly, were developed nearly concurrently). At the forefront of the theoretical underpinnings of optimal control, are, of course, L.S. Pontryagin, who is responsible for Pontryagin's Maximum Principle (PMP) (sometimes the 'M' is Minimum) [193], and R. Bellman, who

is responsible for Dynamic Programming (DP) [21, 22]. The PMP approach to optimal control hinges on the (first-order) necessary conditions for optimality which are obtained via the calculus of variations. One such condition is that the optimal control input (generally denoted with a *, e.g., $\mathbf{u}^*(t)$) maximizes (or minimizes, whichever the case may be) the *Hamiltonian*, $\mathscr{H}$, of a system over the time interval of interest, e.g., $t \in [t_0, t_f]$. The first order necessary conditions for optimality appear also in [40, 136] for a variety of problems (i.e., state constraints, control constraints, fixed final time, free final time, initial boundary conditions specified, final boundary conditions specified, etc.). Note that these variational methods have long been applied to problems of interest, such as pursuit-evasion (e.g., [125]). On the other hand, the principle of optimality which governs DP is the idea that an optimal state trajectory is comprised of the optimal state trajectory in the current stage and the remainder of the optimal state trajectory. DP expresses the recursive relationship between what is optimal now and what is optimal for the remainder of the trajectory. The recursion resolves at termination, meaning that if one can determine what the optimal action was to end at a particular terminal state, one can back-propagate a stage further and thereby chain together a sequence of the optimal state and control trajectory. Interestingly, these two approaches were explicitly compared in the context of energy management for a parallel hybrid electric vehicle [273]. It is important to note that DP is built upon a discrete state and action space, however, as the discretization becomes finer and finer, approaching a continuum, one obtains the HJB equation which is strongly connected to the Hamiltonian. Lastly, note that the Markov Decision Process (MDP) is the stochastic version of (discrete) DP [196], which has been applied to multiplayer pursuit-evasion scenarios, e.g., in [28].

2.3.1 *Direct and Numerical Methods*

Optimal control solution approaches based on the first order necessary conditions for optimality, such as the PMP, are referred to as *indirect*

methods. Solution of the HJB, also, may be considered to be indirect – one obtains an extremal trajectory via the solution of this separate equation. Indirect methods typically rely heavily upon analytical derivation of the optimality conditions for a specific problem. Often one is faced with the challenge of guessing either the initial adjoint variables (which represent the components of the derivative of the Value function w.r.t. the states) or the terminal states and then solving a two-point boundary value problem (TPBVP). The optimal control $\mathbf{u}^*(t)$ associated with the optimal trajectory is essentially a byproduct of this process.

Direct methods, on the other hand, optimize over the control input directly. Generally, the time history of the control vector $\mathbf{u}(t)$ is assumed to have some model governed by a set of parameters or it is discretized at a set of points in time. One of the most popular direct methods is collocation, wherein the state and control trajectories are discretized at a set of points in time and treated as the variables over which to optimize. In collocation, the system dynamics are treated as a constraint – i.e., the dynamics must be respected at each of the discrete collocation points (but may be violated in between the points) [134, 198]. Collocation is made possible by the advancement of general-purpose nonlinear program solvers which are capable of handling large numbers of optimization variables and general nonlinear constraints (e.g., [132], or MATLAB's `fmincon`). Early works such as [31, 34, 241, 242] provided a framework for collocation which has matured in more recent times to include mesh refinement techniques among other advancements [55, 198]. A popular form of collocation, known as pseudospectral methods, approximate the continuous state and control trajectories with Lagrange interpolating polynomials which are constrained to satisfy the system dynamics at specially chosen collocation points [70]. These collocation points correspond to the roots of a Gauss-Legendre polynomial [31]. The advantage of pseudospectral methods is that fewer collocation points are necessary to obtain an approximation with some desired accuracy (compared, e.g., with an even discretization scheme); the promise is of lower computational burden on the nonlinear program solver. In [261], even collocation and pseudospectral methods are compared in the context of the TAD scenario wherein the Attacker

and Defender strategies are given and the Target wishes to maximize its distance from the Attacker at the time it is intercepted by the Defender. Additionally, collocation was used to solve an optimal engagement zone avoidance scenario wherein a vehicle sought to reach a destination in minimum time while avoiding a dynamic keepout zone [263]. In [94] pseudospectral methods are compared with differential dynamic programming, which is iteratively linearizes the system around the trajectory and then uses LQR concepts to alter the trajectory in such a way to reduce the cost [236].

It is also worth noting some recent advances in other direct methods: sequential convex programming and augmented Lagrangian methods. Sequential convex programming bears some similarity to Differential Dynamic Programming (DDP) [93] and iterative LQR (iLQR) [236] in that a trajectory is iteratively *convexified* and improved. It has been successfully applied to, e.g., powered descent guidance for Mars landing [1]. Augmented Lagrangian algorithms such as AL-iLQR and AL-TRO [127] are moreso direct extensions of DDP and iLQR which improve the numerical robustness of DDP and allow initialization of DDP with an infeasible trajectory.

2.3.2 *Path Planning*

In many cases, such as in [50], the result of solving an optimal control problem is a single open-loop state and control trajectory. Such a control may be implemented, but it may not give the desired result if the state of the system deviates at all from the computed open-loop state trajectory. Obtaining the solution for a different set of initial or terminal conditions may require re-running the nonlinear program altogether.

However, in other cases, there are properties of the optimal control whose usefulness transcend that of a single optimal trajectory. An excellent example is that of the famed Dubins vehicle – a nonholonomic agent with constant speed and bounded turn rate. Using the first order necessary conditions for optimality, it can be shown that time-optimal paths for such an agent to reach a specified point are lie in a set of 6 can-

didate paths of the type *CSC* or *CCC* where *C* and *S* denote max-turn-rate circular arcs and straight-line segments, respectively [224]. A similar analysis has been done for the case of a variable speed agent who is capable of slowing down to make tighter turns [266]. Thus, the optimal control analysis, in both of these cases, reveals properties of time-optimal paths which can be leveraged in the slightly more abstract (higher-level) task of path planning. Examples include computing the shortest Dubins path to circles and targets moving on a circle [154, 156] as well as constructing minimum-time tours for, e.g., a traveling salesman problem [6, 239]. In fact, the result of [224] is so powerful that it is possible to compute instantaneous Dubins paths (time-optimal) to a moving target in real time [7].

Part of the purpose of the present study is to draw an analogy with the field of differential games and the utility of their solutions beyond simply obtaining the equilibrium strategies.

2.4 DIFFERENTIAL GAMES

This section is devoted to the area of differential games and describes a body of literature which is most relevant to the proposed approach both in terms of methodology as well as application.

2.4.1 *A Brief History*

Around the time of the genesis of Differential Game Theory (circa 1950's), some ground work for game-theoretic concepts had already been laid in the context of static (i.e., single stage, discrete) games by J. Nash [172] (and even earlier by J. von Neumann and O. Morgernstern). One of Nash's central contributions was the proof of the existence of saddle-point equilibrium strategies in 2-person games, later dubbed the Nash equilibrium. The extension of this important concept into dynamic games is explored more in Section 2.4.3.

R. Isaacs is unanimously considered to be the father of Differential Game Theory and was a contemporary of R. Bellman at the RAND

Corporation (circa 1948). His works on DGT appeared first in a series of internal RAND reports (e.g., [129]) which were later compiled in his seminal book *Differential Games: A Mathematical Theory with Applications to Optimization, Control, and Warfare* [130]. As the title suggests, his book introduces the underlying theory of differential games (although, the book lacks rigorous mathematical proofs for some of the statements) while also applying the theory to a vast number of interesting applications and problems. Not least among these problems are pursuit-evasion (including pursuit with two pursuers) and target guarding, which are central to the present study. Isaacs' works, including [131] also grappled with the subject of singular surfaces, which may the most challenging aspect of applying DGT – the subject of singular surfaces is discussed in Section 2.4.5. Reference [38] extensively covers the works of R. Isaacs and their relation to the history and development of DGT.

Later, in the 1970's, J. V. Breakwell (along with many others) applied DGT to problems of pursuit-evasion [37], role determination in 1v1 aerial combat [177], and surveillance-evasion [145]. His student, T. Merz, is credited with the solution of the famous Homicidal Chauffeur Differential Game (HCDG) in his PhD thesis [164]. In the HCDG, a nonholonomic (turn-constrained) pursuer seeks to capture a slower, holonomic evader. The solution involves many different types of singular surfaces and varies widely over the parameter space. Another of Breakwell's students, P. Bernhard, is credited with the solution of the Isotropic Rocket Differential Game (IRDG), another 1v1 pursuit-evasion game but with different dynamics [24]. [1] A semi-autobiographical summary of the early works by Breakwell, Merz, and Bernhard is given in [26].

---

1 Interestingly, Bernhard attributes the initiation of the research topic to none other than R. Kalman of the Kalman Filter fame, who, like J. Von Neumann, was of Hungarian descent.

2.4.2 *Major Problem Areas: A Meta-Survey*

As mentioned previously, this literature survey is, by no means, exhaustive in any particular topic. This Section points to several excellent reviews of the literature which are particularly relevant to the present study. The major problem areas include pursuit-evasion (of course), target guarding, surveillance-evasion, and "three body" problems such as the Target-Attacker-Defender Differential Game (TADDG). There are large bodies of literature focused on each of these problem areas – variations on the cost functional, motion model (i.e., whether holonomic or nonholonomic), information structure, constraints, fixed opponent strategies, definition of capture (whether point-capture or finite capture radius), definition of termination (whether by reaching a goal state or a specified terminal time), number of agents on either side of the conflict, capabilities of the agents (generally, the relative maximum speeds of the players), and so on partly explain the quantity of works on these topics.

Many of the results found in the literature referenced in the previous Section on Differential Game Theory applied to pursuit-evasion is summarized in a chapter of T. Başar's book [19]. Reference [191] also covers pursuit-evasion differential games with a focus on two major examples: the HCDG [164] and J. Shinar's space interception problem [212]. The survey also includes a detailed description of the extensive development of differential games (particularly having to do with pursuit-evasion) that took place in the Soviet Union from 1960-1980. A significant portion of the development was centered on multi-pursuer single-evader differential games and special formulations and methods for handling cooperation among a large group; this is the subject of S. Kumkov's survey [139]. With regards to multi-pursuer single-evader games, the reader is also referred to Appendix A which shows such a formulation appearing in the popular web comic XKCD [169].

I. Weintraub's introduction to pursuit-evasion differential games summarizes many pursuit-evasion differential game results, particularly for agents with simple motion [262]. In addition to surveying a large cross-section of the literature, it also covers the two-cutters and fugitive

ship differential game [89, 107, 183, 187, 188] (even including the solution verification step), as well as the TADDG (c.f. [104, 184]) in a tutorial manner. The latter problem pits an Attacker against a Target-Defender team; the former wishes to intercept the Target, while the latter team seeks to have the Defender intercept the Attacker far from the Target. At the time of writing, [262] contains the most comprehensive survey of TADDG-related literature (although the number of works on this and related problems continues to grow rapidly).

The dissertation of D. Oyler [182] also contains a survey of pursuit-evasion differential games, grouping the works into categories such as environment type (2-D, 3-D, n-D, obstacles, boundaries), player dynamics (simple motion, HCDG-esque (i.e., assymetrical), and "two cars", meaning both agents are turn-constrained [8, 112, 162]), termination/-payoff (capture, colocation, proximity, target guarding, visibility), team configuration (multi- or single- pursuer or evader), and even solution method used (dynamic programming, viscosity solution, dominance regions, etc.). Oyler's approach to surveying the literature is systematic; no attempt is made here to be as exhaustive. The contributions of [182] lie in techniques for solving 1v1, Mv1, and many-versus-multi (MvN) pursuit-evasion games in environments containing obstacles.

Literature on the surveillance-evasion problem, wherein an agent attempts to escape the sensing region of its opponent is largely contained in the book by J. Lewin [144]. One of the problems covered therein (also in [145]) has a similar setup to the HCDG but with a cost/reward functional of min max escape time. Recently, the problem area is seeing regained interest in works such as [42, 86, 260, 264].

Finally, regarding target guarding (where the target is stationary), was formally specified originally in Isaacs' book [130, Example 1.9.2]. An important variant of this problem is perimeter patrol, wherein the Defender is constrained to move along the perimeter of the target area (echoing a similar kind of constraint which appears in the *Lady in the Lake* differential game [19]). The review by Shishika [216] summarizes a body of literature focused on multi-agent perimeter guarding. In these works, the Attacker(s) score by reaching an undefended point on the target boundary, and the Defender(s) effect capture by being coinci-

dent with an Attacker. The type of analysis ranges from rigorous differential game-theoretic approaches to heuristics and algorithms for coordination and cooperation among large teams of agents. Some recent and important developments include the solution of the 1v1 target guarding differential game for general compact and convex targets [142] and the solution for perimeter patrol along arbitrary convex targets [215]. It is also necessary to explicitly connect the area of target guarding to a body of literature under the moniker *reach-avoid games*; the former may be seen as a particular case of the latter. Relevant works include [44, 45, 157, 209, 269].

2.4.3   *Solution Concepts*

This section gives a brief description of what is meant by "solution of a differential game". Even the term differential game has been slightly abused in the literature at times; by "differential game" it is meant that there are two or more players with their own cost functional with which they wish to minimize (or maximize in the case of a utility functional), the states, control and time are continuous, and termination of the game is well-defined. The solution, then, generally comes in the form of the players' equilibrium strategies, the Value function (i.e., the equilibrium value of the cost functionals as a function of the initial state), and a region over which the solution exists. In the sense of Isaacs [130], the solution also entails "filling the state space with equilibrium trajectories" which, in most cases, are back-propagated from the usable part of the terminal surface – these equilibrium trajectories obey the HJI, which ultimately means that the Value remains constant along the trajectory. Note the similarity between Bellman's and Isaac's approaches which essentially hinge upon back-propagation.

There are three additional aspects of the formulation which must be specified up front: the information structure (who knows what at which times), the control structure of each player, and which type of equilibrium is sought. The first two aspects are defined clearly in [178, 179]. Therein, distinctions are made between open-loop, closed-loop,

Stackelberg open-loop, Stackelberg closed-loop information or control structures and demonstrated for some simple examples. It is assumed that the players make use of whatever information they have available in computing their current control. In the case of open-loop, a player observes the initial state of the system and determines their entire control trajectory. Whereas, in closed-loop, a player observes the instantaneous state and its strategy is typically a function of the state (i.e., it is a state-feedback control strategy). Finally, the Stackelberg setting implies there is a leader, who announces their strategy, and a follower who gets to respond accordingly. As shown in [178] the structures can be mixed (i.e., one agent implements closed-loop while the other implements open-loop); however, when the agents have the same structure, there are particular equilibrium definitions corresponding to each type. Some of the equilibrium definitions appear in [52] and elsewhere, and will be used later on. For certain classes of differential games (or under certain assumptions) the various notions of equilibrium may be coincident. For example, [199] describes a class of games for which the open loop Nash equilibrium is coincident with the closed loop Nash equilibrium, [73] describes a different class for which the open loop Nash equilibrium is a degenerate feedback Nash equilibrium, and [201] shows the coincidence of the feedback Nash equilibrium with the Stackelberg equilibrium under certain assumptions.

The case of limited or asymmetric information (i.e., where one agent knows something the other does not) is difficult. In some cases, the scenario may be treated as a stochastic game [272], otherwise the scenario may be decomposed into phases pertaining to what knowledge is available to each agent, e.g., [217]. Especially in the latter case, some assumptions must be made, such as what should an agent do if it cannot sense its opponent. This particular question is heavily related to a body of literature known as adversarial search; one such related work is [158]. On the discrete-time side, there have been some efforts at formalizing and solving partially observable stochastic games [66], which may be considered to be an extension of the Partially Observable Markov Decision Process (POMDP), which is an extension of the MDP.

2.4.4   *Numerical Approaches*

Numerical approaches to solving differential games (which are distinct from AI approaches for learning strategies in adversarial scenarios) largely fall into the areas of numerical viscosity solutions, reachability analysis, and mixed direct-indirect approaches.

The article by M. Falcone [72] is one of many works on the subject of numerical viscosity solutions to the HJI equation. These methods may be used in cases where direct analytical solution of HJI is infeasible. The viscosity solution refers to a pair of smooth functions which tightly upper- and lower-bound the Value function (which, itself may be nonsmooth, e.g., along singular surfaces).

The reachability analysis literature is based on the numerical computation of reachable sets (or sometimes level sets of some measure), the theoretical support for which appears in [157]. An algorithm for computing reachable sets appears in [166]. Interestingly, Mitchell *et al.* utilize the solution of a terminal value HJI PDE in order to compute reachable sets. Elsewhere, as in [232], the relationship is reversed and numerically computed reachable sets are used to determine the equilibrium capture point in a pursuit-evasion game. Note that the approach used in [232] bears a resemblance to the approach in [181, 182] which made use of *isochrones* (time-dependent forward-reachable sets).

Finally, some numerical approaches leverage the information available from the first-order necessary conditions for equilibrium, which may generally be referred to as direct-indirect methods. Notable examples include works by Horie and Conway (e.g., [48, 126]). In those works, the semi-direct collocation with nonlinear programming (semi-DCNLP) method is desribed and applied to pursuit-evasion scenarios. In semi-DCNLP, the necessary conditions for equilibrium for one of the players (e.g., pursuer) are included as constraints to the nonlinear program which solves for the other player's (e.g., evader) equilibrium control via collocation.

2.4.5 *Singularities*

Much could be said about singularities in differential games; this section briefly describes the singularities themselves, some important works, as well as some more applied examples which prominently feature singularities. Most often, singularities appear along manifolds in the state space and are thus referred to as *singular surfaces* in these cases. A concise definition is reproduced below

> "A singular surface is a manifold on which (i) the equilibrium strategies are not uniquely determined by the necessary conditions, or (ii) the [V]alue function is not continuously differentiable, or (iii) the [V]alue function is discontinuous."
>
> — Tamer Başar [19].

Singular surfaces present the following challenges in obtaining what Isaacs referred to as the solution "in the large" [130] (meaning obtaining the Value function and equilibrium strategies across the whole state space): i) discovery, ii) identification characterization of incoming (or outgoing) equilibrium trajectories, and iv) identifying which player has the authority to push the state of the system off or keep the state of the system on the surface. In some cases, when the equilibrium trajectories do not admit closed-form analytic solutions, numerical computation is necessary to obtain the singular surface itself [2]. All known singular surface typer with continuity in the Value function (transition line, dispersal line, equivocal line, universal line, focal line, and switching envelope) are depicted in [19, Fig. 11].

Proper identification and characterization of all singular surfaces within the solution to a differential game is of paramount importance due to the fact that singular surfaces partition the state space into regions of fundamentally different equilibrium behavior [88]. A simple example is a dispersal surface wherein a turn-constrained agent may be able to turn left *or* right (at max turn rate) and achieve the same Value in equilibrium (thereby partitioning the state space into "turn left" and "turn right" regions). Some surfaces, such as the equivocal

surface in the HCDG solution [164], present additional complexities, e.g., the requirement that one of the agents know the instantaneous control input of its opponent in order to stay on the surface. Failure to properly characterize the singular surfaces may have dire consequences, such as backwards-integrating trajectories into regions for which they are no longer optimal/equilibrium.

It must be noted that much of the work addressing singular surfaces was done by the early practitioners of differential games, not least of which was Isaacs, himself [130, 131]. Additionally, Breakwell, Merz, Bernhard, and Lewin contributed greatly to their study [24, 25, 36, 144, 164]. Lastly, the work of A. Melikyan (another Russian differential games theorist) contributed additional necessary conditions for optimality which must be satisfied along universal surfaces and equivocal surfaces [160] as well as some properties of singular paths in the vicinity of equivocal, focal, and dispersal surfaces [159, 161].

The following works focus specifically on singular surfaces in games which are particularly relevant to the present study. Reference [88] identifies a dispersal and universal surface in the solution of the TADDG wherein the Defender is turn-constrained. In [233] the differential game of cooperative evasion from [85] is made into an optimal control problem by fixing the strategies of the two-evader team; in this problem, two unique dispersal surfaces are identified (one is analytic while the other is obtained numerically). References [165, 186] both analyze the dispersal surface within the two-Pursuer, single-Evader differential game of min max capture time. In the former, it is shown that the classical solution for the Pursuers admits chattering when the Evader stands still on the dispersal surface. The latter work proposes a solution to this phenomenon based on the saddle point of the initial instantaneous loss rate, which suggests the Pursuers to implement PP for a brief period after which the solution becomes unique or else the Evader will suffer a large penalty. Interesting singularities appear also in games with multiple evaders. In [37, 85] a singularity appears when the Pursuer is equidistant to two Evaders. A similar singularity is present in a target guarding game with one pursuer and two Evaders wherein the Value of pursuing Evader 1 and then 2 is equal to that of pursuing Evader

2 and then 1 [109]. Note, for this problem, the singularity due to sequence is distinct from another dispersal surface in which the players' equilibrium headings are non-unique; there are likely configurations for which these two singularities coincide. The singularity due to sequence is very much an open research problem.

2.5 RELATED WORKS AND KEY THEMES

All of the aforementioned works are relevant to the present study, however, this section of the literature survey focuses on some more specific themes within the literature which warrant additional work and form the basis of some of the research objectives stated in Chapter 3. Each of themes appear below as a subsection with a small collection of relevant works; they are: parameter exploration, team composition, role selection, discrete decision making, multistage scenarios, and applications of skirmish-level solutions. The last of these provides some foundation for the utility of skirmish-level solutions in more complex scenarios.

2.5.1 *Parameter Exploration*

Exploration of the parameter space is an important undertaking in many games as often there are regions of fundamentally different equilibrium behavior in different parameter regimes. Like singular surfaces, this theme (at least as it pertains to differential games), is best exemplified by the works of Merz and Lewin. In the HCDG, 12 different "exceptional lines" (singularities) are present (or not) in the solution in over 20 regions/subregions – not only are the equilibrium behaviors different in the various regions, but also the number and type of singularities [164]. A similar style of parameter-space partitioning appears in [146] for a surveillance-evasion differential game with a conic surveillance region. In [229] a reinforcement learning approach is applied to a multi-versus-multi aerial engagement; many simulations are performed for varying weapon effectiveness parameter values for each team. The team with an advantage in weapon effective-

ness generally improved compared to having equal weapon effectiveness, as expected. Note for AI-based approaches, such as this one, that a change in the model parameters may warrant a computationally-expensive re-training of the control policy. Another form of analysis for which parameter exploration is prominent is that of competitive analysis (e.g., [15], which analyzes a 1-D target guarding problem). A competitive policy is one which can guarantee some finite competitive ratio which is the ratio of optimal and guaranteed performance. There, fundamental limits are obtained as functions of the parameters which describe necessary conditions for achieving various levels of competitive ratio. Finally, note the importance of parameter exploration as it pertains to balance in video game applications. In [12], a Koopman-based learning approach is used to tune four parameters for a Starcraft 2 minigame. There, a particular set of parameters is sought which achieves the desired level of balance according to some measure.

2.5.2   *Team Composition*

Team composition may be considered to be a particular instance of parameter exploration, particularly pertaining to teams of heterogenous agents (who are heterogeneous either in role or capability). The proposed extensions to [12] include more complex minigames with heterogenous teams (specifically, Roaches and Hydralisks versus Marines and Tanks). In addition to the much larger unit parameter set, there is the additional question of how many of each agent to utilize. A similar question is explored in a perimeter defense scenario in [218], where, in addition to perimeter-constrained Defenders, there are patrolling vehicles who orbit the perimeter at some distance at constant speed, which further constrain the trajectories of incoming Attackers. Team composition is also important in many biological systems, e.g., the allocation of bees to guard against hornets [234], and also in the types of guard bees deployed at the hive entrance [227].

2.5.3 *Role Selection*

Role selection appears in the literature in several ways. The first is in 1v1, so-called "symmetric" [118] engagements such as in the Game of Two Cars (G2C) (again, studied by Merz) [164]. Here, the role selection may be better described as where/when to act as a pursuer or as an evader (c.f. [112, 163]). This question was also investigated by Olsder and Breakwell in [177].

Elsewhere, role selection takes the form of whether to be the leader or follower among a team of two or more agents as in the cooperative two-Pursuer HCDG [32]. The distinction between leader and follower occurs repeatedly in nature as well (c.f. [225, 230]).

Still, in other works, the roles an agent can take are less well-defined and are implicitly tied to, e.g., the sequence in which a Pursuer pursues the Evaders. Examples include [85] wherein one Evader is being pursued and its teammate performs a flanking maneuver to inflict a cost upon the Pursuer, [37] wherein the Evader who is initially targeted by the Pursuer has a component of velocity which draws the Pursuer further from its partner, and likewise in [109]. In this examples the first Evader may be considered to be bait, or otherwise sacrificial.

2.5.4 *Discrete Decision Making*

In general, discrete decision making could refer to a variety of problems under the classification of AI (c.f., e.g., [202]), including the MDP. Here, discrete decision making in adversarial scenarios is loosely defined as either 1) which agent should perform a task (or, conversely, which task an agent should perform) or 2) which "game to play" in a scenario with multiple outcomes.

The first definition is essentially a slight generalization of role selection, though, here, there is one clear task which a team of agents wish to accomplish and the discrete decision pertains to which of the individual agents will complete it. For example, in [16] a single Evader is pursued by multiple evaders in a relay pursuit scheme; the decision faced

by the Pursuers is which one of the pursuers should actively pursue the Evader in the current time instant (while all other pursuers remain stationary, e.g., to preserve energy). There, a Voronoi-like partitioning is utilized to aid in that decision – the active Pursuer is selected as the one in whose cell of the partition the Evader is currently in. Another example (which aligns with the converse description) is [121] where a single Pursuer seeks to capture one of many evaders

Examples regarding the second definition, i.e., which "game to play", are more sparse. Recently, Shishika *et al.* analyzed a partial information target defense game in which the scenario may end in i) capture of the Attacker by the Defender, ii) breach of the target by the Attacker, or iii) escape of the Attacker into a region un-sensed by the Defender. There, the dominance region of the slower Attacker (i.e., the Apollonius circle) is used to determine which outcomes are possible given the current configuration, and policies are prescribed in the various regions. Perhaps the best example, however, is in the literature on engage or retreat games wherein a mobile Attacker must decide between engaging a target who may inflict damage on it, or retreat to some pre-specified safe zone [83, 84, 87, 90, 221]. The solution of these games hinges upon the fact that the equilibrium choice (i.e., whether the Attacker should engage or retreat, or whether the Defender should retaliate or cooperate) at the start of the scenario must remain optimal throughout the playout, otherwise it was not the equilibrium choice in the first place. Consequently, for, e.g., retreat to be the equilibrium choice, it must not become advantageous for the Attacker to switch to engagement *at any point along the retreat trajectory*, which gives rise to a corresponding path constraint based upon the Value of engagement.

2.5.5 *Multistage Scenarios*

Multistage refers to games or scenarios in which there are discrete events which naturally leads to analyzing a pre-event stage and post-event stage. Capture-the-flag is one such example in which the discrete event may be one of the players capturing the flag (after which they must re-

turn it to their home region); such a scenario was analyzed from a differential game theoretic standpoint, e.g., in [102]. There are also examples from the target guarding literature in which there is a distinction between the players' behaviors when they are far away versus when they are in close proximity [81, 98]. Another example, which has already been briefly mentioned, is that of tasks being done in sequence (e.g., sequential capture of Attackers in [109]). Surveillance-evasion can be thought of as happening in stages corresponding to times at which the Evader enters and exits the surveillance region. For example, [264] considers the post-entrance / pre-exit stage in which the surveilling agent wishes to maximize the time at which the Evader exits. Finally, in [213] Shinar *et al.* analyze a pursuit-evasion game in which the Pursuer can switch to a different set of dynamics once during the game. These scenarios may benefit from a Dynamic Programming interpretation in which the optimal action in the current stage is partly governed by what is optimal in subsequent stages.

2.5.6 *Application of Skirmish-Level Solutions*

In this last section, some examples of ways in which skirmish-level (1v1, 2v1, 1v2, etc.) solutions may be applied to more complex scenarios in an effort to highlight their importance and utility. First, it must be noted that in cases of "symmetric" engagements (where both sides are capable of capturing or destroying the other) the skirmish-level solution to the 1v1 scenario with *a priori* specified roles (i.e., one agent is the Pursuer and the other is the Evader) is a necessary component to the overall solution (e.g., the G2C [162, 163]). Also, pursuit-evasion solutions have been applied to collision avoidance (i.e., the worst possible case is when the other vehicle is deliberately attempting to collide) (see, e.g., [69, 166, 240]). Decomposition of a scenario into smaller sub-games is another area in which skirmish-level solutions are useful. For example, in the engage or retreat scenarios described previously, one must obtain (and use) the solution of the Game of Engagement in order to solve the overall problem [84]. Also, in the partial information target

guarding game in [217], the dominance region of the Evader actually corresponds to a generic 1v1 target guarding game [103]. Similarly, 1v1 solutions are also useful for games with sequential capture, e.g., [80, 150, 231, 274].

Task assignment generally applies to larger multi-versus-multi (or MvN) games wherein the Values of skirmish-level games are used as a cost metric in order to make decisions on which agents or sub-teams should "play" against each enemy agent or sub-team. The best assignment is one which minimizes some overall cost metric. This process is a heuristic approach to what would otherwise be an intractable problem (i.e., to find the equilibrium for such a large joint state and action space). However, it benefits from the guarantees provided by the equilibrium solutions to the skirmish-level games. Examples in the literature include [45, 100, 148, 153, 168, 214, 219, 269].

Another interesting usage is in the embedding of the solutions (or features of the equilibrium behaviors) in AI and learning processes. For example, in [170] the optimal control solution obtained via solution of the HJB is used as a "hot start" in an RL scheme within a more complex autonomous racing application. Another example is that of [8] wherein an evolutionary algorithm is used to design a controller for the G2C; the controller switches between various behaviors that are based off of the true equilibrium behaviors (e.g., turn with max-turn-rate towards the opponent). Other examples include [210, 228, 229].

Finally, there is demonstrated utility of skirmish-level solutions in higher-level planning tasks. In [155], the Value of the TADDG [185] is used as a utility metric over which a Defender agent seeks to maximize over an entire path in coordination with a Target vehicle. Reference [209] analyzes a multi-Pursuer single-Evader reach-avoid game wherein the Evader's policy incorporates the Value of each 1v1 engagement as a measure of risk.

For now, what is important is not finding the
answer, but looking for it.

— Douglas Hofstadter, Gödel, Escher, Bach:
An Eternal Golden Braid

3

RESEARCH OUTLINE

With the context of Chapter 2 laid out, this chapter specifies the research objectives motivating all of the work contained herein. Section 3.1 provides a statement of the research objectives which are closely aligned with the key themes discussed in Section 2.5. The contributions include the formulation and differential game-theoretic or optimal control solution of novel adversarial scenarios, demonstration of how the solutions may be used or applied, techniques for addressing singularities, approximation techniques for games, and a framework for solving multi-stage scenarios. Sections 3.2 and 3.3 discuss the challenges and strengths of the approach. Finally, Section 3.4 contains an outline of the technical content in Parts II and III.

3.1 OBJECTIVES

The overarching goal of this work is to demonstrate a *repeatable* process for obtaining equilibrium behaviors for fundamental, atomic engagements and using them to reason about more and more complex scenarios. Thus, there is an emphasis on small team sizes (1v1, 1v2, 2v1, etc.), obtaining strategies with provable guarantees, and computability (ultimately, to make possible implementation in real-time, on-board applications). Then, the solutions are composed into scenarios involving one or more of the following elements: more agents (i.e., going from 1v1 to 2v1 to Mv1), more objectives, more stages, and the inclusion of discrete events or decision points. The following specific Research Objectives are aimed at addressing the technical challenges associated

with the process described above as well as its demonstration. Their order of appearance roughly corresponds to the order in which they're addressed in the technical content (which is organized by scenario). A summary of the objectives appears in Table 3.2 in Section 3.4.

Research Objective 1 (Approximations). Develop approximations which aid in the computability of differential game and optimal control solutions.

A single open-loop control trajectory may only be useful for a particular set of initial conditions for which the trajectory was computed. Some optimal control problems and differential games may not admit closed-form analytic solutions (i.e., feedback strategies). Often the optimal or equilibrium strategy is also a function of which region the state of the system lies. The manifolds which describe these boundaries may also be non-analytic, which adds additional computational burden in implementing the optimal or equilibrium control. This objective seeks to improve the usability of computationally burdensome solutions through the development of approximations; suitability for real-time, on-board implementation is the ultimate goal.

Research Objective 2 (Parameters). Demonstrate the importance of parameter exploration in adversarial scenarios.

Many cases were cited in Section 2.5.1 in which an exhaustive characterization of a differential game solution over the entire parameter space was undertaken. Often, however, this characterization over the parameter space is missing or incomplete.

Research Objective 3 (Dispersal). Develop techniques for computing DSs in differential games of interest.

Here the intent is to demonstrate an approach proposed by Isaacs [130] for computing a DS in a game by finding the locus of crossings of back-propagated equilibrium trajectories.

Research Objective 4 (Skirmish Utility). Demonstrate the utility of skirmish-level differential game solutions in larger scale, more complex adversarial scenarios.

Part of the purpose of this objective is to expand on what has already been done (e.g., the works cited in Section 2.5.6). Additionally, successful completion of this research objective serves to further motivate finding skirmish-level solutions for novel games. Some notional examples are the application of skirmish-level solutions i) as sub-solutions for other differential games, ii) to obtain bounds for intractable games, and iii) to design systems with provable guarantees.

Research Objective 5 (More Agents). Analyze the effect of increasing the number of agents in differential games.

Of particular interest is the cooperation of multiple Pursuers in the classical differential game of min max capture time with simple motion and point capture. This objective seeks to address the scalability and applicability of the concepts from the 1v1 and 2v1 [89, 107, 187] solutions to the same problem with three or more Pursuers.

Research Objective 6 (New Scenarios). Obtain the differential game solution (i.e., Value function and equilibrium control strategies) for new adversarial scenarios.

There is particular interest in engagements with stationary turn-constrained agents (i.e., turrets). The most notable example from the literature is the Turret Defense Differential Game (TDDG) by Akilan and Fuchs [2] wherein an Attacker moving with simple motion seeks to collide with a Turret (or Defender); an integral cost functional is specified based on the look angle of the Turret. Such a motion model for one of the agents can be seen as a special case of the more general (but far more difficult) one in which the turn-constrained agent also moves (e.g., the HCDG, G2C, and Suicidal Pedestrian Differential Game (SPDG) [68]).

Research Objective 7 (Embedded Optimal Control). Explore methods for numerically solving optimal control problems embedded within differential games.

Optimal control is required, e.g., in cases where an opponent strategy is known or assumed, and as a component of more complex scenar-

ios (like the optimal constrained retreat trajectories in [84]). Thus, various optimal control solution methodologies may or may not be appropriate depending on the application and nature of the solution. This objective seeks to compare selected methods for scenarios of interest.

Research Objective 8 (Multi-Objective). Analyze new scenarios in which agents are free to choose from multiple strategic objectives through the application and extension of [84].

To date, the engage or retreat game [83, 84, 90] remains one of the most well-studied examples of games with multiple outcomes (i.e., requiring discrete decisions to be made). Section 2.5.4 covered some additional examples. The Defender model considered in the works by Fuchs is "simple", meaning the Defender has a discrete control signal, $u(t) \in \{0,1\}$, which turns on or off an integral cost term for the Attacker. This objective is aimed at applying the engage or retreat framework to a more complex Defender model.

Research Objective 9 (Discrete Events). Demonstrate a systematic approach for solving differential games and optimal control problems which have discrete events.

Particularly within the differential game literature, there is a lack of examples containing discrete events (which naturally leads to thinking of multi-stage differential games). Section 2.5.5 mentions the most prevalent example, which pertains to sequences of similar tasks (e.g., sequential capture of multiple Evaders). Here, the goal is to consider more general events and to apply a DP-inspired approach to solving such games.

3.2 CHALLENGES OF THE PROPOSED APPROACH

The proposed approach is centered on DGT and thus most of the challenges are general challenges within existing DGT literature. They can be summarized as restrictive assumptions on available information, relatively low fidelity dynamic models, and relatively small numbers of agents.

First, as is the case in (non-stochastic) optimal control, full state information is required. Moreover, it is assumed that both players within a differential game have the same knowledge of key parameters (e.g., vehicle speeds, maximum turning rates, etc.). Even the concept of equilibrium, itself, assumes that the players know the cost/payoff functional of all other players. In the case of zero-sum differential games, this means that the players agree on the cost/payoff functional upon which they minimize and maximize, respectively. Thus, in addition to knowing an adversary's capability, one must also know the adversary's preference. Another implicit assumption in DGT is that both agents have the same model for the system and its dynamics (an assumption which was relaxed in the context of a discrete pursuit-evasion game in [141]). This is indeed the setup for the examples in [130] and for many examples that have been worked out since. Having access to partial state information would necessitate either state estimation or treatment in a stochastic framework. As mentioned in Chapter 2, POMDPs have been applied to games in discrete type [66], but POMDPs, themselves, are generally computationally intensive to solve and not known to be particularly scalable. In practice, however, DGT-based solutions may still be useful if paired with some state estimation scheme (in the case of not having full state estimation). Furthermore, the robustness guarantee of the saddle-point equilibrium applies even if one's opponent has a different dynamic model or cost functional; only the knowledge of their capability is strictly necessary. Again, if some margin w.r.t. the Value of the game is allowed, estimation may be used even for the adversary capabilities.

The second challenge is in the difficulty of solving a differential game with complex dynamics. Many examples for which DGT has been applied examine players who possess simple motion (i.e., a holonomic vehicle with single integrator kinematics). Compared with the 1v1 pursuit-evasion differential game of min max capture time with simple motion agents and faster Pursuer, the HCDG, wherein just the Pursuer is turn-constrained (i.e., a non-holonomic vehicle with double integrator kinematics) – a modest increase in the complexity of the kinematic model – is vastly more difficult to solve. The full solution of the G2C

(where both agents are turn-constrained) is considered to be an open problem. However, in lieu of working with more complex dynamics, it has been shown that solutions based on simple motion may still be of practical use. For example, in [120] the simple motion solution of the TADDG [97, 99, 184] was implemented within the Air Force simulation environment AFSIM using 3-DOF vehicle models for each agent. The simple-motion-based differential game solution was shown to outperform traditional guidance laws such as PP and PN. Simple motion is also a useful approximation in cases where the minimum turning radii of the vehicles are very small w.r.t. the initial separation distances (e.g., Beyond Visual Range (BVR) engagements, c.f., [105, 108]).

The last significant challenge of traditional DGT is that it suffers from the well-known curse of dimensionality [27]. There is one exception in what are called *potential games*. In potential games, there is an inherent decoupling among all of the players' goals and the equilibrium can be found by solving an optimal control problem for each agent [167]. However, the scenarios of interest to the present study do not possess the required structure to be potential games. Often, the addition of a single additional agent to either team can drive up the complexity of the differential game solution immensely. An excellent example is the increase in complexity in the 1v1 pursuit-evasion game of min max capture time with simple motion, faster pursuer, and point capture. The 1v1 solution is covered briefly in Section 1.1, while the two-Pursuer version (and its singularity) is the subject of references [107, 165, 183, 186, 187]. As will be shown in Chapter 7 extending to the three-Pursuer version presents even more challenges (some of which are still not fully resolved). To combat this, a many-versus-many scenario's true equilibrium solution may be approximated via decomposition (as covered in Section 2.5.6).

3.3   STRENGTHS OF THE PROPOSED APPROACH

In an effort to justify continuing with a DGT-based approach, this Chapter concludes with a discussion of the strengths. They can be summa-

rized as follows: equilibrium strategies are robust to any adversarial input, closed-form solutions are amenable to real-time, onboard implementation, solution can be characterized across many parameter combinations, and the state space partitioning can be used to address higher-level questions.

Perhaps the most important property associated with a differential game solution is the robustness of each player's equilibrium strategy. Implementation of the equilibrium strategy guarantees that a player can do no worse than the Value of the game, regardless of anything the adversary does. Therefore, there is no need to assume an adversary strategy. An extreme case of this is where the strategy of one side is optimized against a fixed opponent strategy iteratively, back and forth, until no (or very little) improvement can be made (see, e.g., [223]). There is also no need to assume what knowledge the adversary has or what rationality level their controller is based upon (see, e.g., [117, 206]). The robustness of the equilibrium strategy applies also to the case in which an opponent is playing according to a different cost functional – thus knowledge of the adversary's exact preferences are not necessary.

Next, when the differential game solution is obtained analytically in closed-form it has extremely low computational complexity. In these cases, implementation of the equilibrium strategy is suitable to implement in computation-power-constrained, real-time applications. Simple motion models yield closed-form analytic solutions to differential games much more often than more complex models thus giving them more practical value in this area. Even if the system dynamics of a real system must be simplified, the resulting simple-motion-based differential game solution provides coarse guidance information.

A differential game solution is generally obtained as a function of the system parameters (e.g., agent speeds, weapon effectiveness, turn-rate, etc.). Thus the solution may apply to a variety of different vehicles or engagements. This is in contrast with reinforcement learning and other AI approaches wherein controllers are often trained on a specific set of parameters. In those cases, it would be necessary to re-train the algorithm on each new parameter setting of interest. Even then, having

a trained controller for one or more settings of parameters does little to help a designer understand how the solution varies across a continuum of parameters. This pitfall can, perhaps, be avoided via judicious non-dimensionalization of the problem and/or proper selection of features.

Similar to the discussion on parameters, above, the differential game solution provides a characterization of the entire state space. In particular, it yields regions of guaranteed win for one or other agent. These regions would not be straight-forward to obtain if only the control policy is available – especially if that control policy is not analytic. Moreover, the worst-case *outcome* of the game is given by the Value function over the whole state space. If only the control policy is available (as in some RL-based approaches) then the outcome *for a single set of initial conditions* must be obtained through simulation. Even then, the outcome has no guarantee attached to it, as it does with DGT. These regions are often useful in task allocation schemes for multi-versus-multi scenarios – feasibility (e.g., whether a particular Pursuer *can* catch an Evader) may be given directly by the associated 1v1 win region. Even the Value associated with a 1v1 subgame may be used to assess the utility of a particular assignment.

3.4   ORGANIZATION

The technical content is, as mentioned previously, organized by scenario. All of the scenarios considered are listed in Table 3.1 along with the chapter in which the scenario is addressed. Generally, the scenarios fall under Pursuit-Evasion or Turret Defense, which comprise Parts II and III, respectively. Within each part, the beginning is focused on the smallest (in terms of number of agents) and most restrictive (in terms of assumptions) instantiation of the scenario. From there, the complexity increases from chapter to chapter as the agents have more freedom in their control strategy, more agents are added to either side, or more choices are made available to the agents.

Part II is focused on pursuit-evasion scenarios in two-dimensional, unbounded environments wherein the agents have simple motion (i.e.,

single integrator kinematics). Most of the content pertains to the question of *when* capture will occur, and all of the chapters involve scenarios with a single Evader. Game of Time refers to the zero-sum differential game of capture time wherein the Evader is the maximizer and the Pursuer is the minimizer. In the final chapter, concepts developed within the capture time chapters are applied to a border defense problem wherein the Pursuers aim to keep the Evader from reaching a goal region.

Part III is focused on two turret defense scenarios: the TDDG and the Turret Engage or Retreat (TEoR). Regarding the former, a distinction is made between two different turret models: a kinetic turret which must be aligned with its target's position to shoot at it (TDDGK) and a wide-beam turret which may affect its target to greater or lesser degree as a function of its look angle (TDDGW, as in [2]). The latter is focused on Engage or Retreat scenarios (c.f. 2.5.4) wherein the Defender agent is a Turret with bounded turn rate. When the numbers of agents are specified in Table 3.1 the convention is (number of Turrets) versus (number of Attackers).

Table 3.1: Adversarial Scenarios addressed in this dissertation

Scenario	Description	Chapter(s)
PEPP	Pursuit-Evasion against PP	5, 6
MP1E GoT	many-Pursuer Game of Time	7, 8
MP1E BD	many-Pursuer Border Defense	9
TDDGK	TDDG kinetic Turret	11
TEoRK	TEoR kinetic Turret	13
TDDGW	TDDG wide-beam Turret	14
TEoRW	TEoR wide-beam Turret	15

Table 3.2 provides the mapping of research objectives to scenarios. Many of the objectives are addressed in more than one scenario.

Table 3.2: Mapping of Research Objectives to Scenarios

RO #	Description	Scenario
1	Approximations	PEPP, TEoRW
2	Parameters	PEPP, MP1E GoT, TDDGW, TEoRK
3	Dispersal	TDDGW, TEoRK
4	Skirmish Utility	PEPP, MP1E GoT, MP1E BD, TDDGK 1v2
5	More Agents	MP1E GoT, MP1E BD, TDDGK 1v2, 2v1
6	New Scenarios	TDDGK
7	Embedded OC	TEoRW
8	Multi-Objective	TEoRW, TEoRK, TDDGK 1v2
9	Discrete Events	TEoRK, TDDGK 1v2

Part II

PURSUIT-EVASION

4

INTRODUCTION TO PURSUIT-EVASION

4.1 BACKGROUND

The subject of one object pursuing another by aiming directly at it has long captured the attention and imagination of mathematicians, engineers, and theoretical biologists alike. The work in this Part begins with a single Pursuer and a single Evader, each with specified strategies (i.e., not a differential game; not even an optimal control problem). The Evader takes a constant heading while the Pursuer employs PP. In subsequent chapters, the Evader employs a control which maximizes the time until capture against two Pursuers. It will become clear that computing the solution to the latter is far more computationally involved than the case of constant Evader heading. Later, two differential games are addressed which pit the Evader against many Pursuers: one pertaining to capture time, and one pertaining to reaching a target. Although pursuit-evasion was mentioned in Chapter 2, some additional background is given here regarding PP, the two-Pursuer game, and the many-Pursuer game.

PURE PURSUIT   Generally, the curve traced by the Pursuer's motion is referred to as a *pursuit curve* [18, 30]. In its purported original incarnation, the problem was formulated by Bouguer as a pursuer (a pirate ship) attempting to capture an evader (a merchant ship) using PP [33]. It was assumed that the Evader's motion is a straight line perpendicular to the initial LOS, i.e., a broadside shot or *abeam*. Since then, many works have been published on this problem and its many variants. Most notably, the more general case in which the Evader moves on a straight line that is *not* perpendicular to the initial LOS was only recently solved in closed form by Eliezer and Barton [64, 65]. No attempt at a comprehensive survey is made here; much of the history

and pertinent works are described in [133, 171][220, Chap. 3]. Nahin's book [171] re-examines the "classic" pursuit problem (in which the Evader's path is a straight line) as well as considering more complicated Evader paths (such as a circle) and cyclic pursuit (in which ring of Pursuers pursue one another). Kamimura and Ohira's book [133] contains many of the same problems as well as some three-dimensional variants (such as helical Evader paths); it also contains chapters on collective motion and group pursuit-evasion, which heavily emphasize simulation over rigorous analysis. PP is the subject of a full chapter in [220]; the book, which gives a general treatment of missile guidance laws, illustrates one of the important physical applications of pursuit curves, namely air-to-air combat. One very recent, noteworthy contribution to this field is [110] wherein Gard proved that the Evader can lead the Pursuer to any point in $\mathbb{R}^n$ while maintaining some desired minimum separation distance for the case of equal speeds.

TWO PURSUERS   In the Two Cutters and Fugitive Ship differential game [130], the Pursuers, acting as a team, wish to capture the Evader in the shortest time possible, while the Evader, as in this case, wishes to delay capture, thus making it a zero-sum differential game. Isaacs' treatment of the problem was geometrical, based on the fact that all three agents' saddle point equilibrium strategies result in straight-line paths [130]. The geometric solution to the game was then validated in [107] by expressing the Value function associated with the geometric solution analytically and showing that it satisfies the HJI equation. Reference [183] sought to show the applicability of the two-pursuer one-evader solution to scenarios with more than two pursuers. In [187], the solution to the Game of Kind (i.e., whether capture by the first, or second, or both Pursuers simultaneously) is given using a reduced state space. The consequences of the DS in the game were considered in [186]. The case of a finite capture radius was then considered in [89, 188]. In the former, the solution is based on a system of two nonlinear equations obtained by analytic retrograde integration of the optimal dynamics; this solution also covers the case of Pursuers with different

speeds. The latter provides a closed-form solution for the case where all three agents have the same speed.

MANY PURSUERS   References [95] and [128] looked at the case of multiple Pursuers seeking to capture an Evader. The concept of a dynamic Voronoi diagram was used to define the closed domain in which Evader capture occurred by these several Pursuers. Reference [63] finds the intercept set of Evaders, assuming that the Evaders' goals were known to the Pursuers. A related work, where the objective is to rescue certain agents with interference of obstacles was modeled in the Prey, Protector, and Predator game [181]. A full solution to the M-on-N (or MPNE) game, even for the case of simple motion on an unbounded plane, is of interest; however, due to the curse of dimensionality, this is seemingly intractable. The same is still true for the case of MP1E, and much of the existing literature has focused on numerical techniques and/or sequential pursuit (as in [16, 231]). Note that *many* other works have been written on the subject of multiple Pursuers and a single Evader (c.f. 2.4.2). The works vary in terms of what motion model is used, the cost functional specified, the type of solution sought, and variations in the environment. For example, some have considered obstacles and bounded environments [46, 128, 181, 182]. Alternatively, so-called games of approach, wherein the Evader is faster than the Pursuers, use the payoff/cost of minimum distance to the Evader achieved by any Pursuer [190].

4.2   PRACTICAL SIGNIFICANCE

The application of differential game theory to problems of military interest is immediately obvious in the context of pursuit-evasion games. One could imagine that the solutions to simple pursuit-evasion games in the plane may one day be expanded to equilibrium strategies for a real-world aerial dog-fighting scenario. For now, the scenarios explored in this part are most closely related to air-to-air, BVR engagements taking place between kinetic munitions and/or UAVs. The prac-

tical implication of BVR is that the minimum turning radius of the vehicles may be negligible w.r.t. the initial engagement distances and thus may be disregarded. What is important, then are the reference course angles, or headings, taken by the agents throughout the engagement. There is also some merit to focusing on multiple Pursuers and one Evader as multiple missiles are regularly fired to increase the probability of kill (see, e.g. [143]).

Concerning the last chapter on target guarding, border surveillance and defense is a task for which UAVs are well-suited to carry out as it requires persistent vigilance. Additionally, their relative speed compared to ground vehicles and the possibility of varied terrain points to the utility of UAVs for this purpose. One possible scenario pertaining to border surveillance and defense is that of an intruder, perhaps performing surveillance for the enemy, which, having gathered some information, is now attempting to escape back to safety through the border to deliver the acquired information. Thus there is interest in engaging the intruder with a number of UAVs with the intent of capturing the intruder or at least destroying it within the border. Conceptually, the border may represent a physical border and thus capture of a vehicle outside the border could be considered an act of war, or perhaps the border represents the sensor range, outside of which the intruder's position can not be tracked. The goal, then, is for the UAVs to cooperate in capturing the intruder based on their positions at the time the laser fence is crossed by the intruder.

4.3 CONTRIBUTIONS

The contributions of the chapters in this part may be summarized as

- PEPP capture time for non-point capture

- PEPP simultaneous capture solution for 2 Pursuers

- PEPP approximations and capture time bounds

- PEPP optimal control solution

- MP1E GoT computationally efficient geometric policy with Evader performance guarantees

- MP1E GoT robust pursuit policy with Pursuer performance guarantees

- MP1E GoT, MP1E BD process for discarding redundant Pursuers

As indicated in Table 3.2, these contributions pertain to Research Objectives 1, 2, 4, and 5 – namely, approximations for computationally-intensive solutions, characterization of solutions over the parameter space, utilization of skirmish-level solutions in more complex scenarios, and adding more agents into the scenarios. Lastly, although not included in this part, Appendix B contains an example of the solution of the 1v1 game of min max capture time being utilized within a larger two-Pursuer, many-Evader scenario, further supporting Research Objective 4.

5

PURE PURSUIT AND CONSTANT EVADER COURSE

5.1 INTRODUCTION

This chapter is the first pertaining to pursuit-evasion against PP (PEPP) and it covers the case in which the Evader implements a constant heading (i.e., holds course). The results partly address Research Objectives 1, 2, and 5 by providing bounds for capture times, approximations of necessary and sufficient conditions for simultaneous capture, considering all ranges for the main parameters (speed ratio, capture radius, and initial distance), and considering more Pursuers. This material is based upon the paper [252], which is in preparation.

Much of the work on PEPP has been focused on the case in which the Evader's path is a straight line and capture occurring when the Pursuer and Evader are coincident (point capture) [18, 64, 65, 220]. This chapter continues in a similar vain but with a focus on non-point capture; that is, the Pursuer may effect capture within some specified distance. This chapter pertains to pursuit-evasion scenarios taking place in an unbounded, obstacle-free two-dimensional environment, i.e., the realistic plane. The agents' speeds are fixed and non-zero. It is assumed throughout that the Evader implements a constant heading in the inertial (global) Cartesian frame – in some cases, the Evader heading is assumed to be given, and in others it may be considered to be a decision made at the initial time instant. The Pursuer(s) strategy is PP; in the case of multiple Pursuers, the speeds are assumed to be equal, however the results in this chapter can be readily extended to the unequal speed case and/or the unequal effector range case. The Pursuer(s) may be endowed with an effector of radius l. Table 5.1 lists the different scenarios considered; μ is the ratio of Evader and Pursuer speeds and d_0 is the initial distance between the agents.

Table 5.1: Taxonomy of Pursuit-Evasion Scenarios

#	Setting			Description
1	$\mu < 1$	$l = 0$	$d_0 > 0$	Point Capture
2	$\mu < 1$	$l > 0$	$d_0 > l$	Capture (slow Evader)
3	$\mu > 1$	$l > 0$	$d_0 > l$	Capture (fast Evader)
4	$\mu > 1$	$l > 0$	$d_0 < l$	Escape from surveillance

The contributions of this chapter are as follows (with parentheses indicating to which Scenario(s) the result applies). Aspect angle refers to the angle that the Evader's heading makes w.r.t. the Pursuer's LOS.

- Solution for the Pursuer's separation distance as a function of aspect angle (1–4)

- Analytic determination of head-on versus tail-chase final configuration (2, 3)

- Necessary and sufficient condition for capture/evasion (3)

- Solution for capture time (2, 3) and escape time (4)

- Proof of the existence of a solution for simultaneous capture (1)

- Necessary condition for simultaneous capture (1, 2) and escape (4)

- Sufficient condition for simultaneous capture (2)

Insomuch as is possible, the results are closed-form, analytic expression. Otherwise, attention is given to numerical implementation. Note that the results herein also apply to the interesting case of equal speeds.

The remainder of the chapter is organized as follows. Section 5.2 specifies the problem setup. Section 5.3 provides the analysis leading to the determination of the final configuration. Section 5.4 contains the results pertaining to final time (i.e., when capture or escape occurs). Section 5.5 examines the case where there are two Pursuers wherein simultaneous capture (or escape) is possible. Finally, Section 5.6 contains the conclusion.
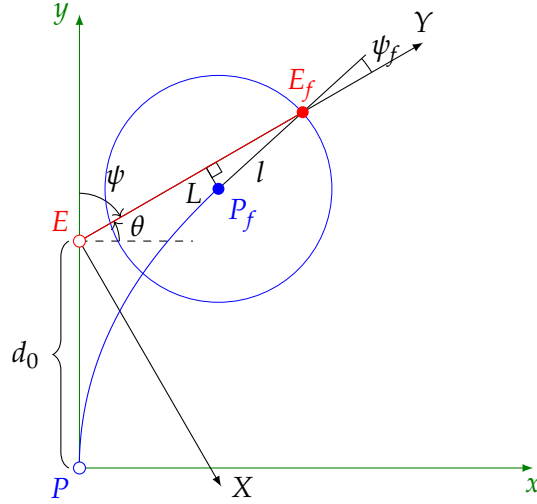
Figure 5.1: Schematic illustration of the scenario corresponding to Scenario 2: slow Evader and finite capture radius.

5.2 PRELIMINARIES

Let the agents' positions be specified by $E \equiv (x_E, y_E)$ and $P \equiv (x_P, y_P)$, $E, P \in \mathbb{R}^2$. In the case that there are $M > 1$ Pursuers, the ith Pursuer position is denoted P_i, where $i \in 1, \ldots, M$. Without loss of generality, the Pursuer(s') speed is fixed to 1 and the Evader's speed is μ. Additionally, w.l.o.g., consider a Cartesian frame (x, y) whose origin is the Pursuer's initial position and whose positive y-axis is aligned with the line segment $\overline{PE}$ at initial time (the green axes in Fig. 5.1)[1]. The kinematics, which arise from P employing the strategy of PP are

$$\begin{bmatrix} \dot{x}_E \\ \dot{y}_E \\ \dot{x}_P \\ \dot{y}_P \end{bmatrix} = \begin{bmatrix} \mu \cos(\theta) \\ \mu \sin(\theta) \\ \frac{x_E - x_P}{d} \\ \frac{y_E - y_P}{d} \end{bmatrix} \tag{5.1}$$

where $d \equiv \sqrt{(x_E - x_P)^2 + (y_E - y_P)^2}$ is the instantaneous distance between the agents, and θ is the Evader's heading angle w.r.t. the positive x-axis. The Pursuer's effector range is l – in the case of capture, this corresponds to the capture radius; in the case of surveillance, this corresponds to observation range. Final time, t_f, is defined as the time at

---

1 We utilize this convention in some portions of the text to match the work of [18]

which the distance between P and E is equal to the effector range, i.e., $d(t_f) = l$.

The rotated Cartesian frame, (X, Y), (shown in black in Fig. 5.1) is used here in the expression of the so-called *pursuit curve* [18]. In this frame, the origin is the Evader's position at the initial time and the Y-axis is aligned with the Evader's (constant) direction of motion, or course. The kinematics in this frame are

$$
\begin{bmatrix} \dot{X}_E \\ \dot{Y}_E \\ \dot{X}_P \\ \dot{Y}_P \end{bmatrix} = \begin{bmatrix} 0 \\ \mu \\ -\dfrac{X_P}{\sqrt{X_P^2 + (Y_E - Y_P)^2}} \\ \dfrac{Y_P}{\sqrt{X_P^2 + (Y_E - Y_P)^2}} \end{bmatrix}.
\tag{5.2}
$$

Two useful expressions are given in [18], namely the tangent function

$$
\frac{dY_P}{dX_P} = \frac{1}{2} \left(w \left(\frac{X_P}{X_{P_0}} \right)^{\mu} - \frac{1}{w} \left(\frac{X_P}{X_{P_0}} \right)^{-\mu} \right),
\tag{5.3}
$$

and the solution of P's trajectory:

$$
Y_P(X_P) = \frac{\mu X_{P_0}(1 + \mu \sin \theta)}{\cos \theta (1 - \mu^2)} +
$$
$$
\frac{X_{P_0}}{2} \left[\frac{w}{1 + \mu} \left(\frac{X_P}{X_{P_0}} \right)^{1+\mu} - \frac{1}{w(1 - \mu)} \left(\frac{X_P}{X_{P_0}} \right)^{1-\mu} \right],
\tag{5.4}
$$

where $w \triangleq \frac{1 - \sin \theta}{\cos \theta}$ and $X_{P_0} \equiv X_P(t_0) = d_0 \cos \theta$. Note that, at all time, the tangent to the Pursuer's trajectory points to E, by construction [18, 220]. The transformation to the (x, y) frame is

$$
\begin{bmatrix} x_P \\ y_P \end{bmatrix} = \begin{bmatrix} \sin \theta & \cos \theta \\ -\cos \theta & \sin \theta \end{bmatrix} + \begin{bmatrix} 0 \\ d_0 \end{bmatrix}.
\tag{5.5}
$$

Lastly, consider a rotating frame which is defined by the distance from P to E and E's heading w.r.t. P's LOS. The kinematics are

$$
\begin{bmatrix} \dot{d} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} \mu \cos \psi - 1 \\ -\dfrac{\mu}{d} \sin \psi \end{bmatrix}
\tag{5.6}
$$

where, w.l.o.g., $\psi \in [0, \pi]$. The term ψ is referred to as the aspect angle, and is generally not constant even though the Evader's heading is constant in the inertial frame (unless $\psi = 0$). It is clear, with the

convention for ψ, that $\dot{\psi} < 0$ for all t since $\sin \psi > 0$ and $d > 0$. This monotonicity holds whether the Evader is slow or fast. Regarding distance, d, if the Evader is slow ($\mu < 1$) then $\dot{d} < 0$ for all t which means the Pursuer is always getting closer; but if the speed ratio $\mu > 1$ then there is a range of ψ for which $\dot{d} > 0$.

Proposition 5.1. *A Pursuer remains on the same half-plane in which it started w.r.t. the Evader's heading and initial position.*

Proof. In order for the Pursuer to cross into the opposite half-plane there must be a time at which $\psi = 0$. It is clear from (5.6) that $\dot{\psi} = 0$ and thus the Pursuer will remain in a trailing configuration as long as the Evader holds course (i.e., P never crosses into the opposite half-plane). $\qquad\square$

It is for this reason that the range $\psi \in [0, \pi]$ is sufficient.

5.3 FINAL CONFIGURATION

Before addressing the determination of final time for a particular problem instance, it is important to determine whether the Evader has a component of velocity *towards* or *away* from the Pursuer at final time, i.e., whether the scenario ends in a head-on or tail-chase configuration. To do this, the solution of (5.6) is obtained. First, divide the two differential equations:

$$\frac{\mathrm{d}d}{\mathrm{d}\psi} = \frac{\dot{d}}{\dot{\psi}} = \frac{\mu \cos \psi - 1}{-\frac{\mu}{d} \sin \psi}$$
$$= \frac{d\left(\frac{1}{\mu} - \cos \psi\right)}{\sin \psi}, \tag{5.7}$$

which can be simplified as follows

$$\frac{1}{d}\mathrm{d}d = \left(\frac{1}{\mu \sin \psi} - \frac{1}{\tan \psi}\right)\mathrm{d}\psi$$

$$\int_{d_0}^{d_f} \frac{1}{d}\,\mathrm{d}d = \int_{\psi_0}^{\psi_f} \frac{1}{\mu \sin \psi}\,\mathrm{d}\psi - \int_{\psi_0}^{\psi_f} \frac{1}{\tan \psi}\,\mathrm{d}\psi$$

$$[\ln d]_{d_0}^{d_f} = -\frac{1}{\mu}\left[\ln \left(\cot \psi + \csc \psi\right)\right]_{\psi_0}^{\psi_f} - \left[\ln \left(\sin \psi\right)\right]_{\psi_0}^{\psi_f}$$

$$\ln d_f - \ln d_0 = -\frac{1}{\mu}\ln \left(\cot \psi_f + \csc \psi_f\right)$$

$$+ \frac{1}{\mu}\ln \left(\cot \psi_0 + \csc \psi_0\right)$$

$$- \ln \sin \psi_f + \ln \sin \psi_0.$$

Taking the exponential of both sides gives

$$e^{\ln d_f - \ln d_0} = e^{-\frac{1}{\mu}\ln\left(\cot \psi_f + \csc \psi_f\right) + \frac{1}{\mu}\ln\left(\cot \psi_0 + \csc \psi_0\right) - \ln \sin \psi_f + \ln \sin \psi_0},$$

which becomes

$$\frac{e^{\ln d_f}}{e^{\ln d_0}} = e^{-\frac{1}{\mu}\ln\left(\cot \psi_f + \csc \psi_f\right)} e^{\ln \frac{1}{\mu}\left(\cot \psi_0 + \csc \psi_0\right)} \frac{e^{\ln \sin \psi_0}}{e^{\ln \sin \psi_f}}$$

$$\frac{d_f}{d_0} = \left(\cot \psi_f + \csc \psi_f\right)^{-\frac{1}{\mu}} \left(\cot \psi_0 + \csc \psi_0\right)^{\frac{1}{\mu}} \frac{\sin \psi_0}{\sin \psi_f},$$

which can be rearranged to

$$d_f \left(\cot \psi_f + \csc \psi_f\right)^{\frac{1}{\mu}} \sin \psi_f = d_0 \left(\cot \psi_0 + \csc \psi_0\right)^{\frac{1}{\mu}} \sin \psi_0. \qquad (5.8)$$

This expression is the solution of the system (5.6). Given the initial and final aspect angle and either distance, (5.8) yields a closed-form, analytic solution for the unknown distance. Unfortunately, if either ψ_0 or ψ_f is unknown, then (5.8) must be solved numerically.

Proposition 5.2. *The curve*

$$d_\perp \left(\psi; l, \mu\right) = l \left(\cot \psi + \csc \psi\right)^{-\frac{1}{\mu}} \csc \psi \qquad (5.9)$$

partitions the state space into a region of Pursuer initial positions, (d, ψ), which end in head-on collision $(\cos \psi_f < 0)$, and a region for which the scenario ends in a tail-chase $(\cos \psi_f > 0)$. More precisely:

$$\text{sign}\left(\cos \psi_f\right) = \begin{cases} -1 & \text{if } d < d_\perp \\ \text{undef.} & \text{if } d = d_\perp \\ 1 & \text{if } d > d_\perp. \end{cases} \tag{5.10}$$

Proof. Substituting $d_f = l$ and $\psi_f = \frac{\pi}{2}$ into (5.8) immediately gives (5.9). This configuration corresponds to the Evader's heading being perpendicular to the Pursuer's at the moment that $d \to l$, which is the case that is neither head-on collision nor tail-chase. If $\psi_f = \frac{\pi}{2}$ then $\cos \psi_f = 0$, hence $\text{sign}\left(\cos \psi_f\right)$ is undefined as in (5.10). The remainder of the proof is broken into two cases depending on the magnitude of μ.

SLOW EVADER   in this case, $\mu < 1$ and both ψ and d are monotonically decreasing w.r.t. time. Thus, if, for the same ψ_0, $d_0 < d_\perp$, the Pursuer must take a shorter time to capture the Evader. Since ψ is monotonic, it must be the case that $\psi_f > \frac{\pi}{2}$ which implies that $\cos \psi_f < 0$. The case where $d_0 > d_\perp$ follows by similar logic.

FAST EVADER   in this case, $\mu > 1$ and only ψ is monotonically decreasing w.r.t. time. From (5.6), it is clear that

$$\text{sign}(\dot{d}) = \begin{cases} -1 & \text{if } \cos \psi < \frac{1}{\mu} \\ 0 & \text{if } \cos \psi = \frac{1}{\mu} \\ 1 & \text{if } \cos \psi > \frac{1}{\mu}. \end{cases}$$

Therefore, the distance becomes monotonically *increasing* once $\cos \psi > \frac{1}{\mu}$. Capture, if it occurs at all, must take place while $\cos \psi < \frac{1}{\mu}$. Thus capture can only occur in the portion of the trajectory for which d is monotonically decreasing. Because of this, the same logic used to prove the case where $\mu < 1$ can be applied to this case as well.   □

Remark 1. For the escape from surveillance scenario (Scenario #4), the final configuration is always tail-chase because for E to escape the observation range of P it must have $d = l$ and $\dot{d} > 0$, which means $\cos \psi > \frac{1}{\mu} > 0$.

Proposition 5.3. *For the fast Evader capture scenario ($\mu > 1$, $d > l$), the curve*

$$d_{c/e}(\psi; l, \mu) = l \left(\cot \psi^\dagger + \csc \psi^\dagger\right)^{\frac{1}{\mu}} \sin \psi^\dagger \cdot (\cot \psi + \csc \psi)^{-\frac{1}{\mu}} \csc \psi, \tag{5.11}$$

where

$$\psi^\dagger \triangleq \cos^{-1} \frac{1}{\mu}, \tag{5.12}$$

partitions the state space into a region of Pursuer initial positions, (d, ψ), which end in capture and a region for which evasion occurs. In the latter case, $t_f \to \infty$, $\psi \to 0$, and $d \to \infty$. Stated differently, $d < d_{c/e}$ is a necessary and sufficient condition for capture.

Proof. Two things are necessary for capture: the first is that the distance from E to P is equal to the latter's effector range, l, and the second is that P is actually *closing* in on E (i.e., $\dot{d} < 0$). If $d = l$ but $\dot{d} \geq 0$ then the Evader is escaping since d will increase monotonically from there on out. This is akin to Isaacs' notion of the *Usable Part* (UP) of the terminal surface [130].

The limiting case for escape occurs when $\dot{d} = 0$, which, from (5.6) occurs when $\psi = \psi^\dagger$. For any $\psi < \psi^\dagger$ it would be the case that $\dot{d} > 0$. Then (5.11) follows directly from substituting $d_f = l$ and $\psi_f = \psi^\dagger$ into (5.8). The fact that $d < d_{c/e}$ follows from monotonicity arguments akin to those in the proof of Proposition 5.2. □

Proposition 5.4. *Ignoring termination, the minimum inter-agent distance in the fast Evader case ($\mu > 1$) is*

$$d_{\min}(d, \psi; \mu) = d \left(\cot \psi + \csc \psi\right)^{\frac{1}{\mu}} \sin \psi \cdot \left(\cot \psi^\dagger + \csc \psi^\dagger\right)^{-\frac{1}{\mu}} \csc \psi^\dagger, \tag{5.13}$$

if $\psi > \psi^\dagger$, and $d_{\min} = d$ otherwise.

Proof. Eq. (5.13) is obtained by substituting $d_f = d_{\min}$ and $\psi_f = \psi^\dagger$ into (5.8). This corresponds to the minimum distance because, from (5.6) $\dot{d} = 0$ when $\psi = \psi^\dagger$ and the fact that d is monotonically decreasing while $\psi > \psi^\dagger$ and monotonically increasing while $\psi < \psi^\dagger$. If $\psi < \psi^\dagger$, then d is monotonically decreasing and thus the minimum distance is the current distance. □

Corollary 5.1. *For the fast Evader capture scenario ($\mu > 1$, $d > l$), the capture condition, (5.11), is equivalent to $d_{\min} < l$.*

Proof. The result follows directly from Propositions 5.2–5.4. □

Following are some examples which illustrate the aforementioned curves and regions for both the slow and fast Evader scenarios. In all of these examples, the Evader's heading is aligned with the positive x-axis. Fig. 5.2 shows the curve $d_\perp$ for a slow Evader along with example tail-chase and head-on Pursuer trajectories. Then Fig. 5.3 shows a similar plot, but for a fast Evader; additionally, the curve $d_{c/e}$ is shown. In the case where P begins outside the capture region (labeled "Miss Example"), the trajectory is shown up to the time at which $d = d_{\min}$. Finally, Fig. 5.4 shows a map of $d_{\min}$ as a function of P's initial position.
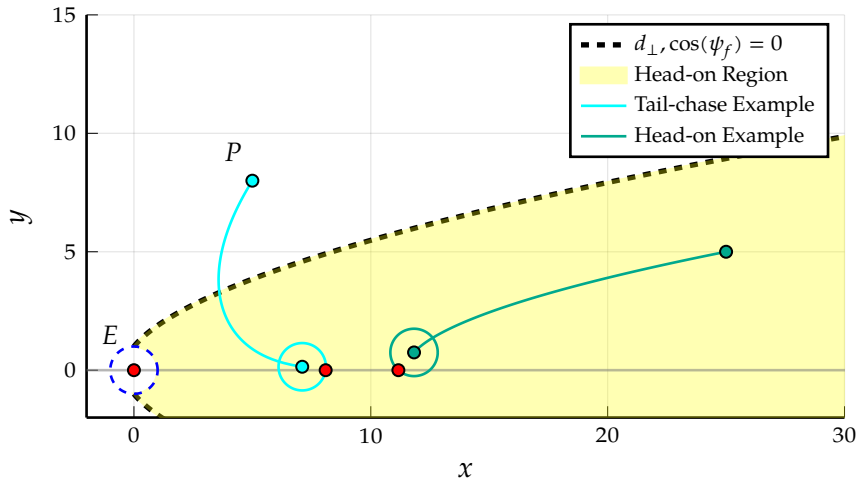
Figure 5.2: Slow Evader example; $\mu = 0.8$, $l = 1$.
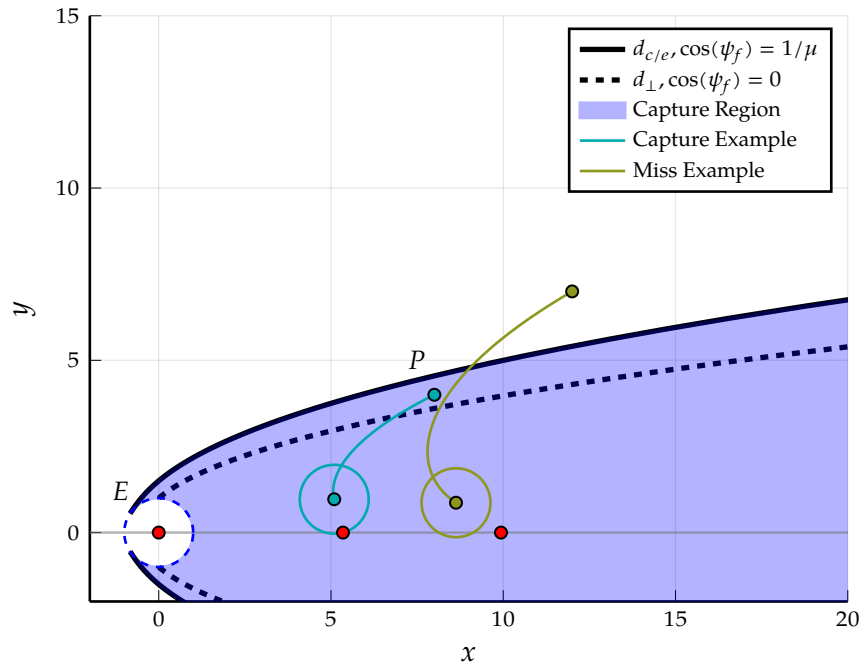
Figure 5.3: Fast Evader example; $\mu = 1.2$, $l = 1$.
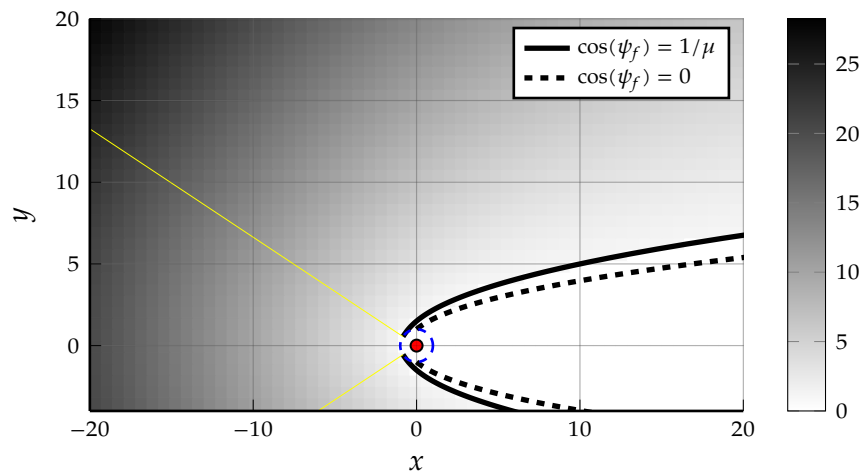
Figure 5.4: Miss distance as a function of P's initial position; $\mu = 1.2$, $l = 1$. The yellow lines mark out a cone inside which $d_{\min} = d_0$.

5.4 FINAL TIME

In this section, the final time, t_f, associated with each of the scenarios is characterized. The final configuration is an important component in the determination of the final time in the case of non-point capture ($l > 0$). For point capture, however, there is no need to make a distinction between head-on and tail-chase; indeed the capture time for a slow, fixed-course Evader and point capture is well-known, appearing in [18, 152, 220] and many other places. Given the pursuit curve in (5.4), the final time associated with Scenario 1 is found by substituting in $X_{P_f} = 0$ and $Y_{E_f} = Y_{P_f}$ and noting that E travels from $(X_E, Y_E) = (0,0)$ to $(0, Y_{E_f})$ in t_f time at speed μ. Thus,

$$
\begin{aligned}
t_f &= \frac{Y_{E_f}}{\mu} = \frac{Y_P (X_P = 0)}{\mu} \\
&= \left(\frac{1 + \mu \sin \theta}{1 - \mu^2} \right) d
\end{aligned}
\tag{5.14}
$$

In Scenarios 2–4 the Pursuer's effector range is $l > 0$. Let the final lateral separation in the Evader-aligned frame, (X, Y), be defined as $X_{P_f} \equiv L$ (see Fig. 5.1). As in every other point along P's trajectory, the

line-of-sight $\overline{P_f E_f}$, is tangent to the pursuit curve, $Y_P(X_P)$ [220]. This observation leads to a derivation for L as follows:

$$\tan \psi_f = \left.\frac{dX_P}{dY_P}\right|_{t=t_f}$$

$$\cot \psi_f = \left.\frac{dY_P}{dX_P}\right|_{t=t_f}$$

$$\frac{\sqrt{1 - \sin^2 \psi_f}}{\sin \psi_f} =$$

$$\frac{l\sqrt{1 - \frac{L^2}{l^2}}}{L} =$$

$$l^2 \left(1 - \frac{L^2}{l^2}\right) = L^2 \left(\left.\frac{dY_P}{dX_P}\right|_{t=t_f}\right)^2$$

$$l^2 - L^2 =$$

$$\implies L = \frac{l}{\sqrt{1 + \left(\left.\frac{dY_P}{dX_P}\right|_{t=t_f}\right)^2}}. \tag{5.15}$$

Evaluating (5.3) at $X_P = L$ and substituting into the above gives

$$L = \frac{l}{\sqrt{1 + \frac{1}{4}\left[w^2 \left(\frac{L}{X_{P_0}}\right)^{2\mu} + \frac{1}{w^2} \left(\frac{L}{X_{P_0}}\right)^{-2\mu} - 2\right]}}$$

$$= \frac{l}{\sqrt{\frac{1}{4}\left[w^2 \left(\frac{L}{X_{P_0}}\right)^{2\mu} + \frac{1}{w^2} \left(\frac{L}{X_{P_0}}\right)^{-2\mu} + 2\right]}}$$

$$\implies L = \frac{2l}{w \left(\frac{L}{X_{P_0}}\right)^{\mu} + \frac{1}{w} \left(\frac{L}{X_{P_0}}\right)^{-\mu}} \tag{5.16}$$

Proposition 5.5. *For $d > l > 0$, $\mu < 1$ (i.e., Scenario 2), (5.16) has a unique fixed point on the interval $L \in [0, \min\{d\cos\theta, l\})$.*

Proof. Concerning the range for the solution, it must be the case that $L < l$ because $L > l$ would require P being further from E than the capture radius. Also, it must be the case that $L = X_{P_f} < X_P = d\cos\theta$ because $X_P(t)$ is monotonically decreasing (which follows from the fact that P is employing the strategy PP) and $d > l$. Two conditions were used to derive (5.16): (*i*) $d = l$ and (*ii*) the tangency equation,

(5.3) is satisfied. The latter is true by construction, since P is always aiming at E's instantaneous position. As for the former, if $d > l$, then at some point it must be the case that $d = l$ because d is monotonically decreasing when $\mu < 1$ (see (5.6)). Additionally, there can only be one such instance, again, because of d's monotonicity. Therefore, a solution (and fixed point) to (5.16) exists and is unique. □

Proposition 5.6. *For $d > l > 0$, $\mu > 1$ (i.e., Scenario 3), (5.16) has a unique fixed point on the interval $L \in [0, \min\{d\cos\theta, l\})$ if and only if $d < d_{c/e}$.*

Proof. When $d < d_{c/e}$, capture must occur due to Proposition 5.3. The distance between P and E must be monotonically decreasing along the entire trajectory since, in order for capture to be possible, $\psi > \psi^\dagger$ for all t. Therefore, by similar arguments as in the preceding proof, the premise must be true. □

Proposition 5.7. *For $d < l$, $\mu > 1$ (i.e., Scenario 4), (5.16) has a unique fixed point on the interval $L \in [0, \min\{d\cos\theta, l\})$.*

Proof. As in Scenario 3, d is monotonically decreasing w.r.t. time while $\psi > \psi^\dagger$ and monotonically increasing thereafter. Since $d < l$, eventually there must be a time at which $d = l$, and this instance must be unique, and, as in previous proofs, this implies that (5.16) is satisfied. □

Remark 2. When $0 < L \ll 1$, (5.16) suffers from some numerical instability due to the $(L/X_0)^{-\mu}$ term. This is often the case for Scenario 4, especially when μ is close to 1. In this case, it is generally easier to numerically solve (5.8) for ψ_f and then compute $L = l\sin\psi_f$.

In the general case, since the solution of L exists and is unique, any root finding method is suitable for its computation. However, in some special circumstances, as shown in the following, the problem structure may be further exploited to aid in the computation of L.

Proposition 5.8. *If either (i) $d > l$ and E can be captured or (ii) $d < l$ with $\mu > 1$, and the speed ratio μ is a rational number (that is, $\mu = \frac{r}{q}$ where $r, q \in \mathbb{N}^+$), then the solution of (5.16) may be obtained via the rooting of a polynomial.*

Proof. Define $\mathscr{L} \equiv \left(\frac{L}{X_{P_0}} \right)^{\frac{1}{q}}$ and substitute into (5.16)

$$\mathscr{L}^q X_{P_0} = \frac{2l}{w\mathscr{L}^r + \frac{1}{w}\mathscr{L}^{-r}}$$

$$\implies \mathscr{L}^{q+r} X_{P_0} w + \mathscr{L}^{q-r} X_{P_0} \frac{1}{w} - 2l = 0. \qquad (5.17)$$

Thus $\mathscr{L}$ may be obtained as the solution of a $(q + r)$th-order (sparse) polynomial, and the associated final lateral separation is $L = \mathscr{L}^q X_{P_0}$.

$\square$

If, for example, the speed ratio is $\mu = \frac{1}{2}$, it comes down to solving a cubic equation.

Proposition 5.9. *For $d > l > 0$ (i.e., Scenarios 2 and 3), the time to capture (if E can be captured) is*

$$t_f = \frac{1}{\mu} \left(Y_P(L) + \text{sign}\,(d - d_\perp)\, \sqrt{l^2 - L^2} \right), \qquad (5.18)$$

where L is obtained as the solution of (5.16), and Y_P is given by (5.4).

Proof. As in the point capture case, the final time is found by substituting the final conditions into the pursuit curve, (5.4), and observing that E traverses a distance Y_{E_f} in t_f time:

$$t_f = \frac{Y_{E_f}}{\mu} = \frac{1}{\mu} \left(Y_{P_f} \pm \sqrt{l^2 - L^2} \right).$$

The $\pm$ in the above expression is due to the fact that $Y_{E_f} - Y_{P_f}$ can be positive or negative. In the case of tail-chase, $Y_{E_f} - Y_{P_f} > 0$, and *vice versa* for head-on collision. Recall, from Proposition 5.2 (particularly (5.10)), that the question of final configuration is determined by whether $d \lessgtr d_\perp$. Hence, the sign $(d - d_\perp)$ term in (5.18) yields the proper sign for $Y_{E_f} - Y_{P_f}$.

$\square$

Proposition 5.10. *For $d > l > 0$ (i.e., Scenarios 2 and 3), the time to capture (if E can be captured) is bounded above by*

$$\bar{t} = \begin{cases} (5.14) & \text{if } \mu < 1 \\ \frac{1}{\mu}\left(Y_P\left(d_{\min}\sin\psi^\dagger\right) + \text{sign}(d - d_\perp)\frac{d_{\min}}{\mu}\right) & \text{if } \mu > 1 \end{cases}, \qquad (5.19)$$

where Y_P is given by (5.4), $d_{\min}$ is given by (5.13), and $\psi^\dagger$ is given by (5.12).

Proof. For $\mu < 1$, the distance, d, is always decreasing monotonically, so since $l > 0$, the time to capture (i.e., (5.18)) must be less than the time to drive $d \to 0$, which is given by the point capture time, (5.14). For $\mu > 1$, the distance, d, is always decreasing monotonically until $\psi = \psi^\dagger$ (as described in the proof of Proposition 5.4). The premise is that capture is possible, so $d_{\min} < l$ from Corollary 5.1, so, again, the time to capture must be less than the time to drive $d \to d_{\min}$. □

This upper bound on capture time may be especially useful in the context of engagements between many agents wherein computational complexity becomes a concern. For example, a particular Evader may only need to consider the one or two Pursuers with the smallest upper bound, after which a more precise computation may be performed (similar to how certain Pursuers could be ignored in [245]). Additionally, for $\mu < 1$, the minimum upper bound, $\min_i \bar{t}_i$, provides an upper bound on the value of the game of min max capture time (c.f. [253]).

Corollary 5.2. *For $d < l$, $\mu > 1$ (i.e., Scenario 4), the time to capture is*

$$t_f = \frac{1}{\mu}\left(Y_P(L) + \sqrt{l^2 - L^2}\right), \qquad (5.20)$$

where L is obtained as the solution of (5.16), and Y_P is given by (5.4).

Proof. The proof is similar to the preceding proofs. Eq. (5.20) is a specialization of (5.18) since the final configuration is always tail-chase for Scenario 4. This is because, for escape, it must be the case that $\dot{d}_f > 0$, otherwise E is *entering* P's effector range. Since $\mu > 1$, it must be the case that $\psi_f < \psi^\dagger$ in order for $\dot{d}_f > 0$, and since $\psi^\dagger < \frac{\pi}{2}$, this must be a tail-chase configuration. □

Using (5.18), the capture loci (i.e., the final Evader position, $E_f(\psi_0)$) are plotted for various effector ranges, l, in Fig. 5.5. As expected, as the Pursuer's effector range increases, the locus shrinks towards E's initial position. The curve corresponding to $l = 0$ is obtained analytically via (5.14). A similar figure appears in [220, Fig. 3.6], but for point capture with varying speed ratios.
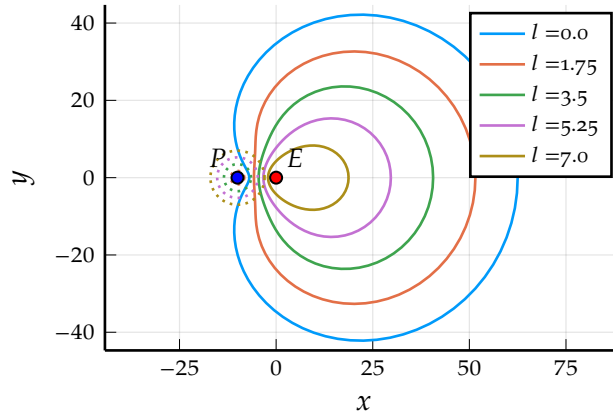
Figure 5.5: Capture loci (final E positions) for constant-heading Evader trajectories for varying effector ranges, l; here $\mu = 0.8$.

5.5 SIMULTANEOUS CAPTURE AND ESCAPE

In this section, a second Pursuer is considered whose speed is equal to that of the first Pursuer[2]; the problem setup is shown in Fig. 5.6 (note that $\alpha \in \left[0, \frac{\pi}{2}\right]$). For Scenarios 1 and 2, it is assumed that E wishes to delay capture for as long as possible while employing a constant-heading strategy. The point capture version of this problem has been analyzed, e.g., in [152]. A finite-capture-range version was treated in [250], however, there, the Evader was not constrained to a constant heading and its optimal control was computed (i.e., $\max_{\theta(t)} t_f$). For capture by a single Pursuer, E's optimal heading was indeed constant (PE, in fact). However, in the case of simultaneous capture by both Pursuers, E's optimal heading was not constant (except when E's position is on the bisector of the angle $\angle P_1 E P_2$). The solution developed here thus pro-

---

2 Unequal Pursuer speeds can also be handled by keeping track of, e.g., μ_1 and μ_2 throughout the derivation, and similarly for unequal effector ranges via l_1 and l_2.

vides a lower bound for the optimal capture time, which, as will be shown, requires far less computation than the optimal, in general. Note that the solution for the differential game version of the two-Pursuer problem with finite capture range is given in [188]. For Scenario 3, conditions for the existence of an evasive heading are developed, which readily extend to the case of many Pursuers. For Scenario 4, it is assumed that E wishes to escape from within the Pursuers' effector range in minimum time. The single-Pursuer case with P implementing the control $\arg\max_{u_P(t)} t_f$ is solved in [264].
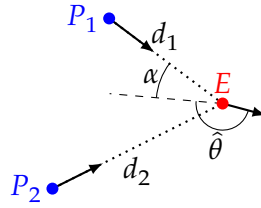
Figure 5.6: Schematic illustration of the two-Pursuer problem.

Let t_{f_1} and t_{f_2} be defined as the times at which $d_1 = l$ and $d_2 = l$, respectively. Obviously, whether considering capture or escape, simultaneity implies that $t_{f_1} = t_{f_2}$. Also define $\hat{\theta}^*$ as the associated optimal Evader heading for each Scenario.

Proposition 5.11. *The optimal Evader heading for Scenarios 1–4 lies in the range $\hat{\theta}^* \in [\pi - \alpha, \pi + \alpha]$.*

Proof. In all Scenarios, E must try to make $\dot{d}_1$ and $\dot{d}_2$ as large as possible. Headings outside the stated range are worse for both distance rates than headings within the range. A similar argument is given in [250] and analyzed in more detail there. □

It may be the case that one or other Pursuer is sufficiently far such that they have no effect on the outcome. Then, it is always best for E to employ the strategy PE against the nearer Pursuer, as in [250]. The following Theorems establish conditions under which simultaneous capture is optimal – they are based on checking if, while employing the

strategy PE from a Pursuer, the other Pursuer captures E first. Define the hypothetical single-Pursuer capture time associated with PE as

$$t_{i_{PE}} = \left| \frac{d_i - l}{1 - \mu} \right|, \qquad i = 1, 2. \tag{5.21}$$

Theorem 5.1. *For $l = 0$, $\mu < 1$ (i.e., Scenario 1), simultaneous capture is optimal iff*

$$\frac{1 + \mu \cos (2\alpha)}{1 + \mu} < \frac{d_1}{d_2} < \frac{1 + \mu}{1 + \mu \cos (2\alpha)}. \tag{5.22}$$

Proof. First, it is required that $t_{1_{PE}} > t_{f_2} (\hat{\theta} = \pi - \alpha)$ and $t_{2_{PE}} > t_{f_1} (\hat{\theta} = \pi + \alpha)$ because, otherwise, it would be optimal to be captured by only one Pursuer (c.f. [250]). Substituting (5.14) and (5.21) (with $l = 0$) into the first condition gives

$$\frac{d_1}{1 - \mu} > \left(\frac{1 - \mu \cos \left(\hat{\theta} - \alpha \right)}{1 - \mu^2} \right) d_2$$

$$\frac{d_1}{d_2} > \frac{1 + \mu \cos (2\alpha)}{1 + \mu}$$

A similar substitution for the second condition yields the inverse of this expression. Hence (5.22) must be satisfied in order for simultaneous capture to be optimal. □

Note, an expression similar to (5.22) appears in the two-Pursuer differential game version of the problem wherein the Pursuers select their instantaneous headings so as to minimize capture time [187].

Theorem 5.2. *For $l = 0$, $\mu < 1$ (i.e., Scenario 1), if (5.22) is satisfied, then the (unique) optimal Evader heading is*

$$\hat{\theta}^* = \sin \left(\frac{d_2 - d_1}{\mu d_1} \right) - \gamma, \tag{5.23}$$

where,

$$\sin \gamma = \left(\frac{d_2}{d_1} - 1 \right) \cos \alpha, \qquad \cos \gamma = \left(\frac{d_2}{d_1} + 1 \right) \sin \alpha. \tag{5.24}$$

Proof. Theorem 5.1 says that simultaneous capture is optimal since (5.22) is satisfied. Thus there must be a $\hat{\theta}^*$ for which $t_{f_1} = t_{f_2}$. Define $\theta_1 =$

$\frac{3\pi}{2} - \hat{\theta} - \alpha$ and $\theta_2 = \hat{\theta} - \frac{\pi}{2} - \alpha$. Substituting them into (5.14) and setting the capture times equal gives

$$\left(\frac{d_2}{d_1} - 1\right) \cos \hat{\theta} \cos \alpha + \left(\frac{d_2}{d_1} + 1\right) \sin \hat{\theta} \sin \alpha = \frac{d_2 - d_1}{\mu d_1}.$$

Then, using the definition for γ, (5.24), along with the angle sum identity, gives (5.23). Note that, because (5.22) is satisfied, $\sin \gamma, \cos \gamma \leq 1$. Moreover, the solution is unique because t_{f_1} and t_{f_2} are monotonically decreasing and increasing, respectively, over the range specified in Proposition 5.11 (which can be easily shown by taking the derivative of t_{f_1} and t_{f_2} w.r.t. $\hat{\theta}$). $\qquad \square$

The process is the same as above for Scenario 2, where $l > 0$.

Corollary 5.3. *For $l > 0$, $\mu < 1$ (i.e., Scenario 2), simultaneous capture is optimal iff*

$$\begin{aligned} t_{1_{PE}} &> t_{f_2}(\hat{\theta} = \pi - \alpha; l > 0), \\ t_{2_{PE}} &> t_{f_1}(\hat{\theta} = \pi + \alpha; l > 0). \end{aligned} \qquad (5.25)$$

However, the computation of $t_{f_1}(\hat{\theta} = \pi + \alpha)$ and $t_{f_2}(\hat{\theta} = \pi - \alpha)$ is more difficult (in that it requires the solution of (5.16)). A more computationally efficient check involves making use of the capture time associated with point capture at the cost of being a weaker condition.

Theorem 5.3. *For $l = 0$, $\mu < 1$ (i.e., Scenario 2), simultaneous capture is optimal if*

$$\frac{1 + \cos(2\alpha)}{1 + \mu} < \frac{d_1 - l}{d_2 - l} < \frac{1 + \mu}{1 + \cos(2\alpha)}. \qquad (5.26)$$

Note, this condition is sufficient, but not necessary.

Proof. First, note that the capture time associated with $l > 0$ must be less than the capture time for $l = 0$ from Proposition 5.10. So, instead of checking for PE against Pursuer i and computing the actual capture time for Pursuer j, the closed-form, analytic expression for the point capture case, (5.14), may be used. Since $t_{f_{l=0}} > t_{f_{l>0}}$, it is sufficient to show that $t_{i_{PE}} > t_{f_{j,l=0}}$ for $i, j \in \{1, 2\}$, $i \neq j$. Then (5.26) is obtained similarly as in the proof of Theorem 5.1. $\qquad \square$

Fig. 5.7 shows, for particular Pursuer positions, the regions for which conditions (5.25) and (5.26) are satisfied. The sufficient region covers

much of the simultaneous capture region, especially for Evader positions near the Pursuers.
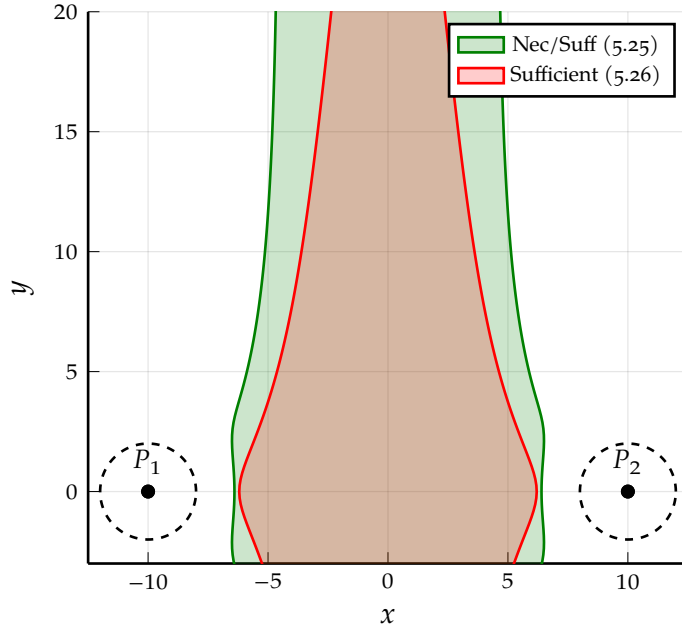
Figure 5.7: Example coverage of the sufficient condition for optimal simultaneous capture, (5.26), compared to the necessary and sufficient condition, (5.25); the Pursuer positions are fixed and the regions correspond to possible Evader positions, $\mu = 0.8$ and $l = 2$.

For Scenario 3, the situation is somewhat different since there is a range of headings for which the Evader can guarantee evasion against a particular Pursuer. In the event that E has no safe headings to take when there are two Pursuers, then Corollary 5.3 applies. The interesting feature of this case thus becomes the Evader's range of safe headings. Recall that the limiting case for capture when $\mu > 1$ corresponds with $\psi_f = \psi^\dagger$ and $d_f = l$ since this corresponds to the minimum distance (c.f. 5.4). Then, given the initial position of P and E, the safe range of Evader headings is given by $\Psi_s \equiv [-\psi_s, \psi_s]$, where ψ_s is the (numerical) solution of (5.8). If there are *many* Pursuers, the safe range associated with each Pursuer, Ψ_{s_i}, may be computed, and the overall safe range is given by their intersection:

$$\Psi_s = \cap_{i=1}^{M} \Psi_{s_i}. \tag{5.27}$$

Of course, if $\Psi_s = \emptyset$, then capture is guaranteed. Fig. 5.8 shows an example with $M = 8$ Pursuers in which the overall safe range, Ψ_s, is nonempty. The Evader need only choose a heading in the clear region in order to guarantee evasion.
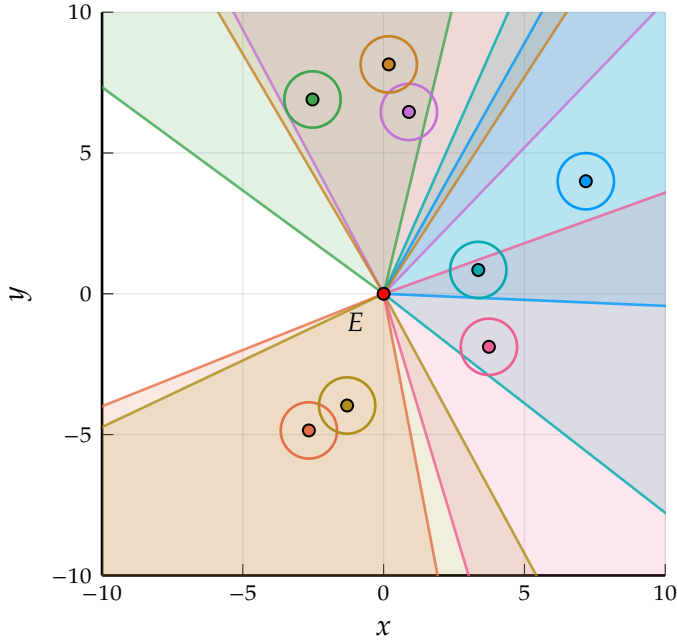
Figure 5.8: Fast capture example with many Pursuers in which the safe range of Evader headings (shown as the clear cone) is nonempty. For the purposes of visualization, the complement of Ψ_{s_i} is shown. Here, $\mu = 1.2$ and $l = 1$.

Finally, for Scenario 4, recall that the Evader seeks to minimize the amount of time it takes to drive $d > l$; the necessary and sufficient condition for the optimality of simultaneous escape is as follows.

Corollary 5.4. *For $l > 0$, $\mu > 1$ (i.e., Scenario 4), if $d_1, d_2 < l$ then simultaneous escape is optimal iff*

$$t_{1_{PE}} < t_{f_2}(\hat{\theta} = \pi - \alpha; l > 0),$$
$$t_{2_{PE}} < t_{f_1}(\hat{\theta} = \pi - \alpha; l > 0).$$

(5.28)

Unlike in Scenario 2 there is no obvious relaxation of this condition, since point capture is not generally possible when $\mu > 1$. The general case in which E is inside the effector range of one Pursuer but outside the other's is more complex, hence, above it is assumed that E is within l of both Pursuers. Fig. 5.9 shows an example of simultaneous escape.
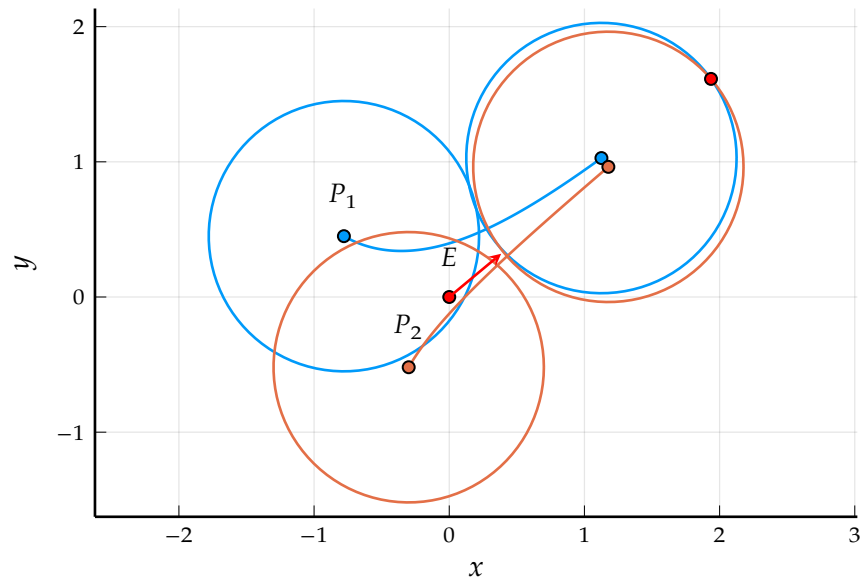
Figure 5.9: Simultaneous escape example with the numerically computed optimal Evader heading; $\mu = 1.2$ and $l = 1$.

5.6 CONCLUSION

In this chapter, the results for the classical pursuit-evasion problem have been extended into the realm of non-point capture. These results are thus one step closer to real-world situations, such as air-to-air combat, wherein a projectile need only come within some finite distance to destroy its target. The curves governing the final configuration (tail-chase or head-on) and whether capture happens (for $\mu > 1$) were obtained in closed form. An analytic expression for the minimum distance attained by the Pursuer was also derived. The final time for finite effector range ($l > 0$) may be obtained via any standard root-finding method, or (for rational μ) as the rooting of a sparse polynomial. Some of the results were applied directly to the case of $M \geq 2$ Pursuers. Another extension may be to utilize the safe range for Evader headings as a constraint in an Evader path planning algorithm. The next chapter will consider what happens when the Evader optimizes its instantaneous heading to maximize capture time.

Any mathematical study, such as this, must be judged,
ultimately upon its intrinsic content, and not by the
density of high-sounding pseudo-abstractions with
which a text may so easily be salted.

— Richard Bellman, Dynamic Programming [21]

6

PURE PURSUIT AND OPTIMAL EVADER

6.1 INTRODUCTION

The PEPP scenario is expanded in this chapter – here, the Evader is no longer assumed to hold course, but, rather, seeks to delay capture for as long as possible. This chapter partly addresses Research Objective 4 (demonstrate utility of skirmish-level solutions) in that the 1v1 pursuit-evasion solution derived in Section 1.1 becomes a sub-solution for the two Pursuer problem for a significant portion of the state space. The contents of this chapter are based upon [250].

The biggest difference in this chapter compared with the aforementioned two-Pursuer works (c.f., Chapter 2) is the fact that, rather than playing the game, the Pursuers' strategy is fixed (to PP). Although DGT is a generalization of optimal control [130], it is sometimes the case that fixing the strategy of one of the sides can make the analysis more challenging. For example, if it is assumed that the Pursuer employs PP, the optimal control analysis (which yields the same optimal action for the Evader) is not trivial (c.f. [5]) due to the nonlinearity of the dynamics induced by the Pursuers' PP state-feedback strategy. This exercise of optimizing an agent's strategy against a particular opponent's strategy, even when the game scenario has been solved, is useful when that opponent strategy is so widely used or well-known, as is the case for PP. For example, in [5], a Target and Defender cooperate against an Attacker who employs PP. Similarly, [99, 258] consider the same scenario but with an Attacker who employs PN (with finite capture radius and

point capture, respectively). The scenario described in this chapter was also considered in [152] wherein it was assumed that the Pursuers employ either PP or a fixed heading strategy. There, in the PP case, point capture was considered and the optimal control problem was solved using numerical pseudo-spectral (collocated) methods. Here, the analytical approach is emphasized.

The main contributions of this chapter are (1) the synthesis of the optimal evasion control law against dual pure pursuit with finite capture radius, (2) characterization of the disjoint regions of the state space corresponding to the different capture cases, (3) proofs regarding the set of terminal Evader headings resulting in optimal dual capture, and (4) a backwards shooting numerical method for solving the TPBVP arising where indirect optimization is employed. The chapter is organized as follows. Section 6.2 contains the optimal control problem formulation and Section 6.3 contains a derivation of the optimality conditions. Section 6.4 characterizes the solutions for both solo and dual capture. Section 6.6 concludes the chapter with remarks on the utility of this solution.

6.2   PROBLEM FORMULATION

Let $E = (x_E, y_E)$, $P_1 = (x_1, y_1)$, and $P_2 = (x_2, y_2)$ denote the Evader, Pursuer 1, and Pursuer 2 and their respective positions in the realistic plane $\mathbb{R}^2$. The agents' velocities are denoted v_E, and $v_1 = v_2 = v_P$, respectively, and $v_E < v_P$. In the realistic plane, the state $\mathbf{x}_G \in \mathbb{R}^6$ has six components corresponding to the coordinates of the three agents in the Euclidean plane. For the remainder of the chapter, a relative state space $\mathbf{x} \in \mathbb{R}^6$ based on the Evader's instantaneous position is utilized. The angle β is the angle of the line EP_1 w.r.t. the realistic x-axis. In later optimality analysis, it will be shown that only the first three state components influence the optimal solution: the Euclidean distances between each Pursuer and the Evader, d_1 and d_2, and the half-angle between the Pursuers w.r.t the Evader, α. Fig. 6.1 shows these key state components (black) along with the remaining state components (blue) used

to relate the relative state to the realistic (global) state. The following
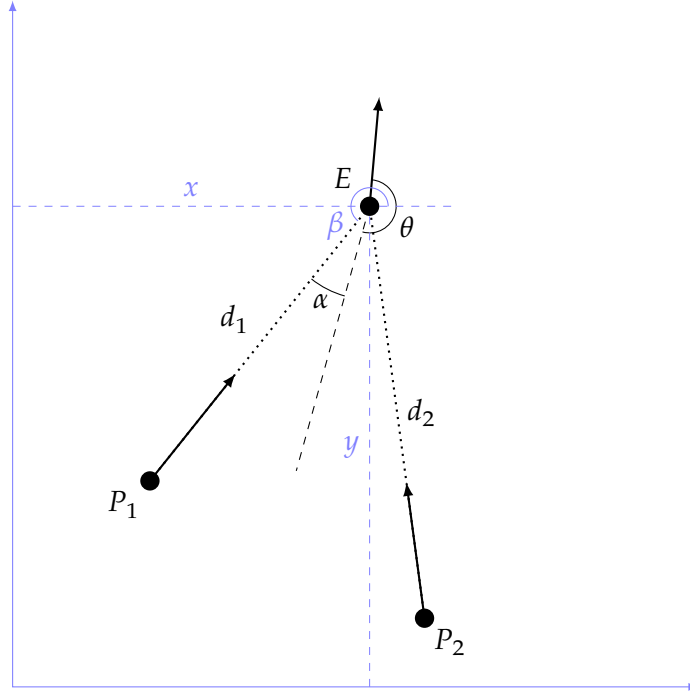
Figure 6.1: Coordinate systems for the evasion scenario with main features and relative states in black and global states in blue.

equations relate the relative state to the global state,

$$
\begin{bmatrix} x_E \\ y_E \\ x_1 \\ y_1 \\ x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x \\ y \\ x + d_1 \cos(\beta) \\ y + d_1 \sin(\beta) \\ x + d_2 \cos(\beta + 2\alpha) \\ y + d_2 \sin(\beta + 2\alpha) \end{bmatrix}.
\tag{6.1}
$$

The dimensional dynamics (denoted with a bar) are given as,

$$
f = \begin{bmatrix} \dot{\bar{d}}_1 \\ \dot{\bar{d}}_2 \\ \dot{\bar{\alpha}} \\ \dot{\bar{x}} \\ \dot{\bar{y}} \\ \dot{\bar{\beta}} \end{bmatrix} = \begin{bmatrix} -v_E \cos\left(\theta + \bar{\alpha}\right) - v_P \\ -v_E \cos\left(\theta - \bar{\alpha}\right) - v_P \\ \frac{v_E}{2}\left(\frac{1}{\bar{d}_1}\sin\left(\theta + \bar{\alpha}\right) - \frac{1}{\bar{d}_2}\sin\left(\theta - \bar{\alpha}\right)\right) \\ v_E \cos\left(\bar{\beta} + \bar{\alpha} + \theta\right) \\ v_E \sin\left(\bar{\beta} + \bar{\alpha} + \theta\right) \\ -\frac{v_E}{\bar{d}_1}\sin\left(\theta + \bar{\alpha}\right) \end{bmatrix}.
\tag{6.2}
$$

6.2.1 *Non-Dimensionalization*

Let the capture radius be denoted d_c, and the ratio of Evader to Pursuer speed be denoted $\mu = v_E/v_P < 1$. It is useful to consider the dimensionless form of the state $\mathbf{x}$ and its dynamics f because, in doing so, the number of parameters is effectively reduced from three (v_E, v_P, and d_c) to one (μ). Let non-dimensional distances be defined by $d = \bar{d}/d_c$ and non-dimensional time be defined by $t = \bar{t}/t_c$, where $t_c = d_c/v_P$ is the amount of time taken by a Pursuer to traverse the capture distance. Then the non-dimensional distance dynamics are obtained from

$$\dot{\bar{z}} = \frac{\mathrm{d}\bar{z}}{\mathrm{d}\bar{t}} = \frac{\mathrm{d}(d_c z)}{\mathrm{d}(t_c t)} = \frac{d_c}{t_c}\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{d_c}{t_c}\dot{z}$$

$$\implies \dot{z} = \frac{t_c}{d_c}\dot{\bar{z}} = \frac{1}{v_P}\dot{\bar{z}}, \tag{6.3}$$

for $z = d_1, d_2, x, y$ and the non-dimensional angular dynamics are obtained from

$$\dot{\bar{\psi}} = \frac{\mathrm{d}\bar{\psi}}{\mathrm{d}\bar{t}} = \frac{\mathrm{d}\psi}{\mathrm{d}(t_c t)} = \frac{1}{t_c}\frac{\mathrm{d}\psi}{\mathrm{d}t} = \frac{1}{t_c}\dot{\psi}$$

$$\implies \dot{\psi} = t_c\dot{\bar{\psi}} = \frac{d_c}{v_P}\dot{\bar{\psi}}, \tag{6.4}$$

for $\psi = \alpha, \beta$. Substituting Eqs. (6.3) and (6.4) into Eq. (6.2) yields the non-dimensional form of the dynamics,

$$f = \begin{bmatrix} \dot{d_1} \\ \dot{d_2} \\ \dot{\alpha} \\ \dot{x} \\ \dot{y} \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} -\mu\cos(\theta+\alpha)-1 \\ -\mu\cos(\theta-\alpha)-1 \\ \frac{\mu}{2}\left(\frac{1}{d_1}\sin(\theta+\alpha)-\frac{1}{d_2}\sin(\theta-\alpha)\right) \\ \mu\cos(\beta+\alpha+\theta) \\ \mu\sin(\beta+\alpha+\theta) \\ -\frac{\mu}{d_1}\sin(\theta+\alpha) \end{bmatrix}. \tag{6.5}$$

6.2.2 *Problem Statement*

The problem for the Evader is to select its heading control $\theta(t)$, $t \in \left[0, t_f\right]$ to maximize the time to capture. The cost functional is,

$$J = \int_0^{t_f}(-1)\,\mathrm{d}t = \int_0^{t_f} L\,\mathrm{d}t = -t_f. \tag{6.6}$$

Note that, due to the non-dimensionalization, the scenario terminates (capture occurs) when either one or both Pursuers come within a non-dimensional distance of 1 to the Evader. The associated boundary condition is given as

$$\phi\left(\mathbf{x}_0, t_f, \mathbf{x}_f\right) = \left(d_{1_f} - 1\right)\left(d_{2_f} - 1\right) = 0, \tag{6.7}$$

where $\mathbf{x}_0$ and $\mathbf{x}_f$ are the initial and final states, respectively. We now express the optimal control problem as

$$\min_{\theta(t)} J, \text{ s.t. } \dot{\mathbf{x}} = f(\mathbf{x}, \theta), \ \phi = 0. \tag{6.8}$$

6.3 OPTIMALITY CONDITIONS

The Hamiltonian is given by,

$$
\begin{aligned}
\mathcal{H} = \lambda_{d_1} & \left(-\mu \cos\left(\theta + \alpha\right) - 1\right) \\
& + \lambda_{d_2}\left(-\mu \cos\left(\theta - \alpha\right) - 1\right) \\
& + \lambda_\alpha \frac{\mu}{2}\left(\frac{1}{d_1}\sin\left(\theta + \alpha\right) - \frac{1}{d_2}\sin\left(\theta - \alpha\right)\right) \\
& + \lambda_x \mu \cos\left(\beta + \alpha + \theta\right) + \lambda_y \mu \sin\left(\beta + \alpha + \theta\right) \\
& - \lambda_\beta \frac{\mu}{d_1}\sin\left(\theta + \alpha\right) - 1 \quad (6.9)
\end{aligned}
$$

where $\lambda \equiv \begin{bmatrix} \lambda_{d_1} & \lambda_{d_2} & \lambda_\alpha & \lambda_x & \lambda_y & \lambda_\beta \end{bmatrix}^\top$ are the adjoint variables, the partial derivatives of the Value function. From the first order optimality conditions, the optimal adjoint dynamics are

$$\dot\lambda_{d_1} = -\frac{\partial \mathcal{H}}{\partial d_1} = \frac{\lambda_\alpha \mu \sin(\theta + \alpha)}{2 d_1^2} \tag{6.10}$$

$$\dot\lambda_{d_2} = -\frac{\partial \mathcal{H}}{\partial d_2} = \frac{\lambda_\alpha \mu \sin(\theta - \alpha)}{2 d_2^2} \tag{6.11}$$

$$\dot\lambda_\alpha = -\frac{\partial \mathcal{H}}{\partial \alpha} \tag{6.12}$$

$$= -\lambda_{d_1} \mu \sin(\theta + \alpha) - \lambda_{d_2} \mu \sin(\theta - \alpha) \tag{6.13}$$
$$- \lambda_\alpha \frac{\mu}{2}\left(\frac{1}{d_1}\cos(\theta + \alpha) + \frac{1}{d_2}\cos(\theta - \alpha)\right)$$

$$\dot\lambda_x = -\frac{\partial \mathcal{H}}{\partial x} = 0 \tag{6.14}$$

$$\dot\lambda_y = -\frac{\partial \mathcal{H}}{\partial y} = 0 \tag{6.15}$$

$$\dot\lambda_\beta = -\frac{\partial \mathcal{H}}{\partial \beta} = 0. \tag{6.16}$$

Because this is a free final time problem with a Lagrange cost functional, the transversality condition gives

$$\lambda^\top(t_f) = \nu \frac{\partial \phi}{\partial \mathbf{x}_f} \tag{6.17}$$
$$= \nu \begin{bmatrix} \left(d_{2_f} - 1\right) & \left(d_{1_f} - 1\right) & 0 & 0 & 0 & 0 \end{bmatrix}$$
$$\implies \lambda_{\alpha_f} = \lambda_{x_f} = \lambda_{y_f} = \lambda_{\beta_f} = 0. \tag{6.18}$$

Since the adjoint variables $\lambda_x, \lambda_y,$ and λ_β are zero at final time (Eq. (6.18)) and their derivatives are zero (Eqs. (6.14)–(6.16)) they are zero for all time and thus the states x, y, and β have no effect on the optimality of the solution; only the $d_1, d_2,$ and α states are pertinent. Substituting

$\lambda_x = \lambda_y = \lambda_\beta = 0$ and using Ptolemy's Trigonometric Identities to expand the cosine/sine of sum terms in Eq. (6.9) yields,

$$\mathcal{H} = -\lambda_{d_1}\mu \left(\cos\theta\cos\alpha - \sin\theta\sin\alpha\right) - \lambda_{d_1} -$$
$$\lambda_{d_2}\mu \left(\cos\theta\cos\alpha + \sin\theta\sin\alpha\right) - \lambda_{d_2} +$$
$$\lambda_\alpha \frac{\mu}{2}\left(\frac{1}{d_1}\left(\sin\theta\cos\alpha + \cos\theta\sin\alpha\right) - \right.$$
$$\left. \frac{1}{d_2}\left(\sin\theta\cos\alpha - \cos\theta\sin\alpha\right)\right) - 1. \quad (6.19)$$

Now define the following two quantities which are the coefficients of the $\cos\theta$ and $\sin\theta$ terms,

$$c_{\cos} = \mu\left(-\lambda_{d_1}\cos\alpha - \lambda_{d_2}\cos\alpha + \frac{\lambda_\alpha}{2}\right. \qquad (6.20)$$
$$\left. \sin\alpha\left(\frac{1}{d_1} + \frac{1}{d_2}\right)\right)$$
$$c_{\sin} = \mu\left(\lambda_{d_1}\sin\alpha - \lambda_{d_2}\sin\alpha + \frac{\lambda_\alpha}{2}\cos\alpha\left(\frac{1}{d_1} - \frac{1}{d_2}\right)\right) \qquad (6.21)$$

and substitute back into (6.19) and simplify to get,

$$\mathcal{H} = c_{\cos}\cos\theta + c_{\sin}\sin\theta - \lambda_{d_1} - \lambda_{d_2} - 1. \qquad (6.22)$$

From PMP, then, the optimal heading is given by $\theta^* = \arg\min_\theta \mathcal{H}$. To minimize $\mathcal{H}$, it must be that the vector $\begin{bmatrix}\cos\theta & \sin\theta\end{bmatrix}^\top$ be antiparallel to the vector $\begin{bmatrix}c_{\cos} & c_{\sin}\end{bmatrix}$, giving,

$$\cos\theta^* = \frac{-c_{\cos}}{\sqrt{c_{\cos}^2 + c_{\sin}^2}}, \quad \sin\theta^* = \frac{-c_{\sin}}{\sqrt{c_{\cos}^2 + c_{\sin}^2}}. \qquad (6.23)$$

The Hamiltonian at final time is given by,

$$\mathcal{H}(t_f) = -\nu\frac{\partial\phi}{\partial t_f} = 0. \qquad (6.24)$$

Since $\mathcal{H}$ is not an explicit function of time, it must also be that $\mathcal{H}(t) = 0$, $\forall t \in [0, t_f]$.

6.4   SOLUTION CHARACTERIZATION

We are interested in both solo and dual capture. Eq. (6.7) is satisfied for all three terminal scenarios: both the solo capture cases (i.e., $d_{1_f} > 1$ and $d_{2_f} = 1$, or vice versa) as well as the dual capture case ($d_{1_f} = d_{2_f} = 1$). We will develop optimal solutions for each of these cases in the following subsections.

6.4.1   Solo Capture

The first case is solo capture where $d_{1_f} = 1$ and $d_{2_f} > 1$, or $d_{1_f} > 1$ and $d_{2_f} = 1$.

Lemma 6.1 (Solo capture trajectories). *The optimal control resulting in solo capture by P_1 is $\theta^*(t) = \pi - \alpha(t)$, $\forall t \in \left[0, t_f\right]$, and for solo capture by P_2 is $\theta^*(t) = \pi + \alpha(t)$, $\forall t \in \left[0, t_f\right]$.*

Proof. Consider the second case: solo capture by P_2, which entails $d_{2_f} = 1$ and $d_{1_f} > 1$. From Eq. (6.17), $\lambda_{d_{1_f}} = 0$, $\lambda_{d_{2_f}} = \nu\left(d_{1_f} - 1\right)$, and $\lambda_{\alpha_f} = 0$. Substituting these terminal adjoint values into Eqs. (6.20) and (6.21) gives,

$$c_{\cos} = -\mu\nu(d_{1_f} - 1)\cos\alpha_f, \quad c_{\sin} = -\mu\nu(d_{1_f} - 1)\sin\alpha_f.$$

Substituting these values into the optimal control Eq. (6.23) gives,

$$\cos\theta_f^* = \operatorname{sign}(\nu)\cos\alpha_f, \quad \sin\theta_f^* = \operatorname{sign}(\nu)\sin\alpha_f.$$

If $\operatorname{sign}(\nu) = 1$ then $\theta_f^* = \alpha_f$, which implies the E heads directly towards P_2, which is clearly suboptimal. Instead, if $\operatorname{sign}(\nu) = -1$ then $\theta_f^* = \pi + \alpha_f$, implying E heads directly away from P_2. Substituting this terminal Evader heading into the adjoint dynamics Eqs. (6.10)–(6.13) gives,

$$\dot{\lambda}_{d_{1_f}} = 0, \; \dot{\lambda}_{d_{2_f}} = 0, \; \dot{\lambda}_{\alpha_f} = 0,$$

which implies that the optimal adjoint values are constant over the trajectory. Then, from the above analysis, the condition $\theta^* = \pi + \alpha$ holds

for all $t \in \left[0, t_f\right]$. By symmetry, the result $\theta^* = \pi - \alpha$ applies for solo capture by P_1. □

Lemma 6.2 (Solo capture trajectory shape). *In the case of solo capture, the Evader's and capturing Pursuer's trajectories are straight lines in the realistic plane.*

Proof. Without loss of generality, consider solo capture by P_1. From Lemma 6.1 the Evader's optimal control is $\theta^*(t) = \pi - \alpha(t)$. The Evader's heading in the realistic plane is given by $\Theta = \beta + \alpha + \theta$. Substituting the optimal control in gives $\Theta = \beta + \pi$. The rate of change of the global Evader heading is $\dot{\Theta} = \dot{\beta}$. From Eq. (6.5) $\dot{\beta} \propto \sin(\theta + \alpha)$. Substituting the optimal control in makes $\dot{\beta} = 0$, thereby making $\dot{\Theta} = 0$. Thus the Evader's heading in the realistic plane is constant, implying a straight-line path. The Evader's heading lies along the LOS $\overline{P_1E}$, and thus P_1's path is also straight. □

Remark 3. Note this is also the solution to the single-Pursuer single-Evader optimal control problem (c.f. [5]), and so the presence of the second Pursuer did not affect the optimal trajectories.

Lemma 6.3 (Closer Pursuer). *Optimal solo capture is always executed by the Pursuer who began closer to the Evader.*

Proof. Without loss of generality, consider solo capture by P_1. From Lemma 6.1 the Evader's optimal control is $\theta^*(t) = \pi - \alpha(t)$. Therefore, from Eq. (6.5) and $\mu < 1$, it is the case that $0 > \dot{d}_1(t) > \dot{d}_2(t)$, $\forall t \in \left[0, t_f\right]$. Since solo capture by P_1 entails $d_{1_f} = 1 < d_{2_f}$ it must be the case that $d_2(t) > d_1(t)$, $\forall t \in \left[0, t_f\right]$. □

There may be initial conditions for which optimal solo capture trajectories do not exist. For example, if E flees from whichever P is closer at initial time (following Lemmas 6.1 and 6.3) and ends up being captured by the other P, then optimal solo capture does not exist. It may also be the case that the optimal solo capture trajectory exists but an optimal dual capture trajectory exists.

Lemma 6.4. *If, for a particular initial condition, both solo capture and dual capture trajectories exist and satisfy all of the optimality conditions, the solo capture trajectory is uniquely optimal.*

Proof. From the definition of ϕ in (6.7) the dual capture candidate solution satisfies all of the optimality conditions for solo capture. However, Lemma 6.1 specifies the optimal control for solo capture ($\theta^* = \pi - \alpha$ for P_1 or $\theta^* = \pi + \alpha$ for P_2). Any other control action would result in a smaller t_f compared to the solo capture candidate solution. Therefore, the solo capture solution is optimal. $\qquad\square$

6.4.2   *Dual Capture*

In the case of dual capture both $d_{1_f}, d_{2_f} = 1$. However, if these terminal distances are substituted into Eq. (6.17) then $\lambda_{d_{1_f}} = \lambda_{d_{2_f}} = 0$. Substituting these values for the terminal distance adjoints (along with all of the other known terminal adjoints) gives $\mathscr{H}(t_f) = -1$ which contradicts Eq. (6.24), which says that $\mathscr{H}(t_f) = 0$. Dual capture, thus, exhibits a singularity, which is also evident by the fact that the terminal surface corresponding to dual capture, $\{d_1, d_2, \alpha \mid d_1 = d_2 = 1\}$, is a line $\in \mathbb{R}^1$. Isaacs states [130] that a non-degenerate terminal surface be of dimension one less than the dimension of the state space. The consequence, here, is that many different trajectories terminate at the same point on the dual capture termination line, even for the same terminal α. In order to proceed, the limiting Evader heading, $\theta^*(t)$, $t \to t_f$ and its relationship to the limiting adjoint values is treated using a procedure described in [89]:

$$\tan \theta^*(t_f) = \lim_{t \to t_f} \tan \theta^* = \lim_{t \to t_f} \frac{-c_{\sin}}{-c_{\cos}}$$

$$= \lim_{t \to t_f} \frac{\sin \alpha_f \left(\lambda_{d_{2_f}} - \lambda_{d_{1_f}} \right)}{\cos \alpha_f \left(\lambda_{d_{2_f}} + \lambda_{d_{1_f}} \right)}.$$

Rearranging this expression for the adjoint variables yields

$$\kappa \equiv \frac{\lambda_{d_{2_f}}}{\lambda_{d_{1_f}}} = \frac{\tan \alpha_f + \tan \theta_f^*}{\tan \alpha_f - \tan \theta_f^*}. \qquad (6.25)$$

Substituting the relation (6.25), along with (6.18) into Eqs. (6.9) and (6.24) and solving for the terminal distance adjoints yields,

$$\lambda_{d_{1_f}} = \frac{1}{-\mu \cos\left(\theta_f + \alpha_f\right) - 1 + \kappa \left(-\mu \cos\left(\theta_f - \alpha_f\right) - 1\right)} \qquad (6.26)$$

or,

$$\lambda_{d_{2_f}} = \frac{1}{\frac{1}{\kappa}\left(-\mu \cos\left(\theta_f + \alpha_f\right) - 1\right) - \mu \cos\left(\theta_f - \alpha_f\right) - 1}. \qquad (6.27)$$

Lemma 6.5. *At $\theta_f = \pi \pm \alpha$ the dual capture solution is equivalent to a solo capture solution with the non-capturing Pursuer's terminal distance approaching 1.*

Proof. From Eq. (6.25), it must be that $\kappa = 0$ when $\theta_f = \pi - \alpha$ and $\kappa = \infty$ when $\theta_f = \pi + \alpha$ which imply $\lambda_{d_{1_f}} \neq 0$ and $\lambda_{d_{2_f}} = 0$, or $\lambda_{d_{1_f}} = 0$ and $\lambda_{d_{2_f}} \neq 0$, respectively. Thus, when one of the terminal distance adjoints is zero and the other is non-zero, then from the analysis in Lemma 6.1 the optimal control is $\theta^*(t) = \pi \pm \alpha(t)$ over the whole trajectory. Therefore, the trajectories are identical to the solo capture case wherein the non-capturing Pursuer's terminal distance approaches 1. $\qquad\square$

Proposition 6.1 (Optimal θ_f for dual capture). *The range*

$$\theta_f \in (\pi - \alpha, \pi + \alpha)$$

produces globally optimal trajectories, and the trajectories produced by $\theta_f \notin (\pi - \alpha, \pi + \alpha)$ are suboptimal.

Proof. From Eq. (6.25) whenever $\theta_f \notin \left[\pi - \alpha_f, \pi + \alpha_f\right]$ the terminal adjoint ratio $\kappa < 0$. This implies that the terminal distance adjoints, $\lambda_{d_{1_f}}$ and $\lambda_{d_{2_f}}$, have different signs; thus one of either $\lambda_{d_{1_f}} > 0$ or $\lambda_{d_{2_f}} > 0$. Suppose, without loss of generality that $\lambda_{d_{1_f}} > 0$. Equation (6.18) states that $\lambda_{\alpha_f} = 0$ – based on the optimal adjoint dynamics, Eq. (6.10), λ_{d_1} is not changing at final time. Thus $\lambda_{d_1} > 0$ for some nonzero time leading up to final time due to the smoothness of Eqs. (6.10)-(6.13). The adjoint variable $\lambda_{d_1} \equiv \frac{\partial V}{\partial d_1}$ where $V = \min_{\theta(t)} J$ is the Value function. Thus if $\lambda_{d_1} > 0$ the Value *increases* as distance from P_1 increases.

Since the cost functional is the negative of final time, increasing the distance from P_1 is a disadvantage to the E, in this case. Contrariwise, when $\theta_f \in \left[\pi - \alpha_f, \pi + \alpha_f\right]$ for dual capture, and in the single capture case, $\lambda_{d_1}, \lambda_{d_2} \leq 0 \forall t$. Thus in the latter, optimal, cases, increasing distance from a Pursuer reduces the Value, which is advantageous for the E. In the absence of turning-rate constraints, increasing distance from a Pursuer (whilst keeping the other distance constant) should always benefit E. Consequently, dual capture with $\theta_f \notin \left[\pi - \alpha_f, \pi + \alpha_f\right]$ must be suboptimal. $\qquad \square$

Remark 4. The terminal Evader headings

$$\theta_f \in \left(\alpha_f, \pi - \alpha_f\right) \cup \left(\pi + \alpha_f, 2\pi - \alpha_f\right),$$

although suboptimal for this scenario (i.e., when both solo and dual capture are possible), are optimal for a scenario in which *only* dual capture is desired. Suppose E is defending some other target against the Pursuers and, after interception, is destroyed. In that scenario, the Evader wishes to collide with the two Pursuers simultaneously.
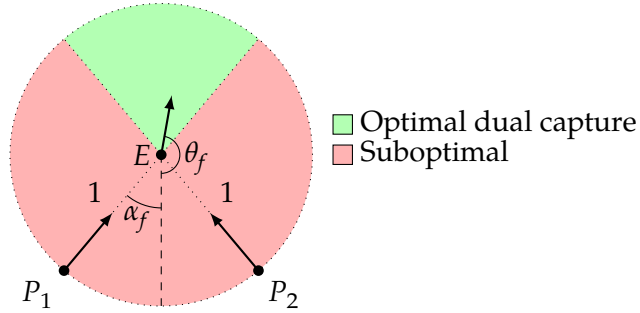
Figure 6.2: Optimality of the terminal Evader heading sectors.

Fig. 6.2 summarizes Proposition 6.1, showing the sectors of θ_f for optimal dual capture (green), and where dual capture is suboptimal (red).

Lemma 6.6 (Symmetric dual capture). *For initial conditions with $d_1 = d_2$, the optimal control is $\theta(t) = \pi$, $\forall t \in \left[0, t_f\right]$, E's trajectory is straight in the realistic plane, and the scenario terminates in dual capture.*

Proof. Suppose $\left(d_{1_f}, d_{2_f}, \alpha_f\right) = \left(1, 1, \alpha_f\right)$ and $\theta_f = \pi$. From Eq. (6.25) $\kappa = 1$ and so $\lambda_{d_{1_f}} = \lambda_{d_{2_f}}$. Substituting into Eq. (6.13) with $\lambda_{\alpha_f} = 0$ (from Eq. (6.17)) gives $\dot{\lambda}_{\alpha_f} = 0$ which implies that $\dot{\lambda}_{d_1} = \dot{\lambda}_{d_2} = \dot{\lambda}_{\alpha} = 0$. So $\lambda_{d_1} = \lambda_{d_2}$, $\forall t \in \left[0, t_f\right]$. From Eqs. (6.20)–(6.23), then, $\theta^*(t) = \pi$, $\forall t \in \left[0, t_f\right]$. Also, Eq. (6.5) implies $d_1(t) = d_2(t)$, $\forall t \in \left[0, t_f\right]$ as a result of this control. Thus any point that starts with $d_1 = d_2$ can be reached (retrogressively) from such a trajectory. The global Evader heading $\Theta = \beta + \alpha + \theta$ thus becomes $\Theta = \beta + \alpha + \pi$. Its time rate of change is $\dot{\Theta} = \dot{\beta} + \dot{\alpha}$, which, from Eq. (6.5), gives $\dot{\Theta} = 0$ since $d_1 = d_2$ and $\theta = \pi$ along the trajectory.   □

Suppose the initial state is s.t. dual capture is optimal. If $d_1 = d_2$, then Lemma 6.6 applies and the Evader's optimal control is $\theta^*(t) = \pi$. In the general case where $d_1 \neq d_2$, however, the optimal control can only be obtained by solving the TPBVP:

$$\theta_f^*, \; \alpha_f^*, \; t_f^* = \underset{\theta_f, \alpha_f, t_f}{\arg\min} \left\| \begin{matrix} d_{1_0} - d_1(0; \theta_f, \alpha_f, t_f) \\ d_{2_0} - d_2(0; \theta_f, \alpha_f, t_f) \\ \alpha_0 - \alpha(0; \theta_f, \alpha_f, t_f) \end{matrix} \right\|, \qquad (6.28)$$

where $\theta_f \in \left[\pi - \alpha_f, \pi + \alpha_f\right]$ and $\left(d_{1_0}, d_{2_0}, \alpha_0\right)$ are obtained using Eq. (6.1), based on the initial Cartesian coordinates of the three agents. Given θ_f, the terminal adjoint values are obtained by Eqs. (6.25)–(6.27). The quantities $\mathbf{x}(0; \theta_f, \alpha_f, t_f)$ are obtained by integrating Eqs. (6.5) and (6.10)–(6.13) backwards from t_f to 0. The trajectory obtained is then converted to the global coordinates via Eq. (6.1) and then shifted and rotated to match the original global configuration. This process is repeated in a backwards shooting numerical solution scheme – NLopt is used with the COBYLA (Constrained Optimization by Linear Approximation) solver to solve Eq. (6.28) based on an initial guess [132, 195]. With the solution to Eq. (6.28) in hand, the Evader may compute $\theta(t)$, the optimal control along every point in the trajectory, from Eq. (6.23).

6.4.3 *Full Solution*

Based on the solution characteristics established above, the state space $\mathscr{C} \equiv \left\{(d_1, d_2, \alpha) \mid d_1, d_2 \geq 1 \text{ and } 0 < \alpha < \frac{\pi}{2}\right\}$ is partitioned into three regions corresponding to the different terminal scenarios:

$$
\mathscr{C} = \cup \begin{cases} \mathscr{R}_1, & \left\{(d_1, d_2, \alpha) \mid d_{1_f} = 1, d_{2_f} > 1\right\} \\ \mathscr{R}_2, & \left\{(d_1, d_2, \alpha) \mid d_{1_f} > 1, d_{2_f} = 1\right\}, \\ \mathscr{R}_{1,2}, & \left\{(d_1, d_2, \alpha) \mid d_{1_f} = 1, d_{2_f} = 1\right\} \end{cases} \tag{6.29}
$$

with the following control modes

$$
\theta^*(t) = \begin{cases} \pi - \alpha(t), & (d_1, d_2, \alpha) \in \mathscr{R}_1 \\ \pi + \alpha(t), & (d_1, d_2, \alpha) \in \mathscr{R}_2 \\ \text{solution to TPBVP}, & (d_1, d_2, \alpha) \in \mathscr{R}_{1,2} \end{cases} \cdot \tag{6.30}
$$

Fig. 6.3 depicts this partitioning of $\mathscr{C}$. The green trajectories correspond to dual capture trajectories with $\theta_f = \pi - \alpha_f$ and the red trajectories correspond to dual capture trajectories with $\theta_f = \pi + \alpha_f$, across the range of α_f. All of the trajectories terminate on the line

$$
\left\{\left(d_{1_f}, d_{2_f}, \alpha_f\right) \mid d_{1_f} = d_{2_f} = 1\right\}.
$$

Given a particular initial condition, the information in Fig. 6.3 is enough to determine the corresponding terminal scenario. Note the two separating surfaces form a boat-hull shape in which the black line forms the bottom and the trajectories along $\alpha = 0$ form the front. A more human-readable representation is given in Fig. 6.4 wherein the initial Pursuer positions are fixed and the Evader's initial position is varied over the realistic plane. The numbered blue points in Fig. 6.4 correspond to initial Evader positions for each of the example simulations contained in the following section.
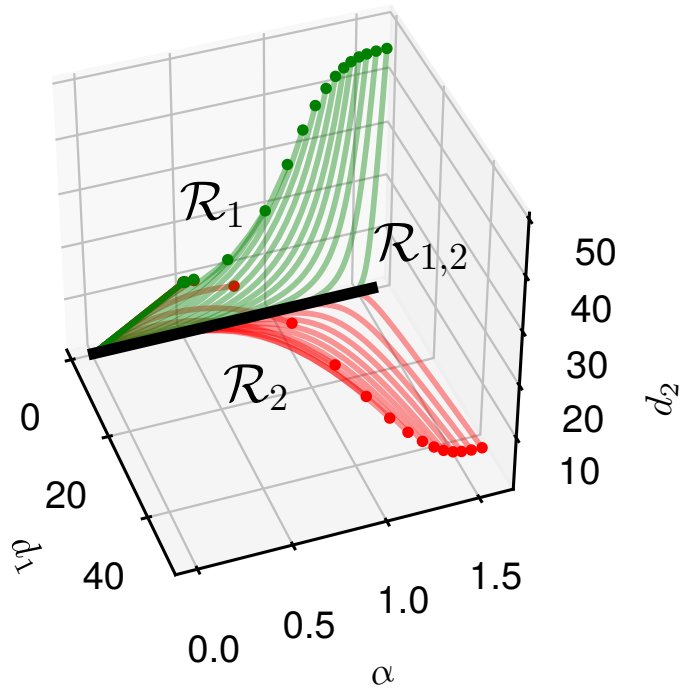
(a)
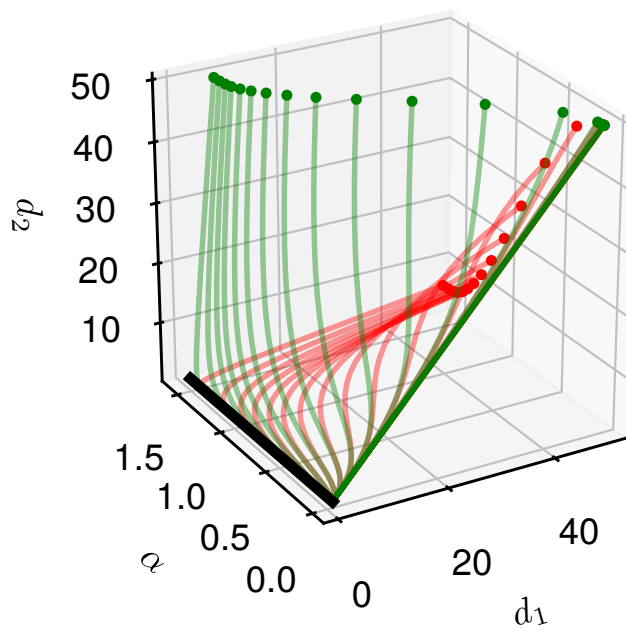
(b) Rotated view of (a)

Figure 6.3: A partitioning of the state space into regions associated with each type of capture.
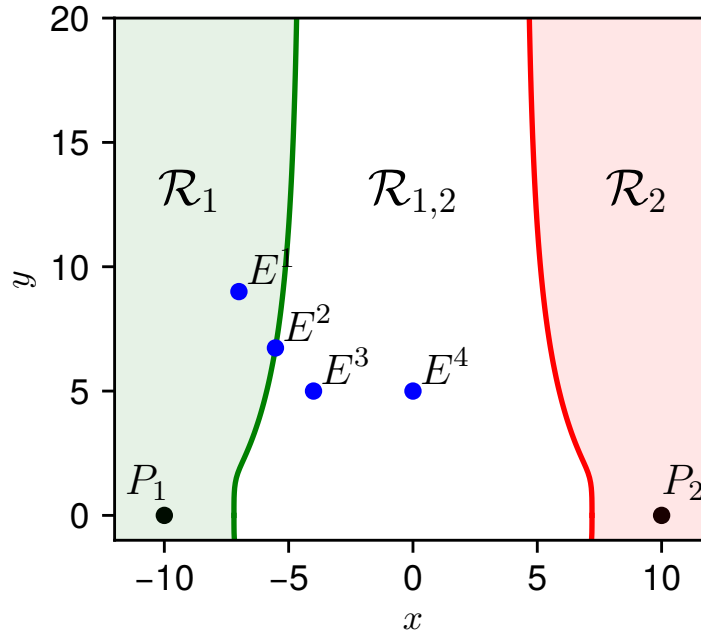
Figure 6.4: A representation of the partitioning of $\mathscr{C}$ in the realistic plane for given Pursuer initial positions for various Evader initial positions.

In lieu of storing some representation of the regions or their partitioning surfaces to determine the optimal capture scenario, Algorithm 1 contains a procedure for computing the optimal control.

Algorithm 1 Optimal Evasion Against Dual Pure Pursuit

Require: $(d_{1_0}, d_{2_0}, \alpha_0)$

  if $d_1 = d_2$ **then**
    $\theta(t) \leftarrow \pi$
  else
    $i \leftarrow \arg\min_{1,2} \overline{PE}$
    Forward shoot assuming E flees from P_i until capture
    if P_i captures **then**
      $\theta(t) \leftarrow \pi \mp \alpha(t)$ $\triangleright$ depending on $i = 1$ or 2
    else
      $\theta(t) \leftarrow$ solution of TPBVP $\triangleright$ Eq. (6.28)

6.5 SIMULATIONS

In this section, four simulations are carried out, demonstrating the solution characteristics described in the previous section. Fig. 6.4 shows the Evader initial positions for each of the simulations. The Pursuers'

Table 6.1: Example simulations parameters and description

#	x	y	Description
1	-7	9	Solo capture by P_1
2	-5.53	6.73	Limiting solo/dual capture
3	-4	5	General dual capture
4	0	5	Symmetric dual capture

initial positions are $(\pm 10, 0)$ for all of the examples. Note that Figs. 6.3 and 6.4 are shown for a speed ratio of $\mu = 0.8$. All of the simulations contained herein are based on the same speed ratio. Table 6.1 summarizes the simulation parameters and identifies the unique feature of each example.

For the first example, the Evader's initial position is s.t. the system state is in $\mathcal{R}_1$ and thus the scenario ends in solo capture by P_1. As proven in Lemma 6.2, the trajectories (in the realistic plane) are straight for E and P_1. Fig. 6.5 shows the trajectories for this example.
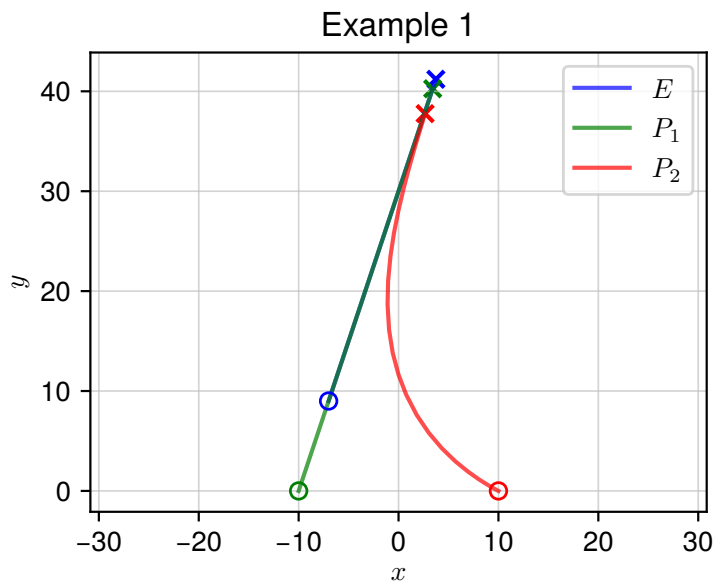
Figure 6.5: Example 1: Solo capture by P_1.

In the second example, the Evader's initial position is s.t. the system state lies on the border between $\mathcal{R}_1$ and $\mathcal{R}_{1,2}$. This is the limiting case of solo/dual capture wherein the Pursuer who is initially further away terminates at exactly the capture distance, but the trajectories for P_1 and E are straight (as in solo capture). Lemma 6.5 proves this behavior. Fig. 6.6 shows the trajectories for this example.
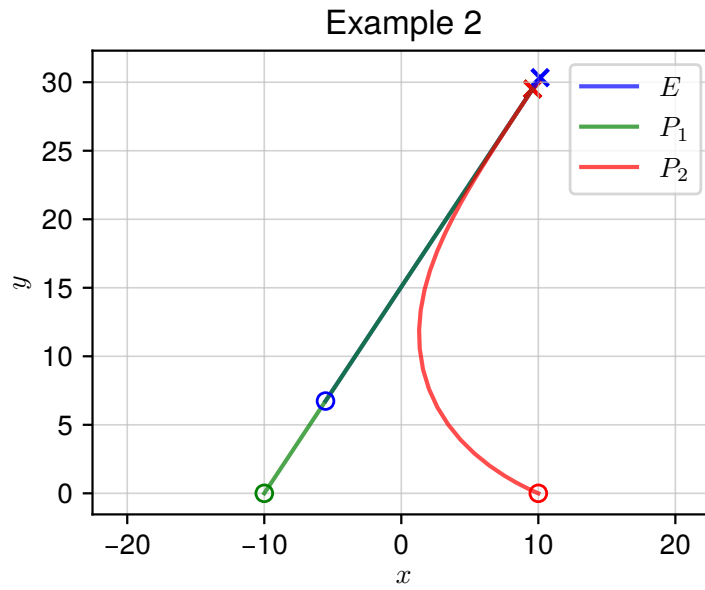
Figure 6.6: Example 2: Limiting solo/dual capture.

In the third example, the Evader's initial position is s.t. the system state lies in $\mathcal{R}_{1,2}$, thus the scenario terminates with both Pursuers capturing the Evader simultaneously. These trajectories were obtained via backwards shooting to the specified initial conditions (c.f. Table 6.1). Fig. 6.7 shows the trajectories obtained for this example.
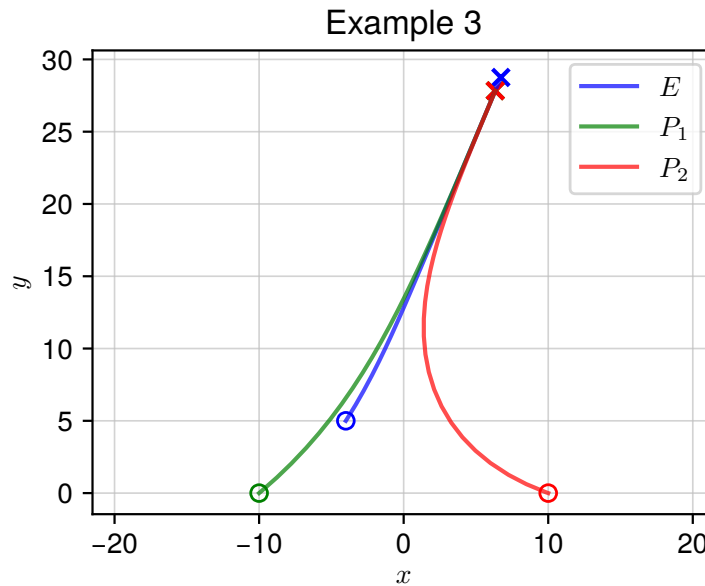
Figure 6.7: Example 3: General dual capture.

In the fourth example, the Evader's initial position is equidistant to the two Pursuers, thus satisfying the conditions of Lemma 6.6. The

resulting Evader trajectory is straight in the realistic plane and the Pursuers' trajectories are symmetric. Fig. 6.8 shows the trajectories for this
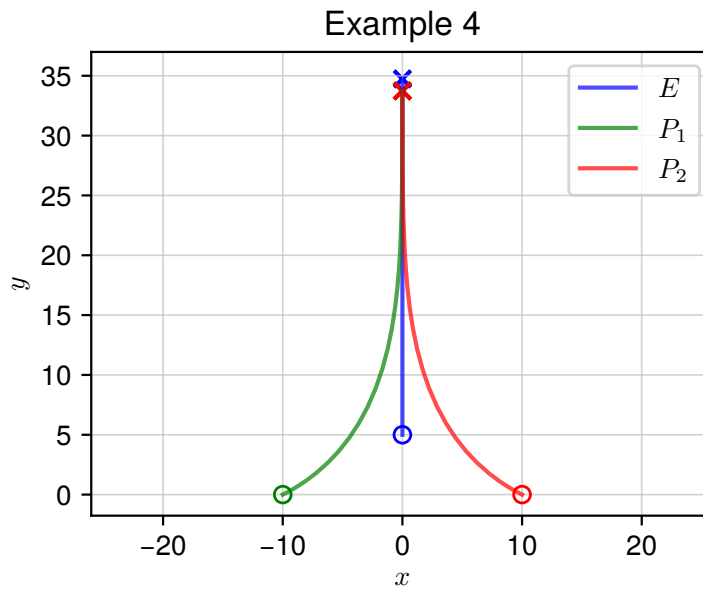
Figure 6.8: Example 4: Symmetric dual capture.

example.

6.6 CONCLUSION

In this chapter the optimal control problem of pursuit-evasion type wherein an Evader seeks to maximize its life in the presence of two faster Pursuers using PP who have a finite capture radius was solved. The optimal control was obtained via PMP and the complete state space was filled with optimal trajectories. Optimal trajectories terminating in an isochronous (dual) capture by both Pursuers produce a singularity which was rectified by analyzing the limiting Evader heading. A partitioning of the state space was generated based off of the solution characteristics separating into regions of capture by P_1 alone, by P_2 alone, and by both simultaneously. In the dual capture case, a TPBVP was posed and a procedure, based on backwards shooting, was described. Compared to the solutions in the previous chapter, wherein the Evader chooses a constant heading, this solution is much more computationally expensive; the former is much more amenable to real-time onboard implementation. The next chapter introduces additional com-

plexity in that more Pursuers are considered and the Pursuers attempt to minimize the time until capture.

7

M PURSUER GAME OF TIME

7.1 INTRODUCTION

Whereas the previous chapters only considered Pursuers who implement the PP policy, this chapter considers more intelligent Pursuers who seek to capture the Evader in minimum time. Because the Evader seeks to delay capture as long as possible, this is no longer an optimal control problem, but rather a zero-sum differential game, i.e., the MP1E GoT. The central contribution of this work is to provide open-loop optimal pursuit and evasion strategies to the MP1E GoT, which partly addresses Research Objectives 4 and 5 (demonstrate utility of 1v1 and 2v1 solutions in this more complex scenario, and adding more Pursuers, respectively). This material is based on [245].

These open-loop strategies, while useful, are not necessarily the solution to the MP1E GoT in the sense of the feedback saddle-point (see [19]). Traditional differential game analysis *a la* Isaacs' method [130] is difficult, as is shown. In order to proceed, a two-person extension to the PMP is employed to establish the necessary conditions for optimality pertaining to the Pursuers' strategy.

The Pursuers' objective is to intercept the Evader in minimum time by cooperating as a team. Works such as [43] highlight the performance advantage of explicit cooperation among the Pursuer team. The agents are assumed to have full access to the state of the system, namely, the positions of each agent. It is shown that, for general initial positions of the agents, cooperation among the Pursuers can significantly reduce the capture time of the Evader compared to operating in isolation. The state of the system is of high dimension due to the number of agents – some work has been done on decomposition methods to ameliorate this issue [74, 75, 77]. However, our approach does not rely on decomposition and therefore considers full cooperation among

Pursuers. This work extends the solution to the one-on-one and two-on-one games by allowing the Evader to stand still when it is advantageous. In these cases, the Evader would only worsen its capture time by moving from this point. Due to the proposed Pursuer strategy which consists of straight-line paths, the solution lends itself to a geometric interpretation with many interesting properties. Several algorithms are presented for the efficient computation of the Evader's region of dominance as well as the optimal capture point under the proposed strategies.

Although the solution presented in this chapter is open-loop optimal, the analysis is a useful step in fully solving and verifying the M-Pursuer one-Evader differential game. The solution of the latter opens up the possibility of analyzing the seemingly intractable (and ambitious) M-Pursuer N-Evader differential game by breaking the game down into instances of M-Pursuer one-Evader games and considering combinations of Pursuer assignments. The advantage, then, is the removal of the burden of a very high dimensional state space in the differential game analysis.

The remainder of the chapter is organized as follows. Section 7.2 contains the problem formulation. Section 7.3 elaborates on the formulation and introduces some solution methods. Section 7.3.3 presents a new geometric approach, Section 7.4 defines several efficient algorithms, and Section 7.5 contains simulation results. Section 7.6 summarizes the results of this chapter.

7.2   PROBLEM FORMULATION

A pursuit-evasion scenario is considered with a single Evader, E, and multiple Pursuers, P_i, for $i = 1, \ldots, M$. The interest is in the case where $M \geq 3$ as the single- and two-Pursuer scenarios have been addressed in [107, 130]. The objective of the Pursuers is to capture the Evader in minimum time, whereas the Evader tries to delay capture as long as possible. The case in which the Pursuers are faster than the Evader is considered. In the case of equal speeds (or a fast Evader) one must first

determine where in the state space capture is even possible (i.e., solve the game of kind) before solving the capture time problem (the game of degree). Accompanying the speed advantage, the Pursuers have an advantage in numbers. There is a marked reduction in capture time in the two-Pursuer scenario compared to having just one Pursuer [107]. The intent of this work is to show even greater improvement when $M > 2$.

The state of the system is determined solely by the positions of each of the agents on the Euclidean plane in two dimensions

$$E = (x_E, y_E),$$
$$P_i = (x_{P_i}, y_{P_i}), \ \forall i \in \{1, \dots, M\},$$

so the state space's dimension is $2(M + 1)$. All the agents have simple motion, meaning their control input at every time instant is their heading angle. For the Evader, the speed is included as a control input bounded by $V_{E_{max}}$; the Pursuers all share a constant velocity, $V_P > V_{E_{max}}$. Thus the kinematics of the system can be written as

$$
\begin{aligned}
\dot{x}_E &= u_E, \\
\dot{y}_E &= v_E, \\
\dot{x}_{P_i} &= u_i && i = 1, \dots, M, \\
\dot{y}_{P_i} &= v_i && i = 1, \dots, M, \\
\text{s.t. } u_E^2 + v_E^2 &\leq V_{E_{max}}, \\
u_i^2 + v_i^2 &= V_P && i = 1, \dots, M,
\end{aligned}
$$
(7.1)

where u_E and v_E are the Evader's velocity components in the x and y direction, and, similarly, u_i and v_i for the ith Pursuer.

Remark 5. Although it appears that the Pursuers have two control variables they only have one since the choice of either u_i or v_i completely determines the other by the final constraint in (7.1).

In the realistic plane, there are $2M + 2$ states, 2 for each agent. However, it is possible to reduce the number of states to $2M$ by considering a relative coordinate system. In this relative coordinate system, the Evader's position is always $(0, 0)$ and thus the $2M$ states correspond

to the x and y coordinate of each Pursuer relative to the Evader. The transformation is given by

$$
\begin{aligned}
x_i &= x_{P_i} - x_E & i &= 1, \dots, M, \\
y_i &= y_{P_i} - y_E & i &= 1, \dots, M,
\end{aligned}
\tag{7.2}
$$

where (x_i, y_i) are the coordinates of the ith Pursuer relative to the Evader. Substituting (7.1) into (7.2) yields the following expressions for the kinematics in the reduced state space

$$
\begin{aligned}
\dot{x}_i &= u_i - u_E & i &= 1, \dots, M, \\
\dot{y}_i &= v_i - v_E & i &= 1, \dots, M, \\
\text{s.t. } u_E^2 + v_E^2 &\le V_{E_{max}}, \\
u_i^2 + v_i^2 &= V_P & i &= 1, \dots, M.
\end{aligned}
\tag{7.3}
$$

Figure 7.1 shows the two coordinate systems and how they are related.

The agents have at their disposal the current state of the system (i.e., full information) but not the current control action of the other agents. One may be tempted to consider a scenario wherein the agents also have access to the history of the system's evolution (including past control actions of all the agents). In this alternate formulation, agents may form some belief about their opponent's next action that is conditioned on the state history and previous control actions. Isaacs discusses this possibility and reasons that, because the agents' control input can change abruptly without notice, it is impossible to rely on any prediction of the opponent's future position [130]. If an agent truly did adhere to some behavior beyond what is specified in the kinematics, then perhaps its opponent *could* improve its performance if it could correctly ascertain the behavior. What is gained in performance is lost in robustness, as is generally the case in the real world.

Let the state of the system be represented by

$$
\mathbf{x} \triangleq ((x_1, y_1), \dots, (x_M, y_M)).
$$
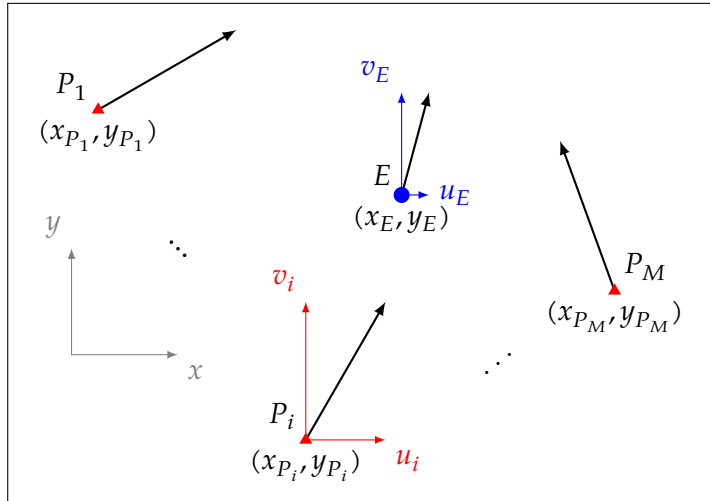
(a) Global
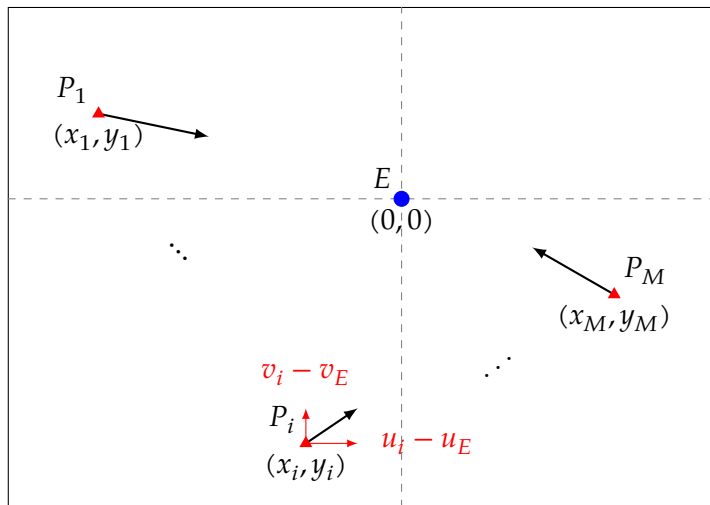
(b) Relative

Figure 7.1: Coordinate systems

Similarly, let

$$\mathbf{u}_E \triangleq (u_E, v_E) \text{ and } \mathbf{u}_P \triangleq (u_1, v_1, \dots, u_M, v_M)$$

represent the control inputs of the Evader and Pursuers, respectively. Note that each part of the state is a function of time, though it will not be notated explicitly. The set of terminal states for the scenario is defined by the requirement of point capture:

$$\Lambda = \{\mathbf{x} \mid \exists i,\ 1 \le i \le M \text{ s.t. } (x_i, y_i) = (0, 0)\}. \tag{7.4}$$

Remark 6. Eq. (7.4) allows one *or more* Pursuers to reach the position $(0, 0)$ at time T.

Another way of denoting the set of the terminal conditions is by defining

$$\Psi(\mathbf{x}) = \prod_{i=1}^{M} (x_i^2 + y_i^2), \tag{7.5}$$

and setting

$$\Lambda = \{\mathbf{x} \mid \Psi(\mathbf{x}) = 0\}. \tag{7.6}$$

The condition $\Psi(\mathbf{x}) = 0$ is akin to a terminal manifold of dimension $2M - 1$.

The terminal time T, or capture time, is the first time such that the system state enters Λ,

$$T = \inf\{t \mid \mathbf{x}(t) \in \Lambda\}. \tag{7.7}$$

Let the set of interceptors be those Pursuers whose positions are $(0, 0)$ at time T, i.e.,

$$\mathcal{I} = \{i \mid (x_i(T), y_i(T)) = (0, 0)\}. \tag{7.8}$$

The terminal time T is also the cost (or payoff) of the game:

$$J(\mathbf{u}_E(\mathbf{x}), \mathbf{u}_P(\mathbf{x})) = \int_0^T d\tau, \tag{7.9}$$

with $\mathbf{u}_E, \mathbf{u}_P$ subject to (7.3). The value function describes the minimax value of the cost function starting from the point $\mathbf{x}_0 = \mathbf{x}(0)$ at time $t = 0$ [19]

$$V(\mathbf{x}_0) = \min_{\mathbf{u_P}} \max_{\mathbf{u_E}} \int_0^T d\tau. \tag{7.10}$$

For such a function to exist the min and max must be interchangeable. The existence of this function or even of a saddle-point in the cost function is not guaranteed, however it will henceforth be assumed that at least the saddle-point exists. That is, the control inputs that satisfies the following is sought

$$J(\mathbf{u}_E, \mathbf{u}_P^*) \leq J(\mathbf{u}_E^*, \mathbf{u}_P^*) \leq J(\mathbf{u}_E^*, \mathbf{u}_P). \tag{7.11}$$

The control policies in (7.9) are state feedback policies, and thus $J(\mathbf{u}_E^*, \mathbf{u}_P^*)$ represents a feedback saddle-point equilibrium. The Isaacs Equation [130] can be written as

$$\min_{\mathbf{u}_P} \max_{\mathbf{u}_E} \sum_{i=1}^{M} V_{x_i}(u_i - u_E) + V_{y_i}(v_i - v_E) + 1 = 0, \tag{7.12}$$

where

$$V_{x_i} = \frac{\partial V}{\partial x_i},$$
$$V_{y_i} = \frac{\partial V}{\partial y_i}.$$

For many problems in [130], the Isaacs Equation (7.12), along with information about the terminal surface, is sufficient to generate game-optimal trajectories. Indeed, this is also the approach taken in [85]. If this is attempted here, the curse of dimensionality bites us and an impasse is reached.

To circumvent the need to deal directly with the value function itself, an analysis based upon open-loop strategies is utilized. As in many of the examples in [19] the explicit dependence of the control policies on the state is dropped

$$\mathbf{u}_P(t, \mathbf{x}) = \hat{\mathbf{u}}_P(t), \qquad \mathbf{u}_E(t, \mathbf{x}) = \hat{\mathbf{u}}_E(t). \tag{7.13}$$

Here, $\hat{\mathbf{u}}_P(t)$ and $\hat{\mathbf{u}}_E(t)$ represent open-loop controls for the Pursuers and Evader, respectively.

7.2.1   *Necessary Conditions for Optimality*

Under the assumptions that the pair $(\mathbf{u}_P^*, \mathbf{u}_E^*)$ provides a saddle-point solution in feedback strategies and the corresponding open-loop representation $(\hat{\mathbf{u}}_P^*, \hat{\mathbf{u}}_E^*)$ provides a saddle-point solution in open-loop policies then Theorem 2 of [19] provides a framework for deriving the necessary conditions for optimality. The procedure is based upon the two-person extension to the PMP. First let the Hamiltonian be given by

$$H(t, \lambda(t), \hat{\mathbf{u}}_P, \hat{\mathbf{u}}_E) = 1 + \lambda^\top \begin{bmatrix} u_1 - u_E \\ v_1 - v_E \\ u_2 - u_E \\ v_2 - v_E \\ \vdots \\ u_M - u_E \\ v_M - v_E \end{bmatrix}, \tag{7.14}$$

where λ is the costate vector in $\mathbb{R}^{2M}$

$$\lambda = (\lambda_{x_1}, \lambda_{y_1}, \dots, \lambda_{x_M}, \lambda_{y_M}).$$

Thus the Hamiltonian becomes

$$H = 1 + \sum_{i=1}^{M} \lambda_{x_i}(u_i - u_E) + \lambda_{y_i}(v_i - v_E), \tag{7.15}$$

where the elements u_i and v_i, $i = 1, \dots, M$, and u_E and v_E are taken to mean the corresponding elements in the open-loop policies $\hat{\mathbf{u}}_P$ and $\hat{\mathbf{u}}_E$, respectively. The costate variables satisfy

$$\dot{\lambda}_{x_i} = -\frac{\partial H}{\partial x_i} = 0, \qquad \dot{\lambda}_{y_i} = -\frac{\partial H}{\partial y_i} = 0, \qquad \text{for } i = 1, \dots, M. \tag{7.16}$$

Eq. (7.14) does not depend explicitly on the state $\mathbf{x}$. This is a result of the fact that the kinematics are only a function of the control inputs. Thus, it must be that $\dot{\lambda} = 0$ which implies the costate variables are con-

stant w.r.t. time (e.g., $\lambda_{x_1}(t) = \lambda_{x_1} = const$). The minimizing controls for the Pursuers can easily be obtained from (7.3) and (7.15) as

$$u_i^* = -\frac{\lambda_{x_i} V_P}{\sqrt{\lambda_{x_i}^2 + \lambda_{y_i}^2}}, \qquad v_i^* = -\frac{\lambda_{y_i} V_P}{\sqrt{\lambda_{x_i}^2 + \lambda_{y_i}^2}}, \qquad (7.17)$$

which are also constant. Because the optimal control policy for the Pursuers is constant, their state trajectories are straight lines in the global coordinate system. The only caveat is that the optimal control policy is defined only if $\lambda_{x_i} \neq 0$ or $\lambda_{y_i} \neq 0$. When $\lambda_{x_i} = \lambda_{y_i} = 0$ the implication is that the value of the game is not sensitive to Pursuer i. When the ith Pursuer participates in the capture of the Evader (i.e., $x_i(T) = y_i(T) = 0$), either one of λ_{x_i} or λ_{y_i} must be nonzero.

7.3 SOLUTION

Now consider the original coordinate system, and without loss of generality consider the starting position of the Evader to be the origin and the Evader's speed to be unity. Let the ratio of Evader's max speed to Pursuers' speed be given by $\alpha = V_{E_{max}}/V_P < 1$. The necessary conditions for optimality derived above help to justify the following:

Proposition 7.1. *The optimal trajectories of the agents are straight line paths.*

Moreover, under optimal play, the Pursuers' heading should not change; that is, their trajectory is comprised of a single straight line segment. Thus, the differential game problem of min-max capture time is re-posed as finding the coordinates $I = (x_I, y_I)$ that maximize the Evader's life assuming all the Pursuers head directly to that point starting at $t = 0$. Then the terminal time (which is also the cost/payoff of the game in the preceding section) can be written simply as

$$\bar{T}(\mathbf{x}, I) = \min_i \alpha \sqrt{(x_{P_i} - x_I)^2 + (y_{P_i} - y_I)^2}, \qquad (7.18)$$

which is the smallest time for any of the Pursuers to reach the designated intercept point. There is a subtle shift here from the traditional game-theoretic framework wherein the agents have absolutely no knowl-

edge of the opponents' present or future control action. Here, the understanding is that the point I represents a designated intercept point that both the Evader and Pursuers have knowledge of. Thus (7.18) is the capture time assuming that all the agents head directly to the designated (or 'agreed upon') point. Suppose the Evader were to choose the point I, then (assuming I could be reached safely by the Evader) (7.18) gives the worst-case capture time from the perspective of the Evader. Conversely, if the Pursuers were to select point I, it is in the Evader's best interest to flee from this point rather than aim towards it. Therefore, the point I is treated as if it is the choice of the Evader:

$$
\begin{aligned}
T(\mathbf{x}) &= \max_{(x_I, y_I)} \bar{T}(\mathbf{x}, I) \\
&= \max_{(x_I, y_I)} \min_i \alpha \sqrt{(x_{P_i} - x_I)^2 + (y_{P_i} - y_I)^2},
\end{aligned}
\tag{7.19}
$$

wherein the Evader is selecting the best choice among all the worst-case capture times, $\bar{T}$. Eq. (7.19) is *analogous* to the value of the game in (7.10), but not necessarily equivalent. As hinted previously, the order of max-min in (7.19) is not interchangeable. In fact, interchangeability of max and min is *required* to obtain a saddle point solution to a differential game (known as Isaacs Condition) [130].

7.3.1   *Constraints*

Eq. (7.19) is not particularly useful because the Evader could designate an intercept point at infinity and thus the capture time would also be infinite. In other words, (7.19) is unconstrained. This should be amended by stating (1) the Evader must be able to reach point I in T time and (2) the Evader must be able to reach point I *safely*, that is, capture of the Evader *en route* to I ought not be possible. Constraint (1) is easily formulated as [246]

$$
x_I^2 + y_I^2 \leq T^2,
\tag{7.20}
$$

since the Evader's speed and starting position are one and $(0, 0)$, respectively. The need for (2) is due to the fact that if capture is possible prior to the Evader reaching I, then the associated terminal time
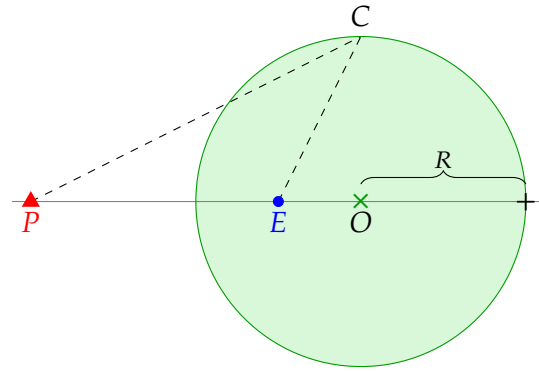
Figure 7.2: AC for a single Pursuer.

T is meaningless in the sense that it is no longer analogous to a game-theoretic solution. To formalize (2), the concept of the AC is useful. In the context of pursuit and evasion, the AC is the locus of points such that the ratio of distances to the Evader and Pursuer is equal to their speed ratio [130]. Thus, the AC defines the points that can be reached simultaneously by the Pursuer and Evader assuming they each head directly to the point at their maximum speeds. Inside the circle (i.e., the Apollonius *disk*) are points in which the Evader can reach before the Pursuer under the same assumptions. When the speed ratio is unity, the AC becomes the orthogonal bisector of the segment $\overline{EP}$. Figure 7.2 displays the AC. Note, the optimal capture point for this configuration is marked with a +. The center of the AC, O, marked with a green ×, lies on the line passing through the Evader and Pursuer. The definition given above stipulates that, for any point C on the circle, the ratio of $\overline{EC}/\overline{PC} = \alpha$. This relation allows one to express the location of the circle's center as well as its radius

$$\overline{EO} = \frac{\alpha^2}{1 - \alpha^2}\overline{PE}, \tag{7.21}$$

$$R = \frac{\alpha}{1 - \alpha^2}\overline{PE}. \tag{7.22}$$

Putting the circle center associated with Pursuer i into the global coordinate system gives

$$x_{O_i} = -\frac{\alpha^2}{1-\alpha^2}x_{P_i},\tag{7.23}$$

$$y_{O_i} = -\frac{\alpha^2}{1-\alpha^2}y_{P_i},\tag{7.24}$$

$$R_i = \frac{\alpha}{1-\alpha^2}\sqrt{x_{P_i}^2 + y_{P_i}^2}.\tag{7.25}$$

The set describing the Apollonius disk associated with Pursuer i is

$$\mathcal{D}_i = \left\{(x,y) \mid (x - x_{O_i})^2 + (y - y_{O_i})^2 \leq R_i^2\right\}.\tag{7.26}$$

For a point to be safely reachable by the Evader, that is, no Pursuer can reach the point or any point along the Evader's path before the Evader, the point and the Evader's path to the point must lie inside the Apollonius disk associated with each Pursuer

$$I \in \cap_{i=1}^{M}\mathcal{D}_i,$$
$$(x_I - x_{O_i})^2 + (y_I - y_{O_i})^2 \leq R_i, \ i = 1,\ldots,M.\tag{7.27}$$

Let the intersection of Apollonius disks be represented by $\mathcal{S} = \cap_{i=1}^{M}\mathcal{D}_i$; $\mathcal{S}$ is referred to as the Evader's *safety region*. Eq. (7.27) is only a necessary condition, however, it does not guarantee safety. First of all, it only stipulates that the point I is inside every Apollonius disk. It is possible for the Evader to take a path which may leave the set $\mathcal{S}$, in which case there exists a valid Pursuer path which leads to capture at the boundary, $\partial\mathcal{S}$. Moreover, one must recompute $\mathcal{S}$ at each instant of time as the game is played; thus under optimal play, $\mathcal{S}$ shrinks over time as the Pursuers approach the Evader. To guarantee safety, then, the point I must lie inside the instantaneous region $\mathcal{S}(t)$ for all $t \leq T$; at T the region $\mathcal{S}$ collapses to a point.

Property 1. A safe Evader path is one in which, given the initial positions of all the agents, $\mathbf{x}_0$, and their velocities, V_E and V_P, there does not exist a Pursuer path that captures the Evader *en route* to its destination.

Lemma 7.1. $\mathcal{S}$ *is a convex set.*

Proof. $\mathcal{S}$ is the intersection of Apollonius disks, and each Apollonius disk is a convex set, thus for any two points $p_1, p_2 \in \mathcal{S}$,

$$p_1, p_2 \in \mathcal{D}_i, \qquad i = 1, \dots, M,$$

from the definition of set intersection and,

$$\mu p_1 + (1 - \mu)p_2 \in \mathcal{D}_i, \qquad i = 1, \dots, M, \ 0 \leq t \leq 1,$$

by convexity of the $\mathcal{D}_i$. Therefore, by set intersection, all of these points are also in $\mathcal{S}$, implying convexity [49]. □

Corollary 7.1. *Any straight-line path starting inside and ending inside $\mathcal{S}$ lies entirely in $\mathcal{S}$.*

Theorem 7.1. *Given an Evader and Pursuers at P_i, $i = 1, \dots, M$ and point $I \in \mathcal{S}$, if the Evader travels directly towards I at maximum speed then the Evader will reach I safely. There does not exist a Pursuer path that can intercept the Evader before $\bar{T}$ according to (7.18).*

Proof. By construction, the Evader's initial position of $(0,0)$ lies inside $\mathcal{S}$. Every point along the straight-line path from $(0,0)$ to I is inside $\mathcal{S}$ from Corollary 7.1. The Evader can reach I under the prescribed Evader policy at or before the time that any Pursuer can reach it by construction of the Apollonius disks $\mathcal{D}_i$ and the fact that $I \in \mathcal{S}$. This implies that the point I lies inside $\mathcal{S}(t)$ for $t \leq t_1$ where

$$t_1 = \sqrt{x_I^2 + y_I^2} \leq T,$$

is the time at which the Evader reaches I. The last inequality is enforced by (7.20). For $t_1 \leq t \leq T$ the Evader's position $E = I$, thus $I \in \mathcal{S}$ for $t_1 \leq t \leq T$ which completes the proof. □

Theorem 7.1 specifies an Evader policy which is guaranteed to be safe in the sense that (7.18) gives a worst-case capture time. However, this policy is not unique and there may be many alternative policies which are also safe [232, 246].

Proposition 7.2. *Given an Evader at E and Pursuers at P_i, $i = 1, \dots, M$ and point $I \in \mathcal{S}$, if the Evader travels directly towards I at speed s then the Evader will reach I safely, where*

$$s \triangleq \frac{1}{\bar{T}} \sqrt{x_I^2 + y_I^2}. \tag{7.28}$$

That is, there does not exist a Pursuer path that can intercept the Evader before $\bar{T}$ according to (7.18).

Another policy, which is analogous to the policy in Theorem 7.1 is for the Evader to head to I at maximum speed and then switch heading between 0 and π infinitely often while remaining at maximum speed [246]. In fact, allowing the Evader to modulate its heading continuously alleviates entirely the need to consider its speed as a control variable. Oddly enough, a straight-line path is not strictly necessary to guarantee the Evader can reach I safely (nor is it sufficient). An example scenario is included in Section 7.5 wherein the Evader takes a safe, non-straight-line path to I corresponding to the arg max in (7.19). For the remainder, the Evader is restricted to employ the policy prescribed in Theorem 7.1 since it is safe and also consistent with Proposition 7.1.

7.3.2 *Linear Program with Quadratic Constraints*

The multiple Pursuer single Evader game has been re-posed as solving (7.19) subject to the constraints (7.20) and (7.27) with the understanding that all the agents head directly to I corresponding to the arg max of (7.19). In order to solve this new problem, it is formulated as a linear program:

$$\max_{\mathbf{z}} \mathbf{c}^\top \begin{bmatrix} \mathbf{z} \\ \mathbf{s} \end{bmatrix}, \tag{7.29}$$

with slack variables $\mathbf{s} = \begin{bmatrix} s_1 & s_2 & \dots & s_M \end{bmatrix}^\top$ subject to the constraints

$$g_i(\mathbf{z}) = 0, \qquad i = 1, \dots, M, \tag{7.30}$$

$$-g_E(\mathbf{z}) \leq 0, \tag{7.31}$$

$$-s_i \leq 0, \qquad i = 1, \dots, M, \tag{7.32}$$

where $\mathbf{c}^\top = \begin{bmatrix} 1 & 0 & ... & 0 \end{bmatrix}_{1\times(M+3)}$, $\mathbf{z}^\top = \begin{bmatrix} m & x & y \end{bmatrix}$. The functions g are defined as

$$g_i(\mathbf{z}) = \frac{m}{\alpha^2} - (x - x_{P_i})^2 - (y - y_{P_i})^2 + s_i, \qquad i = 1, ..., M, \quad (7.33)$$

and

$$g_E(\mathbf{z}) = m - x^2 - y^2. \tag{7.34}$$

Note that the problem is about the maximization of a linear function with quadratic constraints. Furthermore, the constraints (7.30) are active whether or not the point (x, y) is on ∂S or $S \cap \overline{\partial S}$ (i.e., on the boundary or interior of S). The set of Pursuers which intercept the Evader at (x, y) is given as,

$$\mathcal{I} = \{i \mid s_i = 0, \ i = 1, ..., M\} \tag{7.35}$$

hence, s_i represents the remaining time required for Pursuer i to reach (x, y) after the Evader and interceptors have reached it. Thus, (7.33) with (7.32) is analogous to the AC constraint introduced in (7.27). In other words, $s_i > 0$ holds for any point inside Pursuer i's AC. Similarly, (7.34) and (7.31) correspond to the reachability constraint (7.20). This constraint is always satisfied, however, by the definition of the AC and the fact that (7.30) and (7.32) constrain the point to be inside S.

This linear program is now amenable to solution using a generic numerical optimization scheme. Consider the following example, whose numerical results are shown in Figure 7.3.

Example 1.

$$M = 4$$
$$P_1 = (\cos(-\pi/6), \sin(-\pi/6))$$
$$P_2 = (\cos(7\pi/6), \sin(7\pi/6))$$
$$P_3 = (0, 1)$$
$$P_4 = (0, -0.5)$$
$$V_P = 1.5$$

In Figure 7.3 and those to follow the Pursuers are indicated with triangles and colored according to whether or not they are in the set $\mathcal{I}$

(a) Local maximum, $(x, y) = (0, 0.25)$ and $T = 0.5$.

(b) Maximum, $(x, y) = (0.43, 0.25)$ and $T = 0.577$.
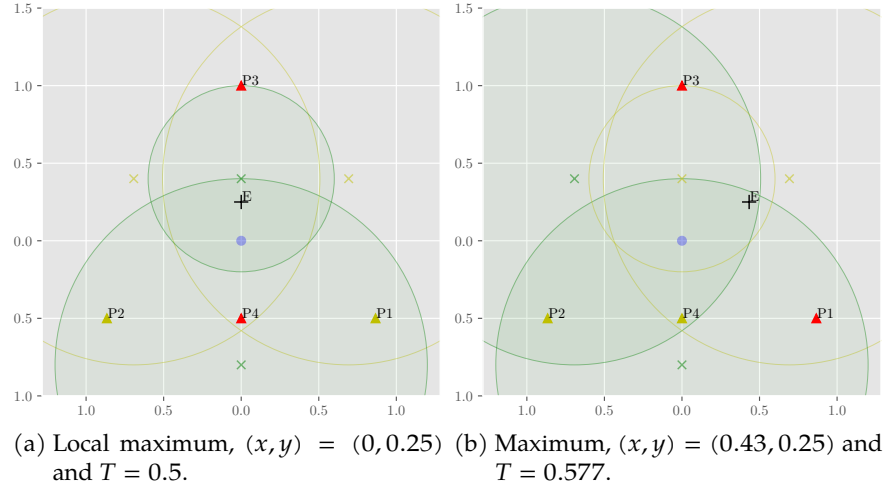
Figure 7.3: Numerical results for Example 1

(red if so, yellow if not). The initial position of the Evader is marked by a blue circle and the capture point is marked with + and E. The ACs for those Pursuers in $\mathcal{I}$ are shown as translucent green and for those not in $\mathcal{I}$ just the border is marked with yellow. Finally, the circle centers are marked with ×. As previously mentioned, the set I is computed from (7.35). When $s_i = 0$, the constraints g_E and g_i define the AC associated with Pursuer i. Thus adding the AC constraints, (7.27), will not change the optimization problem. In fact, once the AC constraints are introduced, the slack variables can be removed from the problem formulation and use $g_i(\mathbf{z}) \leq 0$.

For the scenario in Example 1 there is a local maxima (shown in Figure 7.3a) and two global maxima which have equal value. One of the global maxima is shown in Figure 7.3b, and the other is symmetric about the y axis. The matter of local maxima and the possibility for multiple global maxima presents a practical issue. In general, it is not known how many maxima or local maxima may be present in (7.29). One may need to initialize the numerical optimization procedure many times to uncover all the different maxima, and even then an upper bound on this number is not known at this time. The two initial conditions used for Example 1 are $\mathbf{z} = \begin{bmatrix} 0.1 & 0.1 & 1 \end{bmatrix}^\top$ and $\begin{bmatrix} 0 & 0.1 & 1 \end{bmatrix}^\top$, but these may not be sufficient to find the true global maximum. In

summary, the numerical optimization process is blind to any special structure of the problem and sensitive to the presence of local maxima.

7.3.3 Geometric Approach

The linear program introduced in the previous section must search over a continuous space; it was mentioned that the linear program does not make use of any special structure to search the space intelligently. Consider the following as a motivating example for finding some additional useful information embedded in the problem:

Example 2.

$$M = 4$$
$$P_1 = (\cos(-\pi/6), \sin(-\pi/6))$$
$$P_2 = (\cos(7\pi/6), \sin(7\pi/6))$$
$$P_3 = (0, 1)$$
$$V_P = 1.5$$

Note, this example is the same as Example 1 but with P_4 removed.

Now, the numerical optimization of the linear program is repeated as before, but with initial $z = \begin{bmatrix} 0, & 0, & 1 \end{bmatrix}^\top$ and $\begin{bmatrix} 0.502, & 0.29, & 1 \end{bmatrix}^\top$. Figure 7.4 shows the results of the numerical optimization. Like Example 1 there are a number of local maxima and global maxima. In this case, there are three local maxima, one of which is shown in Figure 7.4a. The other two local maxima are radially symmetric w.r.t the origin. Now, the global maximum is unique and lies in $\mathcal{S} \cap \overline{\partial \mathcal{S}}$, that is, in the interior of the region $\mathcal{S}$ whereas before the global maximum was near (essentially on) the border $\partial \mathcal{S}$. The former case is easy to understand – it corresponds to the furthest distance from the Evader to a point on $\partial \mathcal{S}$. In the latter case it becomes optimal for the Evader to not move at all.

Remark 7. Simultaneous capture by two Pursuers, an example of which is shown in Figure 7.4a, occurs at the intersection of ACs associated with the intercepting Pursuers.

(a) Local max, $(x, y) = (0.502, 0.29)$ and $T = 0.58$.

(b) Max, $(x, y) = (0, 0)$ and $T = 0.667$.
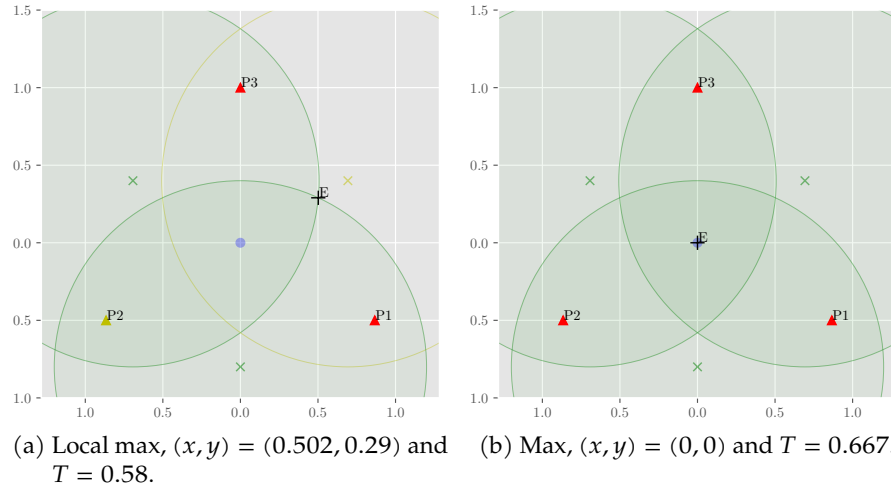
Figure 7.4: Numerical results for Example 2 highlighting some special properties of the solution.

Figure 7.4b, on the other hand, depicts simultaneous capture by three Pursuers.

When Isaacs first posed the two cutters and fugitive ship problem [130] he posited that the optimal strategy would be for all the agents to head to the further of the two AC intersection points. Indeed, this was proven to be true in certain regions of the state space [107]. Of course, the game may also degenerate to capture by a single Pursuer, which also holds true for the multiple Pursuer case as shall be explored later on. The geometry and solution of the two Pursuer game is driven by the fact that optimality dictates straight-line paths. From Proposition 7.1 and its preceding analysis, the analysis becomes primarily geometric in which the ACs are of chief importance. However, from Figure 7.4b it is plain that the optimal intercept point $(0, 0)$ has little to do with the ACs themselves. Up to now only the division of the state space into regions dominated by the Evader and a particular Pursuer have been considered. Now consider the region dominated by a particular Pursuer versus all the other Pursuers; that is, the region of the state space where Pursuer i can reach before any other Pursuer $j \neq i$. This partitioning of the state space (which, at the moment, leaves out any consideration of the Evader) is precisely a Voronoi diagram [58]. Because the Pursuers share the same velocity, the diagram is comprised of straight line segments which partitions the entire xy plane. Figure 7.5 shows
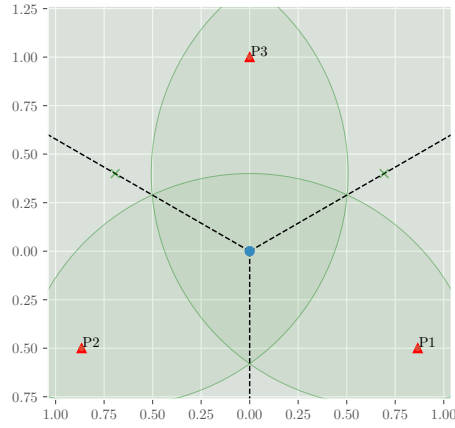
Figure 7.5: Example 2 with the Pursuer Voronoi diagram overlaid. The vertex is colored blue and the cell borders are shown as dashed lines.

the same setup but with the Pursuer Voronoi diagram overlaid.

Remark 8. The AC intersections lie on the edges of the Pursuer Voronoi diagram. Thus, for simultaneous capture by two Pursuers to be optimal it must occur on an edge of the Pursuer Voronoi diagram.

Remark 9. The optimal capture point for Example 7.4b which represents simultaneous capture by three Pursuers coincides with the vertex of the Pursuer Voronoi diagram.

This last observation, in particular, drives the remainder of the analysis and allows us to generalize the multiple Pursuer single Evader game to any number of Pursuers. Also, note that the application of Voronoi diagrams to the analysis of pursuit-evasion games is not novel and has been explored extensively in the literature (c.f. [16, 46, 128, 182, 192]). The Voronoi diagram, in 2D space, defines a tessellation in which each agent resides in their own cell defining points in space they can reach before any other agent. This construct is particularly useful in the present context as it has been established that the optimal trajectories ought to be comprised of constant-heading paths. Define two different Voronoi diagrams, each parameterized as a set of vertices, a set of edges, and a set of agent positions (which correspond to so-called generator points) [246],

$$\mathbb{V}_E = \left(\mathcal{U}_E, \ \mathcal{E}_E, \ \{E, P_1, \dots, P_M\} \right),$$
$$\mathbb{V}_P = \left(\mathcal{U}_P, \ \mathcal{E}_P, \ \{P_1, \dots, P_M\} \right).$$

Note, the only difference between the two is that $\mathbb{V}_E$ includes the Evader as a generator point, and $\mathbb{V}_P$ use only the Pursuer positions as generator points. The latter describes a partitioning wherein the Pursuers can reach points in their own cell before any other Pursuer. Because the Pursuers share the same velocity the edges $\mathcal{E}_P$ are segments of the perpendicular bisectors between neighboring Pursuers. The same is not true, however, for the edges of the Evader's cell in $\mathbb{V}_E$, since E is slower than the neighboring Pursuers. Typically, when different weights (velocities) are involved one may consider $\mathbb{V}_E$ to be a multiplicatively-weighted Voronoi diagram [58]. Note that the Evader's cell in $\mathbb{V}_E$ is exactly the safety region, $\mathcal{S}$, which has already been defined using the ACs. Now let $\mathcal{S}$ be parameterized as

$$\mathcal{S} = (\mathcal{V}_{\mathcal{S}}, \mathcal{E}_{\mathcal{S}}), \tag{7.36}$$

where $\mathcal{V}_{\mathcal{S}}$ is an ordered set of vertices of $\mathcal{S}$, and $\mathcal{E}_{\mathcal{S}}$ is an ordered set of arcs [246]. The ordering of these two sets is such that the ith edge in $\mathcal{E}_{\mathcal{S}}$ connects the $i - 1$th and ith AC intersection in $\mathcal{V}_{\mathcal{S}}$.

7.3.4 *Types of Solutions*

In the introduction it was noted that for the case of $M > 2$ there is a new type of solution over the previously derived solutions to the one-on-one and two-on-one scenarios. First, let us briefly recount these known solutions as they are still solutions to the M-on-one scenario for particular configurations. As example, consider a case with $M > 2$ wherein one Pursuer is very close to the Evader and the other $M - 1$ Pursuers are very (let us say "infinitely") far away. Then the solution, obviously, degenerates to the solution of the one-on-one scenario between the close Pursuer and the Evader. The solution, in this case, is given by pure pursuit: the Evader's and Pursuer's heading should be along the line-of-sight [130]. Capture occurs on the point on the AC which is antipodal to the Pursuer.

Hugo Steinhaus and Rufus Isaacs each proposed the case of two Pursuers against one Evader. Isaacs referred to this scenario as the "two

cutters and fugitive ship problem" [130]. In his book, Isaacs posited that the solution of the game was for all three agents to head to the intersection of ACs furthest from the Evader. Note, however, that Isaacs did not mention the cases in which the two-on-one scenario degenerates to one-on-one. Nonetheless, using the geometric intuition of Isaacs, the Value function and saddle-point strategies were derived only recently in [107].

As will be shown in the following sections, it is not necessary for some Pursuers to be "infinitely far" away from the Evader for the scenario to degenerate to two-on-one or one-on-one. Instead, the solutions to all of these sub-scenarios are retained as candidate solutions. Note that each of the candidates described above necessarily occur on the boundary of the Evader's dominance region $\partial \mathcal{S}$. When $M \geq 3$ new candidate solution must be considered, one in which capture occurs in the interior of $\mathcal{S}$, as opposed on the boundary $\partial \mathcal{S}$. It is claimed that when the solution to (7.19) occurs in the interior of $\mathcal{S}$, the point necessarily corresponds to a vertex of $\mathbb{V}_P$, the Voronoi diagram for the Pursuers [246]

$$(x^*, y^*) \notin \partial \mathcal{S} \implies (x^*, y^*) \in \mathcal{V}_P, \tag{7.37}$$

and thus implies simultaneous capture by three or more Pursuers. Now, with all possible candidate solutions described, one of the main results in [246] is included which states that the solution of (7.19) is among this finite set of candidates comprised of points representing capture by a single Pursuer $(\mathcal{P}_{\mathcal{S}}^1)$, capture by two Pursuers simultaneously $(\mathcal{V}_{\mathcal{S}})$, and capture by three or more Pursuers simultaneously $(\mathcal{V}_{P_{\mathcal{S}}})$.

Theorem 7.2.

$$(x^*, y^*) \in \mathcal{P}_{\mathcal{S}}^1 \cup \mathcal{V}_{\mathcal{S}} \cup \mathcal{V}_{P_{\mathcal{S}}}, \tag{7.38}$$

where

$$(x^*, y^*) = \arg \max_{(x,y) \in S} \min_{i \in \{1, \dots, M\}} \|(x,y) - P_i\| \alpha_P,$$

$$\mathcal{P}_S^1 = \left\{ x_i, y_i \middle| x_i, y_i = R_i(1 + \alpha_P) \frac{E - P_i}{\|E - P_i\|}, \right.$$

$$\left. x_i, y_i \in \mathcal{S}, \ i = 1, \dots, M \right\},$$

$$\mathcal{U}_{P_S} = \mathcal{U}_P \cap \mathcal{S}.$$

(7.39)

Proof. The interested reader is referred to [246].   □

7.3.5 *Categories of Pursuers*

One of the interesting consequences of Theorem 7.2 is the fact that not all the Pursuers have an effect on the playout of the game. For example, a Pursuer that is very far away from the Evader compared to the other Pursuers may not be able to reach the Evader before capture occurs. This is certainly the case when the AC of a Pursuer completely contains the Evader's safety region, $\mathcal{S}$. As mentioned in [246] the set of Pursuers can be broken up into four disjoint sets. The membership of the Pursuers is a function of the optimal capture point I (i.e., the solution to the first equation in (7.39)). The first set contains those Pursuers who reach I at $\bar{T}$, according to (7.18). Earlier, this set was referred to as "interceptors" using the notation $\mathcal{I}$. The set of Pursuers whose AC completely contains the safety region $\mathcal{S}$ is given by

$$\mathcal{I}^- = \left\{ i \mid \partial \mathcal{S} \cap \partial \mathcal{D}_i = \emptyset, \ i = 1, \dots, M \right\}. \tag{7.40}$$

If capture is constrained to occur in $\mathcal{S}$, then these Pursuers have no effect on the game. They can be discarded completely; this set is referred to as "fully redundant". Then the set of Pursuers who share an edge of the Evader's cell of $\mathbb{V}_E$ is given by

$$\mathcal{I}_S = \{1, \dots, M\} \setminus \mathcal{I}^-. \tag{7.41}$$

The set $\mathcal{I}_S \setminus \mathcal{I}$ contains Pursuers which neighbor the Evader in $\mathbb{V}_E$ and can be further broken down by recomputing the solution (i.e., optimal intercept point I and capture time T) and checking if the new solu-
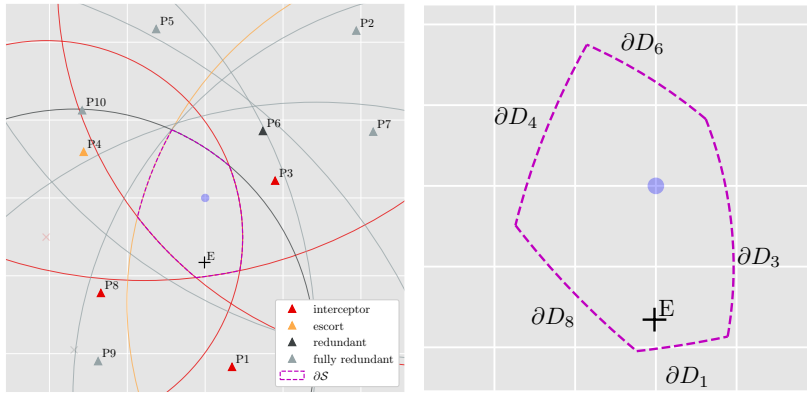
(a) Full scenario with solution shown. (b) Detail of the Evader's region of dominance with each arc segment labeled with the corresponding Apollonius disk.

Figure 7.6: Categories of Pursuers. The color of each Pursuer with its AC denote which category it belongs to. The boundary of the Evader's safety region is rendered in dashed magenta.

tion changed. Let the new solution of the game with the ith Pursuer removed be denoted as I^i and T^i for the intercept point and capture time, respectively. If $T^i = T$ and $I^i = I$, removal of Pursuer i had no effect on the game and thus it is "redundant"; these Pursuers are denoted as $\mathcal{I}_R$. Alternatively, it could be $T^i > T$ or $I^i \neq I$, and thus removal of Pursuer i led to a decrease in the Pursuers' performance or a change in the optimal intercept point, respectively. These Pursuers are referred to as "escorts", denoted $\mathcal{I}_E$. It is necessary for the escorts to "play the game" by heading towards I; alternatively they could simply implement pure pursuit. Their purpose is to control the dynamic shape of $\mathcal{S}$ so that no new AC intersection becomes feasible or advantageous for the Evader. Because they do not actually participate in capture, it is difficult to say precisely what is optimal for these Pursuers [246]. One interesting research question is what an escort could or should do when it is an escort for more than one Evader in an M on N pursuit-evasion scenario.

Figure 7.6 contains an example wherein all the categories are represented. In this example, simultaneous capture is achieved by P1, P3, and P8. If P4 were removed from the game, the Evader could flee to the Apollonius intersection directly to its left and reach it safely, thus

increasing capture time. Therefore, P4 must play the game by shaping $\mathcal{S}$. P6 neighbors E in $\mathbb{V}_E$ which is homologous to stating that part of the boundary of its AC coincides with the boundary of $\mathcal{S}$. Removal of P6 from the game would make other AC intersections safe/feasible for the Evader, however these points are suboptimal. Thus, P6 does not have any effect on the solution provided P1, P3, P4, and P8 behave optimally.

Fortunately, the set $\mathcal{I}^-$ can be determined prior to solving (7.39). In Figure 7.6, the safety region $\mathcal{S}$ lies entirely inside the interior of the fully redundant Pursuers' ACs. Thus, in this case $M = 10$, five may be discarded immediately. The following section presents an algorithm for computing the set $\mathcal{I}_{\mathcal{S}}$, which are the Pursuers who are neighbors of the Evader in $\mathbb{V}_E$. This reduced set includes all the Pursuers who could possibly affect the solution. Interestingly, in this example, using the linear program method yields a different capture point wherein P3, P4, and P8 are interceptors and the capture time is 0.4% worse for the Evader.

7.4   GEOMETRIC ALGORITHMS

For general initial conditions, the size of the set of Pursuers who are neighbors of E in $\mathbb{V}_E$, namely $|\mathcal{I}_s|$, is usually close to four. This could be true even when m is very large (say, in the thousands). However, $|i_s|$ may be large when the positions of the pursuers are highly correlated. The extreme case would be when the pursuers lie on a ring (i.e., are equidistant) from the evader. In that case, $|i_s| = m$ and all the AC will contribute edges to s. For the algorithm to follow, this will yield worst-case performance. In this section, two algorithms are briefly described and summarized: an algorithm which simultaneously computes i_s and s, and another which computes the optimal intercept point (according to (7.39)) given s. The paper [246] contains a complete specification of these algorithms including some supporting lemmas and analysis of the computational complexity.

First note that s corresponds to the evader's cell of $\mathtt{v}_e$, which is the multiplicatively-weighted voronoi diagram consisting of all the agents [246]. Part of the motivation for computing s quickly is the fact that existing approaches (like computing the entire multiplicatively-weighted voronoi diagram) can be $o(m^2)$ [11]. Note also that to construct s, the union of m disks is the essential computation, for which there is an algorithm which takes $\theta(m \log m)$ time [39]. The algorithm described here differs in that no geometric transformations are used. The algorithm which computes s and i_s is referred to as EVADERCELL. The EVADERCELL algorithm takes as input the positions of all the agents and the speed ratio α. First, the ACs for each Pursuer are computed. Then the Pursuers are ranked according to the minimum distance from the Evader to a point on their AC. This measure is useful because it takes into account both the distance from the Pursuer to the Evader as well as its speed. The latter piece makes this algorithm applicable in the case where the Pursuers have different velocities. However, for the present case of equal-speed Pursuers, Pursuers can be simply ranked according to distance from the Evader. Note, for computational speed this ranking is achieved using a heap, in lieu of a full sort, since, in the general case, not all the Pursuers require consideration. EVADERCELL constructs $\mathcal{S}$ iteratively, considering a single AC in each iteration. Initially, the region $\mathcal{S}$ is initialized as the AC of the closest Pursuer. Then, as Pursuers (and their ACs) are dequeued from the heap EVADERCELL computes the intersection of the new AC with the current region $\mathcal{S}$. Eventually, the closest point on a Pursuer's AC will be further from the Evader than the furthest point on $\mathcal{S}$, then it is known that this Pursuer, and all subsequent Pursuers in the heap, fall into the *fully redundant* category [246]. Therefore, the construction of $\mathcal{S}$ is complete and $\mathcal{I}_{\mathcal{S}}$ is simply the set of Pursuers that had been dequeued previously. Thus, EVADERCELL returns the ordered sets of vertices and arc segments $(\mathcal{V}_{\mathcal{S}}, \mathcal{E}_{\mathcal{S}})$ and $\mathcal{I}_{\mathcal{S}}$.

Now, several of the lemmas from [246] are summarized which affirm the correctness of EVADERCELL.

1. The Pursuer closest to E is in $\mathcal{I}_{\mathcal{S}}$ and has an AC which comes the closest to E.

2. The closest distance of the Pursuers' AC s is monotonically non-decreasing with each iteration (due to the queue/heap property).

3. The point on $\partial\mathcal{S}$ furthest from the Evader is either at the intersection of two AC s or corresponds to the solution to one of the one-on-one games.

4. The distance from the Evader to the point on $\partial\mathcal{S}$ furthest away is monotonically non-increasing with each iteration.

5. None of the remaining Pursuers' AC s intersect $\mathcal{S}$ once a Pursuer whose AC 's minimum distance to the Evader is greater than the furthest point from the Evader on $\partial\mathcal{S}$ is reached.

6. There can be at most one one-on-one optimal intercept point on $\partial\mathcal{S}$.

7. If a one-on-one optimal intercept point lies on $\partial\mathcal{S}$, it is the optimal intercept point for the overall game.

Once $\mathcal{S}$ has been constructed computing the optimal intercept point, which is the solution to (7.39), is very straightforward. This process is referred to as MPURSUER1EVADER [246]. The MPURSUER1EVADER algorithm is based entirely off of Theorem 7.2, and therefore proceeds by computing all the candidate solutions and comparing the distance from the nearest Pursuer to each candidate. From Theorem 7.2 the optimal intercept point is

$$(x^*, y^*) = \underset{(x,y)\in\mathcal{P}_\mathcal{S}^1\cup\mathcal{V}_\mathcal{S}\cup\mathcal{V}_{P_\mathcal{S}}}{\arg\max} \min_{i\in\{1,\dots,M\}} \|(x,y) - P_i\|\alpha_P, \qquad (7.42)$$

where $\mathcal{P}_\mathcal{S}^1$ is the set of single Pursuer optimal intercept points in $\mathcal{S}$, $\mathcal{V}_\mathcal{S}$ are the vertices of the Evader's region of dominance (given by EVADER-CELL), and $\mathcal{V}_{P_\mathcal{S}}$ is the set of vertices of the Pursuer-only Voronoi diagram inside $\mathcal{S}$. Due to the seventh item in the previous paragraph, the first thing to do is check whether any of the one-on-one optimal intercept points are on $\partial\mathcal{S}$. Because the Pursuers share the same speed, it is either the case that the solution to the one-on-one game between the Pursuer closest to the Evader and the Evader is on $\partial\mathcal{S}$ or none of the

one-on-one solutions are on $\partial \mathcal{S}$. Thus this step is trivial. If it is not the case that the optimal intercept point of a one-on-one game is on $\partial \mathcal{S}$, then one must turn to the other candidates. As mentioned previously, the points $\mathcal{U}_{\mathcal{S}}$ are given by EVADERCELL. The last subset of candidate solutions to compute, then, is $\mathcal{U}_{P_{\mathcal{S}}}$ for which the Voronoi vertices of $\mathbb{V}_P$ must first be computed. Because it has been assumed that EVADERCELL is used prior to this point one can essentially forget about any of the Pursuers not in the set $\mathcal{I}_{\mathcal{S}}$. To compute the Voronoi vertices between the Pursuers Fortune's Algorithm [79] is employed over the Pursuers in $\mathcal{I}_{\mathcal{S}}$. Then, to get $\mathcal{U}_{P_{\mathcal{S}}}$ one simply checks to see if each Voronoi vertex is inside or outside of $\mathcal{S}$. Finally, the distance from each candidate in $\mathcal{U}_{\mathcal{S}} \cup \mathcal{U}_{P_{\mathcal{S}}}$ to its nearest Pursuer is computed (RHS of (7.42)), and the optimal intercept point is the candidate which maximizes this distance. The terminal time is simply the travel time of the Pursuer nearest that point via a straight-line path.

7.5   RESULTS

The goal of this investigation is to consider full cooperation among the Pursuers. Alternatively, the Pursuers could, for example, act entirely independently of one another. In the latter case, each Pursuer may implement its best strategy in the sense of one-on-one, that is PP. The approach detailed here and in [246] is referred to as the geometric (G) policy. Note that it was not claimed that G policy, if implemented by the Pursuers and the Evader, is a saddle-point pair of strategies in the sense of a feedback Nash equilibrium. Therefore, some numerical simulations are included here in order to demonstrate some of the merits of the G policy in comparison to other strategies, namely PP. The comparison of these two policies is done using a discrete time numerical simulation in which agents evaluate their respective control policies at each time step. For those agents implementing the G policy, the optimal intercept point (x^*, y^*) is computed using the current positions of the agents, and then those agents' heading would aim towards this point. Also, if the Evader, implementing G, reaches the intercept point

(a) Pursuers implement G, $t_f = 0.57$ (b) Pursuers implement PP, $t = 0.86$
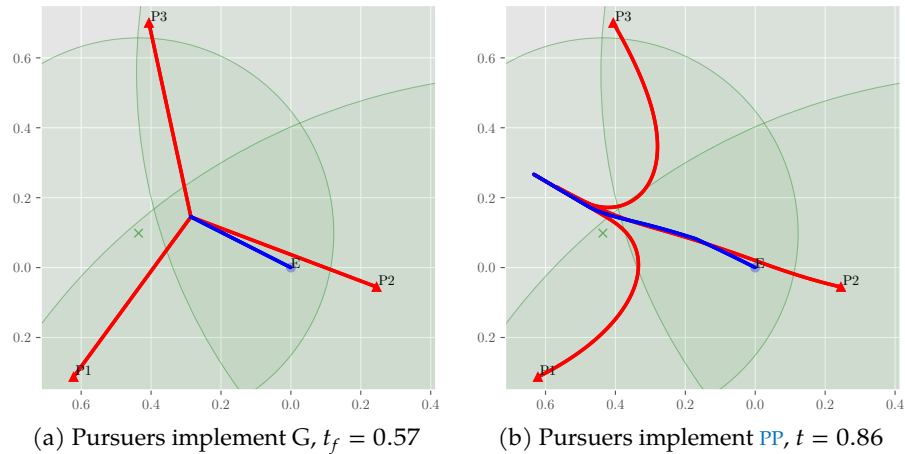
Figure 7.7: Simulation results; Evader implements G and Pursuers implement G (a) or PP (b).

before any Pursuer, then it will stand still (provided that point remains the solution to (7.42)).

Figure 7.7 displays the trajectories produced by each policy pair. In these simulations, the time step Δt is 0.001, the Pursuers speed is 1, and the Evader speed is 0.8 (and $\alpha = 0.8$ as well). From $t = 0$ the solution to (7.42) is the Voronoi vertex created by P1, P2, and P3. Thus, all three Pursuers are required to achieve the capture time of 0.57 seen in Figure 7.7a. Consequently, if the Pursuers act independently and implement PP, the Evader, implementing G, can increase the capture time to 0.86, an increase of over 50%. Observe that in Figures 7.7a and 7.7b P2's trajectory is quite similar. In the latter, however, P2 is not able to catch all the way up to E because P1 and P3 do not block the Evader from passing between them. The drastic difference between P1 and P3's headings with respect to E and with respect to the Voronoi vertex at $t = 0$ is what allows the Evader to capitalize on their behavior. These results are included merely to highlight the benefits of cooperation and to demonstrate how the G policy achieves it.

The next set of examples demonstrate the non-uniqueness of the Evader's trajectory in cases when the solution to (7.19) is a Voronoi vertex, v^*. In this case, all the Evader must do is take a safe path to v^* (according to Property 1) in order to guarantee capture at time t^*. Meanwhile, the Pursuers will implement the G policy – that is, at every

(a) Random Evader path until v^* near ∂S

(b) Zoomed (a)

(c) Spiral Evader path until v^* reachable in one timestep

(d) Zoomed (c)
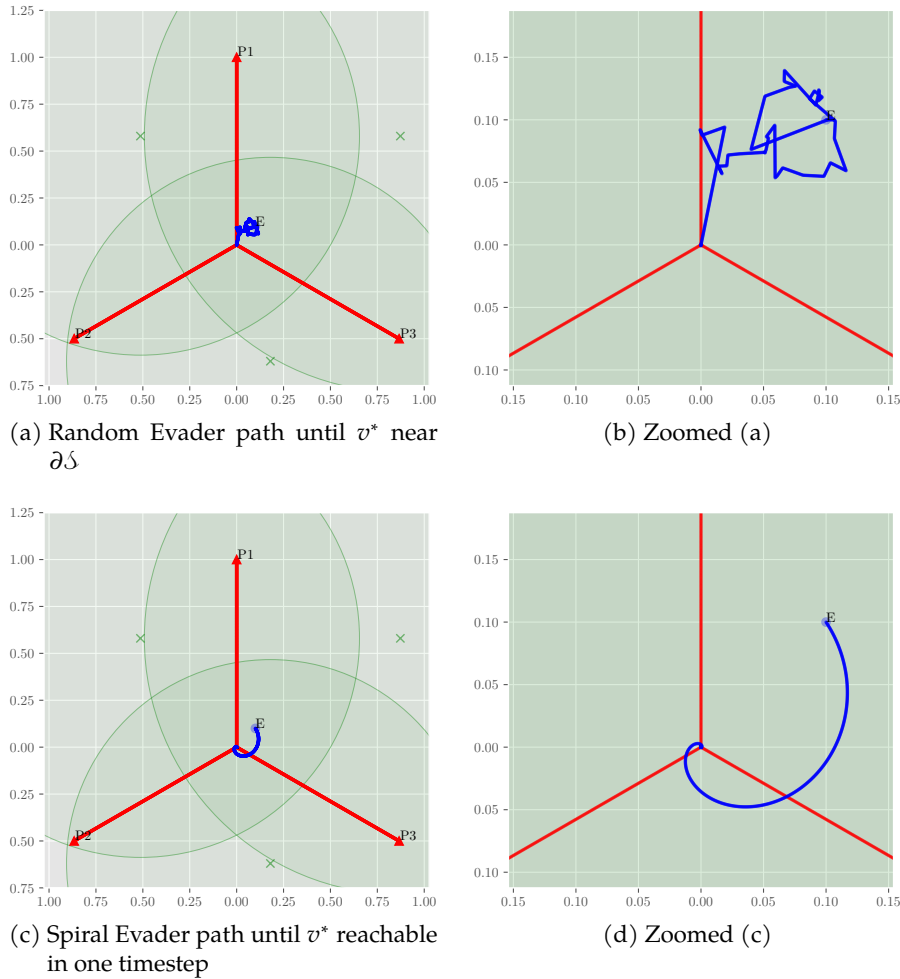
Figure 7.8: Examples of 'optimal' Evader trajectories resulting in capture at the Voronoi vertex demonstrating non-uniqueness of the path. Evader begins at $(0.1, 0.1)$, $V_P = 1$, $\alpha = 2/3$, and $t^* = 1\ s$.

time step, they will call MPursuer1Evader to compute their headings. As a consequence of the Evader taking a safe path, the point (x^*, y^*) will remain invariant and the Pursuers' trajectories will be straight. Figure 7.8 shows the trajectories for two different simulations. In Figures 7.8a and 7.8b, the Evader's policy is a random walk with some additional logic to ensure v^* remains inside the instantaneous safe region S. In Figures 7.8c and 7.8d, the Evader's policy is such that the Evader spirals in towards v^*; once the Evader can reach v^* within a single time step, it simply heads there at maximum speed. It is difficult, in general, to guarantee the safety of a path, *a priori*, and hence the Evader policy prescribed by Theorem 7.1 wherein the Evader takes a straight-line path at maximum speed to v^* and then stands still is useful.

The selection of the timestep Δt certainly has an effect on the play-out of the game, and is an important consideration not only for simulation but also for real-world control implementations. For sufficiently large timesteps, capture may not even be possible! However, for the scenarios included in this section note that the results do not change appreciably for $\Delta t = 0.01$ versus $\Delta t = 0.001$. For even smaller timesteps ($\Delta t = 1e - 4$), the difference is even smaller. Therefore, as $\Delta t \to 0$, the capture time asymptotically approaches the value it would have in continuous time. The results presented for $\Delta t = 0.001$ are sufficiently close to this asymptotic value.

7.6 CONCLUSION

In this chapter, the problem of multiple agents pursuing a single Evader wherein all the agents have simple motion was considered. This problem is a direct extension of the two-Pursuer one-Evader problem originally posed by Isaacs [130] and verified formally in [107]. The intent is to exploit the benefits of cooperation among a team of three or more Pursuers. Intuitively, the presence of additional Pursuers was shown to reduce the capture time of the Evader.

The initial problem formulation and analysis highlights the difficulty in analyzing this problem using the techniques of Isaacs. Part of the issue is the curse of dimensionality brought about by including additional Pursuers. A general strategy in differential games is to reduce the state-space into at most three dimensions, which cannot be done here. In lieu of a full verified feedback-optimal differential game solution a route of analysis was pursued which yielded open-loop optimal strategies for the Pursuers and Evader. These strategies are open-loop in the sense that they depend only on the initial conditions and are not necessarily robust to all other choice of opponent strategy when implemented as a state-feedback policy. Isaacs' methods pertain to analysis of regular solutions; special care must be given to singularities wherein Isaacs' methods may not yield an optimal control action for one or more agents. It is likely, with the increased number of agents,

that there are more singular surfaces in the M-Pursuer one-Evader differential game, requiring a more careful analysis.

Subsequent analysis relies on the Pursuers taking straight-line paths to the capture point, which is suggested by the derived necessary conditions for optimality as well as the solutions to both the one-Pursuer and two-Pursuer versions of the game. Then, a linear program was posed to determine which capture location is optimal. However, a strong geometric interpretation of the problem was observed which resulted from the fact that Pursuers take straight-line paths. Thus, a geometric solution was prescribed based on two Voronoi diagrams which allows searching over a discrete set of candidate solutions, as opposed to a continuous space as in the linear program case. Finally, algorithms were presented to compute the solution efficiently while also providing the precise shape of the Evader's region of dominance. The next chapter elucidates some of the properties of the policy presented in this chapter and introduces a pursuit policy which provides an upper bound on the time until capture.

ROBUST POLICIES FOR THE M PURSUER GAME

8.1 INTRODUCTION

This chapter presents research results pertaining to Research Objectives 4 and 5 which have to do with demonstrating the utility of skirmish-level solutions and adding agents to one or other team in a differential game, respectively. These results are for a differential game of pursuit-evasion between three (or more) Pursuers and a single Evader where the cost functional is the capture time, the Pursuers are faster, the agents move with simple motion, and termination is defined as point capture (i.e., the MP1E GoT). The remainder of the chapter is based off of the paper [253], which is an extension of the previous chapter.

In this chapter, the fact that the full solution to the MP1E GoT is still at large is highlighted. The geometric policy proposed by [246] is shown to be a Global Stackelberg Equilibrium strategy pair. For many games, it is the case that the Global Stackelberg Equilibrium is coincident with the Feedback Nash Equilibrium (i.e. the "full solution" of the game). However, it is shown that it is not the case for the MP1E GoT: the geometric policy violates the saddle-point condition necessary to be a Feedback Nash Equilibrium. This is shown through a counterexample which brings to light the deficiency of the geometric policy. One redeeming property of the geometric policy, though, is that it is robust from the Evader's perspective. That is, the Evader can do no worse if it implements this policy. A new pursuit policy is proposed which carries the same sort of robustness but from the Pursuers' side. The deficiencies of the geometric policy are addressed in this chapter to an extent, however, it is clear that the Feedback Nash Equilibrium strategies have yet to be discovered.

The remainder is organized as follows. Section 8.2 introduces the MP1E differential game as well as the Evader's Safe Region, a geometric construction used throughout. Section 8.3 summarizes the geometric policy introduced by [246] and gives a proof that it corresponds to the Global Stackelberg Equilibrium. Capturability is discussed in Section 8.4 as it is a pertinent topic when discussing non-feedback-equilibrium strategies. Section 8.5 proposes a new pursuit policy which has robustness guarantees. Conclusions are contained in Section 8.8.

8.2 TECHNICAL PRELIMINARIES

The pursuit-evasion scenario considered here is defined by the following kinematics

$$
\begin{aligned}
\dot{x}_E &= V_E \cos \phi \\
\dot{y}_E &= V_E \sin \phi \\
\dot{x}_i &= V_P \cos \psi_i \\
\dot{y}_i &= V_P \sin \psi_i, \ i = 1, \ldots, M
\end{aligned}
\tag{8.1}
$$

where E denotes the Evader and the subscripts i denote the ith Pursuer, of which there are M, V_k, $k = E, P$ are the speeds, ϕ is the Evader's heading angle, and ψ_i $i = 1, \ldots, M$ are the Pursuers' heading angles. Let

$$
\mathbf{x} = (x_E, y_E, x_1, y_1, \ldots, x_M, y_M)
$$

be the state of the system. Thus $\mathbf{x} \in \mathbb{R}^{2M+2}$. The size of $\mathbf{x}$ may be reduced to $2M$ states by considering a relative coordinate system which is fixed to the Evader. This reduction is of little consequence: with $M = 3$, one still ends up with six states. Few differential games with more than two states have been solved analytically (see, e.g., [96, 107, 164]). Therefore, the kinematics in (8.1) are retained for the remainder. It is assumed that $V_P > V_E$ since, as will be shown in the sequel, capture may be guaranteed for certain pursuit strategies. Let $\alpha = V_P/V_E < 1$ denote the speed ratio.

The game is over when one or more Pursuers captures the Evader; the final time T is

$$T = \min \{ t \mid \exists\, i \text{ s.t. } (x_i(t), y_i(t)) = (x_E(t), y_E(t)) \} \qquad (8.2)$$

This capture criterion is commonly referred to as point capture, as opposed to scenarios where the Pursuers have a non-zero capture radius. The Pursuer team seeks to minimize the time to capture while the Evader seeks to maximize the time. The scenario is modeled as a two-player zero-sum differential game, wherein the Pursuers, cooperating as a single entity, and Evader seek to minimize/maximize the following cost/payoff, respectively,

$$J(\mathbf{u}_E(\mathbf{x}), \mathbf{u}_P(\mathbf{x})) = \int_0^T d\tau \qquad (8.3)$$

where $\mathbf{u}_E = \phi$ is the Evader's control policy and $\mathbf{u}_P = (\psi_1, \dots, \psi_M)$ is the Pursuers' control policy in state-feedback form. The Value function describes the minimax value of the cost/payoff, (8.3), when starting from some point $\mathbf{x}_0$ in the state space

$$V(\mathbf{x}_0) = \min_{\mathbf{u}_P} \max_{\mathbf{u}_E} \int_0^T d\tau \qquad (8.4)$$

Note that the min and max are interchangeable [130]. Such a function, if it exists, is continuous and continuously differentiable in $\mathbf{x}$ and satisfies the so-called Isaacs equation [20]:

$$\min_{\mathbf{u}_P} \max_{\mathbf{u}_E} \left[\frac{\partial V}{\partial \mathbf{x}} f(\mathbf{x}, \mathbf{u}_P, \mathbf{u}_E) + g(\mathbf{x}, \mathbf{u}_P, \mathbf{u}_E) \right] = -\frac{\partial V}{\partial t} \qquad (8.5)$$

where $\frac{\partial V}{\partial \mathbf{x}}$ is a vector of derivatives of the Value function w.r.t. each state, $f(\mathbf{x}, \mathbf{u}_P, \mathbf{u}_E) = \dot{\mathbf{x}}$ are the kinematics (8.1), and g is the integrand of the cost/payoff in (8.3). Thus it must be that $g = 1$ and $\frac{\partial V}{\partial t} = 0$. Solving a differential game entails solving the Game of Kind (determining regions of the state space in which one or the other player wins) and solving the Game of Degree (i.e. determining the Value of the game in each region of the state space). For pursuit-evasion, the Game of Kind asks whether capture is inevitable or escape is inevitable. Since

the speed ratio has been restricted to $\alpha < 1$, the Evader cannot guarantee escape; however, this does not mean that capture is guaranteed for every pursuit strategy. In solving the Game of Degree, usually a candidate Value function is derived and then shown to satisfy (8.5). This is not trivial in practice. The benefit of having a full solution is knowledge of the saddle-point strategies $\mathbf{u}_P^*$ and $\mathbf{u}_E^*$ which satisfy

$$J(\mathbf{u}_E, \mathbf{u}_P^*) \leq J(\mathbf{u}_E^*, \mathbf{u}_P^*) \leq J(\mathbf{u}_E^*, \mathbf{u}_P) \qquad \forall \mathbf{u}_E \in U_E, \; \mathbf{u}_P \in U_P \quad (8.6)$$

where U_E and U_P are the sets of admissable control strategies for the Evader and Pursuers, respectively. Condition (8.6) means that the saddle-point strategies are robust to *any* admissable opponent strategy. In this chapter, no attempt is made to obtain $\mathbf{u}_E^*$ and $\mathbf{u}_P^*$ directly, but rather, in the sections to follow, policies for the Evader and Pursuers are proposed which exhibit one-sided robustness.

Now the Apollonius disk is defined which is a geometric construct utilized in all of the policies described hereafter. For agents with simple motion and zero capture radius, the Apollonius disk represents the region in the realistic plane where the Evader can reach before a particular Pursuer. The Apollonius *circle* is the boundary of this region, and it is defined as the locus of points in which the Evader and Pursuer can reach simultaneously by taking straight-line paths at maximum speed, respectively. For each Pursuer $i = 1, \ldots, M$, the Apollonius disk center, C_i, and radius, R_i are,

$$x_{C_i} = \frac{1}{1 - \alpha^2} x_E - \frac{\alpha^2}{1 - \alpha^2} x_i \tag{8.7}$$

$$y_{C_i} = \frac{1}{1 - \alpha^2} y_E - \frac{\alpha^2}{1 - \alpha^2} y_i \tag{8.8}$$

$$R_i = \frac{\alpha}{1 - \alpha^2} \sqrt{(x_i - x_E)^2 + (y_i - y_E)^2} \tag{8.9}$$

The following set defines the Apollonius disk for Pursuer i,

$$D_i = \left\{ (x, y) \in \mathbb{R}^2 \mid \left(x - x_{C_i} \right)^2 + \left(y - y_{C_i} \right)^2 \leq R_i^2 \right\} \tag{8.10}$$

The intersection of all the Pursuers' Apollonius disks defines a region in which the Evader can reach before *any* of the pursuers. This region is referred to as the Safe Region,

$$SR = \cap_{i=1}^{M} D_i \tag{8.11}$$

Remark 10. The Safe Region can not be empty; the Evader position, (x_E, y_E), is inside each Apollonius disk by the definitions (8.7)–(8.10) and so, at the very least, the Evader is inside SR.

Remark 11. The Safe Region collapses to a single point – the Evader position – when one or more Pursuers are coincident with the Evader or the speed ratio is infinite.

The Boundary of the Safe Region (or BSR) will often be utilized in the sequel. Let the BSR be parameterized by a set of circular arcs corresponding to segments of ACs, $\mathcal{C}$, and a set of vertices, $\mathcal{A}$, corresponding to endpoints of the arcs (which are AC intersections),

$$BSR = (\mathcal{C}, \mathcal{A}) \tag{8.12}$$

Whether BSR refers to its set of points or as its parameterization should be clear from context.

8.3   ROBUST EVADER POLICY

The robust Evader policy presented here is based upon Chapter 7. The symbol G (for "geometric") is used to denote the policy. Note that the G policy specifies an intercept point which is optimal, in some sense, for both the Evader and Pursuers and thus, although it is referred to as a robust Evader policy, it may also be implemented by the Pursuers. The main idea in Chaper 7 is that the Evader can safely travel to any point inside SR (defined as above) and thus the agents should take straight-line paths to a point, I, which satisfies

$$I = \arg\max_{(x,y)} \min_i (x_i - x)^2 + (y_i - y)^2$$
$$\text{s.t. } (x,y) \in SR, \qquad i \in 1, \dots, M. \tag{8.13}$$

The solution for the optimal intercept point is described in Sections 7.3.3 and 7.3.4 and the finite set of candidate points is stated precisely in (7.38).

8.3.1 *Properties of the G Policy*

This section provides a more detailed characterization of the policy and identifies in what sense the policy is optimal. In order to do so, the following are defined:

Definition 1. A pair of strategies forms a subgame perfect equilibrium (or *time consistent* equilibrium) if the strategies are in equilibrium for every subgame of the game's playout [180]. That is, in order for the strategy pair to be subgame perfect, it must never become advantageous for one or other agent to switch strategies at any point along the game's trajectory.

Definition 2 (Global Stackelberg Equilibrium (GSE)). An equilibrium over open-loop strategies wherein the leader selects a control action (i.e. a control trajectory from $t = 0$ through the game's termination) from a certain class of behaviors and announces the strategy to the follower.

The role of the follower is to compute its best response to the announced leader strategy. The leader can also compute the follower's best response and should therefore choose a strategy which maximizes its reward [59]. Also, the leader's control is purely a function of the initial conditions.

Definition 3 (Feedback Stackelberg Equilibrium (FSE)). An equilibrium over feedback strategies in which, at each instant in time, the leader selects a control action and announces the strategy to the follower.

Note that the GSE is time consistent, which is not always the case for FSE [201]. In the open-loop case, if the leader is allowed to plan at $t_1 > 0$, then there is no benefit to adhere to the promised plan [201]. The terms time consistent and subgame perfect will be used interchangeably. If

one were to apply the GSE policies over a finite timestep Δt, then as $\Delta t \to 0$ the FSE trajectories would be recovered.

Definition 4 (Feedback Nash Equilibrium (FNE))**.** An equilibrium over feedback strategies corresponding to the saddle point in (8.4).

The FNE is what is traditionally meant by the solution to the differential game. For detailed description of these equilibrium concepts, see [52].

In the context of the MP1E GoT, the Evader may be thought of as the leader since the onus is on the Pursuers to capture him. If the Pursuers were the leader, then, upon announcing their strategy, the Evader could choose from a myriad of trajectories which do not collide with the announced strategy. Define the class of leader behaviors as either a straight-line (constant heading) path, or a straight-line path to a point followed by stopping (dithering) at the point. In general, the follower's best response should satisfy the necessary conditions for optimality that are (usually) derived via the Hamiltonian. Thus the Pursuers' best response ought to consist of straight-line paths [246].

Theorem 8.1. *The G policy is a Global Stackelberg Equilibrium of the multiple pursuer single evader game under the kinematics in* (8.1)*, and the class of Evader behaviors consisting of straight-line paths which may or may not terminate at a point.*

Proof. Because of the restriction on the Evader's behaviors, the Evader selection of control strategy is equivalent to selecting a point $I \in SR$. For any point $(x, y) \notin SR$, there exists a Pursuer strategy consisting of a straight-line path in which capture occurs on BSR. This follows from the definition of the Apollonius disk, (8.10). The point I, since it is in SR may be reached by one or more Pursuers at or later than the Evader. Thus the Pursuers' best response is to head directly to I; any deviation will delay their arrival to I and the capture of E. The objective in (8.13), which characterizes the G policy, is akin to capture time (Evader reward) under such a pursuit policy. Therefore, (8.13) represents a maximization over the leader's reward given the follower's best response. □

Corollary 8.1. *The G policy is a robust Evader policy. That is,*

$$J(\mathbf{u}_E^G, \mathbf{u}_P) \geq \max_{(x,y) \in SR} \min_i \frac{1}{V_P} \sqrt{(x_i - x)^2 + (y_i - y)^2} \qquad \forall \mathbf{u}_P \in U_P$$

$$(8.14)$$

Proof. The result follows from the fact that the Pursuers' best response is a straight-line path to the intercept point. The geometry of the problem prevents the Pursuers from doing any better. □

Let the right-hand side of (8.14) be abbreviated as $LB(\mathbf{x})$ (for lower bound). Note, the G policy bears a striking resemblance to the so called open-loop policy proposed by [150] for the single pursuer multiple evader game. There, the Evaders jointly select their headings (assuming the worst case), announce their headings to the Pursuer, and then commit to those headings for the duration of the game. Liu indicates that this policy is conservative from the Evaders' perspective. Analogously, Corollary 8.1 indicates that the G policy is conservative from the Evader's perspective for the MP1E GoT.

Reference [201] investigated the criteria for which the FNE coincides with a Stackelberg equilibrium (GSE or FSE) and analyzed several cases where FNE and FSE coincide as well as cases where FNE and FSE do not coincide. The MP1E GoT falls under the latter category: the FNE does not coincide with the Stackelberg equilibria.

Theorem 8.2. *The G policy, although it is a global Stackelberg equilibrium, is not a feedback Nash equilibrium. That is, GSE $\neq$ FNE, which means the G policy does not constitute a solution to the multiple pursuer single evader differential game under the kinematics in (8.1).*

Proof. (by contradiction). Suppose $GSE = FNE$; that is, suppose that the G policy, implemented continuously in time by both teams, constitutes a feedback Nash equilibrium. Then the strategy pair $\mathbf{u}_E^G, \mathbf{u}_P^G$ must satisfy the saddle point condition,

$$J(\mathbf{u}_E, \mathbf{u}_P^G) \leq J(\mathbf{u}_E^G, \mathbf{u}_P^G) \leq J(\mathbf{u}_E^G, \mathbf{u}_P), \qquad \forall \mathbf{u}_E \in U_E, \ \mathbf{u}_P \in U_P \quad (8.15)$$

Table 8.1: Initial conditions and parameters for proof of Theorem 8.2

State/Paremeter	Value
E	$(0,0)$
P_1	$(0,1)$
P_2	$\left(\cos\frac{7\pi}{6}, \sin\frac{7\pi}{6}\right)$
P_3	$\left(\cos-\frac{\pi}{6}, \sin-\frac{\pi}{6}\right)$
V_P	1
α	0.8

Consider the case where the initial state and parameters are as shown in Table 8.1. The solution to (7.39) is $(0,0)$, the initial position of the Evader and the Voronoi vertex v. The predicted capture time w.r.t. this solution is $J(\mathbf{u}_E^G, \mathbf{u}_P^G) = 1$. Let the Evader strategy $\mathbf{u}_E^\downarrow$ be defined as $\mathbf{u}_E^\downarrow(t) = \frac{3\pi}{2}$. Under $\mathbf{u}_E^\downarrow$ the Evader's heading angle, ϕ, is downwards for all time, regardless of the state of the system. Note that $\mathbf{u}_E^\downarrow$ falls under the class of behaviors defined in Proposition 8.1 and is thus trivially in the set of admissable controls U_E. Let v and a represent the instantaneous position of the Voronoi vertex and lower AC intersection, respectively. Let the Pursuers implement the G policy – thus, the Pursuers aim towards $(0,0)$ initially. Because the Evader is traveling downward (away from v and towards a) there comes a time $0 < t_1 < 1$ wherein the Value (capture time) associated with all agents heading to a and all agents heading to v is equal. This situation is akin to a dispersal surface in differential games wherein the solution is non-unique. Let the distance between P_2 and v be written as $d_{P_2 v}$, and similarly for the point a. If the Pursuers continue along their original trajectory for some small δt then the point a will become the solution to (7.39) and Pursuers P_2 and P_3 have an incentive to switch their aim point to a. Figure 8.1 shows the configuration at this time. With the Pursuers aimed at a, the distance $d_{P_2 v}$ changes as,

$$
\begin{aligned}
\dot{d}_{P_2 v} &= -V_v \cos \overline{\psi}_2 - V_P \sin\left(\frac{\pi}{2} - \overline{\psi}_2\right) \\
&= -V_v \cos \overline{\psi}_2 - V_P \cos \overline{\psi}_2
\end{aligned}
\tag{8.16}
$$

Similarly, the distance from P_1 to v changes as,

$$
\dot{d}_{P_1 v} = V_v - V_P
\tag{8.17}
$$
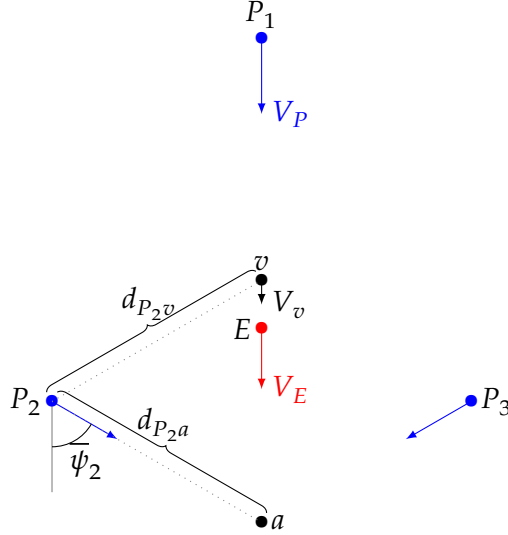
Figure 8.1: Configuration at t_1 wherein the points a and v are equidistant from P_2.

Now, since v is the Voronoi vertex and the Pursuers share the same speed, the distances to each of the Pursuers must remain the same, and thus the velocity at which each Pursuer approaches v must also be equal. Setting (8.16) equal to (8.17) yields,

$$V_v - V_P = -V_v \cos \overline{\psi}_2 - V_P \cos \overline{\psi}_2 \tag{8.18}$$

Collecting terms and solving for V_v,

$$V_v = \frac{V_P(1 - \cos \overline{\psi}_2)}{(1 + \cos \overline{\psi}_2)} \tag{8.19}$$

Here, $\overline{\psi}_2 = \pi/3$, and thus $V_v = 1/3$ and $\dot{d}_{P_2v} = -2/3$. The distance d_{P_2a} changes as,

$$\dot{d}_{P_2a} = -V_P = -1 \tag{8.20}$$

because the point a is stationary when P_2, P_3, and E are aimed towards it. Therefore it must be that,

$$\dot{d}_{P_2v} > \dot{d}_{P_2a} \tag{8.21}$$

After some infinitesimally small amount of time δt it must be that $d_{P_2v} > d_{P_2a}$ and thus Pursuers P_2 and P_3 have incentive to switch their aim

point *back* to the Voronoi vertex v. When P_2 and P_3 are aimed at v, (8.21) is reversed and thus the Pursuers' aim point chatters between a and v until a time t_2 when v exits the Safe Region *SR*. This "fast switching" induced by non-optimal play by one player has been observed in the two-pursuer one-evader differential game [186]. There, a dispersal surface is present when the three agents are collinear and the Evader lies between the two Pursuers. If the Evader stands still, the Pursuers' optimal behavior is to switch aim points between the two AC intersections infinitely fast. In this case, the capture time is equal to the Value of the game. Here, however, the two aim points a and v move at different rates; the point a is governed by the positions of P_2, P_3, and E while the point v is solely a function of the Pursuers' positions (see (8.19)). Each point is stationary when the Pursuers are aimed towards it. The result of the fast switching for $t_1 < t < t_2$ is that $t_f > 1$, that is, the Evader increased its capture time under the $\mathbf{u}_E^{\downarrow}$ policy:

$$J(\mathbf{u}_E^{\downarrow}, \mathbf{u}_P^G) > J(\mathbf{u}_E^G, \mathbf{u}_P^G) \tag{8.22}$$

(8.22) contradicts the saddle point condition (8.6). Therefore, the strategy pair $\mathbf{u}_E^G, \mathbf{u}_P^G$ is not a feedback Nash equilibrium.   □

Remark 12. A full treatment of this scenario requires solutions to differential equations with discontinuous right-hand sides (i.e. in the sense of Filippov) due to the fast switching behavior of the Pursuers [76].

The scenario described above is simulated numerically with a timestep of 0.005 wherein, at each time step, the Pursuers aim at the current solution to (7.39). Figure 8.2 shows the trajectories generated from this strategy pair along with a plot of time to capture versus simulation time. If the G policy were truly the solution to the MP1E GoT, one would expect the time to capture to remain on or below the line from $(0, 1)$ to $(1, 0)$ in Figure 8.2b if the Pursuers implement G. In other words, under game-optimal play, one expects $\frac{\partial t_f}{\partial t} = -1$. From time t_1 to t_2 the points v and a remain nearly equidistant from P_2 and P_3 as their heading chatters back and forth. The actual capture takes place at $t = 1.2$ because there is a loss of Pursuer performance during this window of time. This loss is caused by a dilemma of choosing between

(a) Trajectories

(b) Time to capture
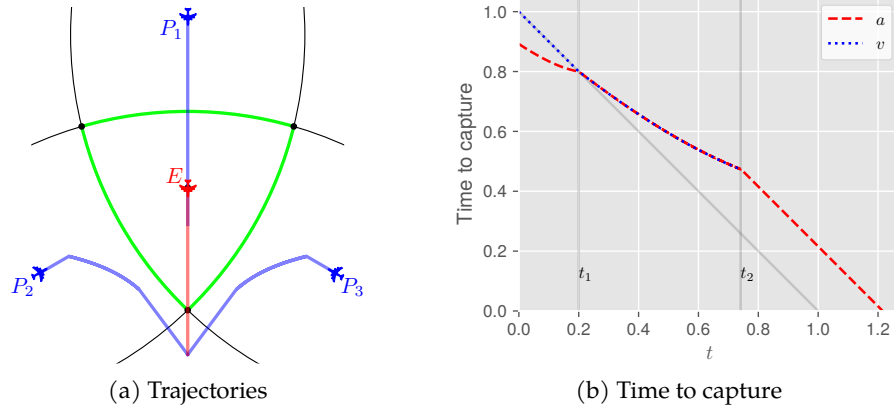
Figure 8.2: Simulation of $\mathbf{u}_E^\downarrow$, $\mathbf{u}_P^G$ starting from a symmetric configuration.

a and v which is perpetuated from t_1 to t_2. Each time P_2 and P_3 switch headings a small loss ε is incurred. In the limit as the timestep $\Delta t \to 0$ the Pursuers incur an infinite number of these small losses. One may be tempted to think that the sum of these losses $\sum \varepsilon \to 0$ as $\Delta t \to 0$. However, even if one considers the trajectory of P_2 and P_3 in a Filippov sense in which they become smooth as opposed to piecewise continuous, it is clear that the trajectories will still be curved. Thus the result in Figure 8.2b changes negligibly for very small time steps.

Consider, again, Figure 8.1 which shows the position of all the agents at time t_1 when $d_{P_2 a} = d_{P_2 v}$. Under the Evader strategy $\mathbf{u}_E^\downarrow$ it is clear that capture must occur on the line $x = 0$ at a $y < E_y(t_1)$. The only way to recover the initially predicted Value of $t_f = 1$, P_2 and P_3 must either commit to heading to v or $a(t_1)$ for the remainder of the game. Clearly, committing to v is a poor choice since the Evader is heading away from it. Figure 8.3 shows the case wherein at time t_1 Pursuers P_2 and P_3 make a single switch and aim at a. The issue with this strategy is that at any time $t_1 < t < t_2$ the Evader could switch to heading towards v and guarantee a capture time $t_f > 1$. This is evidenced by the fact that the line corresponding to v in Figure 8.3b lies above the line corresponding to a.

The situation encountered when $d_{P_2 v} = d_{P_2 a}$, as mentioned previously, is something like a singular surface. Here, the Pursuers can only recover the predicted capture time (associated with (7.39)) if they know which point, a or v, the Evader will choose. Guessing wrong for

(a) Trajectories   (b) Time to capture
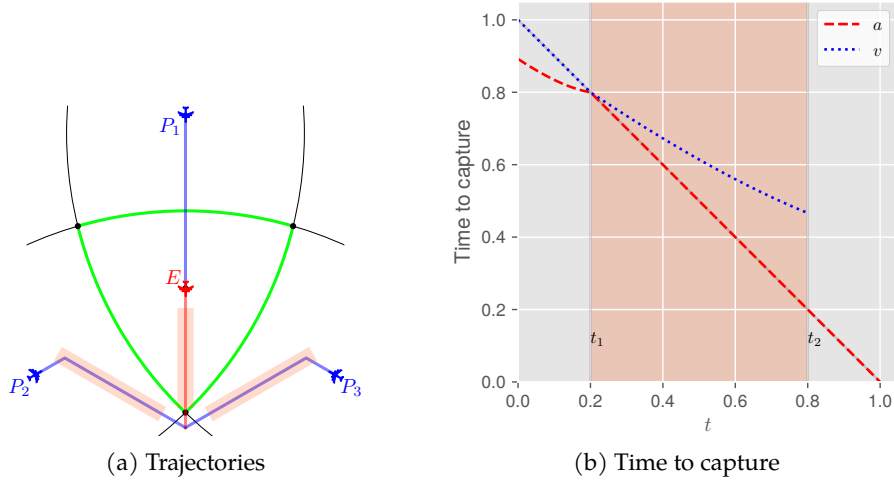
Figure 8.3: Simulation of $\mathbf{u}_E^\downarrow$, $\mathbf{u}_P^G$ until t_1, after which Pursuers aim at a. High-lighted segments indicate portions of the trajectory which are not subgame perfect.

a short time does not eliminate this dilemma (see, e.g., Figure 8.2b). Another case where the Evader's control action is required for the Pursuer to play optimally is found in the HCDG. There, using the classical parameters, an equivocal surface is present wherein the Evader has the authority to stay or leave the surface and the Pursuer must know the Evader's choice in order to play optimally on the surface [164]. In the HCDG, the Pursuer can force the system off of the equivocal surface (not to return under optimal play) by choosing a suboptimal control for a short time. However, in MP1E, even if the Pursuers take some suboptimal action when $d_{P_2 v} = d_{P_2 a}$ they cannot prevent the system from reentering such a configuration.

Consider another 3P1E scenario that is not symmetric and the solution to (7.39) is the point a, the lower AC intersection. Figure 8.4 contains the trajectories and time-to-go for the points a and v under the strategy pair $\mathbf{u}_E^G$, $\mathbf{u}_P^G$. The capture time predicted by the G policy is 1.22. Interestingly, at t_1, when a and v are nearly equidistant from P_2 and P_3, all the agents switch to the current Voronoi vertex $v(t_1)$. Afterwards, the agents have no further incentive to switch and capture occurs at precisely $t_f = 1.22$ as predicted, albeit not at the predicted location. Suppose the agents adhere to their initial headings, that is they aim towards a for the duration of the game. Figure 8.5 shows the results

(a) Trajectories

(b) Time to capture
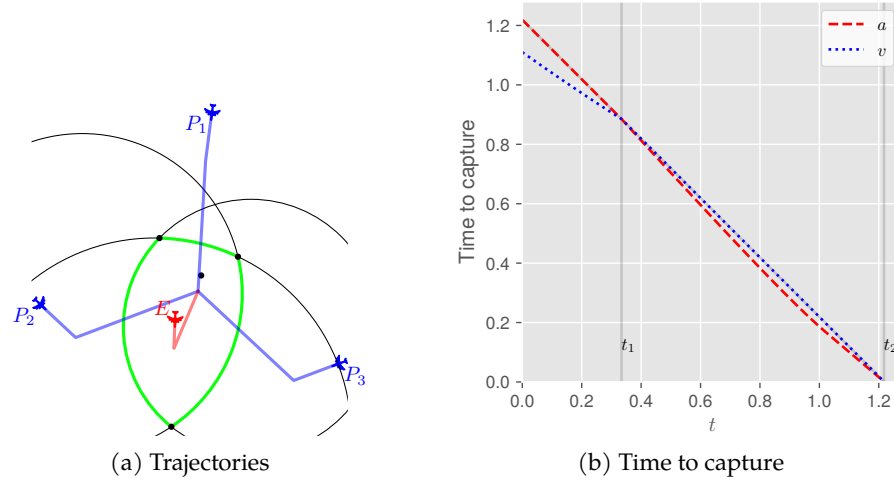
Figure 8.4: Simulation of $\mathbf{u}_E^G$, $\mathbf{u}_P^G$ starting from an asymmetric configuration; $t_f = 1.22$.

of this scenario; the trajectories in Figure 8.5a correspond to the trajectories predicted by the G policy. The capture time t_f is unchanged, however in Figure 8.5b it is clear there is an incentive to switch to v at some time $t_1 < t < t_2$, thus these trajectories are not subgame perfect. The behaviors shown in Figures 8.2 and 8.3 are a symptom of the fact that the pair $\mathbf{u}_E^G$, $\mathbf{u}_P^G$ is not a feedback Nash equilibrium.

8.4 CAPTURABILITY

Let R be the distance between P and E. Then if $R(0) > 0$ and $\dot{R} < 0$ for all $t > 0$, then capture is guaranteed.

$$\dot{R} = -V_P \cos \psi + V_E \cos \phi \tag{8.23}$$

where ψ and ϕ are defined relative to the line-of-sight (PE). Now, the Evader can choose any heading; $\phi = 0$ maximizes (8.23), thus,

$$\dot{R} \leq -V_P \cos \psi + V_E \tag{8.24}$$

The critical Pursuer heading, ψ_c thus occurs when the RHS is equal to zero,

$$-V_P \cos \psi_c + V_E = 0$$

(a) Trajectories

(b) Time to capture
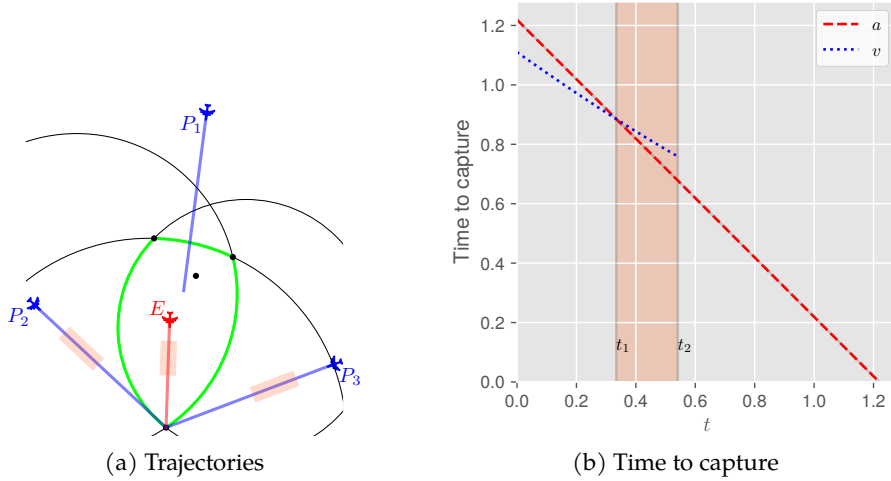
Figure 8.5: Simulation with same initial conditions as Fig. 8.4 but the agents head towards a for the duration of the game; $t_f = 1.22$. Highlighted segments indicate portions of the trajectory which are not subgame perfect.

yielding,

$$\psi_c = \cos^{-1} \alpha$$

Thus $\dot{R} < 0$ when $\psi < \cos^{-1} \alpha$. In the limit as $\alpha \to 1$ the range of Pursuer headings which guarantees capture collapses to $\psi = 0$ The line passing through P that is tangent to the AC has an angle of $\sin^{-1} \alpha$ w.r.t. the line PE. When $\alpha < \frac{\sqrt{2}}{2}$ it must be that $\sin^{-1} \alpha < \cos^{-1} \alpha$ and thus $\dot{R} < 0$ for any pursuit policy in which the Pursuer aims towards a point on the AC. The canonical 2P1E Pursuer policy [130] and the G policy both fall into this category. However, unlike the G policy, the 2P1E Pursuer policy is a game-optimal policy (in the sense of FNE).

Proposition 8.1. *Obtaining regular solutions to the Game of Degree over the whole state space is sufficient to guarantee capturability in pursuit-evasion differential games.*

Garcia *et al.* [107] proved that the pursuit policy put forth by Isaacs is indeed the solution for the whole state space – thus capture is guaranteed as long as $\alpha < 1$. Because the G policy is *not* the solution to the MP1E GoT (according to Theorem 8.2), rigorously proving that G guarantees capture for the case that $\alpha > \frac{\sqrt{2}}{2}$ is difficult. These two issues suggest the need for another Pursuit policy for MP1E with more desirable properties.

8.5   ROBUST PURSUIT POLICY

The difficulty in obtaining the solution to the MP1E differential game is, in part, due to the curse of dimensionality [245]. Reference [148] also notes that the traditional process of obtaining the retrograde partial differential equations (c.f. [130]) is difficult because the terminal states are unknown. To combat this issue, [148] proposes a hierarchical approach for multiplayer pursuit-evasion differential games which is conservative from the Pursuers' standpoint. The process is based on exploiting the solutions to games involving only a subset of the agents. Because the canonical 2P1E pursuit policy has been proven to be solution to the 2P1E differential game [107] the following is proposed,

$$i^*, j^* = \underset{i,j \in \{1,\dots,M\}}{\arg\min} \ V^{2P1E}(E, P_i, P_j) \tag{8.25}$$

$$\psi_{i^*,j^*} = \psi_{i^*,j^*}^{2P1E} \tag{8.26}$$

$$\psi_k = \tan^{-1} \frac{y_E - y_k}{x_E - x_k}, \qquad k \notin \{i^*, j^*\} \tag{8.27}$$

where $V^{2P1E}(E, P_i, P_j)$ is the Value function of the corresponding 2P1E game between E and the P_i, P_j Pursuer team starting from their current positions. Let $\mathbf{u}_P^R$ denote the policy described in (8.25) to (8.27); this policy is referred to as the R (for robust pursuit) policy. In the R policy, the Pursuers compare all possible 2P1E games and choose to play the game which yields the smallest capture time from the current positions. The headings for the chosen team, i, j, are given by the 2P1E game whereas all of the other agents aim LOS (PP). The solutions to the 2P1E games are given by (7.39) (note $\mathcal{U}_{\mathcal{P}_{\mathcal{B,S,R}}} = \emptyset$). An explicit derivation of the 2P1E Value function and heading angles is given by [107]. In general, these solutions are either the further of the AC intersections between P_i and P_j's circles or the 1P1E solution for one or the other Pursuer. The idea is that as $\mathbf{u}_P^R$ is implemented continuously in feedback fashion the Pursuer assignments (i.e. whether each Pursuer is aiming LOS or cooperating with another Pursuer in the 2P1E game) may switch to whatever is most advantageous at that time according to (8.25).

Theorem 8.3. *The R policy is a robust Pursuer policy. That is,*

$$J(\mathbf{u}_E, \mathbf{u}_P^R) \leq V^{2P1E}(E, P_{i^*}, P_{j^*}) \qquad \forall \mathbf{u}_E \in U_E \qquad (8.28)$$

where i^*, j^* are given by (8.25).

Proof.

Case 1 (No switches occur): The team i^*, j^* selected at time $t = 0$ remains the best team according to (8.25) for all $0 \leq t \leq t_f$. Then the other pursuers, $k \notin \{i^*, j^*\}$ had no effect on the game and it is as if the scenario is an instance of the 2P1E game. In this case (8.28) is given by the fact that the headings in (8.26) are the saddle-point strategies for the Pursuers P_{i^*}, P_{j^*} in the 2P1E game.

Case 2 (One or more switches occur): Let the initial team assignment be i_0^*, j_0^*. Let t_1 be the time in which the first switch occurs, $0 < t_1 < t_f$, and the new team assignment be i_1^*, j_1^*. At t_1 it must be, from (8.25),

$$V^{2P1E}(E, P_{i_1^*}, P_{j_1^*}) < V^{2P1E}(E, P_{i_0^*}, P_{j_0^*})$$

otherwise, a switch would not have occurred. For the remainder of the game $t_1 < t \leq t_f$, the scenario falls into either of these two cases.

Thus, robustness of R is guaranteed by the robustness of the 2P1E solution and the method of assignment, (8.25). □

Theorem 8.4. *The Value of the multiple pursuer single evader game under the kinematics in* (8.1) *and feedback Nash equilibrium strategies $\mathbf{u}_E^*$, $\mathbf{u}_P^*$, if they exist, satisfies*

$$LB(\mathbf{x}) \leq J(\mathbf{u}_E^G, \mathbf{u}_P^*) \leq J(\mathbf{u}_E^*, \mathbf{u}_P^*) \leq J(\mathbf{u}_E^*, \mathbf{u}_P^R) \leq V^{2P1E}(E, P_{i^*}, P_{j^*})$$

$$(8.29)$$

where i^*, j^* are given by (8.25). *In other words, the Value of the game is bounded.*

Proof. The first inequality follows directly from (8.14) in Corollary 8.1. The middle two inequalities follow from Definition 4. The last inequality follows directly from (8.28) in Theorem 8.3. □

Although the true solution to the [MP1E GoT] is not known, the robustness properties of the G and R policies yield a bound on the game optimal capture time. It is also true that if the Evader implements G and the Pursuers implement R the following is satisfied,

$$LB(\mathbf{x}) \leq J(\mathbf{u}_E^G, \mathbf{u}_P^R) \leq V^{2P1E}(E, P_{i^*}, P_{j^*}) \tag{8.30}$$

It is possible for the upper and lower bounds to be equivalent. The following theorem identifies the conditions under which this is true as well as the implications of such a scenario. Let the capture point associated with the 2P1E game between E, P_{i^*}, and P_{j^*} be denoted I^{2P1E}.

Theorem 8.5. *If $I^{2P1E} \in SR$, the strategy pair $\mathbf{u}_E^G$, $\mathbf{u}_P^R$ is a feedback Nash equilibrium of the multiple pursuer single evader game under the kinematics in* (8.1)*.*

Proof. Let i^*, j^* be the solution to (8.25) corresponding to I^{2P1E}. Pursuers P_{i^*} and P_{j^*} thus aim at I^{2P1E} under the R policy according to (8.26). Now it will be shown that I^{2P1E} is the solution to (7.39). This statement is predicated on the point I^{2P1E} being in SR. Suppose there is another point $I^\dagger = (x^\dagger, y^\dagger) \in SR$ such that,

$$\min_i (x_i - x^\dagger)^2 + (y_i - y^\dagger)^2 > \min_i (x_i - x^{2P1E})^2 + (y_i - y^{2P1E})^2$$

that is, the point $I^\dagger$ is further away from the nearest Pursuer than I^{2P1E}. By virtue of the fact that $I^\dagger \in SR$, the Evader can safely reach $I^\dagger$ by aiming directly at it and achieve a better capture time than if it had aimed at I^{2P1E} (i.e. the Value of the 2P1E game, $V^{2P1E}(E, P_{i^*}, P_{j^*})$). Then $LB > V^{2P1E}$; but this statement contradicts (8.29) in Theorem 8.4. Therefore, since no point in SR can yield a better capture time for the Evader, the point I^{2P1E} must be the solution to (7.39). Thus the lower bound is equal to the upper bound,

$$LB(\mathbf{x}) = V^{2P1E}(E, P_{i^*}, P_{j^*}) \tag{8.31}$$

Under the G policy, then, the Evader aims at the point I^{2P1E}. Thus the strategies $\mathbf{u}_E^G$, $\mathbf{u}_P^R$ yield the same behavior for E, P_{i^*}, and P_{j^*} as under the game optimal 2P1E strategies,

$$J(\mathbf{u}_E^G, \mathbf{u}_P^R) = J(\mathbf{u}_E^*, \mathbf{u}_P^*) \tag{8.32}$$

where i^*, j^* are given by (8.25). □

It is unclear, however, what the relationship between $J(\mathbf{u}_E^G, \mathbf{u}_P^R)$ and $J(\mathbf{u}_E^*, \mathbf{u}_P^*)$ is in general.

8.6 CONVEX POLICY

In this section another pursuit policy is introduced, the C policy (for convex) It may not have the same robustness properties of R, but is an attempt at smoothing out the chattering behavior of G in configurations like that of Figure 8.1. The Pursuers' loss observed in Figure 8.2b is induced (1) by the curvature in P_2 and P_3's paths and (2) by an effective slowing down while the Voronoi vertex v and AC intersection a are similar in value (i.e. the Pursuers headings chatter). In the case of a discrete time implementation, this latter piece may be understood to be the result of vector addition. Figure 8.6 illustrates the idea that
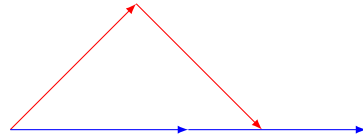
Figure 8.6: Red path: slower effective speed due to chattering, Blue path: smoothed trajectory obtained by taking a convex combination of headings.

the effect of slowing down due to chattering can be combated by using a combination of *headings*. Thus the C policy for the Pursuers is proposed as an augmentation of the G policy. Let $d(x,y)$ be the distance from the nearest Pursuer to the point (x,y),

$$d(x,y) = \min_i \sqrt{(x_i - x)^2 + (y_i - y)^2} \tag{8.33}$$

and, let $d^* = d(x^*, y^*)$ be the distance from the nearest Pursuer to the aim point suggested by the G policy be written where (x^*, y^*) is the solution to (7.39). Then, define the following set of points,

$$C = \left\{ (x_k, y_k) \mid (x_k, y_k) \in S_{BSR} \cap \mathcal{A} \cap \mathcal{U}_{P_{BSR}}, \ |d(x_k, y_k) - d^*| < \varepsilon \right\}$$
(8.34)

which are the candidate solutions to (7.39) whose distance to the nearest Pursuer is within some neighborhood ε of d^*. The C policy, then, for the ith Pursuer is to take a convex combination of headings to each of the points in C in which the weight associated with a point (x_k, y_k) is inversely proportional to d_k.

$$\mathbf{u}_i^C = \frac{\sum_{k=1}^{|C|} \frac{1}{d_k} \psi_{i,k}}{\sum_{k=1}^{|C|} \frac{1}{d_k}}, \qquad i = 1, \dots, M$$
(8.35)

where $\psi_{i,k}$ is the angle of the point $p_k \in C$ relative to P_i,

$$\psi_{i,k} = \tan^{-1} \frac{y_{p_k} - y_i}{x_{p_k} - x_i}, \qquad i = 1, \dots, M, \ k = 1, \dots, |C|$$
(8.36)

Remark 13. In practice it is simpler to apply the weighting in (8.35) to the points p_k directly to get the convex combination of *aim points*, and then converting to a heading.

The selection of ε is of some importance. As $\varepsilon \to 0$ only the solution to (7.39), (x^*, y^*), is considered, and all other candidates ignored; in this case, the C policy is identical to the G policy. It is possible for there to be more than one solution to (7.39), however this almost never occurs in reality. For such a configuration to occur, the agents would either need to begin as such or implement their control in continuous time with infinite precision. For very large ε the set C is equivalent to the set of candidates $S_{BSR} \cup \mathcal{A} \cup \mathcal{U}_{P_{BSR}}$. Pursuer performance in this case can be very poor, especially in the case when E is between and nearly collinear with two of the Pursuers. Therefore, in order to improve upon the G policy, the neighborhood should generally be set such that,

$$0 < \varepsilon \ll 1$$
(8.37)

For discrete time implementations, the size of ε should be larger for large Δt. This is because, in a single time step, the system state $\mathbf{x}$ may very well jump across the neighborhood wherein the solution to (7.39) is different. As previously mentioned, there are no obvious analytical properties of the C policy (e.g. in the way of robustness), but, as will be shown in the sequel, its performance is the best out of any of the pursuit policies presented herein for some particular Evader behaviors.

8.7 RESULTS

Because none of the policies presented in this chapter are feedback Nash equilibrium strategies, it is interesting to compare the performance of the different policies via simulation. The limitation of this approach is that one may never be certain whether a particular policy performs better than another for all cases. A framework for comparing the merits or de-merits of non-FNE strategies via simulation is out of the scope of the present work, although it represents an interesting area of research in the case that the FNE strategies are not known (or do not exist). Here, in order to compare the G, R, and C policies initial conditions like those in Figure 8.2a are utilized and the Evader's control input ϕ is restricted to be constant over the course of a single playout:

$$\mathbf{u}_E^{\phi}(t) := \phi_c, \qquad \phi_c \in [0, 2\pi] \qquad (8.38)$$

Then, ϕ is swept from 0 to 2π and simulate the game for each constant evader heading. For all of the simulations in this section the following settings are used: $V_P = 1$, $\alpha = 0.9$, and $\Delta t = 5e - 3$. Figure 8.7 depicts the location in which the Evader is captured for each of the pursuit policies. The BSR shown (and hereafter mentioned) is the BSR corresponding to $t = 0$ – the BSR changes as a function of the instantaneous position of the agents. Because the Evader is implementing a constant heading policy, the best performance the Pursuers can achieve is capture on the BSR. Capture on the BSR can only be achieved if the intercepting Pursuer takes a constant-heading path to the intercept point. In general, the Pursuers would need to know ϕ_c from $t = 0$ in order to
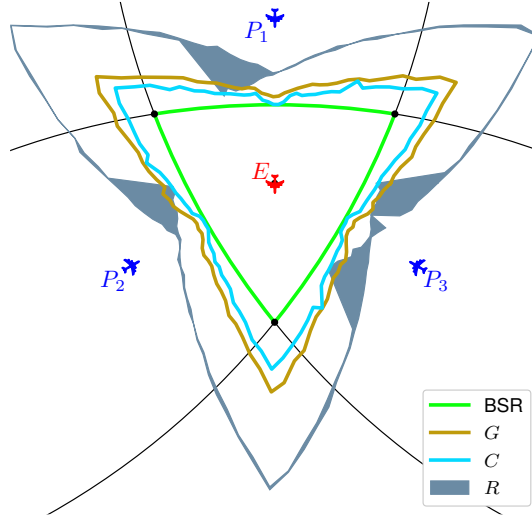
Figure 8.7: Capture locations under Evader constant-heading policy $\mathbf{u}_E^\phi$ against Pursuers' geometric policy $\mathbf{u}_P^G$, convex policy $\mathbf{u}_P^C$, and robust policy $\mathbf{u}_P^R$.

accomplish this. Nonetheless, the closer the intercept points are to the *BSR* the "better", in this case.

Figure 8.8 shows the capture times associated with the playouts in Figure 8.7. The upper bound for the capture time is given by the two-

(a)

(b) zoomed
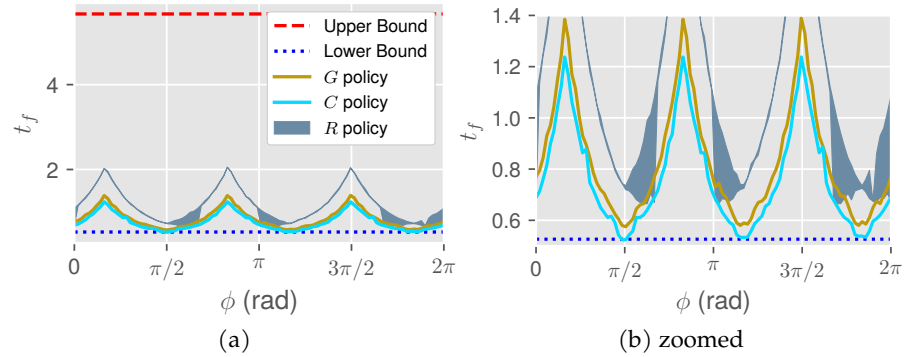
Figure 8.8: Capture times under Evader constant-heading policy $\mathbf{u}_E^\phi$ against Pursuers' geometric policy $\mathbf{u}_P^G$, convex policy $\mathbf{u}_P^C$, and robust policy $\mathbf{u}_P^R$.

on-one game with the smallest capture time, $V^{2P1E}(E, P_{i^*}, P_{j^*})$. The lower bound for the capture time is simply the smallest time required for a Pursuer to each the Evader if the Evader were to take a collision course with that Pursuer:

$$t = \min_i \frac{\overline{EP_i}}{V_E + V_P}$$

Note, *this* lower bound is not to be confused with the Evader's robust bound, defined in (8.14). The value of the Evader's robust bound is 1 and corresponds to the Evader standing still at $(0,0)$, which is the Voronoi vertex at $t = 0$. If this bound were equal to the Value of the game (i.e. if G were the Evader's FNE strategy), then one would expect an FNE pursuit strategy to achieve capture times at or below 1 for these playouts. The fact that none of the pursuit policies meet this condition for all $\phi_c \in [0, 2\pi]$ means that G is not the Evader's FNE strategy and/or none of the pursuit strategies presented herein are FNE pursuit strategies. It is likely that both are true, and likely that the inequalities in (8.30) are strict equalities.

The thickness in the R line in the above plots is due to the fact that at $t = 0$ all three Pursuer pairings yield the same two-on-one capture time. Whenever it is true that the Pursuer team pairing i^*, j^* is not unique the simulation selects a pairing from among the minima at random. Thus, for the R policy, the simulations were repeated five times and the minimum and maximum capture times recorded and displayed. From Figure 8.8 it is clear that the C policy always out performs the G policy which almost always outperforms the R policy. This is, perhaps, due to fact that the G and C policies consider explicit cooperation among multiple (> 2) pursuers whereas the R policy addresses cooperation between 2 pursuers (since the other $M - 2$ pursuers operate independently). So although the G and C policies do not come with any robustness guarantees from the Pursuers' perspective, they typically perform much better. This is especially comforting considering this configuration was used to disprove that the G policy is an FNE pursuit strategy (see Theorem 8.2)! Note that, for this particular scenario, the Pursuers' robust (upper) bound is quite high compared to the capture times, even under the R policy. Thus, if one were to choose a pursuit policy, the relative benefit of robustness must be weighed against typical performance. One should expect that multi-pursuer cooperation will improve the performance of the Pursuer team. However, analysis of cooperation among multiple (> 2) pursuers is more challenging and the FNE is not easy to obtain.

Lastly, Figure 8.9 shows the agents' trajectories for the case where $\phi_c = \frac{3\pi}{2}$. The capture times corresponding to the trajectories in Fig-
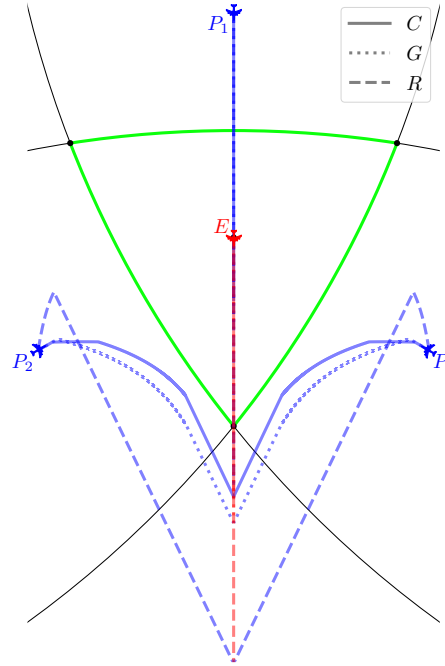
Figure 8.9: Comparison of Pursuer trajectories under the robust (R), geometric (G), and convex (C) policies against an evader implementing a constant heading of $\phi_c = \frac{3\pi}{2}$.

ure 8.9 are given in Table 8.2 along with the relative performance penalty w.r.t. the C policy. Notice the trajectories generated by the C policy and

Table 8.2: Capture times for different pursuit policies

Policy	t_f	Penalty
C	1.28	0%
G	1.40	10%
R	2.19	64%

how P_2 and P_3 tend towards the y-axis much moreso than G or R and thus capture E much sooner. Under the R policy, pursuers P_2 and P_3 begin by aiming to the intersection of their AC intersections which is further away from the Evader (far above the plot shown). They stay on this course until they become collinear with the Evader. After this point, the game plays out as a normal two-on-one game between P_2, P_3, and E (as these three agents are now headed towards the furthest AC intersection).

8.8 CONCLUSION

In this chapter the Multiple Pursuer Single Evader Differential Game comprised of agents with simple motion and a slow Evader was presented. It was shown that the recently proposed geometric policy based on comparing distances to relevant single-pursuer solutions, AC intersections, and Voronoi vertices is a Global Stackelberg (i.e., open-loop) Equilibrium when both the Evader and Pursuers implement it. This policy was also shown to *not* be a Feedback Nash Equilibrium strategy pair. The true Feedback Nash Equilibrium strategies correspond to the solution of this differential game; if they exist, they are not currently known. Despite the geometric policy not being a Feedback Nash Equilibrium strategy, the fact that the policy is robust from the perspective of the Evader was identified. Similarly, a new pursuit strategy was proposed based on the solution to differential games between subsets of the agents which is robust from the Pursers' perspective. Alterations to the geometric policy were proposed and shown to improve performance for a particular test scenario. In the next chapter, these many-Pursuer concepts are applied to a scenario with a different cost functional which is based on distance to a target.

M PURSUER BORDER DEFENSE

9.1 INTRODUCTION

This chapter applies concepts from the MP1E GoT to a multiple-Pursuer, single-Evader border defense differential game (MP1E BD). Here, the Pursuers (henceforth referred to as the UAVs) seek to prevent the Evader from reaching a specified goal (henceforth referred to as the border). In the case that the Evader cannot be guaranteed to reach the border, a zero-sum differential game is posed in which the cost functional is the terminal distance to the border at the time of capture (i.e., the Game of Distance). The results partly address Research Objectives 4 and 5 by demonstrating the utility of the 1v1 MP1E BD solution in a scenario with many Pursuers. This chapter is based on [251].

Two questions (or games) may be considered in this scenario: (1) is escape possible for the Evader (the game of kind) and (2) how close can the Evader get to the border before being captured (the game of degree). In the remainder of the paper, both games will be addressed. Section 9.2 expresses the problem formulation of the game of degree. Section 9.3 describes the game of kind and characterizes the win regions for the single and multiple Pursuer cases. Section 9.4 discusses the possibility for a dispersal surface in the game and some practical consequences. Lastly, Section 9.5 concludes this chapter.

9.2   GAME OF DEGREE

Here the game of degree is considered wherein escape through the border is not possible for Evader. Consider M Pursuers and an Evader, with simple motion in the $\mathbb{R}^2$ plane:

$$
\begin{aligned}
\dot{x}_E &= \alpha \cos \phi, \\
\dot{y}_E &= \alpha \sin \phi, \\
\dot{x}_{P_i} &= \cos \psi_i, \qquad i = 1, \dots, M \\
\dot{y}_{P_i} &= \sin \psi_i, \qquad i = 1, \dots, M
\end{aligned}
\tag{9.1}
$$

where $\alpha = V_E/V_P < 1$ is the speed ratio constant; note, the Pursuers share the same velocity $V_{P_i} = V_P$, $i = 1, \dots, M$. The admissible controls are given by $\phi, \psi_i \in [-\pi, \pi]$. The cost/payoff function is the terminal distance, at the time of interception, between the Evader and the border. The Pursuer strives to capture the Evader and maximize this separation while the Evader wants to minimize it. The cost/payoff function can be written as follows:

$$
J = \sqrt{(x_{E_f} - x_F)^2 + (y_{E_f} - y_F)^2}.
\tag{9.2}
$$

where the point $F : (x_F, y_F)$ is defined as the point on (any segment of) the border which is the closest to the terminal position of the Evader.

Interception is defined by point capture, that is, the game terminates when the state of the system enters the set

$$
\mathcal{C} := \left\{ \mathbf{x} \,\middle|\, \exists i \text{ s.t. } \sqrt{(x_{P_i} - x_E)^2 + (y_{P_i} - y_E)^2} = 0 \right\}
\tag{9.3}
$$

Figure 9.1 specifies the coordinate system as well as the payoff/cost for a given interception point (x_{E_f}, y_{E_f}). The main point of defining the objective in (9.2) this way is robustness to uncertainty, both in terms of the opponents' strategy as well as the environment. The Pursuers want to keep the Evader as far from the border as possible in case a Pursuer is incapacitated or blown off-course by wind, which could allow the Evader to get closer to the border or even escape. For some applications, whether or not escape occurs is the only thing that mat-
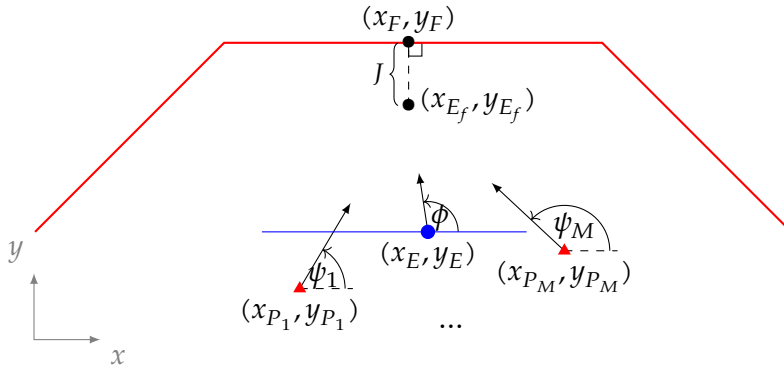
Figure 9.1: Definition of scenario, coordinate system, and payoff/cost. Red triangles represent Pursuers, the blue circle is the Evader, the blue line is the laser fence, and the red line is the border.

ters. However, by solving the Game of Degree, the Pursuers give themselves the best shot at capturing the Evader in the face of uncertainty in the environment and in the Evader's implemented control.

9.2.1   *Single Pursuer Solution and Verification*

In this differential game one can then define an AC using the instantaneous separation between P (Pursuer) and E (Evader) and the speed ratio parameter $\alpha = V_E/V_P < 1$. Let C denote the center of the circle. The three points: P, E, and C are located on the same line, see Fig. 9.2. Also in this figure one can see that the Borderline separates the plane into two regions. The game is played in the region of the Euclidean plane $\mathcal{G}$ where both the Pursuer and the Evader are initially located. This region is further divided into the region of win for the Pursuers and the region of win of the Evader.

Define $d = \sqrt{(x_E - x_P)^2 + (y_E - y_P)^2}$ as the distance between players P and E. Also, let r be the radius of the AC. Then, r is given by

$$r = \frac{\alpha}{1 - \alpha^2}d. \tag{9.4}$$

Assume that the parameters (m_i, n_i) of each border segment are given, where the segment i is given by $y = m_i x + n_i$. Also assume that the $\mathbf{x} \in \mathcal{R}_P$, that is, in the region of win of the Pursuer. $\mathcal{R}_P$ can be deter-
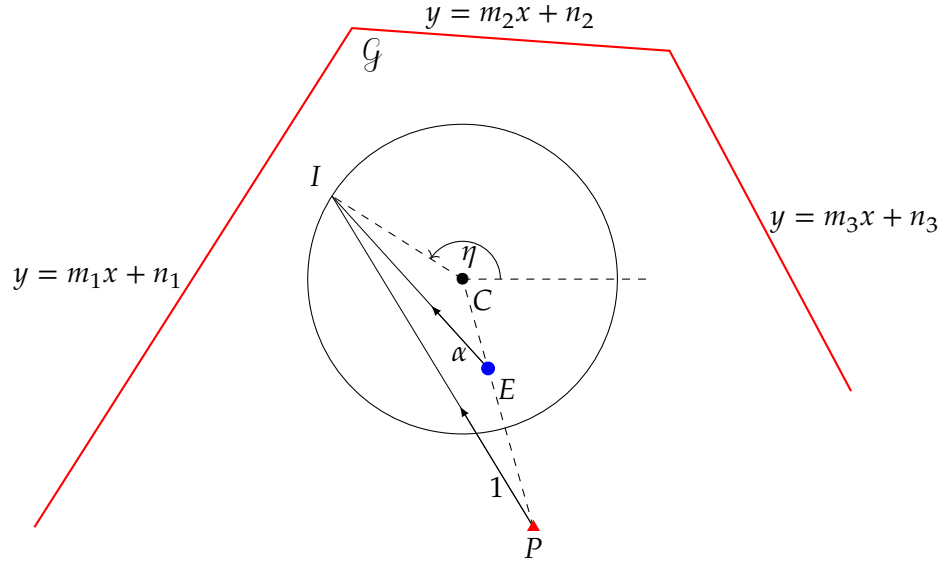
Figure 9.2: Border defense scenario 1 Pursuer - 1 Evader

mined by analyzing whether or not the AC intersects any segment of the border.

Proposition 9.1. *Consider the differential game of border defense* (9.1)–(9.3). *The Value function is given by* $V(\mathbf{x}) = V_{i^*}(\mathbf{x})$ *where* $i^* = \arg\min_i V_i(\mathbf{x})$ *and*

$$
V_i(\mathbf{x}) =
$$
$$
\frac{1}{1-\alpha^2}\Bigg[\left(x_E - \alpha^2 x_P - m_i\frac{y_E - \bar{m}_i x_E - \alpha^2(y_P - \bar{m}_i x_P) - (1-\alpha^2)n_i}{m_i^2 + 1}\right)^2
$$
$$
+ \left(y_E - \alpha^2 y_P - m_i^2\frac{y_E - \bar{m}_i x_E - \alpha^2(y_P - \bar{m}_i x_P) + \frac{1-\alpha^2}{m_i^2}n_i}{m_i^2 + 1}\right)^2\Bigg]^{1/2}
$$
$$
- \frac{\alpha}{1-\alpha^2}\left[(x_E - x_P)^2 + (y_E - y_P)^2\right]^{1/2}
$$

(9.5)

where $\mathbf{x} = \begin{bmatrix} x_E & y_E & x_P & y_P \end{bmatrix}^\top$.

Proof. The optimal cost/payoff is obtained when E is intercepted by P at the point on the AC closest to the border. For each segment of the border, the closest point between the circle and the border segment is given by the orthogonal line to the border segment that passes by the

center of the circle C as shown in Fig. 9.3. The distance between points I_i and F_i is given by

$$V_i(\mathbf{x}) = \sqrt{(x_c - x_F)^2 + (y_c - y_F)^2} - r \qquad (9.6)$$

where the center of the circle is explicitly given by:

$$x_c = \frac{1}{1 - \alpha^2} \left(x_E - \alpha^2 x_P \right), \qquad y_c = \frac{1}{1 - \alpha^2} \left(y_E - \alpha^2 y_P \right) \qquad (9.7)$$

In order to write the Value function directly in terms of the state $\mathbf{x}$ the point $F : (x_F, y_F)$ must be evaluated in terms of the state $\mathbf{x}$. The equation of the orthogonal line to segment i passing through C is $y = \bar{m}_i x + \bar{n}_i$ where $\bar{m}_i = -\frac{1}{m_i}$ and

$$\begin{aligned} \bar{n}_i(\mathbf{x}) &= y_c - \bar{m}_i x_c \\ &= \frac{1}{1 - \alpha^2} \left(y_E - \alpha^2 y_P - \bar{m}_i \left(x_E - \alpha^2 x_P \right) \right) \end{aligned} \qquad (9.8)$$

Now solve the following linear equation to determine x_F

$$m_i x_F + n_i = \bar{m}_i x_F + \bar{n}_i$$

The coordinate x_F is explicitly given by:

$$x_F(\mathbf{x}) = \frac{\bar{n}_i - n_i}{m_i - \bar{m}_i} = \frac{m_i}{1 - \alpha^2} \cdot \frac{y_E - \bar{m}_i x_E - \alpha^2 (y_P - \bar{m}_i x_P) - (1 - \alpha^2) n_i}{m_i^2 + 1} \qquad (9.9)$$

and y_F is given by:

$$y_F(\mathbf{x}) = m_i x_F + n_i = \frac{m_i^2}{1 - \alpha^2} \cdot \frac{y_E - \bar{m}_i x_E - \alpha^2 (y_P - \bar{m}_i x_P) + \frac{1 - \alpha^2}{m_i^2} n_i}{m_i^2 + 1} \qquad (9.10)$$

Substitute (9.4), (9.7), (9.9), (9.10) into (9.6) to obtain eq. (9.5). The problem of finding the closest point on the border to the AC has been simplified by obtaining the closest point of each segment of the border to the AC. Then, the solution is given by $i^* = \arg\min_i V_i(\mathbf{x})$. $\qquad \square$
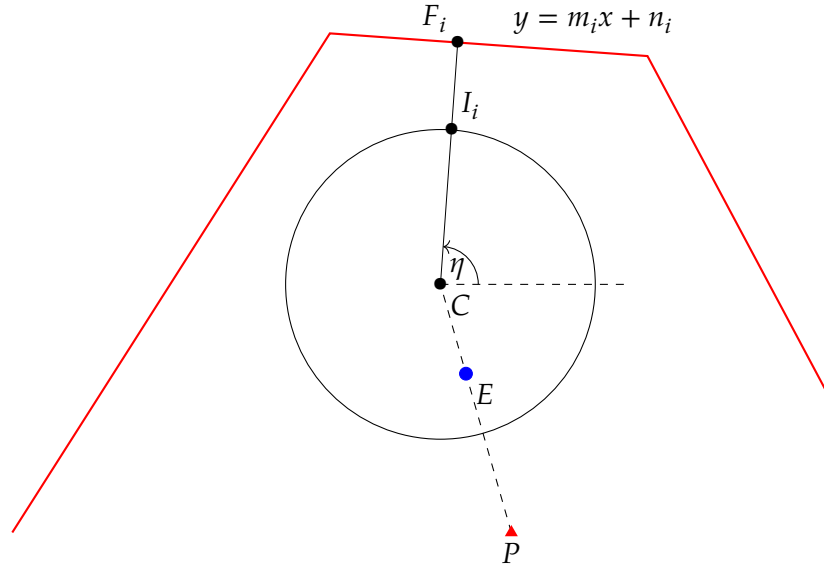
Figure 9.3: Derivation of $V(\mathbf{x})$

In the next theorem the subscript i is removed for simplicity, that is, it is assumed that the optimal segment i has been determined.

Theorem 9.1. *Consider the differential game of border defense* (9.1)–(9.3). *The saddle point strategies of the Evader and the Pursuer are*

$$
\begin{aligned}
\cos \phi^* &= \frac{x_I - x_E}{\sqrt{(x_I - x_E)^2 + (y_I - y_E)^2}} \\
\sin \phi^* &= \frac{y_I - y_E}{\sqrt{(x_I - x_E)^2 + (y_I - y_E)^2}} \\
\cos \psi^* &= \frac{x_I - x_P}{\sqrt{(x_I - x_P)^2 + (y_I - y_P)^2}} \\
\sin \psi^* &= \frac{y_I - y_P}{\sqrt{(x_I - x_P)^2 + (y_I - y_P)^2}}.
\end{aligned}
\tag{9.11}
$$

In addition, $V(\mathbf{x})$ is continuous and continuously differentiable (except at singular surfaces) and $V(\mathbf{x})$ satisfies the HJI equation.

Proof. One can obtain the partial derivatives of V with respect to each element of the state. They are given by

$$
\begin{aligned}
\frac{\partial V}{\partial x_E} &= \frac{m}{(1-\alpha^2)(m^2+1)} \cdot \frac{m(x_c-x_F)-(y_c-y_F)}{\sqrt{(x_c-x_F)^2+(y_c-y_F)^2}} \\
&\quad - \frac{\alpha}{1-\alpha^2} \cdot \frac{x_E-x_P}{\sqrt{(x_E-x_P)^2+(y_E-y_P)^2}} \\
\frac{\partial V}{\partial y_E} &= \frac{1}{(1-\alpha^2)(m^2+1)} \cdot \frac{-m(x_c-x_F)+(y_c-y_F)}{\sqrt{(x_c-x_F)^2+(y_c-y_F)^2}} \\
&\quad - \frac{\alpha}{1-\alpha^2} \cdot \frac{y_E-y_P}{\sqrt{(x_E-x_P)^2+(y_E-y_P)^2}} \\
\frac{\partial V}{\partial x_P} &= \frac{\alpha^2 m}{(1-\alpha^2)(m^2+1)} \cdot \frac{-m(x_c-x_F)+(y_c-y_F)}{\sqrt{(x_c-x_F)^2+(y_c-y_F)^2}} \\
&\quad + \frac{\alpha}{1-\alpha^2} \cdot \frac{x_E-x_P}{\sqrt{(x_E-x_P)^2+(y_E-y_P)^2}} \\
\frac{\partial V}{\partial y_P} &= \frac{\alpha^2}{(1-\alpha^2)(m^2+1)} \cdot \frac{m(x_c-x_F)-(y_c-y_F)}{\sqrt{(x_c-x_F)^2+(y_c-y_F)^2}} \\
&\quad + \frac{\alpha}{1-\alpha^2} \cdot \frac{y_E-y_P}{\sqrt{(x_E-x_P)^2+(y_E-y_P)^2}}.
\end{aligned}
\tag{9.12}
$$

Define the following

$$
\cos\eta = \frac{x_F-x_c}{\sqrt{(x_F-x_c)^2+(y_F-y_c)^2}}, \qquad \sin\eta = \frac{y_F-y_c}{\sqrt{(x_F-x_c)^2+(y_F-y_c)^2}} \tag{9.13}
$$

so that $\tan\eta = \bar{m}$ as it can be seen in Fig. 9.3. Also define

$$
\cos\lambda = \frac{x_E-x_P}{\sqrt{(x_E-x_P)^2+(y_E-y_P)^2}}, \qquad \sin\lambda = \frac{y_E-y_P}{\sqrt{(x_E-x_P)^2+(y_E-y_P)^2}} \tag{9.14}
$$

where λ is the LOS angle from P to E.

Then, equations (9.12) can be written as follows

$$
\begin{aligned}
\frac{\partial V}{\partial x_E} &= \frac{1}{1-\alpha^2}\left[\frac{m}{m^2+1}(-m\cos\eta+\sin\eta)-\alpha\cos\lambda\right] \\
\frac{\partial V}{\partial y_E} &= \frac{1}{1-\alpha^2}\left[\frac{1}{m^2+1}(m\cos\eta-\sin\eta)-\alpha\sin\lambda\right] \\
\frac{\partial V}{\partial x_P} &= \frac{\alpha}{1-\alpha^2}\left[\frac{\alpha m}{m^2+1}(m\cos\eta-\sin\eta)+\cos\lambda\right] \\
\frac{\partial V}{\partial y_P} &= \frac{\alpha}{1-\alpha^2}\left[\frac{\alpha}{m^2+1}(-m\cos\eta+\sin\eta)+\sin\lambda\right]
\end{aligned}
\tag{9.15}
$$

Note that $\bar{m} = \tan \eta = \frac{\sin \eta}{\cos \eta}$. Then $m = -\frac{1}{\bar{m}} = -\frac{\cos \eta}{\sin \eta}$ and $m^2 + 1 = \frac{1}{\sin^2 \eta}$. Additionally, $m \cos \eta - \sin \eta = -\frac{1}{\sin \eta}$ and (9.15) can be written in the simplified form:

$$
\begin{aligned}
\frac{\partial V}{\partial x_E} &= -\frac{1}{1 - \alpha^2}(\cos \eta + \alpha \cos \lambda) \\
\frac{\partial V}{\partial y_E} &= -\frac{1}{1 - \alpha^2}(\sin \eta + \alpha \sin \lambda) \\
\frac{\partial V}{\partial x_P} &= \frac{\alpha}{1 - \alpha^2}(\alpha \cos \eta + \cos \lambda) \\
\frac{\partial V}{\partial y_P} &= \frac{\alpha}{1 - \alpha^2}(\alpha \sin \eta + \sin \lambda)
\end{aligned}
\tag{9.16}
$$

Note that the coordinates of the interception point I can be written as follows: $x_I = x_c + r \cos \eta$ and $y_I = y_c + r \sin \eta$. One can write

$$
\begin{aligned}
\frac{x_I - x_E}{r} &= \alpha \cos \lambda + \cos \eta \\
\frac{y_I - y_E}{r} &= \alpha \sin \lambda + \sin \eta \\
\frac{x_I - x_P}{r} &= \frac{1}{\alpha} \cos \lambda + \cos \eta \\
\frac{y_I - y_P}{r} &= \frac{1}{\alpha} \sin \lambda + \sin \eta
\end{aligned}
\tag{9.17}
$$

By multiplying and dividing each equation in (9.11) by $1/r$ gives

$$
\begin{aligned}
\cos \phi^* &= \frac{\alpha \cos \lambda + \cos \eta}{\sqrt{(\alpha \cos \lambda + \cos \eta)^2 + (\alpha \sin \lambda + \sin \eta)^2}} \\
\sin \phi^* &= \frac{\alpha \sin \lambda + \sin \eta}{\sqrt{(\alpha \cos \lambda + \cos \eta)^2 + (\alpha \sin \lambda + \sin \eta)^2}} \\
\cos \psi^* &= \frac{\frac{1}{\alpha} \cos \lambda + \cos \eta}{\sqrt{(\frac{1}{\alpha} \cos \lambda + \cos \eta)^2 + (\frac{1}{\alpha} \sin \lambda + \sin \eta)^2}} \\
\sin \psi^* &= \frac{\frac{1}{\alpha} \sin \lambda + \sin \eta}{\sqrt{(\frac{1}{\alpha} \cos \lambda + \cos \eta)^2 + (\frac{1}{\alpha} \sin \lambda + \sin \eta)^2}}
\end{aligned}
\tag{9.18}
$$

In general, the HJI equation is given by

$$
-\frac{\partial V}{\partial t} = \frac{\partial V}{\partial \mathbf{x}} \cdot \mathbf{f}(\mathbf{x}, \psi^*, \phi^*) + g(t, \mathbf{x}, \psi^*, \phi^*)
\tag{9.19}
$$
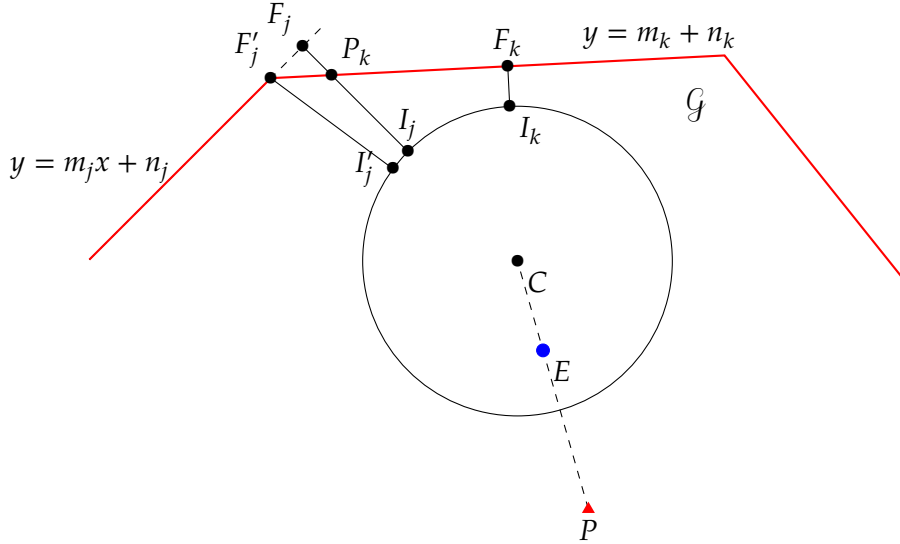
Figure 9.4: Convex Border

In this problem it must be that $\frac{\partial V}{\partial t} = 0$ and $g(t, \mathbf{x}, \psi^*, \phi^*) = 0$. Then, the following is computed,

$$
\frac{\partial V}{\partial x_E} \alpha \cos \phi^* + \frac{\partial V}{\partial y_E} \alpha \sin \phi^* + \frac{\partial V}{\partial x_P} \cos \psi^* + \frac{\partial V}{\partial y_P} \sin \psi^*
$$

$$
= -\frac{\alpha}{1 - \alpha^2} \cdot \frac{(\cos \eta + \alpha \cos \lambda)^2 + (\sin \eta + \alpha \sin \lambda)^2}{\sqrt{1 + \alpha^2 + 2\alpha(\cos \lambda \cos \eta + \sin \lambda \sin \eta)}}
$$

$$
+ \frac{\alpha^2}{1 - \alpha^2} \cdot \frac{\frac{1}{\alpha^2}(\alpha \cos \eta + \cos \lambda)^2 + \frac{1}{\alpha^2}(\alpha \sin \eta + \sin \lambda)^2}{\frac{1}{\alpha}\sqrt{1 + \alpha^2 + 2\alpha(\cos \lambda \cos \eta + \sin \lambda \sin \eta)}}
$$

$$
= \frac{\alpha}{1 - \alpha^2} \cdot \frac{-\left(1 + \alpha^2 + 2\alpha(\cos \lambda \cos \eta + \sin \lambda \sin \eta)\right) + 1 + \alpha^2 + 2\alpha(\cos \lambda \cos \eta + \sin \lambda \sin \eta)}{\sqrt{1 + \alpha^2 + 2\alpha(\cos \lambda \cos \eta + \sin \lambda \sin \eta)}}
$$

$$
= 0
$$

$$
(9.20)
$$

In summary, the Value function $V(\mathbf{x})$ was obtained, it is continuous and continuously differentiable (outside dispersal surfaces), and it satisfies the HJI equation. □

Recall that the optimal strategies, as obtained in Proposition 9.1, were obtained from candidate strategies corresponding to each segment of the border. Each candidate solution is obtained by finding the closest point on the segment to the AC whose radius and center are given by (9.4) and (9.7), respectively. The closest point to the circle is obtained by means of the orthogonality condition in Proposition 9.1.
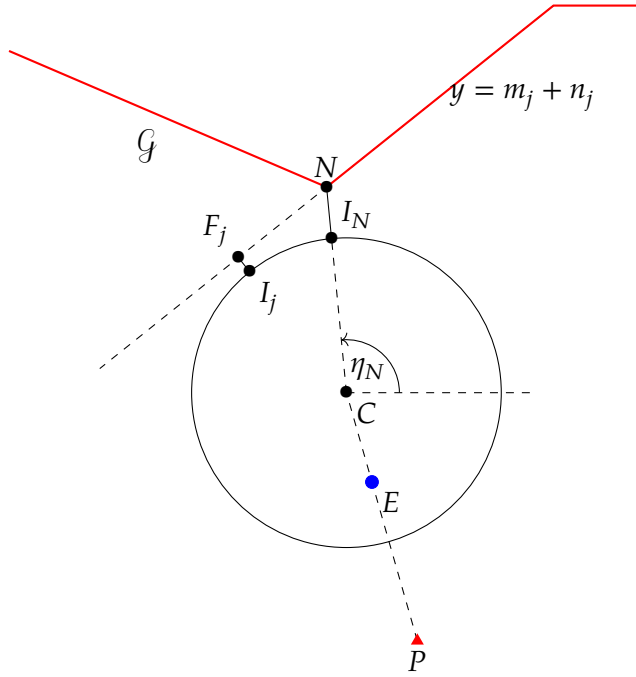
Figure 9.5: Non-convex Border

When considering the candidate solution in the Euclidean plane and when the border is convex, there exist two cases: the closest point on segment j is such that $F_j \in \partial\mathcal{G}$ and the case where the point on the extended segment is such that $F_j \notin \mathcal{G}$. The second case is illustrated in Fig. 9.4 by drawing the extended line corresponding to segment j.

Corollary 9.1. *If the closest point on segment j is such that $F_j \notin \mathcal{G}$, there exists $F_k \in \mathcal{G}$ such that $V_k(\mathbf{x}) < V_j(\mathbf{x})$. Hence, the point F_j is not the solution of the game and need not be considered in order to obtain the optimal strategies.*

Proof. Consider $\mathbf{x} \in \mathcal{R}_p$ and consider the candidate solution $F_j \notin \mathcal{G}$. As it is shown in Fig. 9.4, it holds that $V_j(\mathbf{x}) = \overline{F_j I_j} > \overline{P_k I_j} > \overline{F_k I_k} = V_k(\mathbf{x})$, where P_k is the intersection point between segment k of the border and the line segment $\overline{F_j I_j}$. By the previous relationship, one can also discard any other point on segment j. For instance, since $F_j \notin \mathcal{G}$ is an infeasible solution, the point $F_j' \in \partial\mathcal{G}$ (in segment j) is the closest feasible point to the circle (to the corresponding point I_j' on the circle) and one may be tempted to consider this choice. However, it must be that $\overline{F_j' I_j'} > \overline{F_j I_j}$ by definition of F_j, then, $\overline{F_j' I_j'} > \overline{F_k I_k} = V_k(\mathbf{x})$. $\qquad\square$

9.2.2   *Non-convex Border*

In this section, the existence of a non-convex border is considered. When the border is non-convex, the case where the closest point on the extended segment j is such that $F_j \in \mathcal{G}$ is possible. This is now illustrated in Fig. 9.5. Point $F_j \in \mathcal{G}$ is an infeasible solution since it does not lead the Evader directly into safe haven. However, one can still consider the closest feasible point on segment j which is point $N \in \partial \mathcal{G}$ in Fig. 9.5, the 'corner' of the non-convex border into consideration. Under this scenario, define

$$
V_j(\mathbf{x}) = \sqrt{\left(\tfrac{1}{1-\alpha^2}(x_E - \alpha^2 x_P) - x_N \right)^2 + \left(\tfrac{1}{1-\alpha^2}(y_E - \alpha^2 y_P) - y_N \right)^2}
$$
$$
- \tfrac{\alpha}{1-\alpha^2}\sqrt{(x_E - x_P)^2 + (y_E - y_P)^2}
$$

$$(9.21)$$

where the coordinates of point N are (x_N, y_N).

Corollary 9.2. *Consider* $\mathbf{x} \in \mathcal{R}_p$ *and a non-convex border where each segment j has an associated $V_j(\mathbf{x})$ as defined in* (9.21). *In addition, the case where $i^* = j$ is considered, that is, the closest point on the border to the* AC *is point N. Then, the Value function* (9.21) *is continuous and continuously differentiable (except at singular surfaces) and it satisfies the* HJI *equation.*

Proof. The partial derivatives of the Value function with respect to each element of the state are given by

$$
\begin{aligned}
\frac{\partial V}{\partial x_E} &= \frac{1}{1 - \alpha^2} \cdot \frac{x_c - x_N}{\sqrt{(x_c - x_N)^2 + (y_c - y_N)^2}} \\
&\quad - \frac{\alpha}{1 - \alpha^2} \cdot \frac{x_E - x_P}{\sqrt{(x_E - x_P)^2 + (y_E - y_P)^2}} \\
\frac{\partial V}{\partial y_E} &= \frac{1}{1 - \alpha^2} \cdot \frac{y_c - y_N}{\sqrt{(x_c - x_N)^2 + (y_c - y_N)^2}} \\
&\quad - \frac{\alpha}{1 - \alpha^2} \cdot \frac{y_E - y_P}{\sqrt{(x_E - x_P)^2 + (y_E - y_P)^2}} \\
\frac{\partial V}{\partial x_P} &= -\frac{\alpha^2}{1 - \alpha^2} \cdot \frac{x_c - x_N}{\sqrt{(x_c - x_N)^2 + (y_c - y_N)^2}} \\
&\quad + \frac{\alpha}{1 - \alpha^2} \cdot \frac{x_E - x_P}{\sqrt{(x_E - x_P)^2 + (y_E - y_P)^2}} \\
\frac{\partial V}{\partial y_P} &= -\frac{\alpha^2}{1 - \alpha^2} \cdot \frac{y_c - y_N}{\sqrt{(x_c - x_N)^2 + (y_c - y_N)^2}} \\
&\quad + \frac{\alpha}{1 - \alpha^2} \cdot \frac{y_E - y_P}{\sqrt{(x_E - x_P)^2 + (y_E - y_P)^2}}.
\end{aligned}
\tag{9.22}
$$

In this case, define

$$
\begin{aligned}
\cos \eta_N &= \frac{x_N - x_c}{\sqrt{(x_N - x_c)^2 + (y_N - y_c)^2}}, \\
\sin \eta_N &= \frac{y_N - y_c}{\sqrt{(x_N - x_c)^2 + (y_N - y_c)^2}}
\end{aligned}
\tag{9.23}
$$

and one can write (9.22) as follows

$$
\begin{aligned}
\frac{\partial V}{\partial x_E} &= -\frac{1}{1 - \alpha^2} \left(\cos \eta_N + \alpha \cos \lambda \right) \\
\frac{\partial V}{\partial y_E} &= -\frac{1}{1 - \alpha^2} \left(\sin \eta_N + \alpha \sin \lambda \right) \\
\frac{\partial V}{\partial x_P} &= \frac{\alpha}{1 - \alpha^2} \left(\alpha \cos \eta_N + \cos \lambda \right) \\
\frac{\partial V}{\partial y_P} &= \frac{\alpha}{1 - \alpha^2} \left(\alpha \sin \eta_N + \sin \lambda \right).
\end{aligned}
\tag{9.24}
$$

The coordinates of point I_N in Fig. 9.5 can be written as follows: $x_{I_N} = x_c + r \cos \eta_N$ and $y_{I_N} = y_c + r \sin \eta_N$. In addition, one can also obtain

similar expressions to (9.17)-(9.18) but in terms of η_N. Finally, the HJI equation is satisfied in a similar form to (9.20).   □

9.2.3 *Multiple Pursuer Solution*

The results from the previous section are now extended to the case of multiple Pursuers. This is done primarily by making use of the solution to the minimum capture time problem for the two-Pursuer [107, 130] and multiple-Pursuer [246] scenarios. In the minimum capture time two-Pursuer one-Evader problem, depending on the initial conditions, the solution is for all three agents to head to the further of the two AC intersections. For that game, it is also possible for the solution to degenerate to capture by a single Pursuer; in this case, the presence of the second Pursuer does not affect the outcome of the game. For more than two Pursuers, [246] characterized the feasible region for the Evader; that is, the set of points reachable by the Evader before any of the Pursuers. Let the Apollonius disk (i.e. the AC and its interior) for the ith Pursuer be given as

$$D_i = (r_i, \ (x_{C_i}, y_{C_i})) \tag{9.25}$$

where r_i is the radius of the AC from (9.4) and (x_{C_i}, y_{C_i}) is the center of the AC from (9.7). Then the Evader's feasible region is defined as

$$S = \cap_{i=1}^{M} D_i, \qquad i = 1, \dots, M \tag{9.26}$$

where M is the number of Pursuers. Then let ∂S be the (inclusive) boundary of the region.

Proposition 9.2. *Optimal capture must occur in S*

For any point outside of S, it cannot be guaranteed that the Evader can reach the point before being captured (see [46, 246]). Up to a point, increasing the number of Pursuers has the effect of reducing the Evader's feasible region which is generally advantageous for the Pursuers.

For the purposes of computing the multiple Pursuer solution it is useful to reparameterize the border using an ordered set of points with a line segment joining consecutive points. Let B_i represent one such point, then the border is comprised of line segments $B_{i,i+1}$ for $0 < i < K$ where K is the total number of corners of the border, including endpoints. Thus the border, B, is defined as a set of points along these line segments:

$$B \equiv \big\{ (x,y) \mid \exists i \text{ s.t. } (x,y) \in B_{i,i+1} \big\} \tag{9.27}$$

It is also assumed that the border is open, that is $B_K \neq B_1$, and free of any self intersections (i.e. no two segments cross).

In Section 9.2.1 it was shown that the closest point on a circle to line, say $B_{j,j+1}$, can be found by drawing a line segment perpendicular to $B_{j,j+1}$ that terminates at the circle's center. The closest point I_j is then the intersection between this line segment and the circle. Similarly, F_j is the other endpoint of this line segment, incident to $B_{j,j+1}$. Repeating for all border segments $0 < j < K$ for a Pursuer (say, Pursuer i) yields a set of candidate capture points $\mathcal{I}_i = \{I_1, \dots, I_{K-1}\}$ and their corresponding projection onto the border segment $\mathcal{F}_i = \{F_1, \dots, F_{K-1}\}$. The optimal capture point for the single-Pursuer problem is then given as

$$I^* = \arg\min_{I_j \in \mathcal{I}_i} \|I_j - F_j\| \tag{9.28}$$

For the multiple Pursuer case, the above procedure is repeated for each Pursuer, but there are some additional caveats. As was shown in [107, 130, 246] optimal capture for the multiple Pursuer case may occur at the intersection of ACs. For the minimum capture time problem, it was also shown in [246] that capture may occur in the interior of S (i.e. in $S \cap \overline{\partial S}$) with three or more Pursuers, but that is not the case here:

Lemma 9.1. *Under optimal play, capture occurs on ∂S.*

Proof. The distance from a point in any shape (e.g. S) to a point outside that shape (e.g. $(x,y) \in B$) is minimized at the boundary of that shape (e.g. ∂S). Since the Evader's cost is the distance from the point of capture to the closest border segment, the Evader would incur a loss

by choosing a point on the interior of S. Points beyond ∂S cannot be optimal capture locations as stated in Proposition 9.2. □

One consequence of Lemma 9.1 is that the closest point on a Pursuer's AC does not correspond to a candidate capture point if that point is not in ∂S. Thus it is also required for all I in $\mathcal{I}_i$ that $I \in \partial S$. Second, since capture may occur at an intersection of ACs, one must also consider each of these points as candidate solutions by computing the smallest distance from each AC intersection to each border segment. This is done similarly as before: draw a line segment which is perpendicular to the border segment whose endpoints are the border segment and the AC intersection; the length of this line segment is the cost/payoff associated with this candidate solution. Let the set of AC intersections be denoted as $\mathcal{I}_C$. Lastly, the orthogonality condition may yield a point F_j which lies beyond one of the endpoints of the line segment. In these cases, F_j may be moved to the nearest end point (either B_j or B_{j+1}). The corresponding I_j (for an AC) is the point on the circle that lies on the line connecting the circle center to F_j. This is especially pertinent in the case of a non-convex border as these candidates may indeed be optimal. Algorithm 2 gives a sketch of the solution process.

9.2.4 *Examples*

The first example, shown in Fig. 9.6, utilizes the same setup as in Fig. 9.1 with $\alpha = 0.6$. Essentially, the Value of the game is determined by comparing the lengths of all of the dotted gray lines emanating from the candidates. These are the lines of shortest length connecting the candidate to each border segment. It is clear by inspection of Fig. 9.6 that Algorithm 2 is inefficient: there are some lines which could not possibly be the shortest distance. Algorithm 2 contains a nested for-loop which has complexity $O(MK)$, since the number of AC intersections is at most M (i.e. $|\mathcal{I}_C| \leq M$). Using the parameterization of S given in [246], Algorithm 2 may be redesigned to achieve better worst-case performance (in a computational sense). This exercise is left for future work.

Algorithm 2 MP1E Border Solution

---

procedure MP1E-BORDER($\mathbf{x}$, α, B)
    $S \leftarrow$ Compute Evader's feasible region   ▷ see (9.26) and [246]
    $\mathcal{I}_C \leftarrow$ vertices of S ▷ see [246]
    if $S \cap B$ **then**
        return   ▷ escape is possible
    $d \leftarrow \infty$
    $I^* \leftarrow (0,0)$ ▷ dummy initialization
    for all D_i ACs and I_k vertices $\in \mathcal{I}_C$ **do**
        for $j = 1, \dots, K-1$ **do**   ▷ for each border segment
        Compute F ▷ Eqs. (9.9) and (9.10) for ACs
        if F lies off of $B_{j,j+1}$ **then**
            $F \leftarrow \arg\min_{h \in j,j+1} \overline{FB_h}$
        if AC, D_i **then**
            $I \leftarrow$ Compute the intersection of $\overline{FC}$ and ∂D_i
        elseVertex, $I_k \in \mathcal{I}_C$
            $I \leftarrow I_k$
        if $I \in \partial S$ **then**
            $d' \leftarrow \|I - F\|$
            if $d' < d$ **then**   ▷ update solution
                $d \leftarrow d'$
                $I^* \leftarrow I$
            else if $d' = d$ **then**
                $I^* \leftarrow [I^*, I]$ ▷ non-unique solution
    return d, I^* ▷ Value and capture point(s)
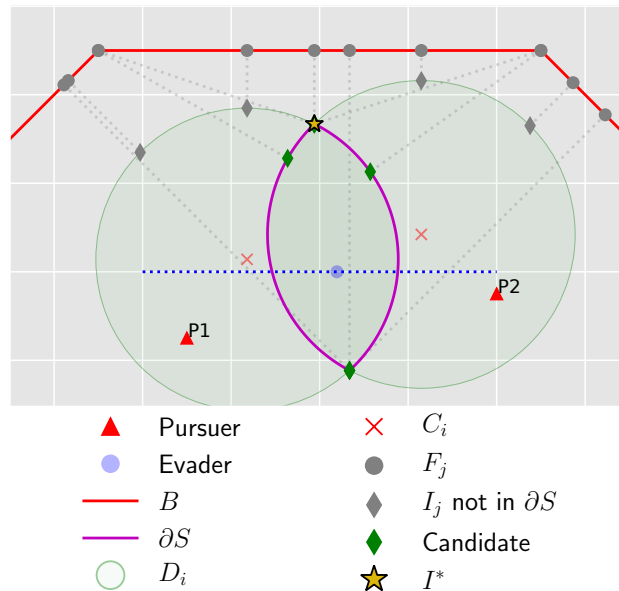
Figure 9.6: 2P1E scenario highlighting salient features in the solution process.
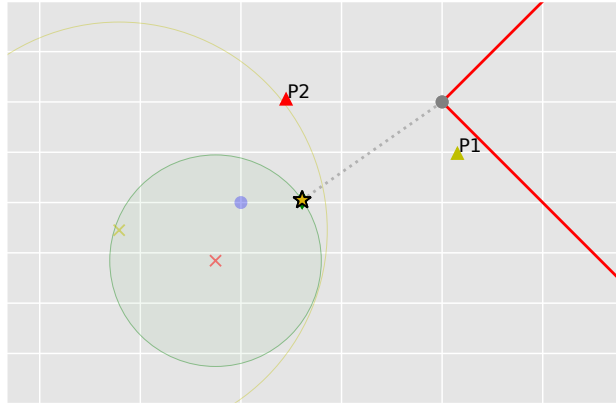
Figure 9.7: 2P1E scenario highlighting a non-convex border as well as degeneracy to capture by a single Pursuer. P1 and its AC are shown in yellow to signify that P1 does not participate in capture.

The next example highlights the case where the border is non-convex w.r.t the Evader as well as degeneracy to capture by a single Pursuer. Note the optimal capture point I^* is closest to to the corner of the border B_2. Incidentally, the line F_1I^* is not perpendicular to either border segment. In Fig. 9.7 Pursuer P1 is not able to reach the optimal capture point at the same time as the Evader and P2, and thus P1 and its AC are shown in yellow. It just so happens that the AC for P2 is contained entirely inside that of P1, and so ∂S is simply P2's AC. Because of the shape of the border there is only one candidate solution. Here, P1 is free to do anything (including flee from the Evader) without affecting the outcome of the game. These conditions are not necessary for degeneracy to capture by a single Pursuer, however. Indeed it is possible for ∂S to be determined by several Pursuers but still capture is undertaken by a single Pursuer. In the latter case, the non-capturing Pursuers may still need to aim for the optimal capture point. The non-capturing Pursuers cannot *reduce* the capture time. That is, they cannot improve the Pursuers' objective. However, if the non-capturing Pursuer(s) were to, e.g., flee from the Evader, it could be the case that a different point on ∂S becomes optimal. Such a switch may be beneficial (but never detrimental) to the Evader. Reference [246] contains a more complete categorization of the Pursuers in a MP1E scenario, which applies here as well. For the remainder, however, only a distinction between capturing (red) and non-capturing Pursuers (yellow) is made.
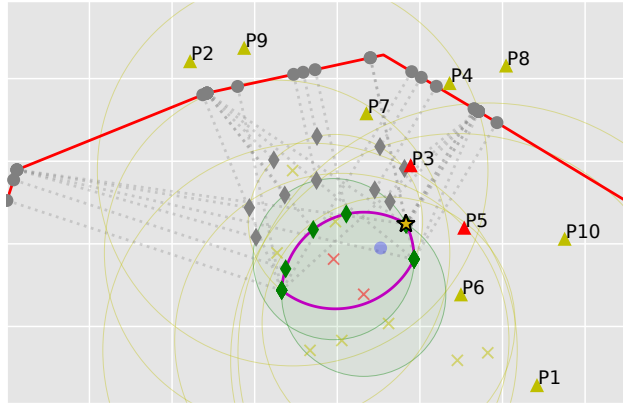
Figure 9.8: 10P1E example highlighting that capture is carried out by two Pursuers.
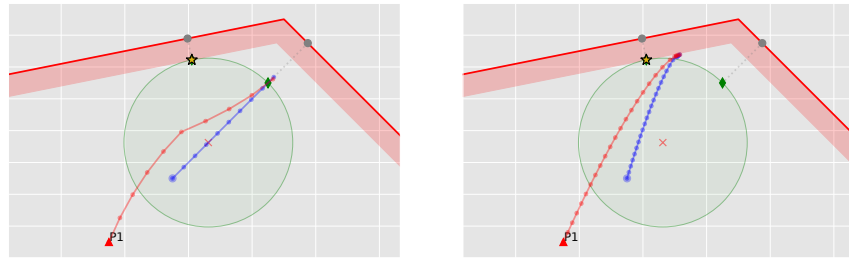
(a) Evader chooses the suboptimal candidate; $V = 0.068$ and $J = 0.177$

(b) Evader mixes pure evasion with heading to nearest border; $V = 0.068$ and $J = 0.073$

Figure 9.9: Optimal pursuit strategy against two different suboptimal Evader strategies demonstrating robustness.

Figure 9.8, shows that even for large M (e.g. 10, in this case) capture is almost always carried out by two or one Pursuers.

Remark 14. It is clear that only those Pursuers whose ACs are a part of ∂S are pertinent to the solution of the game. Thus Pursuers for which the following holds are discarded prior to invoking Algorithm 2,

$$\partial D_i \cap \partial S = \emptyset \tag{9.29}$$

The last example in this section, shown in Fig. 9.9, demonstrates the robustness of the policy from the perspective of the Pursuer team. For this example, the agents' controllers are implemented in a discretized manner in order to visualize the response of the players to this type of implementation constraint. At each discrete timestep, the Pursuer recomputes the optimal capture point via Algorithm 2 and chooses its

heading accordingly. In Fig. 9.9a the Evader chooses to head straight to the border segment on the right, a choice which has some merit as it initially aims directly away from the Pursuer and directly towards a border segment. Initially, the Pursuer, unaware of the Evader's implemented heading, heads towards the optimal capture point which is associated with the border segment on the left. At the fifth timestep, there is a discrete switch to heading to a capture point associated with the border segment on the right. Although the suboptimal Evader strategy allowed the Evader to reach a point beyond S, the Evader was not allowed to enter the red shaded region. In fact, the saddle-point property of the Value function ((9.5)) guarantees this robustness. The same is true for the scenario in Fig. 9.9b where the Evader balances fleeing directly from the Pursuer to avoid capture and heading directly towards the nearest border. Without solving the differential game, this latter strategy may seem like a good heuristic approach, however, the Evader's cost is greater than the Value of the game.

9.3 GAME OF KIND

This section addresses the game of kind within the border defense differential game.

9.3.1 *Single Pursuer*

The solution of the game of kind characterizes the barrier surface $H(\mathbf{x}; \alpha, m_i, n_i)$ that separates the state space into the two regions $\mathcal{R}_P$, the region of win of the Pursuer, and $\mathcal{R}_E$, the region of win of the Evader. If $\mathbf{x} \in \mathcal{R}_P$ the Pursuer, playing optimally, is guaranteed to capture the Evader before the latter reaches the border. On the other hand, if $\mathbf{x} \in \mathcal{R}_E$ the Evader, playing optimally, is guaranteed to escape and reach the border before being captured by the Pursuer.

A visual representation is provided on the Cartesian plane of the function $H(\mathbf{x}; \alpha, m_i, n_i)$ which specifies the surface that separates $\mathcal{R}_P$ from $\mathcal{R}_E$. By fixing the Pursuer coordinates (x_P, y_P) the closed form of

the barrier surface $H(x, y; x_P, y_P, \alpha, m_i, n_i) = 0$ is constructed. In other words, the barrier surface is comprised of the coordinate pairs (x, y) of the possible Evader position with respect to the Pursuer coordinates which guarantee that E will escape P if E plays optimally. Let $\mathcal{R}_P, \mathcal{R}_E \subset \mathbb{R}^2$ denote the region of win of the Pursuer and of the Evader, respectively, in the Cartesian plane when P coordinates are fixed.

Theorem 9.2. *Given the parameters (m_i, n_i) of segment i, for a given speed ratio parameter $0 < \alpha < 1$ and Pursuer coordinates (x_P, y_P), the barrier surface cross section that divides the Cartesian plane into the two regions R_P and R_E with respect to segment i is given by the Pursuer's side branch of the hyperbola*

$$h_{xx}x^2 + h_{yy}y^2 + 2h_{xy}xy + 2h_x x + 2h_y y + h = 0 \tag{9.30}$$

where

$$
\begin{aligned}
h_{xx} &= \left(1 - \alpha^2\right)\left(m_i^2 + 1\right) - 1, \\
h_{yy} &= \left(1 - \alpha^2\right)\left(m_i^2 + 1\right) - m_i^2, \\
h_{xy} &= -m_i, \\
h_x &= \alpha^2\left(x_P + m_i y_P\right) + \left(1 - \alpha^2\right)m_i n_i, \\
h_y &= \alpha^2\left(x_P + m_i y_P\right)m_i - \left(1 - \alpha^2\right)n_i, \\
h &= \left(1 - \alpha^2\right)n_i^2 + \alpha^2\left(1 - \alpha^2\right)\left[2n_i\left(y_P - mx_P\right)\right. \\
&\qquad \left. - \left(m_i^2 + 1\right)\left(x_P^2 + y_P^2\right)\right] - \alpha^4\left(x_P + m_i y_P\right)^2
\end{aligned}
\tag{9.31}
$$

Proof. In order to determine the barrier surface with respect to segment i consider the case where the AC is tangent to segment i. The following condition holds

$$\left(x_c - x_F\right)^2 + \left(y_c - y_F\right)^2 = \frac{\alpha^2}{\left(1 - \alpha^2\right)^2}\left[\left(x - x_P\right)^2 + \left(y - y_P\right)^2\right] \tag{9.32}$$

where the potential position of E is denoted by (x, y). Equation (9.32) can be written as follows

$$\frac{1}{1 - \alpha^2}\left(x^2 + y^2\right) - \frac{\alpha^2}{1 - \alpha^2}\left(x_P^2 + y_P^2\right) + x_F^2 y_F^2 - 2\left(x_c x_F + y_c y_F\right) = 0 \tag{9.33}$$

Further, writing (x_c, y_c, x_F, y_F) in terms of (x, y, x_P, y_P) and simplifying the resulting expression gives

$$\frac{1}{1 - \alpha^2}(x^2 + y^2) - \frac{\alpha^2}{1 - \alpha^2}(x_P^2 + y_P^2) + \frac{n_i^2}{m_i^2 + 1}$$
$$- \frac{\left(x + m_i y - \alpha^2 (x_P + m_i y_P)\right)^2 + 2\left(1 - \alpha^2\right) n_i (y - m_i x + \alpha^2 (m_i x_P - y_P))}{(1 - \alpha^2)^2 (m_i^2 + 1)}$$
$$= 0 \quad (9.34)$$

Multiplying the previous equation by $(1 - \alpha^2)^2 (m_i^2 + 1)$ and grouping the corresponding terms the quadratic equation (9.30) is obtained. The discriminant of the quadratic equation (9.30) is given by

$$D = \begin{vmatrix} \left(1 - \alpha^2\right)\left(m_i^2 + 1\right) - 1 & -m \\ -m & \left(1 - \alpha^2\right)\left(m_i^2 + 1\right) - m_i^2 \end{vmatrix}$$
$$= \left(1 - \alpha^2\right)^2 \left(m_i^2 + 1\right)^2 - \left(1 - \alpha^2\right)\left(m_i^2 + 1\right)^2$$
$$= -\alpha^2 \left(1 - \alpha^2\right)\left(m_i^2 + 1\right)^2$$
$$< 0$$

hence, equation (9.30) represents an hyperbola. □

Corollary 9.3. *Given a point N of a non-convex border, a given speed ratio parameter $0 < \alpha < 1$, and Pursuer coordinates (x_P, y_P), the barrier surface cross section that divides the Cartesian plane into the two regions R_P and R_E with respect to point N is an arc of the circle*

$$(x - x_N)^2 + (y - y_N)^2 = \alpha^2 \left[(x_P - x_N)^2 + (y_P - y_N)^2\right] \quad (9.35)$$

Proof. The proof is similar to the proof for Theorem 9.2 and is omitted.
□

In order to obtain the barrier surface for a collection of segments forming a convex border one needs to compute the intersection of the corresponding hyperbolas / circles to each couple of adjacent segments.

9.3.2   *Multiple Pursuers*

Theorem 9.3 expresses a more general condition that determines the
game of kind, which is applicable for both the single and multiple Pur-
suer cases.

Theorem 9.3. *Given the initial positions of the Evader and Pursuers, **x**, and
their respective speeds, if the border intersects the Evader's feasible region, the
Evader is able to escape:*

$$B \cap S \neq \emptyset \implies \mathbf{x} \in \mathcal{R}_E \tag{9.36}$$

where $\mathcal{R}_E$ is the region of win for the Evader.

Proof. The Apollonius disk defines points that can be reached by the
Evader at or before a particular Pursuer. Since, S is defined as the set
intersection of all of the Apollonius disks (c.f. (9.26)) the points in
S are reachable by the Evader at or before any Pursuer, provided the
Evader takes an appropriate action. If the Evader takes a straight-line
path to the point, then every point along its path is also inside S due
to the convexity of the region. Thus there always exists an *evasive path*
(in the sense of [46]) for the Evader to travel to a point in S. Therefore,
since $B \cap S \neq \emptyset$ there exists $(x, y) \in B$ that is also in S so the Evader
can escape safely, that is, the Pursuers can not guarantee capture.   □

Remark 15. If there is only one such point that meets the condition
stated in Theorem 9.3 the border is tangent to S and capture occurs
precisely at that point on the border.

9.3.2.1   *Example Problem*

Now these results are used to solve an example problem. Consider a
border comprised of a single, infinitely long line with equally spaced
UAVs flying along it in synchronized fashion, such that their spacing
is always constant; the scenario is depicted in Fig. 9.10. The objective
of the design problem is to defend the border by guaranteeing capture
of any single intruder while minimizing cost. Cost, here, is monoton-
ically increasing with the number of UAVs required (i.e. the density
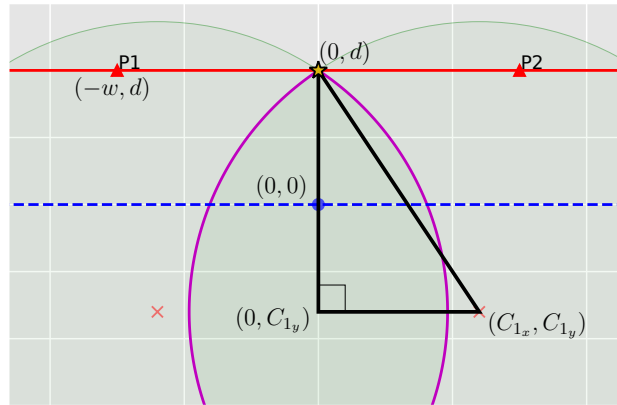
Figure 9.10: Critical point in the example problem

of UAVs along the border), the speed of the UAVs, and the sensing capabilities at the border. The cost may be considered as the monetary cost required to implement the border defense system. Thus the design variables are w, the half-distance between each UAV, α, the ratio of UAV speed to intruder speed, and d the sensor range at the border. The spacing w is essentially half of the inverse of UAV density along the border; smaller w means more UAVs and higher cost. The variable d determines the distance inside the border that the laser fence is located; this distance essentially determines when the pursuit begins. In addition, for closed-loop optimal control in the sense of the differential game solution, all the positions of the agents must be known; thus the border's tracking sensors must have full coverage of the space between the laser fence and the border.

To solve the design problem, the solution of the game of kind for multiple Pursuers is utilized. In particular, the critical point where capture occurs exactly at the border in the worst case is found. This critical point occurs when ∂S is tangent to the border, according to Theorem 9.3 and the subsequent remark. The best the Evader can do is to start halfway in between the two Pursuers; if Evader starts any closer to either Pursuer it will be captured further inside the border. This is due to the fact that simultaneous capture occurs on the perpendicular bisector of $\overline{P1\,P2}$ (c.f. [187]). Thus, in order for the Evader to minimize distance to the border at capture (i.e. maximize I_y), the Evader must begin on the bisector. Any amount of horizontal distance traveled is

akin to wasted time. For interception to occur exactly on the border, the following condition must hold:

$$C_{1_x}^2 + \left(d - C_{1_y} \right)^2 = R_1^2 \tag{9.37}$$

where,

$$\begin{aligned} C_{1_x} &= \frac{w\alpha^2}{1 - \alpha^2} \\ C_{1_y} &= \frac{-d\alpha^2}{1 - \alpha^2} \\ R_1 &= \frac{\alpha}{1 - \alpha^2}\sqrt{w^2 + d^2} \end{aligned} \tag{9.38}$$

Note, (9.38) comes from the definition of the AC. Substituting (9.38) into 9.37 and simplifying gives

$$\alpha = \frac{d}{w} \tag{9.39}$$

Eq. (9.39) also comes from the fact that the time of travel for Pursuers and Evader must be the same. At the critical point, the distance traveled by each Pursuer is w and the distance traveled by Evader is d, so $w\alpha = d$. Therefore, given all but one of the parameters α, w, d one can determine the value of the free parameter from (9.39). If the design parameters do not satisfy (9.39) the system is either over-designed (i.e. more costly than it ought to be) or cannot guarantee capture for all possible Evader initial conditions.

9.4   DISPERSAL SURFACE

9.4.1   *Single Pursuer*

For the single Pursuer case, when $V(\mathbf{x}) = \min_i V_i(\mathbf{x}) = V_k(\mathbf{x}) = V_l(\mathbf{x})$ for $k \neq l$ then a dispersal surface exists since more than one optimal solution exists. In general, there could be any number of optimal solutions where the region S is equidistant to multiple segments of the border and these distances are also the minimum. In such a case the

(a)   (b)   (c)
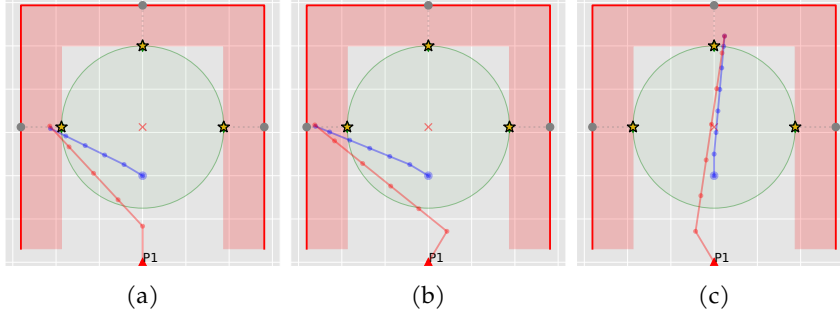
Figure 9.11: Dispersal surface in a 1P1E scenario with three optimal capture
points. Demonstration of agents choosing differently with $\Delta t = 0.05$ and $V = 0.094$.

Table 9.1: Cost/Payoff ($J \cdot 10^3$) for initial control action in Fig. 9.11

P \ E	West	North	East
West	94	71	25
North	75	94	75
East	25	71	94

Value function is continuous but it is not continuously differentiable since $\frac{\partial V_k}{\partial \mathbf{x}} \neq \frac{\partial V_l}{\partial \mathbf{x}}$, for instance

$$
\begin{aligned}
\frac{\partial V_k}{\partial x_E} &= \frac{m_k}{(1-\alpha^2)(m_k^2+1)} \cdot \frac{m_k(x_c-x_{F_k})-(y_c-y_{F_k})}{\sqrt{(x_c-x_{F_k})^2+(y_c-y_{F_k})^2}} - \frac{\alpha}{1-\alpha^2} \cdot \frac{x_E-x_P}{\sqrt{(x_E-x_P)^2+(y_E-y_P)^2}} \\
&\neq \frac{m_l}{(1-\alpha^2)(m_l^2+1)} \cdot \frac{m_l(x_c-x_{F_l})-(y_c-y_{F_l})}{\sqrt{(x_c-x_{F_l})^2+(y_c-y_{F_l})^2}} - \frac{\alpha}{1-\alpha^2} \cdot \frac{x_E-x_P}{\sqrt{(x_E-x_P)^2+(y_E-y_P)^2}} = \frac{\partial V_l}{\partial x_E}
\end{aligned}
$$

$$(9.40)$$

The main consequence is that the Pursuers will see a small decrease of performance if they choose different than the Evader. This is seen clearly in Fig. 9.11 where the Evader is able to breach the red shaded region if the Pursuer aims at a different optima initially. Note that the cost/payoff, J, is equal to V if the agents aim at the same optima initially. In all other cases, $J < V$ representing a loss for the Pursuer. Interestingly, the loss in performance varies with the actual choices of the agents. Table 9.1 gives the cost/payoff for every combination of initial control action. West, North, and East denote aiming at the optima associated with these directions, respectively. Now the choice of initial heading and the values in Table 9.1 may be treated as a matrix game to determine the best choice of initial heading for the agents. The optimal

Table 9.2: Cost/Payoff ($J \cdot 10^3$) for initial control action in Fig. 9.12

$\Delta t = 0.10$			$\Delta t = 0.15$		
P1/P2 \ E	W	E	P1/P2 \ E	W	E
W/W	211	206	W/W	211	201
W/E	208	208	W/E	206	206
E/W	208	208	E/W	206	206
E/E	206	211	E/E	201	211

strategies come from the Nash equilibrium (or equilibria) which may suggest either a pure strategy or mixed strategy is optimal. It is obvious, in this case, that the Pursuer's optimal strategy is to aim North, whereas the Evader chooses either West or East with equal probability. Thus the optimal playout of the full game appears as in Fig. 9.11a or its mirror image. Although the Value of the differential game is 0.094, the presence of the dispersal surface results in a cost/payoff of 0.075, assuming optimal play of this matrix game followed by optimal play in the differential game. The values in Table 9.1 are dependent upon the timestep. Smaller timesteps will yield smaller losses relative to the Value of the game. However, for the scenario shown in Fig. 9.11 the Nash equilibrium, and thus the optimal strategies, of the initial matrix game are independent of the timestep (unless, of course, the timestep is so large that reaching the border can be achieved in a single step).

9.4.2 *Multiple Pursuers*

The possibility for a dispersal surface exists also for the multiple Pursuer case. As in the single Pursuer case, the dispersal surface is the result of multiple segments of the border. That is, this dispersal surface does not arise from the presence of multiple Pursuers. If the border is made up of only one single segment, then the multi-Pursuer case does not contain a singular surface. Consider the scenario shown in Fig. 9.12 which has two Pursuers and two optima. There is a horizontal border segment joining the two vertical segments far above which does not affect the solution of the game. The Pursuers, in addition to having to choose East or West, also must choose whether to agree on a direction or split up. Table 9.2 contains the values of the cost/payoff for two different timestep sizes. For the matrix game with $\Delta t = 0.10$, there is a sin-
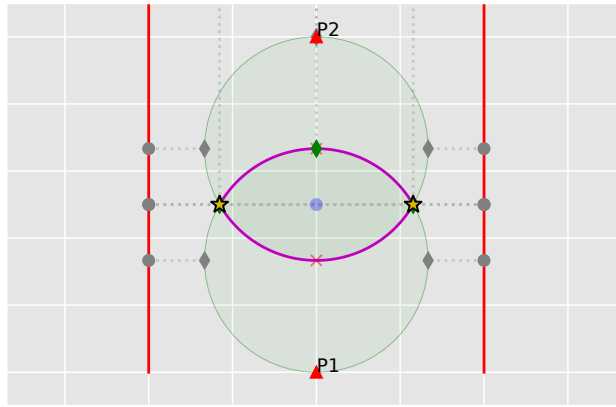
Figure 9.12: Dispersal surface in a 2P1E scenario with two optimal capture points.

gle equilibrium wherein the Pursuers' strategy is to both choose West or both choose East with equal probability, while the Evader's strategy is to choose West or East with equal probability. This is not the case for $\Delta t = 0.15$. Using vertex enumeration (c.f. [13]) to compute the equilibria of the matrix game yields four different equilibria. Now there is the issue that a strategy is only optimal against the corresponding opponent strategy associated with a particular equilibrium. In some scenarios, there may exist a focal point (see [203]) wherein there is a particular equilibrium that is sensible for the players to agree upon. Interestingly, for the present case, all four equilibria have a Pursuer strategy of splitting up! This is intuitive – the potential loss associated with both Pursuers choosing opposite the Evader is much greater for $\Delta t = 0.15$ than it is for $\Delta t = 0.10$; thus, splitting up is the best option here.

Finally, note that the *perpetual dilemma*, as described by Isaacs [130], does not appear to occur in this game. That is to say, after the agents make a choice of initial heading, regardless of the choices made, the system state will move off of the dispersal surface.

9.5   CONCLUSION

In this chapter a multiple Pursuer single Evader differential game with a practical application to border defense was formulated and solved.

An intruder who is attempting to escape through the border with some valuable intelligence was considered; the border is guarded by one or more UAVs whose goal is to capture the Evader (if possible) as far inside the border as possible. For the case of multiple UAVs, or Pursuers, strategies which maximize the cooperation were sought. Treating the scenario as a two-player differential game played by the intruder and the team of UAVs allowed just that. In the game of degree, capture can be guaranteed, and the cost/payoff is the smallest distance from the point of capture to any segment of the border. Using the well-known AC, the Value function for the single Pursuer single Evader game was expressed analytically and shown to satisfy the Hamilton-Jacobi-Isaacs equation. In this way, the solution was verified. The single Pursuer solution to multiple Pursuers was extended by considering simultaneous capture by two or more Pursuers at the intersection of their ACs. These candidate solutions behave much the same way as the single Pursuer candidate solutions. A simple approach was presented to determine the optimal capture point for given initial conditions, which, in turn, determines the Value of the game and the optimal headings for all the agents. For the game of kind, conditions were expressed, which can be easily checked, to determine in which region a given initial condition lies (whether in the Evader or Pursuers' win region). Finally, the dispersal surface in this game was characterized and optimal strategies for the Evader and Pursuers for choosing an initial heading were presented.

Part III

TURRET DEFENSE

INTRODUCTION TO TURRET DEFENSE

10.1 BACKGROUND

In comparison to the field of pursuit-evasion, turret defense is smaller and newer. Turret defense (and the closely related problem of perimeter patrol and defense) has a direct application to many military scenarios. In all of the work in this part, the Turret is pitted against one or more mobile agents referred to as Attackers. This part considers two different models for the Turret: 1) a kinetic Turret which must be aligned the Attacker in order to fire upon it, and 2) a wide-beam Turret which may affect the Attacker more or less as a function of the Attacker's distance and relative azimuth. The circle centered on the Turret with some particular radius is referred to henceforth as the target. The Attacker is successful in engaging the Turret if it is able to contact the target. Note that the kinetic Turret is, in essence, equivalent to a Defender agent who is constrained to move on a circular perimeter (in fact, this description is used in one of the following chapters); generally, the names Turret and Defender are interchangeable. The two Turret models are analyzed in the context of two different scenarios, one which is generally referred to as the TDDG wherein one or more mobile Attackers seek to collide with the Turret, and the other is the TEoR, wherein the Attacker has a choice between engaging the Turret or retreating to some other zone. Some of the work is focused on building up results from 1v1 versions of the problems to 1v2 and 2v1 versions; much of the rest is focused on higher-level decision-making and discrete events. This section contains some additional background on the two models and scenarios.

TURRET DEFENSE DIFFERENTIAL GAME   Turret defense may be considered to be a subclass of target guarding (c.f. [130, 142, 147]), or even

more generally, reach-avoid problems (e.g., [157, 270]). The work on reach-avoid games ranges from numerical computation of the reachable sets for each player to extensions to teams of players via decomposition [44, 45, 157]. Some recent works build up the solution to two-on-one scenarios using a rigorous differential game formulation with the one-on-one solution as a basis. For example, [268] derives the analytical barrier between Attacker and Defenders winning for a reach-avoid game which takes place inside a rectangular domain. A similar result has been obtained for a blocking game where two Defenders seek to prevent an Attacker from reaching a line segment [106]. Finally, several recent works have analyzed reach-avoid games inside a circular domain but with freely moving Defenders [103, 267]. The utility of these rigorous analytical results have been highlighted in task allocation schemes which are capable of handling teams of many agents [269].

Various turret and turret-like defense scenarios have been explored in recent literature. In [2], the authors formulated and solved the TDDG (along with all of its singularities) for the wide-beam Turret model, wherein the cost functional included a state-dependent integral cost. There, a single mobile Attacker sought to balance time-to-target with avoiding the LOS of the Turret; the resulting Attacker trajectories are generally curved in the Cartesian frame. Additionally, the three singular surfaces were analyzed: the Defender (Turret) Universal Surface, the Defender DS, and the Attacker DS wherein the Attacker chooses from a direct and indirect route to termination. Reference [214] analyzed a perimeter patrol scenario wherein termination occurs either when the Attacker reaches the target or when the Defender and Attacker are coincident. The solution characteristics of the single-Attacker, single-Defender and single-Attacker, two-Defender scenarios were then extended to a many-Attacker, many-Defender variant wherein the teams maximize (minimize, respectively) the number of hits on the target. An extension considered a heterogeneous Defender team comprised of uncontrolled and controlled patrollers [218], and also arbitrary convex targets [215].

For the scenarios with either multiple collocated Turrets or multiple Attackers, it is essential to understand how the agents on a team

can cooperate to achieve a common goal. Certain outcomes may only possible through cooperation; victory could even be contingent on the sacrifice of a particular agent. In Chapter 12 a particular sub-case is considered within the two-Attacker TDDG wherein one of the Attackers must sacrifice itself in order for the other Attacker to reach the target unhindered. Because the Attackers essentially have different roles, this problem is also related to other "three-body" problems in the literature, such as the TADDG [101, 184]. There, the Defender/Target team seek to cooperatively maneuver in such a way for the Defender to intercept the Attacker as far from the Target as possible. Another example is the single-pursuer, two-evader cooperative defense scenario presented in [85] wherein one of the Evaders performs a flanking maneuver on the Pursuer to drive up the Pursuer's cost. This scenario is also related to the problem of capture of evaders in succession [37, 109, 150, 270] since the Turret is free to aim at another Attacker once one is neutralized. The role selection portion of the analysis pertains to the determination of which Attacker will be pursued by the Turret first thereby fixing that Attacker to be the sacrificial one. This type of question (i.e., whether to behave as a 'Runner' or 'Penetrator') appears elsewhere in the differential game literature. For example, whether to behave as the Pursuer or Evader in symmetric engagements [118], e.g., the Game of Two Cars [112, 163], and in [177]. Other examples include the determination of which agent is the leader and which is the follower, as in the cooperative HCDG studied in [32].

TURRET ENGAGE OR RETREAT   In the TEoR scenario, the Attacker must decide whether to engage (and destroy) the Defender or retreat to a safe zone. The objective is, therefore, not only to obtain the equilibrium control strategies for the agents (i.e., instantaneous heading and turning direction), but also to determine the optimal choice or intention over the whole state space.

In the case of the kinetic Turret model, D can only deal damage when its turret's aim is fixed upon the A. Thus there are potentially two segments of the scenario: (1) D seeks to "lock on" to A while the latter has some interest in avoiding the turret's LOS, and (2) after lock, A ulti-

mately proceeds with its intent to engage or retreat. Prior to lock-on, A accumulates no cost; afterwards, it has a control- and time-dependent running cost. It is possible that lock-on does not occur at all, in which case, there is no running cost. This scenario is referred to as the Kinetic Turret Engage or Retreat (TEoRK) game.

In the case of the wide-beam Turret model, D can deal *some* amount of damage to A for any look angle. The analysis is simpler than the latter case in the sense that there are fewer distinct outcomes or situations (i.e., there are no "unlocked" cases), but the also more complex in the sense that the solutions are more difficult to compute. This scenario is referred to as the Wide-Beam Turret Engage or Retreat (TEoRW) game.

Many different approaches have been used to ascertain an adversary's intent. For example, in [206], the authors maintain a representation of an agent's belief of its adversary's most likely spatial trajectories as a probabilistic program. The agent's reasoning is "nested" (or recursive) in the sense that it then reasons about its adversary's belief of its own actions. Alternatively, in [117], the approach is based on estimating the rationality level of the adversary. The rationality level refers to the number of levels to which an agent is capable of computing a best response, back and forth. A rationality level of 2 means an agent can compute its best response to its adversary's best response of the initial strategy. The main difficulty in this approach is the estimation of the adversary's rationality level, assuming it is unknown. In lieu of methods based on inference (whether probability-based or estimation-based), this chapter is based upon game theory. The main advantage game theory offers is that, if an equilibrium can be found, then the associated equilibrium strategies yield robust performance regardless of what the adversary implements.

Existing game-theoretic literature has focused on larger scale conflicts with more agents, but in a discrete-time discrete-action space formulation [53, 71], or on pursuit-evasion engagements where there is one clearly specified goal [101, 147]. Of these two areas within the literature, the latter is most applicable to the current study as the primary concern the (continuous) spatial maneuvers of the players in continuous time.

The topic of intent selection has appeared within differential game literature most often under the guise of "role" selection (c.f. 2.5.3). Additionally, determination of intent (and, often, deliberate *signaling* of intent) occurs in natural adversarial scenarios [207]. For example, in [234], Tan *et al.* conclude that Asian honeybees signal their intent to retaliate against hornets should they decide to attack their hive. Additionally, some bird species are known to make alarm calls when predators are detected – the call serves a twofold purpose: to warn and rally nearby birds, and to signal to the predator that its presence is known [23]. In the latter case, the bird is indicating to the predator that it has sufficient time to escape should the predator decide to pursue. The parallel for the scenario of interest is that, even if D decides not to fire upon A, by aiming its turret to lock onto A it is indicating its ability to inflict damage.

More specifically, the work in the TEoR chapters is closely related to literature on engage or retreat differential games [83, 84, 90]. In the initial formulation, the Defender agent had no state and its control was whether or not to "turn on" its (omnidirectional, infinite range) defensive capability. The solution for the simple Defender case is mostly analytic [90]. Reference [84] abstracts the particulars of the agents and gives a more general framework for addressing engage or retreat games. In [248], the Defender was modeled as a turn-constrained turret (as in this part). Rather than having to lock onto A's position, D could inflict a "soft" cost as a function of its look angle. Consequently, there was no discrete switch in the cost or equilibrium behaviors. Generally, the engage or retreat problem is solved by breaking it into two subproblems: the Game of Engagement (GoE), which corresponds to the differential game that plays out when the Attacker intends to collide with the Turret, and Optimal Constrained Retreat (OCR), which corresponds to the optimal control problem that arises when the Attacker intends to retreat. Using this method, it was shown that switching intents (from retreat to engage, or vice versa) within the engage or retreat game does not yield a better utility [84]. To ensure that switching to engagement does not yield a better utility while retreating, a path constraint is imposed.

10.2   PRACTICAL SIGNIFICANCE

One real-world example of target guarding is the protection of a building's perimeter from mobile infiltrators, which may be considered to be people, ground vehicles, air vehicles, or even certain types of munitions. While protection from a single adversary is necessary, a more likely scenario is that multiple adversaries are present. With the proliferation of lower-cost unmanned vehicles and guided munitions, it is not unreasonable to imagine an attack comprised of tens or hundreds of agents. See, for example, the following excerpt from the Air Force 2030 Science & Technology Strategy document [237] (emphasis added):

> Swarms of low-cost, autonomous air and space systems can ... **absorb losses** that manned systems cannot... Low-end systems can restore the agility to attack adversary weaknesses in unexpected ways by **exploiting numbers** and complexity.

Thus the analysis of the two-Attacker case is a step towards analyzing the defense of a static location or asset against "swarms" of Attackers with a directional defensive weapon. There, the Turret must destroy all (or as many as is possible) Attackers in succession; the Attackers, meanwhile, coordinate their attack to maximize successful hits.

Moreover, real-world adversarial conflicts are complex and may evolve through many, fundamentally different, stages. Consider, for example, a ground-based defensive site or sensor. The site may be equipped with a radar system which is responsible for detecting targets [17], which, upon detection, may then estimate the position and pose of the target [82], and subsequently track the target over time [92]. In the case that the site has defensive capability, there may be more stages eventually leading up to the decision to fire its weapons or not. The decision of whether to fire may depend on what the target is *doing*, as well as what the ground-based site believes the target's *intentions* are. If, for example, the defensive site believes the target intends to engage and destroy the site, then it might decide to retaliate in the hopes of damaging the

target prior to its arrival. On the other hand, if the target appears as if it has no interest in engaging the site, then it may be best to hold fire. One could imagine that, if the target truly intends to fly past the site without engaging it that firing upon it could cause the target to reconsider.

The TEoR scenarios considered here have application to defense against risk-conscious attackers (e.g. vehicles, or otherwise expensive munitions) by a stationary platform. The Turret itself may represent a weapons platform of some kind or even a surveillance asset seeking to steer its sensor to maximize observation of the incoming Attacker. Generally speaking, the methodology itself is important because real-world conflicts of warfare rarely entail a single well-defined objective. Therefore, the ability to systematically consider multiple layers of decision-making and control available to both sides is highly desirable.

10.3 CONTRIBUTIONS

The contributions of the chapters in this part may be summarized as

- TDDGK solution for 1v1, 1v2, and 2v1 cases

- TDDGK Attacker(s) and Turret(s) winning regions

- TEoRK solution

- TEoRK engage, retreat, locked, unlocked regions

- TEoRK procedure for computing DSs

- TDDGW parameter exploration

- TDDGW procedure for computing DS

- TEoRW OCR solution

- Framework for treating two-stage differential games (demonstrated on TDDGK 2v1 and TEoRK)

Table 3.2 indicates that these contributions pertain to all nine of the Research Objectives laid out in Section 3.1. The chapter on the TEoR

game have a heavy focus on Research Objective 8, which is concerned with multi-objective scenarios. Additionally, the chapters for which the kinetic Turret model is analyzed heavily emphasize discrete events (i.e., Research Objective 9) – these discrete events largely correspond to the moment at which the Turret's look angle is aligned with an Attacker's position. Finally, a variety of DSs appear in these turret defense scenarios and this part details their construction and computation (which pertains to Research Objective 3).

11

SINGLE ATTACKER AND KINETIC TURRET(S)

11.1 INTRODUCTION

This chapter is the first pertaining to kinetic turret defense, (TDDGK) and it covers the case wherein the Attacker seeks to reach the target circle centered on the Turret while avoiding alignment with its LOS. To start, a single Attacker is pitted against a kinetic turret (the Defender) – in the case that the Attacker can win (i.e., reach the target) a zero-sum differential game is formulated over the terminal angle relative to the Defender's LOS; in the case that the Defender can win a zero-sum differential game is formulated over the Attacker's distance to the target at terminal time. A second Defender is also considered, which is collocated with the first; the presence of this second turret reduces the size of the region from where the Attacker can guarantee reaching the target. The results partly address Research Objectives 5 and 6 (adding more agents to a scenario and solving novel scenarios, respectively). Results for closely related problems exist, however the particular model used here, wherein the Attacker is neutralized upon alignment with the Defender's LOS, is novel. This material is based on the paper [255], which is in review.

This chapter contains the following contributions: (i) the one-on-one Attacker-win and Defender-win scenarios are formulated and solved rigorously using a differential game theoretic approach, verifying the saddle-point equilibrium status of strategies existing in the literature [214]; (ii) analytic expressions for the Value functions are derived for both one-on-one scenarios; (iii) the two-Defender, one-Attacker scenarios are formulated and the equilibrium strategies and Value functions are derived; (iv) the entire state space is partitioned based on all of the different terminal scenarios, and analytic expressions for the separating surfaces are derived; (v) an alternative scenario in which the

Attacker seeks to reach the target in minimum time is solved. The emphasis is on the analysis and proof methods, which are based on differential game theory, in comparison to the geometric methods used previously [214]. Sections 11.2 and 11.3 cover the one- and two-Defender cases, respectively. In each of those sections, both the Attacker-win and Defender(s)-win scenarios are formulated and solved. Section 11.4 concludes the chapter.

11.2 ONE DEFENDER

This section formulates the target guarding problem wherein the Defender (D) is constrained to move along the circular target perimeter and the Attacker (A) moves in the plane with simple motion. Figure 11.1 shows the local coordinate system (black) used in much of the analysis to appear, as well as the global (inertial) (x, y)-coordinate system (green). The following assumptions are made on the problem
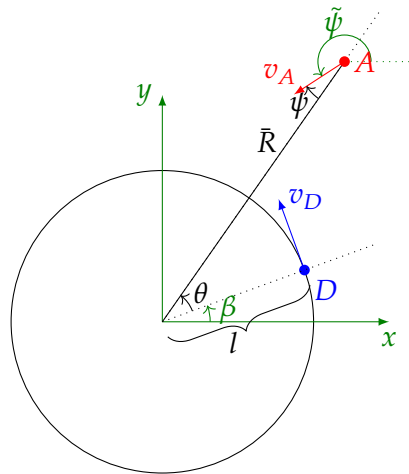
Figure 11.1: Circular perimeter patrol with one Defender and one Attacker.

setup:

Assumption 1. The players' speeds are such that $0 < v_A \leq v_D$, where v_A and v_D are the speeds of the Attacker and Defender, respectively.

Assumption 2. The initial separation angle is such that $\theta(t_0) = \theta_0 \in [0, \pi)$.

Assumption 3. The initial Attacker distance is such that $R(t_0) > 1 -$ that is, A begins outside the target circle.

Assumption 2 will be lifted after the equilibrium strategies have been derived and the symmetry (and attendant singularity) identified. The (dimensional) kinematics, based on Fig. 11.1 are

$$\bar{f}(\bar{\mathbf{x}}, \bar{u}, \bar{t}) = \dot{\bar{\mathbf{x}}} = \begin{bmatrix} \dot{\bar{R}} \\ \dot{\bar{\theta}} \\ \dot{\bar{\beta}} \end{bmatrix} = \begin{bmatrix} -v_A \cos \psi \\ \frac{v_A}{\bar{R}} \sin \psi - \frac{v_D}{l} \\ \frac{v_D}{l} \end{bmatrix}, \tag{11.1}$$

where $\bar{\theta} \in [-\pi, \pi]$ is the angle of A's position w.r.t. D and $\beta \in [0, 2\pi]$ represents the rotation of D about the circle's center w.r.t. a global (x, y)-plane. With the following definitions,

$$R \equiv \frac{\bar{R}}{l}, \quad t \equiv \frac{v_{D_{\max}}}{l}\bar{t}, \quad u_D \equiv \frac{v_D}{v_{D_{\max}}}, \quad \nu = \frac{v_A}{v_{D_{\max}}},$$

where $v_{D_{\max}}$ is the maximum Defender speed and the speed ratio $0 < \nu \leq 1$, the kinematics in (11.1) are non-dimensionalized:

$$f(\mathbf{x}, u, t) = \dot{\mathbf{x}} = \begin{bmatrix} \dot{R} \\ \dot{\theta} \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} -\nu \cos \psi \\ \nu \frac{1}{R} \sin \psi - u_D \\ u_D \end{bmatrix}. \tag{11.2}$$

The Defender control lies in the range $u_D \in [-1, 1]$, and the Attacker control lies in the range $\psi \in [-\pi, \pi]$. Note θ and $\bar{\theta}$ are equivalent, but their time derivatives differ due to the scaling of time.

The *Game of Kind* is defined as the question of whether Attacker can reach the perimeter $(R \to 1)$ with non-zero terminal separation angle (Attacker 'wins') or the Defender can drive $\theta \to 0$ before the Attacker reaches the perimeter (Defender 'wins'). The subscript f refers to conditions at termination (e.g., t_f is the terminal time). In the following sections, the surface separating these two cases is derived and a *Game of Degree* is specified and solved for each case.

Note that if $v_A > v_D$, the Attacker need only come within some distance $l < \bar{R} < l\frac{v_A}{v_D}$ wherein the Attacker has the control authority to force $\theta \to \pi$. Similarly, when $v_A \leq v_D$, if at some point $\theta = 0$ the game is over because the Defender has sufficient control authority to keep $\theta = 0$ regardless of the Attacker's control. It is assumed that if $\theta_f = 0$

the Defender has successfully intercepted the Attacker and thwarted its attack. The question of whether the Attacker wins (i.e. $\theta_f > 0$) or the Defender wins ($\theta_f = 0$) is referred to as the *Game of Kind*.

11.2.1   *Defender Win Scenario*

This section is focused on the *Game of Degree* which takes place when D is able to drive $\theta \to 0$ before A can reach the target. Here, the initial condition of the system lies in the region $\mathscr{R}_D$, which is the region of win for the Defender (see (11.28)). In this case, it is sensible for the agents to play a zero-sum game over the cost functional

$$J_d = \Phi_d\left(\mathbf{x}_f, t_f\right) = -R_f, \tag{11.3}$$

where the subscript f denotes conditions at termination. The negative sign in (11.3) is present so that the Defender is the minimizing player and the Attacker is the maximizing player. That is, the Attacker seeks to get as close as possible to $R_f = 1$ and the Defender seeks to maximize the terminal distance. This game is referred to as the *Game of Distance*, and it is denoted with subscript d, in general. The Value of the game, if it exists, is the saddle-point equilibrium of the cost functional over state-feedback strategies

$$V_d = \min_{u_D(\cdot)} \max_{\psi(\cdot)} J_d = \max_{\psi(\cdot)} \min_{u_D(\cdot)} J_d. \tag{11.4}$$

The terminal constraint for the *Game of Distance* is

$$\phi_d\left(\mathbf{x}_f, t_f\right) = \theta_f = 0. \tag{11.5}$$

The final time, t_f, is the first time for which $\theta(t) = 0$. Thus, the Terminal Surface is defined as the set of states satisfying (11.5)

$$\mathscr{T}_d = \{\mathbf{x} \mid R > 1 \text{ and } \theta = 0\}. \tag{11.6}$$

Assumptions 1 and 2 are retained for this analysis.

11.2.1.1 *First Order Necessary Conditions for Optimality*

The analysis is carried out according to a classical differential game approach [20, 130]. The kinematics remain unchanged from the previous analysis; the Hamiltonian for the *Game of Distance* is

$$\mathcal{H}_d = -\sigma_R v \cos \psi + \sigma_\theta \left(v \frac{1}{R} \sin \psi - u_D \right) + \sigma_\beta u_D, \tag{11.7}$$

where $\sigma \equiv \begin{bmatrix} \sigma_R & \sigma_\theta & \sigma_\beta \end{bmatrix}^\top$ is the adjoint vector for the *Game of Distance*. The Hamiltonian is a separable function of the controls u_D and ψ, and thus *Isaacs' condition* [20, 130] holds:

$$\min_{u_D} \max_{\psi} \mathcal{H}_d = \max_{\psi} \min_{u_D} \mathcal{H}_d, \qquad \forall \mathbf{x},$$

where $u_D \in [-1, 1]$ and $\psi \in [-\pi, \pi]$. The equilibrium adjoint dynamics are given by

$$\dot{\sigma}_R = -\frac{\partial \mathcal{H}_d}{\partial R} = v \sigma_\theta \frac{1}{R^2} \sin \psi, \tag{11.8}$$

$$\dot{\sigma}_\theta = -\frac{\partial \mathcal{H}_d}{\partial \theta} = 0, \tag{11.9}$$

$$\dot{\sigma}_\beta = -\frac{\partial \mathcal{H}_d}{\partial \beta} = 0. \tag{11.10}$$

The terminal adjoint values are obtained from the transversality condition [40, pg. 89]

$$\sigma^\top(t_f) = \frac{\partial \Phi_d}{\partial \mathbf{x}_f} + \eta \frac{\partial \phi_d}{\partial \mathbf{x}_f}$$

$$= \begin{bmatrix} -1 & 0 & 0 \end{bmatrix} + \eta \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$$

$$\sigma_{R_f} = -1$$

$$\implies \sigma_{\theta_f} = \eta \tag{11.11}$$

$$\sigma_{\beta_f} = 0,$$

where η is an additional adjoint variable whose value will be determined later in the analysis. Therefore, with (11.9)–(11.11), the following hold

$$\sigma_\theta(t) = \eta, \qquad \forall t \in \left[t_0, t_f\right] \tag{11.12}$$

$$\sigma_\beta(t) = 0, \qquad \forall t \in \left[t_0, t_f\right]. \tag{11.13}$$

Once again, since $\sigma_\beta(t) = 0$ for all $t \in \left[t_0, t_f\right]$, the state component β has no effect on the equilibrium trajectory or the equilibrium control strategies. The terminal Hamiltonian satisfies [40]

$$\mathcal{H}_d(t_f) = -\frac{\partial \Phi_d}{\partial t_f} - \eta \frac{\partial \phi_d}{\partial t_f} = 0, \tag{11.14}$$

and $\frac{\mathrm{d}\mathcal{H}_d}{\mathrm{d}t} = 0$, so $\mathcal{H}_d(t) = 0$ for all $t \in \left[t_0, t_f\right]$.

The equilibrium control actions of the Attacker and Defender maximize and minimize (11.7), respectively: $\mathcal{H}_d^* = \max_\psi \min_{u_D} \mathcal{H}_d$. In order to maximize (11.7) (with (11.12)), the vector $\begin{bmatrix} \cos\psi & \sin\psi \end{bmatrix}$ must be parallel to the vector $\begin{bmatrix} \sigma_R & \frac{\eta}{R} \end{bmatrix}$, giving

$$\cos\psi^* = \frac{-\sigma_R}{\sqrt{\sigma_R^2 + \frac{\eta^2}{R^2}}}, \qquad \sin\psi^* = \frac{\eta}{R\sqrt{\sigma_R^2 + \frac{\eta^2}{R^2}}}. \tag{11.15}$$

If $\eta < 0$, this implies $\sin\psi^* < 0$ due to (11.15). However, this would mean the Attacker has a component of its motion that points *towards* the Defender due to Assumption 2 (see, e.g., Fig. 11.1). Thus, it must be the case that $\eta > 0$. In order to minimize (11.7) (with (11.12)), the Defender's control must satisfy

$$u_D^* = \operatorname{sign} \eta = 1, \tag{11.16}$$

since $\eta > 0$.

Substituting the equilibrium controls, (11.15) and (11.16), into the Hamiltonian, (11.7), and evaluating at final time with (11.11) and (11.14) gives

$$\mathcal{H}_d^*(t_f) = 0 = \frac{\nu \sigma_{R_f}^2}{\sqrt{\sigma_{R_f}^2 + \frac{\eta^2}{R_f^2}}} + \frac{\nu \eta^2}{R_f^2 \sqrt{\sigma_{R_f}^2 + \frac{\eta^2}{R_f^2}}} - \eta$$

$$\implies \eta = \pm \nu R_f \sqrt{\frac{1}{R_f^2 - \nu^2}}.$$

Since $\eta > 0$, it must be that

$$\eta = \nu R_f \sqrt{\frac{1}{R_f^2 - \nu^2}}. \tag{11.17}$$

11.2.1.2 *Solution Characteristics*

An expression for σ_R is obtained by considering the Hamiltonian at a general time, making the same substitutions as before, with the additional substitution of (11.17):

$$\mathcal{H}_d^*(t) = 0 = \nu \sqrt{\sigma_R^2 + \frac{\eta^2}{R^2}} - \eta$$

$$\implies \sigma_R = \pm \sqrt{\frac{\eta^2}{\nu^2} - \frac{\eta^2}{R^2}}$$

$$= \pm \frac{R_f}{R} \sqrt{\frac{R^2 - \nu^2}{R_f^2 - \nu^2}}.$$

Since $\sigma_{R_f} < 0$ (due to (11.11)) and $\dot{\sigma}_R > 0$ (due to (11.8) with (11.15) and $\eta > 0$) it must be that $\sigma_R(t) < 0$ for all $t \in \left[t_0, t_f\right]$, thus

$$\sigma_R = -\frac{R_f}{R} \sqrt{\frac{R^2 - \nu^2}{R_f^2 - \nu^2}}. \tag{11.18}$$

The retrograde equilibrium kinematics (denoted by $\mathring{\mathbf{x}}^*$, where $\mathring{\mathbf{x}}^* = -\dot{\mathbf{x}}^*$) can be obtained by substituting the equilibrium controls, (11.15)

and (11.16), along with the adjoints, (11.12), (11.13), and (11.18), into (11.2) which yields

$$\mathring{R}^* = \nu\sqrt{1 - \frac{\nu^2}{R^2}}, \qquad \mathring{\theta}^* = 1 - \frac{\nu^2}{R^2}, \tag{11.19}$$

with the following boundary conditions

$$R(t_f) > 1, \qquad \theta(t_f) = 0. \tag{11.20}$$

Note that both $\mathring{R}$ and $\mathring{\theta}$ are monotonically increasing according to (11.19). Consider the differential equation obtained by dividing the equations in (11.19)

$$\frac{\mathrm{d}R}{\mathrm{d}\theta} = \frac{\nu}{\sqrt{1 - \frac{\nu^2}{R^2}}}$$

$$\implies \nu\left[\sqrt{\frac{R^2}{\nu^2} - 1} + \sin^{-1}\left(\frac{\nu}{R}\right)\right]_{R_f}^{R} = \nu\left(\theta - \theta_f\right).$$

Define

$$g(R) = \sqrt{\frac{R^2}{\nu^2} - 1} + \sin^{-1}\left(\frac{\nu}{R}\right), \tag{11.21}$$

$$\implies \nu\left(g(R) - g(R_f)\right) = \nu\left(\theta - \theta_f\right)$$
$$\implies \theta\left(R; R_f, \theta_f\right) = g(R) - g(R_f) + \theta_f, \quad \theta_f \leq \theta < \pi. \tag{11.22}$$

Setting $\theta_f = 0$ in (11.22) (i.e., $\theta(R; R_f, 0)$) describes the equilibrium flow field for the *Game of Distance* (i.e., assuming the Defender can drive $\theta \to 0$ before the Attacker can reach the target). The curve in (11.22) is the involute of a circle of radius ν.

Up until now, θ has been assumed to be in the range $[0, \pi)$, however, the results apply to the range $(-\pi, 0]$ with some slight modification.

Lemma 11.1. *The surface*

$$\mathscr{D} \equiv \{\mathbf{x} \mid \theta = \pi\}, \tag{11.23}$$

is a DS (c.f. [130]) wherein the Defender can choose either $u_D = 1$ or $u_D = -1$ and both choices are optimal. Furthermore, when $\theta < 0$, the equilibrium controls are given by $u_D^* = -1$ and $\sin \psi^* < 0$.

Proof. By definition, points on a DS have two or more associated equilibrium trajectories which yield the same Value. It will shown that (11.23) is indeed a DS by constructing a pair of equilibrium trajectories that integrate back to the same point on $\mathscr{D}$. Consider an initial state on the DS, $\mathbf{x}_{\mathscr{D}} = (R_0, \pi) \in \mathscr{D}$. The system (11.19) describes the evolution of R and θ in backwards time assuming $\eta > 0$. Now, let $\mathbf{x}_f \equiv \left(R_f, \theta_f \right)$ (where $R_f > 1$, $\theta_f \geq 0$) be the terminal state, which, when integrated through the retrograde kinematics (11.19), yields the initial point $\mathbf{x}_{\mathscr{D}}$. A symmetric solution can be constructed by switching the sign of η and θ_f, then integrating the retrograde kinematics back to $\mathbf{x}_{\mathscr{D}}$. Now, let $\eta < 0$; then $\sin \psi^* < 0$ from (11.15), and $u_D^* = -1$ from (11.16). Substitution into the Hamiltonian at final time yields $\eta = -\nu R_f \sqrt{\frac{1}{R_f^2 - \nu^2}}$. Substituting all of these into the Hamiltonian at general time yields the same expression for σ_R as in (11.18). Then, from (11.2), the retrograde kinematics are

$$\mathring{R} = \nu \sqrt{1 - \frac{\nu^2}{R^2}}, \qquad \mathring{\theta} = \frac{\nu^2}{R^2} - 1.$$

Clearly, these are the same kinematics as in (11.19) except the sign of $\mathring{\theta}$ is reversed. These kinematics can be integrated back from the symmetric terminal point $\left(R_f, -\theta_f \right)$ to the point $(R_0, -\pi)$, which is equivalent to $\mathbf{x}_{\mathscr{D}}$. This pair of trajectories emanating (forward in time) from $\mathbf{x}_{\mathscr{D}}$ have the same Value for all terminal cost functionals of R_f and $|\theta_f|$. Note this method for proving the presence of a DS is similar to the one used for a problem with similar dynamics in [2]. □

As a consequence, Assumption 2 may be relaxed, and the state space may be expanded to $\theta \in [-\pi, \pi]$.

Theorem 11.1 (*Game of Distance* Solution). *The equilibrium state feedback control strategies for the Game of Distance are given by*

$$\psi^* = \text{sign}\,(\theta) \sin^{-1} \left(\frac{\nu}{R} \right), \qquad u_D^* = \text{sign}\,(\theta) . \tag{11.24}$$

The Value of the game is

$$V_d(R, \theta) = -R_f = -g^{-1}\left(g(R) - |\theta|\right). \tag{11.25}$$

Proof. The expression for ψ^* is obtained by substituting (11.17) and (11.18) into (11.15), taking into account the sign of θ (due to Lemma 11.1). Similarly, the Defender strategy is given by (11.16), accounting for Lemma 11.1. The corresponding form of (11.22) for the *Game of Distance* is

$$\theta\left(R; R_f\right) = g(R) - g(R_f). \tag{11.26}$$

Thus, (11.25) is obtained by rearranging this expression and solving for R_f, with $g(\cdot)$ defined as in (11.21). Because V_d is defined using the inverse of the function g, it is necessary to show that $g(R)$ is monotonic. Taking the derivative of (11.21) w.r.t. R gives

$$\frac{\mathrm{d}g}{\mathrm{d}R} = \frac{\sqrt{R^2 - \nu^2}}{\nu R}.$$

From Assumption 1, it must be that $0 < \nu < 1$, and from Assumption 3 it must be that $R > 1$ throughout the game. So we have $R > \nu$ and $R, \nu > 0$, which implies that $g(R)$ is monotonic. □

The Value function does not have a closed form analytic expression since g^{-1} cannot be expressed in closed form.

The limiting case for the *Game of Distance* is one in which $R_f \to 1$; thus the surface

$$\theta_{GoK}(R) = g(R) - g(1) \tag{11.27}$$

partitions the state space into regions of win for the Defender and Attacker, respectively,

$$\mathscr{R}_D = \{\mathbf{x} \mid |\theta| \le \theta_{GoK}(R)\} \tag{11.28}$$

$$\mathscr{R}_A = \{\mathbf{x} \mid |\theta| > \theta_{GoK}(R)\}. \tag{11.29}$$

Note that the value domain of $g(R)$ is $[g(1), \infty)$ since $R \ge 1$ and g is monotonic; from (11.28) $|\theta| \le \theta_{GoK}(R)$ in $\mathscr{R}_D$, so the argument to g^{-1}

in (11.25) is $g(R) - |\theta| \geq g(R) - (g(R) - g(1)) = g(1)$ which is in the value domain of $g(R)$.

11.2.2 *Attacker Win Scenario*

In the region of the state space in which the Attacker 'wins' (i.e., can reach $R = 1$ while avoiding $\theta = 0$), a *Game of Degree* is considered wherein the players max/min the terminal separation angle; this is referred to as the *Game of Angle*. The cost/payoff functional is given as

$$J = \Phi\left(\mathbf{x}_f, t_f\right) = \theta_f. \tag{11.30}$$

The Attacker seeks to maximize the terminal separation angle whereas the Defender seeks to minimize. Termination occurs when the Attacker penetrates the target circle,

$$\phi\left(\mathbf{x}_f, t_f\right) = R_f - 1 = 0. \tag{11.31}$$

Theorem 11.2 (*Game of Angle* Solution). *The equilibrium state feedback strategies for the Game of Angle match those of the Game of Distance, i.e., are given by* (11.24). *The Value function is given by*

$$V(R, \theta) = \theta_f = \theta - g(R) + g(1). \tag{11.32}$$

Proof. This proof is based upon showing satisfaction of the sufficient condition for equilibrium via substitution of the proposed equilibrium strategies and Value function into the HJI equation [130],

$$\min_{u_D} \max_{\psi} \left\{ l\left(\mathbf{x}, u_D, \psi, t\right) + \frac{\partial V}{\partial t} + V_{\mathbf{x}} \cdot f\left(\mathbf{x}, u_D, \psi, t\right) \right\}$$
$$= 0, \tag{11.33}$$

where $V_{\mathbf{x}}$ is the vector $\left[\frac{\partial V}{\partial R} \quad \frac{\partial V}{\partial \theta} \quad \frac{\partial V}{\partial \beta} \right]^{\top}$, and l represents an integral cost component. First, note that the cost, (11.30), has no integral component, and thus $l = 0$. Also, the proposed Value function, (11.32) is
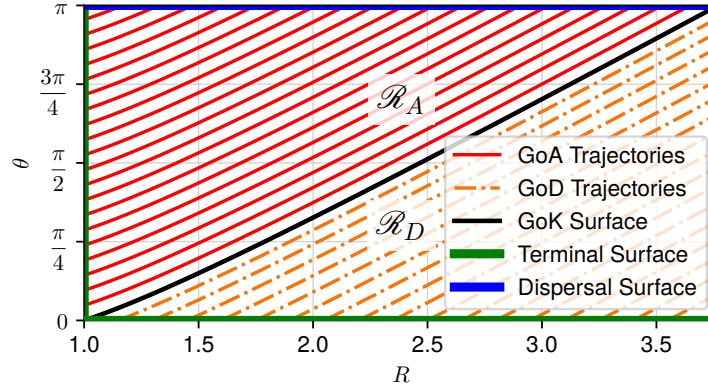
Figure 11.2: Full equilibrium flow field with $\nu = 0.8$

not an explicit function of time and thus $\frac{\partial V}{\partial t} = 0$. The vector $V_{\mathbf{x}}$ is obtained by differentiating (11.32) w.r.t. each state,

$$V_{\mathbf{x}} = \left[\begin{matrix} \frac{-\sqrt{R^2 - \nu^2}}{R\nu} & 1 & 0 \end{matrix} \right].$$

The (forward) equilibrium dynamics, f, are given by the negative of (11.19). Substituting all of these expressions into (11.33) gives

$$\frac{\partial V}{\partial R}\dot{R} + \frac{\partial V}{\partial \theta}\dot{\theta} =$$
$$\left(\frac{-\sqrt{R^2 - \nu^2}}{R\nu} \right)\left(-\nu\sqrt{1 - \frac{\nu^2}{R^2}} \right) + \left(\frac{\nu^2}{R^2 - 1} \right) = 0.$$

The proposed Value function is continuous and continuously differentiable (except on the DS, $\mathscr{D}$), and it satisfies the HJI hyperbolic PDE. $\square$

Remark 16. Theorem 11.1 (as well as Theorems 11.3 and 11.4) can be verified in a similar fashion (i.e., by substituting the respective Value functions into the HJI to show it is satisfied). The analysis would be nearly identical to the above proof and is therefore omitted.

11.2.3   *Full Equilibrium Flow Field*

With the analysis in Sections 11.2.1 and 11.2.2, the entire (usable) state space can be filled with equilibrium trajectories. Figure 11.2 shows (11.22) and (11.26) in the Attacker win and lose regions, respectively.

Lemma 11.2. *The Attacker's equilibrium trajectory is a straight line in the inertial (non-rotating) (x, y)-plane.*

Proof. Consider Fig. 11.1 which shows the Attacker's heading angle, $\tilde{\psi}$, w.r.t. the inertial (x, y)-plane. The following relation holds

$$\tilde{\psi} = \beta + \theta + \pi - \psi$$

Thus, the time derivative of the global Attacker heading angle is given as

$$\dot{\tilde{\psi}} = \dot{\beta} + \dot{\theta} - \dot{\psi}$$

Substituting (11.24) and (11.19) into the above gives

$$
\begin{aligned}
\dot{\tilde{\psi}} &= 1 + \frac{\nu^2}{R^2} - 1 - \frac{\partial}{\partial t} \sin^{-1}\left(\frac{\nu}{R}\right) \\
&= \frac{\nu^2}{R^2} - \left(\frac{-1}{\sqrt{1 - \frac{\nu^2}{R^2}}}\right)\left(\frac{\nu}{R^2}\right)\dot{R} \\
&= \frac{\nu^2}{R^2} + \left(\frac{1}{\sqrt{1 - \frac{\nu^2}{R^2}}}\right)\left(\frac{\nu}{R^2}\right)\left(-\nu\sqrt{1 - \frac{\nu^2}{R^2}}\right) \\
&= 0.
\end{aligned}
$$

Because $\dot{\tilde{\psi}} = 0$, the global Attacker heading angle is constant, and thus the Attacker's path is a straight-line in the inertial (x, y)-plane. □

11.2.4 *Alternative Attacker Win Scenario*

Depending on the particular physical application or interpretation of the scenario, the Attacker may be interested in penetrating the target circle in minimum time. For example, if the Attacker is a munition of some kind, it may not matter how far or close the Defender is at the time of penetration. Formulating a game of min / max terminal angular separation, on the other hand, may make sense when the Attacker is some kind of vehicle or person who seeks to intrude inside the target circle while avoiding, as much as possible, coming into contact with the Defender.

The cost functional considered here is

$$J = \Phi\left(\mathbf{x}_f, t_f\right) = -t_f. \tag{11.34}$$

Once again, the negative sign is used to adhere to the convention established in previous sections in which the Attacker and Defender are the maximizer and minimizer, respectively. Since penetration is assumed to occur, the termination condition is given by (11.31), i.e., when the Attacker reaches the target, and it must be ensured that $\theta(t) \neq 0$ *en route*. This scenario is referred to as the *Game of Min Time*.

Proposition 11.1. *For the zero-sum differential game whose cost is given by* (11.34), *the Defender's equilibrium control is*

$$u_D^* = \text{sign}\left(\theta\right). \tag{11.35}$$

Proof. First, note the proposed Defender control is the same as in both of the *Games of Degree* considered previously (c.f. Theorems 11.1 and 11.2). The Defender's essential goal in all of these scenarios is to align with the Attacker, if possible, or otherwise impede the Attacker somehow. Likewise, the Attacker must avoid alignment with the Defender in order to achieve its objectives. Thus the Defender is always interested in driving $\theta \to 0$ as quickly as possible. From (11.2), the fastest way to achieve this is traversing, at maximum speed, in the direction of the Attacker, i.e., by implementing (11.35). □

Lemma 11.3. *For any scenario in which the Defender implements $u_D = \text{sign}\left(\theta\right)$, it is necessary for $\mathbf{x} \in \mathscr{R}_A$ for all $t \in \left[0, t_f\right]$ in order for the Attacker to achieve penetration, i.e., $R_f = 1$ with $\theta > 0$ for all $t \in \left[0, t_f\right]$.*

Proof. This Defender control is the equilibrium control for the *Game of Distance*. In the *Game of Distance*, which is played when $\mathbf{x} \notin \mathscr{R}_A$, the Attacker seeks to come as close to the target circle before alignment with the Defender occurs. The limiting case is when the Attacker reaches the target circle at the exact moment alignment occurs. If it had gotten there sooner, then it must have been the case that $\mathbf{x} \in \mathscr{R}_A$ since the equilibrium controls are the same as in the *Game of Angle*. In general,

$R_f > 1$, which means that the Defender achieves alignment before the Attacker achieves penetration *under the best possible Attacker control.* □

Lemma 11.3 expresses the necessary condition for the Attacker to penetrate the target circle, which is applicable to any scenario in which the Defender seeks to align with the Attacker. Thus the *Game of Min Time* takes place in $\mathscr{R}_A$. Let us focus on the case where $\theta \in [0, \pi]$. The terminal condition is

$$\phi\left(\mathbf{x}_f, t_f\right) = R_f - 1 = 0, \ \theta \in [0, \pi]. \tag{11.36}$$

Thus the terminal surface is the zero-level set of ϕ and is left-discontinuous at $\theta = 0$. Additionally, the boundary of the terminal surface, $(R, \theta) = (1, 0)$, lies on the boundary of the state space, $\partial\mathscr{R}_A$.

The Hamiltonian is

$$\mathscr{H} = -\lambda_R v \cos \psi + \lambda_\theta \left(\frac{v}{R} \sin \psi - 1\right) + \lambda_\beta, \tag{11.37}$$

where $\lambda \equiv \begin{bmatrix} \lambda_R & \lambda_\theta & \lambda_\beta \end{bmatrix}$. Both λ_θ and λ_β are constant since $\frac{\partial\mathscr{H}}{\partial\theta} = 0$ and $\frac{\partial\mathscr{H}}{\partial\beta} = 0$, respectively. In general, the transversality condition [40] is

$$\lambda^\top\left(t_f\right) = \frac{\partial\Phi}{\partial\mathbf{x}_f} + \mu\frac{\partial\phi}{\partial\mathbf{x}_f}. \tag{11.38}$$

The term $\frac{\partial\phi}{\partial\theta}$ is well-defined when $\theta \in (0, \pi)$, but it is undefined at the "corner point" $(R, \theta) = (1, 0)$. The former case is treated first.

11.2.5   *General Case*

Lemma 11.4. *For the zero-sum differential game whose cost is given by* (11.34), *the equilibrium Attacker control is*

$$\psi^* = 0, \qquad \forall \mathbf{x} \ s.t. \ \theta > \frac{R-1}{v}. \tag{11.39}$$

Proof. Specializing (11.38) gives

$$\lambda^\top(t_f) = 0 + \mu\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \tag{11.40}$$

thus λ_θ, $\lambda_\beta = 0$ for all $t \in [0, t_f]$. Note that it has been assumed that $\theta_f > 0$ which implies that $\frac{\partial \phi}{\partial \theta_f}$ exists and equals 0. Substituting into the Hamiltonian, (11.37), gives

$$\mathscr{H} = -\lambda_R \nu \cos \psi,$$

which is maximized for $\cos \psi^* = -\operatorname{sign}(\lambda_R) = \pm 1$. It is obvious that the Attacker must run *toward* the target circle, hence (11.39) holds.

Now, it must ensured that the assumption $\theta_f > 0$ is valid. Under the equilibrium control strategies $u_D^* = \operatorname{sign}(\theta)$ and $\psi^* = 0$ the R and θ dynamics are

$$\dot{R} = -\nu, \qquad \dot{\theta} = -1, \tag{11.41}$$

and thus $\frac{\mathrm{d}\theta}{\mathrm{d}R} = \frac{1}{\nu}$, implying that the trajectories are straight lines in the (R, θ) plane. Furthermore, the unconstrained equilibrium flowfield is

$$\theta(R) = \frac{1}{\nu}(R - R_0) + \theta_0. \tag{11.42}$$

The critical case occurs when the constraint activates at the precise moment that the Attacker reaches the target (i.e., $\theta_f = 0$). The time to traverse from their initial positions must be equal, giving

$$\theta_c(R) = \frac{R - 1}{\nu}. \tag{11.43}$$

If $\theta > \theta_c$, then $\theta_f > 0$ under equilibrium play, hence the specification in (11.39). $\qquad \square$

However, if $\theta < \theta_c$ and the Attacker aims at the circle center, the Defender can drive $\theta \to 0$ before the Attacker reaches the target.

11.2.6 Corner Case

Now the case wherein the game terminates on "corner point" $(R, \theta) = (1, 0)$ is treated. Recall the fact that $\frac{\partial \phi}{\partial \theta_f}$ is undefined when $\theta_f = 0$ which results in λ_{θ_f} being free. The consequence is that the incoming equilibrium trajectory to the corner point is not unique, unlike elsewhere in the terminal surface. Therefore, a family of trajectories, beginning from

a range of initial conditions all terminate at $(R, \theta) = (1, 0)$. A similar situation arises in pursuit-evasion scenarios with a finite capture radius [89, 271] (see, also, Chapter 6). For notational convenience, let $\lambda_{R_f} \equiv \mu$ and $\lambda_{\theta_f} \equiv \eta$; $\lambda_{\beta_f} = 0$, as before.

Lemma 11.5. *For the zero-sum differential game whose cost is given by* (11.34) *the equilibrium Attacker control is*

$$\psi^* = \sin^{-1}\left(\frac{\kappa}{R}\right), \ \kappa \in [0, \nu],$$
$$\forall \mathbf{x} \in \mathscr{R}_A \ s.t. \ \theta \le \frac{R-1}{\nu}, \tag{11.44}$$

where κ satisfies

$$0 = \frac{1}{\nu}\left(\sqrt{R^2 - \kappa^2} - \sqrt{1 - \kappa^2}\right) - \theta - \sin^{-1}\kappa + \sin^{-1}\left(\frac{\kappa}{R}\right). \tag{11.45}$$

The Attacker trajectory is a straight line terminating at $(R, \theta) = (1, 0)$.

Proof. The Hamiltonian, (11.37), evaluated at final time is

$$\mathscr{H}_f = -\mu\nu\cos\psi_f + \eta\left(\nu\sin\psi_f - 1\right). \tag{11.46}$$

The equilibrium Attacker heading must maximize $\mathscr{H}_f$, thus

$$\cos\psi_f^* = \frac{-\mu}{\sqrt{\mu^2 + \eta^2}}, \qquad \sin\psi_f^* = \frac{\eta}{\sqrt{\mu^2 + \eta^2}}. \tag{11.47}$$

Substituting this terminal Attacker heading back into (11.46) gives

$$\mathscr{H}_f = \nu\sqrt{\mu^2 + \eta^2} - \eta. \tag{11.48}$$

The terminal Hamiltonian value is also specified by [40]

$$\mathscr{H}_f = -\frac{\partial\Phi}{\partial t_f} - \mu\frac{\partial\phi}{\partial t_f} = -(-1) - 0 = 1 \tag{11.49}$$

Substituting back into (11.48) and rearranging gives

$$\sin\psi_f^* = \frac{\nu\eta}{1 + \eta}. \tag{11.50}$$

Define, for convenience, $\kappa \equiv \sin \psi_f$. The value of κ is bounded,

$$\kappa \in [0, \nu]. \tag{11.51}$$

The lower bound is due to the fact that A ought not aim *towards* D; the upper bound is due to the fact that $\kappa > \nu$ would immediately push the state of the system out of $\mathscr{R}_A$.

Because the system is time-autonomous, the value of the Hamiltonian is constant, i.e., $\mathscr{H}(t) = 1$ for all $t \in \left[0, t_f\right]$. Rewriting (11.46) and (11.47) at general time and solving for λ_R^2 gives

$$\lambda_R^2 = \left(\frac{1 + \eta}{\nu}\right)^2 - \frac{\eta^2}{R^2}. \tag{11.52}$$

Similarly, the Attacker heading at general time is

$$\begin{aligned}
\sin \psi &= \frac{\eta}{R\sqrt{\lambda_R^2 + \frac{\eta^2}{R^2}}} \\
&= \frac{\nu \eta}{R(1 + \eta)} = \frac{\sin \psi_f}{R} = \frac{\kappa}{R}
\end{aligned} \tag{11.53}$$

Therefore, Lemma 11.2 holds here as well since κ is a constant; that is, the Attacker's path is a straight line in the non-rotating (x, y)-frame.

Concerning the determination of κ for a general position $R > 1$ and $\theta \in [\theta_{GoK}, \theta_c]$, it is useful to consider the geometry. Let the point I be the point on the target circle in which A will terminate; by construction, this must be the point D terminates as well. A right triangle is formed by $\triangle ATO$ where T is the tangent point of the extension of A's trajectory on a circle of radius κ, and O is the target circle's center. The hypotenuse of $\triangle ATO$ is R. Another right triangle is formed by $\triangle ITO$; its hypotenuse is 1. See Fig. 11.3 for a representation of the geometry. The distance traveled by the Attacker is

$$\begin{aligned}
\overline{AI} &= \overline{AT} - \overline{TI} \\
&= \sqrt{R^2 - \kappa^2} - \sqrt{1 - \kappa^2}.
\end{aligned} \tag{11.54}$$

The Defender must cover an angular distance θ as well as the circular sector between A and I. Define $\rho \equiv \angle AOI$, which is given by

$$\rho = \sin^{-1} \kappa - \sin^{-1}\left(\frac{\kappa}{R}\right). \tag{11.55}$$

Define m as the difference in Attacker and Defender travel times to the point I:

$$m(\kappa) = \frac{1}{\nu}\overline{AI} - (\theta + \rho). \tag{11.56}$$

Then (11.45) is obtained by substituting in (11.54) and (11.55) and setting equal to zero, which represents simultaneous arrival to the point I. The value of κ for which this occurs may be obtained numerically. $\quad\square$
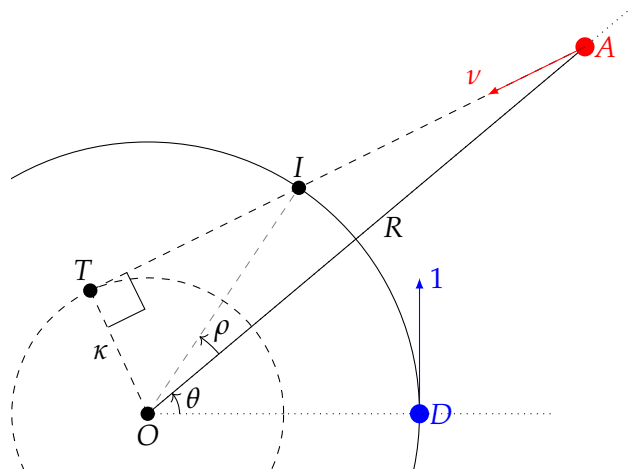
Figure 11.3: Schematic of the *Game of Min Time* scenario in which A takes an "evasive" path in order to arrive at the point I simultaneously with D.

In the *Game of Min Time* nothing is gained by increasing θ_f; the Attacker either heads directly towards the Target without regard for the Defender or performs the minimum "evasion" necessary to reach the Target, whereupon $\theta_f = 0$. The equilibrium kinematics may obtained by substituting (11.44) into (11.2). A closed form expression of the equilibrium trajectory through the (R, θ) space is obtained via a process similar to (11.19)–(11.22) in Section 11.2.1.2:

$$\theta(R; \kappa) = \frac{\sqrt{R^2 - \kappa^2}}{\nu} + \sin^{-1}\frac{\kappa}{R} - \frac{\sqrt{1 - \kappa^2}}{\nu} - \sin^{-1}\kappa \tag{11.57}$$

The full solution of the *Game of Min Time* is depicted in Fig. 11.4.
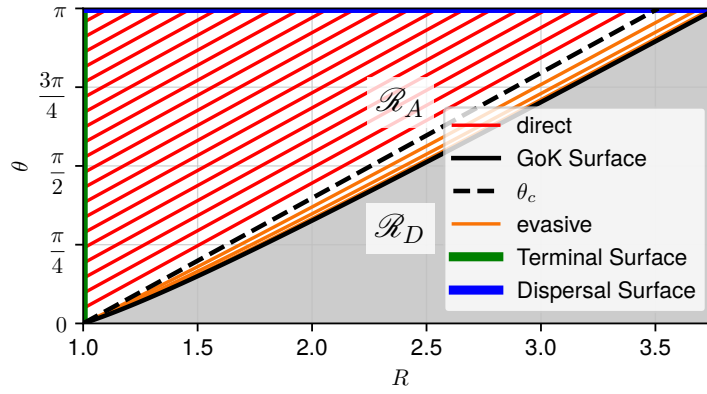
Figure 11.4: Equilibrium flowfield for the *Game of Min Time* with $\nu = 0.8$. The "direct" trajectories are straight lines with slope $\frac{1}{\nu}$. The "evasive" trajectories are described by (11.57) with various κ. The dashed black line is the critical trajectory described by (11.43).

11.3   TWO DEFENDERS

In this section, the circular target guarding game with two Defenders, D_1 and D_2, is considered with the following assumption:

Assumption 4. The two Defenders share the same maximum speed: $v_{D_{1_{max}}} = v_{D_{2_{max}}} = v_{D_{max}}$.
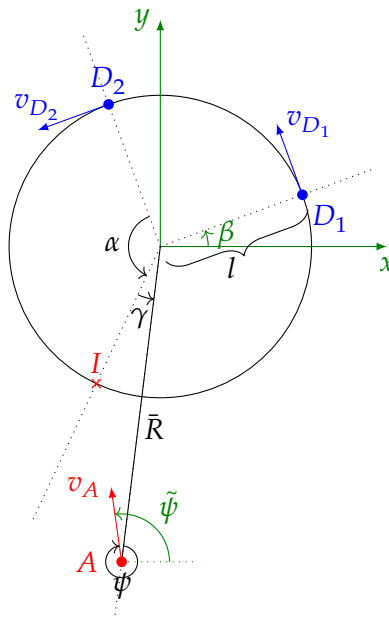
Figure 11.5: Circular perimeter patrol with two Defenders and one Attacker.

The scenario is depicted in Fig. 11.5, and the (nondimensional) kinematics of the system are given as

$$f(\mathbf{x}, \mathbf{u}, t) = \dot{\mathbf{x}} = \begin{bmatrix} \dot{R} \\ \dot{\gamma} \\ \dot{\alpha} \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} -v\cos\psi \\ \frac{v}{R}\sin\psi - \frac{1}{2}\left(u_{D_1} + u_{D_2}\right) \\ \frac{1}{2}\left(u_{D_1} - u_{D_2}\right) \\ u_{D_1} \end{bmatrix}. \tag{11.58}$$

The angle α is measured from D_2 to the angular bisector (on the side of A) of the positions of D_1 and D_2. Similarly, the angle γ is measured as A's angular offset w.r.t. this bisector.

Assumption 5. The relative angular position of the Attacker is bounded such that $-\alpha \leq \gamma \leq \alpha$.

Although Assumption 5 has been imposed, it is of little consequence since the forthcoming solution would still apply for γ outside this range by, for example, switching the designation of D_1 and D_2. Just as in the analysis of the one-on-one game, there are three "games" or questions of interest: 1) can the Attacker reach the target (the *Game of Kind*), 2) what is the equilibrium terminal angular separation between the Attacker and the closest Defender (the *Game of Angle*), and 3) what is the equilibrium terminal distance from the target center (the *Game of Distance*).

Note that the rotation of the system w.r.t. the global x-axis, β, has no effect on the optimality of the trajectories as in the one-on-one analysis and is therefore omitted in the following.

11.3.1 *Game of Degree When Attacker Wins*

Here, the *Game of Angle*, which applies to the scenario when the Attacker is able to reach the target ($R_f = 1$), is treated. The cost functional is given as

$$J = \Phi\left(\mathbf{x}_f, t_f\right) = \alpha_f - |\gamma_f|, \tag{11.59}$$

and the goal is to obtain the Value of the game:

$$V(\mathbf{x}) = \min_{\mathbf{u}(\cdot)} \max_{\psi(\cdot)} J = \max_{\psi(\cdot)} \min_{\mathbf{u}(\cdot)} J. \tag{11.60}$$

This game terminates when the following condition is satisfied

$$\phi\left(\mathbf{x}_f, t_f\right) = R_f - 1 = 0. \tag{11.61}$$

11.3.1.1   *First Order Necessary Conditions for Optimality*

First, form the Hamiltonian as

$$\mathcal{H} = -\lambda_R v \cos \psi + \lambda_\gamma \left(\frac{v}{R} \sin \psi - \frac{1}{2} \left(u_{D_1} + u_{D_2} \right) \right)$$
$$+ \lambda_\alpha \frac{1}{2} \left(u_{D_1} - u_{D_2} \right). \tag{11.62}$$

The equilibrium adjoint dynamics obey [40]

$$\dot{\lambda}_R = -\frac{\partial \mathcal{H}}{\partial R} = \lambda_\gamma \frac{v}{R^2} \sin \psi, \tag{11.63}$$

$$\dot{\lambda}_\gamma = -\frac{\partial \mathcal{H}}{\partial \gamma} = 0, \tag{11.64}$$

$$\dot{\lambda}_\alpha = -\frac{\partial \mathcal{H}}{\partial \alpha} = 0. \tag{11.65}$$

From the transversality condition [40], the equilibrium terminal adjoint values satisfy

$$\boldsymbol{\lambda}^\top (t_f) = \frac{\partial \Phi}{\partial \mathbf{x}_f} + \mu \frac{\partial \phi}{\partial \mathbf{x}_f}$$
$$\lambda_R (t_f) = \mu,$$
$$\implies \lambda_\gamma (t_f) = -\operatorname{sign}(\gamma_f), \tag{11.66}$$
$$\lambda_\alpha (t_f) = 1.$$

Because $\dot{\lambda}_\gamma = \dot{\lambda}_\alpha = 0$ it must be that $\lambda_\gamma(t) = -\operatorname{sign}(\gamma_f)$ and $\lambda_\alpha(t) = 1$ for all $t \in \left[t_0, t_f\right]$.

11.3.1.2   *Solution Characteristics*

Lemma 11.6. *For games terminating with $\gamma_f \neq 0$, the game's Value function and optimal strategies are that of the one-on-one game: $V = \alpha - |\gamma| - g(R) + g(1)$, and $\psi^* = \sin^{-1}\left(-\frac{v}{R}\operatorname{sign}(\gamma_f)\right)$. The second Defender is redundant.*

Proof. Suppose that $\gamma_f < 0$; substituting the corresponding λ_γ and λ_α values into (11.62) gives

$$\mathscr{H} = -\lambda_R v \cos \psi - \frac{v}{R} \sin \psi - u_{D_2}. \tag{11.67}$$

Note that u_{D_1} does not appear in (11.67) and thus D_1 has no effect on the optimality of the trajectory and is therefore redundant. Again, since the final time is free, the Hamiltonian, at terminal time, is subject to (11.14) [40]; that is, $\mathscr{H}(t_f) = 0$. Since (11.58) are autonomous, it must be that $\mathscr{H}(t) = 0$ for all $t \in [t_0, t_f]$. Therefore, (11.67) is identical to the Hamiltonian for the one-on-one case between the Attacker and D_2. Furthermore, the terminal condition is the same, and the cost functional is identical since $\theta = \alpha - |\gamma| = J$, in this case. Thus, the Value function for the one-on-one case, (11.32), and the equilibrium Attacker heading control, (11.24) are the solution for this game (making the appropriate substitution of $\theta = \alpha - |\gamma|$). The $-$ sign in the ψ^* expression, in this case, accounts for the case when $\gamma_f > 0$ in which the game plays out between the Attacker and D_1, by symmetry. In that case, the scenario is a mirror image of Fig. 11.1 and the sign of u_{D_1} is reversed (i.e., D_1 moves clockwise) as is the sign of $\sin \psi^*$. $\square$

Since $\gamma_f \neq 0$ corresponds to either one-on-one game, special attention is given to the case when $\gamma_f = 0$. When $\gamma_f = 0$, the Attacker terminates at a position which is equidistant from the two defenders. Note that, according to (11.66), $\lambda_\gamma(t_f) = \lambda_{\gamma_f}$ is undefined when $\gamma_f = 0$. As before, the Defenders seek to minimize the Hamiltonian, (11.62):

$$\begin{aligned} u_{D_1}^*, u_{D_2}^* &= \underset{u_{D_1}, u_{D_2}}{\arg\min} \mathscr{H} \\ &= \underset{u_{D_1}, u_{D_2}}{\arg\min} u_{D_1} \left(1 - \lambda_\gamma\right) + u_{D_2} \left(-1 - \lambda_\gamma\right). \end{aligned} \tag{11.68}$$

Now, according to (11.68), if $\lambda_\gamma > 1$ or $\lambda_\gamma < -1$ then $u_{D_1}^* = u_{D_2}^*$ which means the Defenders should move in the same direction. However, if this were the case then $\dot{\alpha} = 0$ which is clearly undesirable since α appears in the cost, J. Thus the value of λ_γ is bounded:

$$-1 \leq \lambda_\gamma \leq 1. \tag{11.69}$$

By inspection, it is clear that the Defenders should seek to minimize $\dot{\alpha}$ which occurs for

$$u_{D_1}^* = -1, \quad u_{D_2}^* = 1. \tag{11.70}$$

substituting in (11.66) and (11.70) into (11.62) leads to an expression for λ_R:

$$\mathcal{H}(t) = 0 = \nu\sqrt{\lambda_R^2 + \frac{\lambda_\gamma^2}{R^2}} - 1$$

$$\implies \lambda_R = \pm\sqrt{\frac{1}{\nu^2} - \frac{\lambda_\gamma^2}{R^2}}.$$

Since $\dot{R}_f \propto \cos\psi_f \propto \nu$ it must be that $\lambda_{R_f}, \nu < 0$ in order for the state of the system to penetrate the boundary. In order to maximize the Hamiltonian, it must be that $\sin\psi^* \propto \lambda_\gamma$; thus, from (11.63), $\dot{\lambda}_R(t) < 0$ for all $t \in \left[t_0, t_f\right]$. Therefore, $\lambda_R(t) < 0$ for all $t \in \left[t_0, t_f\right]$, which leads to

$$\lambda_R = -\sqrt{\frac{1}{\nu^2} - \frac{\lambda_\gamma^2}{R^2}}. \tag{11.71}$$

Lemma 11.7. *For games terminating with $\gamma_f = 0$, the equilibrium heading angle is*

$$\psi^* = \sin^{-1}\left(\lambda_\gamma \frac{\nu}{R}\right), \tag{11.72}$$

and is bounded by $-\sin^{-1}\left(\frac{\nu}{R}\right) \leq \psi^* \leq \sin^{-1}\left(\frac{\nu}{R}\right)$.

Proof. Substituting (11.71) with (11.70) into (11.62) gives

$$\mathcal{H} = 0 = -\nu\sqrt{\frac{1}{\nu^2} - \frac{\lambda_\gamma^2}{R^2}}\cos\psi + \lambda_\gamma\frac{\nu}{R}\sin\psi. \tag{11.73}$$

The Attacker seeks to maximize the Hamiltonian, and thus

$$\cos\psi^* = -\sqrt{1 - \frac{\nu^2\lambda_\gamma^2}{R^2}}, \quad \sin\psi^* = \lambda_\gamma\frac{\nu}{R}, \tag{11.74}$$

and $-1 \leq \lambda_\gamma \leq 1$ according to (11.69), hence $-\frac{\nu}{R} \leq \sin\psi^* \leq \frac{\nu}{R}$.   □

Lemma 11.8. *The trajectories corresponding to $\lambda_\gamma = \pm 1$ separate the state space into regions of asymmetric termination ($\gamma_f \neq 0$) and symmetric termination ($\gamma_f = 0$).*

Proof. Suppose $\lambda_\gamma = 1$, then the Attacker's equilibrium strategy is identical to the one-on-one game with D_2 (c.f. (11.24)). The trajectory is a straight line in the global (x, y)-frame since the one-on-one game Attacker trajectories are straight (due to Lemma 11.2). Trajectories with $\lambda_\gamma < 1$ lie on one side of this surface and one-on-one trajectories (against D_2) lie on the other side. □

Lemma 11.9. *Attacker trajectories resulting in symmetric termination ($\gamma_f = 0$) are straight lines in the (x, y)-plane terminating at a point I, where*

$$I = \begin{bmatrix} I_x \\ I_y \end{bmatrix} = \begin{bmatrix} \cos{(\beta_0 - \alpha_0)} \\ \sin{(\beta_0 - \alpha_0)} \end{bmatrix}. \tag{11.75}$$

Proof. Just as in Lemma 11.2, the Attacker trajectory is shown to be a straight line in the inertial frame via direct substitution of the equilibrium strategies. Consider Fig. 11.5 which shows the Attacker's heading angle $\tilde{\psi}$, w.r.t. the inertial (x, y)-plane. It is expressed

$$\tilde{\psi} = \beta + (2\pi - 2\alpha) + \alpha + \gamma - \psi,$$

and its time derivative is

$$\dot{\tilde{\psi}} = \dot{\beta} - \dot{\alpha} + \dot{\gamma} - \dot{\psi}.$$

Substitution of the kinematics, (11.58), and the equilibrium controls, (11.70) and (11.72), gives

$$\dot{\tilde{\psi}} = u_{D_1} - \frac{1}{2}\left(u_{D_1} - u_{D_2}\right) + \frac{v}{R}\sin\psi -$$
$$\frac{1}{2}\left(u_{D_1} + u_{D_2}\right) - \frac{\partial}{\partial t}\sin^{-1}\lambda_\gamma\frac{v}{R}$$
$$= -1 + 1 + \frac{\lambda_\gamma v^2}{R^2} - 0 - \left(\frac{-\lambda_\gamma v}{R^2\sqrt{1 - \frac{\lambda_\gamma^2 v^2}{R^2}}}\right)(-v\cos\psi)$$
$$= \frac{\lambda_\gamma v^2}{R^2} - \frac{\lambda_\gamma v^2}{R^2} = 0$$

Since the Attacker heading in the inertial (x, y)-plane is constant, the Attacker path is a straight line. For symmetric termination, the state of the system lies at $R = 1$ and $\gamma = 0$. The $\gamma = 0$ angle corresponds

to $\beta - \alpha$. Because $u_{D_1}^* = -1$ and $u_{D_2}^* = 1$ (due to (11.74)) it must be that $\dot{\alpha} = -1 = \dot{\beta}$ and thus the angle $\beta - \alpha$ is invariant in the global (x, y)-plane. $\qquad\square$

Lemma 11.10. *For symmetric termination ($\gamma_f = 0$), the separating surface of the Game of Kind in the global (x, y)-plane is given by a circular arc centered I with radius $\nu\alpha_0$ whose bounds are defined by $\sin^{-1}(-\nu)$ and $\sin^{-1}(\nu)$ relative to the $\gamma = 0$ axis.*

Proof. Symmetric termination trajectories terminate at I, defined and according to Lemma 11.9. The limiting case occurs when the Attacker reaches the target circle at the exact moment in which the Defenders reach I (i.e. $\alpha_f \to 0$). Due to (11.70), it must be that $\dot{\alpha} = -1$. Therefore, the Defenders reach $\alpha = 0$ in α_0 time. Symmetric termination trajectories may thus extend from I for a maximum distance of $\nu\alpha_0$; beyond this distance, the Attacker cannot reach the target. The Attacker trajectories are straight, also due to Lemma 11.9, thus the *Game of Kind* surface is a circular arc. The bounds of the circular arc are given directly by the range of ψ_f^* which is obtained by substituting $R = 1$ into (11.72) and applying the bounds stated in Lemma 11.7. $\qquad\square$

The regions $\mathscr{R}_{A_1}$ and $\mathscr{R}_{A_2}$ are the sets of states for which the game terminates with $\gamma_f > 0$ (one-on-one with D_1) and $\gamma_f < 0$ (one-on-one with D_2), respectively (c.f. Lemma 11.6). Similarly, the region $\mathscr{R}_{A_{1,2}}$ is the set of states for which the game terminates with $\gamma_f = 0$ and is completely specified by Lemmas 11.8–11.10. The polar distance at which the *Game of Kind* surface switches from the one-on-one surface, governed by (11.27), and the two-on-one surface, described in Lemma 11.10, is given by

$$R_s = +\sqrt{\nu^2\alpha^2 + 1 + 2\nu\alpha\sqrt{1 - \nu^2}}, \qquad (11.76)$$

which is derived from the Law of Cosines (see Fig. 11.7).

Theorem 11.3. *In the region $\mathscr{R}_{A_{1,2}}$, the equilibrium Attacker heading angle is given by*

$$\psi^* = \sin^{-1}\left(\frac{\sin\gamma}{p}\right), \qquad (11.77)$$

and the associated Value function is

$$V(\mathbf{x}) = \alpha_f = \alpha - \frac{p}{v}, \tag{11.78}$$

where

$$p = +\sqrt{R^2 + 1 - 2R\cos\gamma}.$$

Proof. Consider the triangle formed by the Attacker's position, the target circle center, and the point I as defined in (11.75). By construction, the Attacker starts in $\mathscr{R}_{A_{1,2}}$ and its equilibrium trajectory must terminate at I due to Lemma 11.9. Let the distance traveled from A_0 to I be p, which can be obtained from the Law of Cosines (as defined above). Then, (11.77) can be obtained from the Law of Sines. The time taken to traverse this path is p/v, and $\dot{\alpha} = -1$ (due to (11.70)), thus (11.78) follows. $\qquad\qquad\square$

11.3.2 *Game of Degree When Attacker Loses*

In this section, the *Game of Distance*, which applies to the scenario when A is not able to reach the target before one or both Defenders can align with A (i.e. $\alpha - |\gamma| = 0$), is treated. The cost functional is the same as in the one-on-one case, i.e., (11.3). This game terminates when the following condition is satisfied

$$\phi_d\left(\mathbf{x}_f, t_f\right) = \alpha_f - |\gamma_f| = 0. \tag{11.79}$$

11.3.2.1 *First Order Necessary Conditions*

The Hamiltonian is

$$\mathcal{H}_d = -\sigma_R v \cos\psi + \sigma_\gamma\left(\frac{v}{R}\sin\psi - \frac{1}{2}\left(u_{D_1} + u_{D_2}\right)\right) + \tag{11.80}$$

$$\sigma_\alpha\frac{1}{2}\left(u_{D_1} - u_{D_2}\right),$$

and thus the equilibrium adjoint dynamics are

$$\dot{\sigma}_R = -\frac{\partial \mathcal{H}_d}{\partial R} = \sigma_\gamma \frac{v}{R^2} \sin \psi \tag{11.81}$$

$$\dot{\sigma}_\gamma = -\frac{\partial \mathcal{H}_d}{\partial \gamma} = 0 \tag{11.82}$$

$$\dot{\sigma}_\alpha = -\frac{\partial \mathcal{H}_d}{\partial \alpha} = 0. \tag{11.83}$$

From the transversality condition [40], the terminal adjoint values are

$$\sigma^\top(t_f) = \frac{\partial \Phi_d}{\partial \mathbf{x}_f} + \eta \frac{\partial \phi_d}{\partial \mathbf{x}_f} \tag{11.84}$$

$$= \begin{bmatrix} -1 & 0 & 0 \end{bmatrix} + \eta \begin{bmatrix} 0 & \pm 1 & 1 \end{bmatrix}. \tag{11.85}$$

When $\gamma_f = 0$, however, the derivative $\frac{\partial \phi_d}{\partial \gamma_f}$, and thus σ_{γ_f}, is not defined. Evaluating (11.80) at final time and substituting in the terminal adjoint values gives

$$\mathcal{H}_d(t_f) = v \cos \psi_f + \sigma_{\gamma_f} \left(\frac{v}{R_f} \sin \psi_f - \frac{1}{2} \left(u_{D_1} + u_{D_2} \right) \right) \tag{11.86}$$

$$+ \frac{\eta}{2} \left(u_{D_1} - u_{D_2} \right)$$

The Hamiltonian at terminal time is given by [40]

$$\mathcal{H}_d(t_f) = -\frac{\partial \Phi_d}{\partial t_f} - \eta \frac{\partial \phi_d}{\partial t_f} = 0. \tag{11.87}$$

11.3.2.2   *Solution Characteristics*

Lemma 11.11. *For games terminating with $\gamma_f \neq 0$, the game's Value function and optimal strategies correspond to the one-on-one game (c.f. Theorem 11.1). The second Defender is redundant.*

Proof. The proof is similar to that of Lemma 11.6 in that the Hamiltonian is formed and a particular sign of γ_f is assumed, which results in reduction to the one-on-one Hamiltonian with identical cost and terminal boundary condition. If, for example, $\gamma_f < 0$ then $\sigma_{\gamma_f} = \eta$ and is σ_γ is constant since $\dot{\sigma}_\gamma = 0$. The Hamiltonian would be reduced to

$$\mathcal{H}_d = -\sigma_R v \cos \psi + \eta \left(\frac{v}{R} \sin \psi - u_{D_2} \right),$$

which matches exactly with the one-Attacker one-Defender Hamilto-
nian, (11.7). The cost functional (based on terminal distance, (11.3)) is
the same and thus the two-Defender scenario reduces to the one-De-
fender scenario whenever $\gamma_f \neq 0$, by symmetry. $\qquad\square$

Lemma 11.12. *For games terminating with $\gamma_f = 0$, the equilibrium Attacker
heading angle is*

$$\psi^* = \sin^{-1}\left(\chi \frac{v}{R}\right), \qquad \chi \in [-1, 1]. \tag{11.88}$$

Proof. The proof is similar to that of Lemma 11.7, but with the associ-
ated first order necessary conditions for the *Game of Distance* from the
previous section. Now, if A and D_2 were to play the one-on-one *Game
of Distance* D_2 would move counterclockwise, i.e., $u_{D_2} = 1$. The pres-
ence of D_1 ought not change the control of D_2 – counterclockwise is
still the direction which closes the angular gap the between D_2 and A
the fastest. Therefore, let $u^*_{D_{1_f}} = -1$ and $u^*_{D_{2_f}} = 1$. As before, the De-
fenders must minimize the Hamiltonian, including at final time. Thus
from (11.87) the Defenders' control can be written as

$$
\begin{aligned}
u^*_{D_{1_f}}, u^*_{D_{2_f}} &= \arg\min_{u_{D_{1_f}}, u_{D_{2_f}}} \mathscr{H} \\
&= \arg\min_{u_{D_{1_f}}, u_{D_{2_f}}} \left(-\sigma_{\gamma_f} + \eta\right) u_{D_{1_f}} + \\
&\qquad\qquad\qquad \left(-\sigma_{\gamma_f} - \eta\right) u_{D_{2_f}} \\
\implies u^*_{D_{1_f}} &= -\operatorname{sign}\left(-\sigma_{\gamma_f} + \eta\right) = -1, \\
u^*_{D_{2_f}} &= -\operatorname{sign}\left(-\sigma_{\gamma_f} - \eta\right) = 1.
\end{aligned}
$$

The last two expressions, together, imply

$$-\eta \leq \sigma_{\gamma_f} \leq \eta$$

Since $-\eta \leq \sigma_{\gamma_f} \leq \eta$ define $\sigma_{\gamma_f} \equiv \chi\eta$ for $\chi \in [-1, 1]$. Substitution of
the Defender controls into the terminal Hamiltonian gives

$$\mathscr{H}_d(t_f) = v \cos\psi_f + \chi\eta \frac{v}{R_f} \sin\psi_f - \eta = 0.$$

The Attacker must maximize the Hamiltonian, and thus

$$\cos \psi_f^* = \frac{1}{\sqrt{1 + \frac{\chi^2 \eta^2}{R_f^2}}}, \qquad \sin \psi_f^* = \frac{\chi \eta}{R_f \sqrt{1 + \frac{\chi^2 \eta^2}{R_f^2}}}.$$

The terminal Hamiltonian becomes

$$\mathcal{H}_d(t_f)^* = \nu \sqrt{1 + \frac{\chi^2 \eta^2}{R_f^2}} - \eta = 0.$$

Solving for η:

$$\eta = \nu \sqrt{1 + \frac{\chi^2 \eta^2}{R_f^2}}$$

$$\eta^2 = \nu^2 \left(1 + \frac{\chi^2 \eta^2}{R_f^2} \right)$$

$$\implies \eta = \pm \frac{\nu R_f}{\sqrt{R_f^2 - \chi^2}}$$

Recall $\eta \equiv \sigma_\alpha \equiv \frac{\partial V}{\partial \alpha}$; thus an increase in α should give advantage to the Attacker which implies $\eta > 0$. At general time, the Hamiltonian is

$$\mathcal{H}_d(t) = 0$$

$$= -\sigma_R \nu \cos \psi + \frac{\nu^2 \chi R_f}{R \sqrt{R_f^2 - \chi^2}} \sin \psi - \frac{\nu R_f}{\sqrt{R_f^2 - \chi^2}}$$

Again, the Attacker maximizes the Hamiltonian,

$$\cos \psi^* = \frac{-\sigma_R}{\sqrt{\sigma_R^2 + \frac{\chi^2 \nu^2 R_f^2}{R^2 \left(R_f^2 - \chi^2 \right)}}},$$

$$\sin \psi^* = \frac{\nu \chi R_f}{R \sqrt{R_f^2 - \chi^2} \sqrt{\sigma_R^2 + \frac{\nu^2 \chi^2 R_f^2}{R^2 \left(R_f^2 - \chi^2 \right)}}}.$$

Substituting back into the Hamiltonian and solving for σ_R:

$$\sigma_R = \pm \sqrt{\frac{R^2 R_f^2 - \nu^2 \chi^2 R_f^2}{R^2 \left(R_f^2 - \chi^2 \right)}}.$$

Finally, substitution into the equilibrium Attacker control gives

$$\sin \psi^* = \chi \frac{\nu}{R}.$$

$\square$

Note that the form of the Attacker equilibrium control for this scenario is identical to that of the *Game of Angle* scenario.

Lemma 11.13. *The trajectories corresponding to $\chi = \pm 1$ separate the state space into regions of solo capture ($\gamma_f \neq 0$) and dual capture ($\gamma_f = 0$).*

Proof. The result follows from substitution of $\chi = 1$ or $\chi = -1$ into (11.88). $\square$

Lemma 11.14. *Attacker trajectories resulting in dual capture ($\gamma_f = 0$) are straight lines in the (x, y)-plane terminating at a point I' where*

$$I' = \begin{bmatrix} I'_x \\ I'_y \end{bmatrix} = \begin{bmatrix} R_f \cos (\beta_0 - \alpha_0) \\ R_f \sin (\beta_0 - \alpha_0) \end{bmatrix}. \tag{11.89}$$

Proof. The Attacker equilibrium control, (11.88), is identical in form to that of the *Game of Angle* solution, (11.74), and thus the Attacker path is a straight line for the same arguments as presented in Lemma 11.9. The angle $\beta_0 - \alpha_0$ corresponds to the $\gamma = 0$ axis, which, as in Lemma 11.9, is invariant under equilibrium play. $\square$

Lemma 11.15. *The surfaces separating solo and dual capture are given by the expression*

$$w (\hat{R}) = \pm \frac{\nu^2 \alpha}{R_f}, \tag{11.90}$$

where $\hat{R} = R \cos (\gamma) = R_f + \nu \alpha \sqrt{1 - \frac{\nu^2}{R_f^2}}$ is the polar distance measured along the $\gamma = 0$ axis and w is measured perpendicular to the $\gamma = 0$ axis, and $\hat{R} \in \left[1 + \nu \alpha_0 \sqrt{1 - \nu^2}, \infty \right]$.

Proof. As in Lemma 11.10, the dual capture trajectory terminates in α time (since $\dot{\alpha}^* = -1$ and $\alpha = 0$ in the dual capture scenario). The dual capture trajectories are thus straight lines (due to Lemma 11.14) of length $\nu \alpha_0$ which terminate at I', as defined in Lemma 11.14. Consider the upper limit of ψ_f^*, which is given by (11.88) with $\chi = 1$ to be $\psi^* =$

$\sin^{-1}\left(\frac{v}{R_f}\right)$. The corresponding distance perpendicular to the $\gamma = 0$ axis is $w = \sin \sin^{-1}\left(\frac{v}{R_f}\right) \cdot v\alpha_0 = \frac{v^2\alpha_0}{R_f}$. This w corresponds to a position which is $v\alpha_0\sqrt{1 - \frac{v^2}{R_f}}$ further than R_f, i.e., $\hat{R} = R_f + v\alpha_0\sqrt{1 - \frac{v^2}{R_f^2}}$. Taking the lower limit of ψ_f^* gives the corresponding negative width. $\qquad\square$

The regions $\mathscr{R}_{D_1}$ and $\mathscr{R}_{D_2}$ are defined as the sets of states for which the game terminates with $\gamma_f > 0$ (one-on-one with D_1) and $\gamma_f < 0$ (one-on-one with D_2), respectively (c.f. Lemma 11.11). Similarly, the region $\mathscr{R}_{D_{1,2}}$ is defined as the set of states for which the game terminates with $\gamma_f = 0$ which is completely specified by Lemma 11.10 and Lemmas 11.13–11.13.

Theorem 11.4. *For states in the region $\mathscr{R}_{D_{1,2}}$ the equilibrium Attacker heading angle is*

$$\psi^* = \gamma + \sin^{-1}\left(\frac{R\sin\gamma}{v\alpha}\right) \qquad (11.91)$$

and the Value function is

$$V(\mathbf{x}) = -R_f = v\alpha\frac{\sin\psi^*}{\sin\gamma}. \qquad (11.92)$$

Proof. The result follows from Lemmas 11.12–11.14 via a geometric proof process similar to Theorem 11.3. Consider the triangle $\triangle AI'C$, as shown in Fig. 11.6, where C is the target circle's center. Since the system begins in $\mathscr{R}_{D_{1,2}}$, the scenario terminates with $\alpha_f = \gamma_f = 0$. The time for each Defender to traverse an angle α around the perimeter of the target circle is α, since $\dot{\alpha} = -1$. Therefore, $\overline{AI'} = v\alpha$. Using the Law of Sines, the quantities are related as follows

$$\frac{R_f}{\sin(2\pi - \psi)} = \frac{v\alpha}{\sin\gamma} = \frac{R}{\sin(\psi - \gamma)}$$
$$\implies \frac{R_f}{-\sin\psi} = \frac{v\alpha}{\sin\gamma} = \frac{R}{\sin(\psi - \gamma)}.$$

The second equality may be rearranged to obtain (11.91). Likewise, the first equality may be rearranged to obtain (11.92). $\qquad\square$
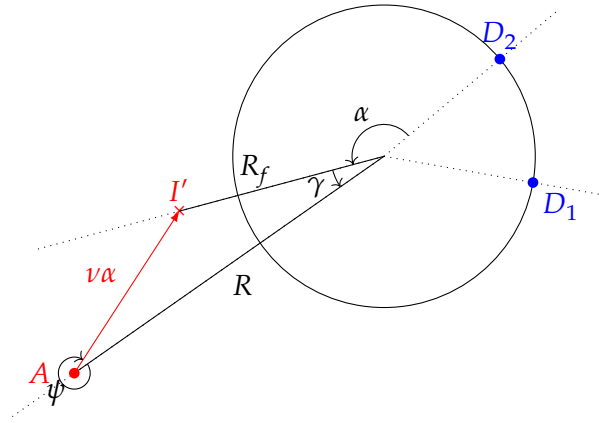
Figure 11.6: Illustration of the derivation of the equilibrium Attacker heading and Value function for the two-Defender *Game of Distance* with symmetric termination.

11.3.3 *Full Solution*

The two Defender game is truly three dimensional (in the reduced state space, i.e., R, γ, α). Although one may obtain the equilibrium flowfield over the whole state space by substituting the equilibrium strategies into the kinematics, it is more illustrative to visualize the solution in the (x, y)-plane for a particular α. Figure 11.7 shows the full solution of the two-Defender one-Attacker game, including all of the separating surfaces, regions, and salient features along with several representative Attacker trajectories. Note that the solution may be generalized to any number of Defenders simply by considering the two Defenders nearest to the Attacker's initial position.

11.3.4 *Alternative Attacker Win Scenario*

As in the one-Defender case, here, the solution of the *Game of Min Time* wherein the cost functional is $\min \max -t_f$, i.e., (11.34) is given. The analysis follows quite closely with those in the preceding sections. Figure 11.8 shows the solution with several representative trajectories.

"Cooperation" among the Defenders, i.e., where neither Defender is redundant, only occurs when the Attacker begins on the purple semi-circular section of $\partial \mathscr{R}_{A_1}$ and $\partial \mathscr{R}_{A_2}$. Otherwise, the Attacker plays the single-Defender version of the game with the nearest Defender. The
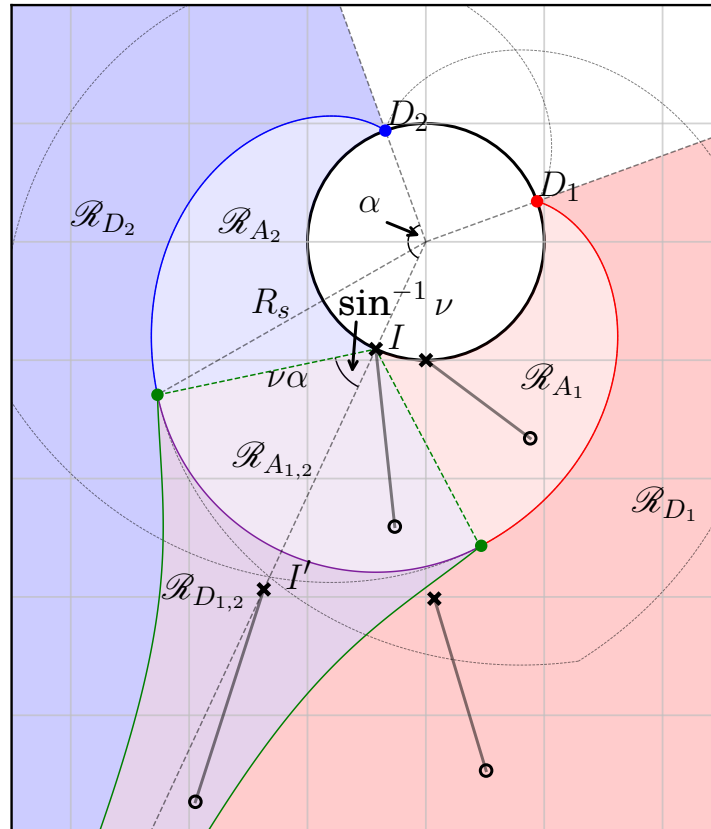
Figure 11.7: Separating surfaces for the two Defender game in the realistic plane for $\alpha_0 = \frac{3\pi}{4}$ and $\nu = 0.8$. Representative Attacker trajectories are shown in the symmetric termination regions and Defender 1 regions. Open black circles denote different Attacker initial positions, black ×'s denote the corresponding terminal Attacker positions.
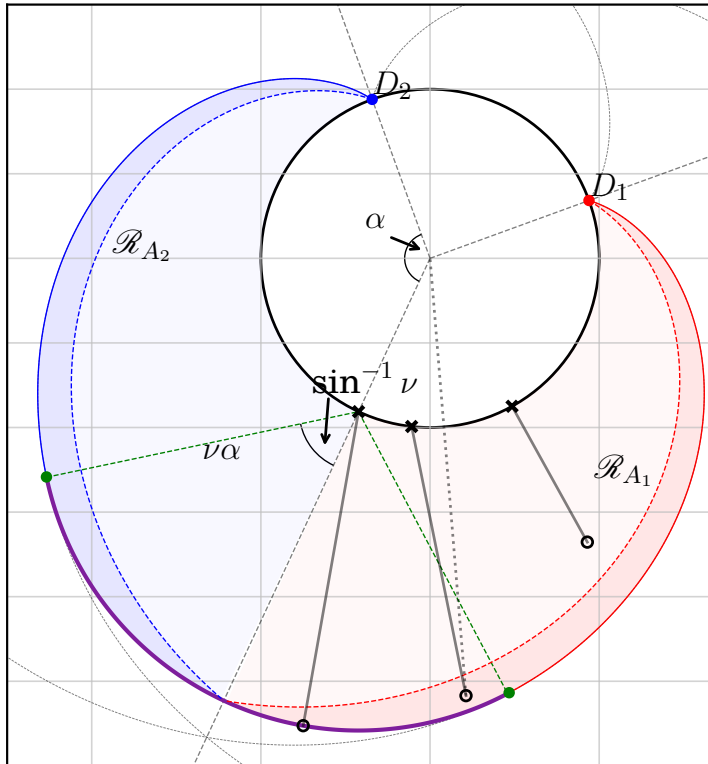
Figure 11.8: Two Defender *Game of Min Time* state space for a particular α with $\nu = 0.8$. The 3 Attacker trajectories, left-to-right, are 1) limiting, symmetric termination, 2) evasive (A cannot aim directly at the target circle center), and 3) direct (A aims at the circle center). Initial conditions in the light shaded regions result in direct trajectories, whereas the dark shaded regions represent initial conditions resulting in evasive trajectories.

main advantage for having two Defenders is that the state space $\mathscr{R}_{A_1} \cup \mathscr{R}_{A_2}$ is strictly smaller than the single-Defender state space, $\mathscr{R}_A$ (which is true for the *Game of Angle* as well).

11.4 CONCLUSION

The problem of guarding a circular target by patrolling its perimeter was considered. The one-Defender one-Attacker and two-Defender one-Attacker scenarios were formulated as zero-sum differential games with different cost/payoff functionals depending on whether the Attacker could reach the target's perimeter before the Defender(s) could 'lock on'. The analysis formally verifies that the Attacker heading strategy given in the literature for the one-Defender scenario is indeed the saddle-point equilibrium strategy for the games posed here [214]. For the two-Defender scenario, the state space was partitioned into regions based on the equilibrium termination condition. Analytic expressions for the separating surfaces between these regions and Value functions for each case were derived. The Attacker strategy in the Defenders-win, symmetric termination region differs from that of [214], partly due to differences in the termination condition and cost/payoff functional. An alternative scenario in which the Attacker seeks to reach the target in minimum time was also solved for both the one- and two-Defender cases. In the next chapter, the one-Defender will be expanded by adding a second Attacker.

<div style="text-align: right; font-size: 3em;">12</div>

TWO ATTACKERS AND A KINETIC TURRET

12.1 INTRODUCTION

This chapter extends the results of the previous to the case of two Attackers against a single Turret. It will become clear that the 1v1 solution is a vital component of the solution for this case, thereby addressing Research Objective 4. Additionally, the introduction of the second Attacker pertains to Research Objective 5. For the cases analyzed in this chapter, one Attacker must sacrifice itself for the other to reach the target; this Attacker's objective thus becomes to provide distraction (thus relating to Research Objective 8). Finally, the moment in which the first Attacker is taken out of action serves as a discrete event in the game which crucially affects the equilibrium strategies (i.e., Research Objective 9). This material is based upon the paper [256], which is in review.

The two-Attacker, single-Turret problem is formulated and solved (for a particular region of the state space) using the framework of DGT. In particular, the case in which neither Attacker can guarantee to reach the target individually, but, through their cooperation, the Attackers can guarantee that one can is addressed. Thus the Turret-Runner-Penetrator Differential Game (TRPDG) is posed and solved, and both the Value function and equilibrium strategies are provided. Within the TRPDG, two cases are considered: the case where the Runner is neutralized before the Penetrator reaches the target, and the case where the Penetrator reaches the target first. Section 12.2 provides a formulation for the overall two-Attacker, single-Turret problem and breaks the general problem up based on how many Attackers can be guaranteed to reach the target. Section 12.4 specifies the TRPDG with Runner neutralization which takes place in the state space region where exactly one Attacker can be guaranteed to reach the target; this version of the game ends when the Runner is neutralized. Section 12.5 provides a formula-

tion and solution for the TRPDG with early penetration; in this version of the game, the Penetrator reaches the target circle before the Runner is neutralized. Section 12.6 considers the case where the Turret is allowed to choose either Attacker to pursue (at any time). Section 12.8 provides some conclusions.

12.2   PROBLEM FORMULATION

In general, there are two scenarios one may consider: upon an Attacker's arrival to the target circle 1) the Turret is destroyed or 2) the Turret is not destroyed. Focus is given to the latter scenario. Concerning a measure of performance, there are two obvious metrics that may be considered: time (e.g., time to neutralize, time to penetrate, etc.) or angular separation (i.e., at the time of penetration), either of which are perfectly valid. The former makes sense if, for example, the Attackers represent some kind of munition and thereby angular separation is not as critical. Meanwhile, the latter may make sense if the Attackers represent intruders that have secondary objectives upon reaching the target circle. Both metrics are considered, e.g., in Chapter 11. In this chapter angular separation is considered to be the metric of interest.

In this formulation, the speed of the two Attackers are equal. Let $\nu < 1$ be the ratio of the Attackers' speed and Turret's maximum turn rate, the latter of which is normalized to 1. Let $\hat{\mathbf{z}} = (x_R, y_R, x_P, y_P, \beta)$ be the state of the system wherein the two Attackers' positions are represented by their 2-D Cartesian coordinates and the Turret's look angle is β w.r.t. the positive x-axis. The subscript R denotes Runner, and the subscript P denotes Penetrator. The target circle has a perimeter of 1; thus the Turret has an angular velocity advantage. The kinematics are thus

$$\hat{f}(\hat{\mathbf{z}}) = \begin{bmatrix} \dot{x}_R \\ \dot{y}_R \\ \dot{x}_P \\ \dot{y}_P \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} \nu \cos \hat{\psi}_R \\ \nu \sin \hat{\psi}_R \\ \nu \cos \hat{\psi}_P \\ \nu \sin \hat{\psi}_P \\ u_T \end{bmatrix}, \tag{12.1}$$

where $\hat{\psi}_R, \hat{\psi}_P$ are the Attackers' headings measured w.r.t. the positive x-axis, and $u_T \in [-1, 1]$ is the Turret's angular velocity input (with

positive u_T corresponding to counterclockwise motion). Alternatively, the Attackers' positions may be expressed in a polar coordinate system centered on the target circle's center. Define $\mathbf{z} = (r_R, \theta_R, r_P, \theta_P, \beta)$ where θ_R, θ_P are measured relative to the Turret's look angle. Also let $A_R \equiv (r_R, \theta_R)$ and $A_P \equiv (r_P, \theta_P)$; the Turret is also denoted T. The associated kinematics are

$$
f(\mathbf{z}) = \begin{bmatrix} \dot{r}_R \\ \dot{\theta}_R \\ \dot{r}_P \\ \dot{\theta}_P \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} -v \cos \psi_R \\ \frac{v}{r_R} \sin \psi_R - u_T \\ -v \cos \psi_P \\ \frac{v}{r_P} \sin \psi_P - u_T \\ u_T \end{bmatrix}
\tag{12.2}
$$

where ψ_R, ψ_P are measured clockwise w.r.t. the line from the respective Attacker to the target circle center. Figure 12.1 depicts the scenario, showing both coordinate systems specified above.
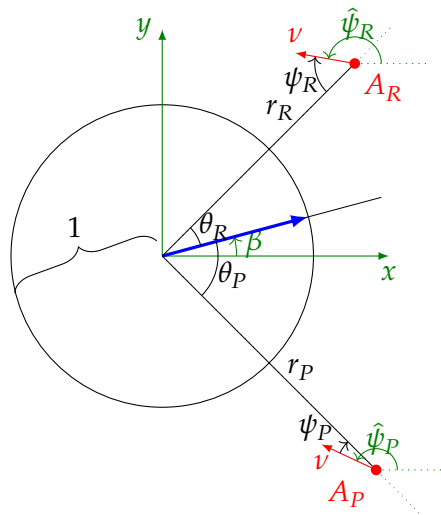
Figure 12.1: Two Attacker Scenario – the green color indicates the Cartesian coordinate system; black represents the polar coordinate system. The Attacker position angles, θ_R and θ_P, are measured w.r.t. T's look angle and are positive in the CCW direction (thus $\theta_P < 0$, as shown).

An Attacker A_i is considered to be neutralized (and removed from the remainder of the playout, if any) if at any time $\theta_i = 0$. Conversely, A_i is said to penetrate the target if it can maneuver all the way to the target circle ($r_i = 1$) while avoiding the Turret's line-of-sight. Ideally, both Attackers would like to penetrate the target without being neutralized.

In the general case, there are three termination cases: (*i*) both Attackers penetrate the target, (*ii*) one Attacker is neutralized and one penetrates, or (*iii*) both Attackers are neutralized. Cases (*i*) and (*iii*) are discussed briefly in the following section. The remainder of the paper focuses on the state space region wherein $A_R, A_P \notin \mathscr{R}_A$ at initial time, where $\mathscr{R}_A$ is the single-Attacker, single-Turret Attacker's win region, defined in (12.5). This is the region in which a single Attacker can be guaranteed to penetrate the target. When $A_R, A_P \notin \mathscr{R}_A$, neither Attacker can guarantee successful penetration of the target by itself; a subset of this region is will be constructed in which, through their cooperation and superiority in numbers, one of the Attackers can successfully penetrate. Let the region of interest for this state space be defined as

$$\Omega := \{\mathbf{z} \mid A_R, A_P \notin \mathscr{R}_A, \ r_R, r_P \geq 1\}. \tag{12.3}$$
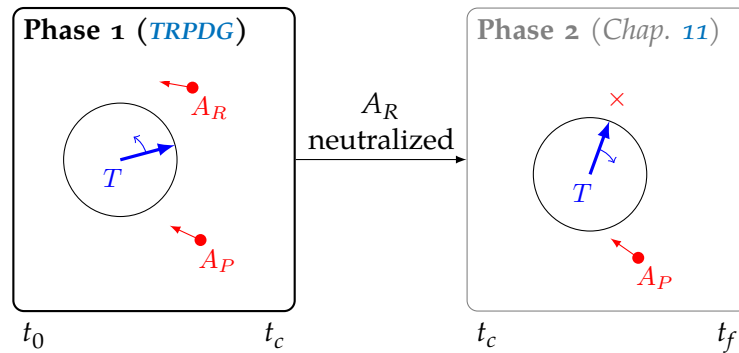
Figure 12.2: Abstract depiction of the scenario; in Phase 1 T pursues A_R while A_P seeks advantageous position for Phase 2, and Phase 2 is the remaining single-Attacker *Game of Angle*.

As mentioned previously, angular separation is the metric of interest, therefore, when it is possible for the Penetrator to reach the target, the cost functional takes the form

$$J(\mathbf{z}; u_T(\cdot), \psi_R(\cdot), \psi_P(\cdot)) = |\theta_P(t_f)|. \tag{12.4}$$

The overall scenario may be broken down into two phases. In Phase 1, without loss of generality, the Turret pursues A_R until it's aim is aligned with A_R's position, at which time A_R is taken out of action.

Thereby, Phase 2 commences, wherein the Turret begins pursuing A_P. Figure 12.2 depicts these two phases.

It may also be the case that the A_P is able to reach the target circle prior to A_R's demise, whereby $t_f \leq t_c$. In this case, there is no Phase 2. This case is referred to as the Early Penetrator (EP) case. The case depicted in Fig. 12.2 is analyzed in Section 12.4, and the Early Penetrator is the focus of Section 12.5. Both of these cases fall under the TRPDG. The following Section covers all other possible cases.

12.3 OTHER CASES

12.3.1 *Both Attackers Win*

Consider the single-Attacker, single-Turret scenario analyzed in [214] and Chapter 11. The region of win for the Attacker, i.e., wherein the Attacker is guaranteed to reach the target circle under optimal play is defined as

$$\mathcal{R}_A \equiv \{(r,\theta) \mid \theta > \theta_{GoK}(r)\}, \tag{12.5}$$

where

$$\theta_{GoK}(r) = \sqrt{\frac{r^2}{\nu^2} - 1} + \sin^{-1}\left(\frac{\nu}{r}\right) - \sqrt{\frac{1}{\nu^2} - 1} - \sin^{-1}\nu \tag{12.6}$$

Lemma 12.1. *In the two-Attacker, single-Turret scenario with kinematics given by* (12.2)*, both Attackers are guaranteed to reach the target circle under optimal play, that is, $r_{R_f} = r_{P_f} = 1$ if and only if $A_R, A_P \in \mathcal{R}_A$.*

Proof. Optimal play is given by the respective single-Attacker, single-Turret equilibrium control [214] (see also Chapter 11)

$$\sin \psi_i^* = \text{sign}\,(\sin \theta_i)\left(\frac{\nu}{r_i}\right).$$

The fact that $A_R, A_P \in \mathcal{R}_A \implies r_{R_f} = r_{P_f} = 1$ is due to each Attacker being able to win individually; the presence of additional Attackers does not aid the Turret in any way – both A_R and A_P are able to win. Now, it will be proven that $r_{R_f} = r_{R_f} = 1 \implies A_R, A_P \in \mathcal{R}_A$. Suppose

$A_i \notin \mathcal{R}_A$, the Turret could choose to implement its one-on-one strategy $u_T = \text{sign}\left(\sin\theta_i\right)$ against A_i and be guaranteed to neutralize A_i with $r_i > 1$. □

12.3.2 One or more Attackers Lose

Let the Turret's one-on-one win region be defined $\mathcal{R}_T = \mathcal{R}_A^c$. The trivial case occurs when $A_i \in \mathcal{R}_A$ and $A_j \notin \mathcal{R}_A$ for $i,j \in \{R,P\}$, $i \neq j$. Clearly, A_i can guarantee a win while T can guarantee neutralization of A_j. The construction and solution of a *Game of Degree* in this region of the state space is left for future work.

When $A_R, A_P \notin \mathcal{R}_A$, there is a region of the state space in which one the Attackers can win and a region in which neither can win. The former is analyzed in this chapter in detail; the analysis of what the agents should do in the latter region is left for future work.

12.3.3 The One-Attacker, One-Turret Differential Game

This subject was covered in the previous chapter, however, the following derivation was absent (for the case of Attacker win), so it is included here because it is useful in later analyses.

Lemma 12.2 (Form of the single-Attacker strategy)**.** *The single-Attacker game, with kinematics* $\dot{\mathbf{z}} = \begin{bmatrix} \dot{r}_P & \dot{\theta}_P & \dot{\beta} \end{bmatrix}^\top$, *Value function* $V_{A_P} = \max_{\psi_P} \min_{u_T} |\theta_{P_f}|$, *and terminal surface* $\phi = r_{P_f} - 1 = 0$ *has an equilibrium the Penetrator strategy of the form*

$$\sin\psi^* = \text{sign}\left(\sigma_\theta\right)\frac{v}{r_P}. \tag{12.7}$$

Proof. The Hamiltonian is

$$\mathscr{H} = -\sigma_r v \cos\psi_P + \sigma_\theta\left(\frac{v}{r_P}\sin\psi_P - u_T\right) + \sigma_\beta u_T, \tag{12.8}$$

and the adjoint dynamics for the θ_P and β states are

$$\dot{\sigma}_\theta = -\frac{\partial\mathscr{H}}{\partial\theta_P} = 0, \qquad \dot{\sigma}_\beta = -\frac{\partial\mathscr{H}}{\partial\beta} = 0. \tag{12.9}$$

At final time $t = t_f$, the transversality condition yields the terminal adjoint value for the β state

$$\sigma_{\beta_f} = \frac{\partial \Phi}{\partial \beta_f} + \mu \frac{\partial \phi}{\partial \beta_f} = 0, \tag{12.10}$$

where $\Phi \equiv |\theta_{2_f}|$. Thus $\sigma_\beta = 0$ for all $t \in \left[0, t_f\right]$. The Penetrator wishes to maximize the Hamiltonian, while the Turret seeks to minimize it, giving

$$\sin \psi_P^* = \frac{\sigma_\theta}{r_P \sqrt{\sigma_r^2 + \frac{\sigma_\theta^2}{r_P^2}}}, \qquad u_T^* = \text{sign}\left(\sigma_\theta\right). \tag{12.11}$$

Substituting (12.10) and (12.11) into (12.8) gives

$$\mathscr{H} = \nu \sqrt{\sigma_r^2 + \frac{\sigma_\theta^2}{r_P^2}} - |\sigma_\theta| \tag{12.12}$$

The terminal Hamiltonian value is

$$\mathscr{H}_f = -\frac{\partial \Phi}{\partial t_f} - \mu \frac{\partial \phi}{\partial t_f} = 0. \tag{12.13}$$

Since the state dynamics are time-autonomous, $\mathscr{H} = 0$ for all $t \in \left[0, t_f\right]$. Substituting into (12.12) and solving for σ_r^2 gives

$$\sigma_r^2 = \frac{\sigma_\theta^2}{\nu^2} - \frac{\sigma_\theta^2}{r_P^2}. \tag{12.14}$$

Substituting into (12.11) yields (12.7). $\qquad\qquad\qquad\qquad\square$

12.4 TRPDG WITH RUNNER NEUTRALIZATION

In this section a differential game is constructed representing Phase 1 in Fig. 12.2 wherein it is assumed that T neutralizes A_R (before A_P is able to penetrate the target). The analysis proceeds in the polar coordinate system, utilizing (12.2), with $\mathbf{z} \in \Omega$. A major assumption is made at this point, which is that the fate (goals and roles) of each Attacker is set *a priori*; and, moreover, these roles cannot be switched during the

playout of the game. Let A_R be the first Attacker to be neutralized by T, regardless of the position of A_P. A complete solution, which involves the agents determining which Attacker T will pursue and neutralize first, necessarily depends on the solution of this simpler problem. This restriction is also motivated by some real-world considerations: often it is costly for a weapon (or targeting) system to switch targets after it has begun tracking a particular target.

Concerning the control signals u_T, ψ_R, and ψ_P, it is assumed that the agents have full state information (i.e., **z** is known) but they do not know the instantaneous control of the adversary. That is, neither T nor A_R, A_P are discriminated. In general, the solution approach utilized throughout the remainder of the paper, which is based on the formation of the first-order necessary conditions for equilibrium, yields an open-loop equilibrium [20, pg. 344]. From the open-loop solution, the closed-loop (state-feedback) strategies may be synthesized [20, pg. 344]. In order to constitute an equilibrium, the proposed strategies would need to satisfy the sufficient conditions (i.e., yield a Value function that is C^1 and satisfy the Hamilton-Jacobi-Isaacs equation) – however, these conditions will be satisfied by construction everywhere except for the singularities. Thus special attention is given to these singularities to ensure the validity of the solution.

Begin by assuming that A_P *can* reach the target circle ($r_P(t_f) = 1$) in Phase 2 (c.f. Fig. 12.2) and that A_P prefers to maximize its angular separation from T at final time. That is, in the second phase of the engagement that begins when A_R is neutralized, A_P plays the *Game of Angle*, as specified in Chapter 11. As such, A_R is referred to as the Runner, and A_P as the Penetrator. Define the region in which A_P can reach the target circle in Phase 2 as

$$\mathscr{R}_{2A} \equiv \{\mathbf{z} \mid A_P(t_c) \in \mathscr{R}_A\}. \tag{12.15}$$

The explicit construction and solution of a *Game of Degree* wherein A_P is also neutralized is left for future work. The *first* phase of the engage-

ment is modeled as a zero-sum differential game over the cost functional

$$
\begin{aligned}
J\left(\mathbf{z}; u_T(\cdot), \psi_R(\cdot), \psi_P(\cdot)\right) &= \Phi\left(\mathbf{z}(t_c), t_c\right) \\
&= V_{A_P}\left(r_P(t_c), \theta_P(t_c)\right)
\end{aligned}
\tag{12.16}
$$

where $\mathbf{z} \in \mathscr{R}_{2A}$ and V_{A_P} is the Value of the *Game of Angle* (i.e., the single-Attacker game, studied in Chapter 11) played between A_P and T starting from $t = t_c$, and t_c is the terminal time of the game, which occurs when A_R is neutralized. Eq. (12.16) is related to (12.4) in that by employing equilibrium strategies in Phase 1, T and A_P set themselves up for the best possible outcome in Phase 2 with regards to $|\theta_P(t_f)|$. The terminal manifold is defined as

$$
\phi\left(\mathbf{z}(t_c), t_c\right) = \theta_R(t_c) = 0.
\tag{12.17}
$$

The Attackers cooperatively seek to maximize J, while the Turret wants to minimize J. Thus the Value function for the TRPDG with Runner neutralization is defined as

$$
V(\mathbf{z}) = \min_{u_T(\cdot)} \max_{\psi_R(\cdot), \psi_P(\cdot)} J\left(\mathbf{z}; u_T(\cdot), \psi_R(\cdot), \psi_P(\cdot)\right).
\tag{12.18}
$$

The Value function of the *Game of Angle* is given in Chapter 11 as

$$
V_{A_P}\left(r_P, \theta_P\right) = |\theta_P| - \theta_{GoK}\left(r_P\right),
\tag{12.19}
$$

where θ_{GoK} is the single-Attacker, single-Turret *Game of Kind* surface defined in (12.6) in Section 12.3. Figure 12.2 depicts the overall scenario broken up into two distinct phases: Phase 1, which terminates at $t = t_c$ when A_R is neutralized, and Phase 2 wherein A_P and T playout the *Game of Angle*. The Value function, V_{A_P}, of Phase 2 determines, in part, the equilibrium strategies in Phase 1.

The notation $\mathbf{z}_c \equiv \mathbf{z}(t_c)$ is used generally. The Hamiltonian is

$$
\mathscr{H} = \lambda_\beta u_T + \sum_{i=R,P} -\lambda_{r_i} v \cos \psi_i + \lambda_{\theta_i}\left(\frac{v}{r_i} \sin \psi_i - u_T\right).
\tag{12.20}
$$

The equilibrium adjoint dynamics are [40]

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial \mathscr{H}}{\partial \mathbf{z}} = \begin{bmatrix} -\frac{v}{r_R^2} \lambda_{\theta_R} \sin \psi_R \\ 0 \\ -\frac{v}{r_P^2} \lambda_{\theta_P} \sin \psi_P \\ 0 \\ 0 \end{bmatrix}, \tag{12.21}$$

and thus λ_{θ_R}, λ_{θ_P}, and λ_β are constant. The transversality condition yields the adjoint values at terminal time [40]

$$\boldsymbol{\lambda}_c^\top = \frac{\partial \Phi}{\partial \mathbf{z}_c} + \mu \frac{\partial \phi}{\partial \mathbf{z}_c} \tag{12.22}$$

$$= \begin{bmatrix} 0 & 0 & \frac{\partial V_{A_P}}{\partial r_{P_c}} & \frac{\partial V_{A_P}}{\partial \theta_{P_c}} & 0 \end{bmatrix} + \mu \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Let the adjoints of A_P's single-Attacker *Game of Angle* be written

$$\boldsymbol{\sigma}^\top \equiv \begin{bmatrix} \sigma_r & \sigma_\theta \end{bmatrix} = \begin{bmatrix} \frac{\partial V_{A_P}}{\partial r_P} & \frac{\partial V_{A_P}}{\partial \theta_P} \end{bmatrix}.$$

Notice that $\lambda_{\beta_c} = 0$ and $\dot{\lambda}_\beta = 0$, thus $\lambda_\beta = 0$ for all $t \in [0, t_c]$. Similarly, $\lambda_{\theta_R} = \mu$ and $\lambda_{\theta_P} = \sigma_\theta$ for all $t \in [0, t_c]$. Substituting the values of λ_β, λ_{θ_R}, and λ_{θ_P}, the Hamiltonian becomes

$$\begin{aligned} \mathscr{H} = {} & -\lambda_{r_R} v \cos \psi_R + \mu \left(\frac{v}{r_R} \sin \psi_R - u_T \right) \\ & -\lambda_{r_P} v \cos \psi_P + \sigma_\theta \left(\frac{v}{r_P} \sin \psi_P - u_T \right). \end{aligned} \tag{12.23}$$

The Hamiltonian is a separable function of the controls u_T and ψ_R, ψ_P, and thus *Isaacs' condition* [20, 130] holds:

$$\min_{u_T} \max_{\psi_R, \psi_P} \mathscr{H} = \max_{\psi_R, \psi_P} \min_{u_T} \mathscr{H}.$$

The following result applies generally to differential games based on these dynamics with a well-defined terminal cost functional and terminal surface; it arises mainly as a consequence of the fact that the Attackers have simple motion (i.e., single integrator dynamics). Most of the later results in this chapter rely heavily on the following:

Lemma 12.3 (Equilibrium Controls are Constant). *For any differential game with unconstrained kinematics described by (12.1) and a Mayer-type cost functional, the equilibrium strategies of all the agents are constant. In particular, each Attacker's equilibrium trajectory is a straight line (in the Cartesian plane), and the Turret's control is either always clockwise or always counterclockwise.*

Proof. Given that the cost functional is of Mayer-type, the Hamiltonian for the system (12.1) is

$$\mathcal{H} = \lambda_\beta u_T + \sum_{i=R,P} \lambda_{x_i} v \cos \hat{\psi}_i + \lambda_{y_i} v \sin \hat{\psi}_i \qquad (12.24)$$

Let $\hat{\boldsymbol{\lambda}} \equiv \begin{bmatrix} \lambda_{x_R} & \lambda_{y_R} & \lambda_{x_P} & \lambda_{y_P} & \lambda_\beta \end{bmatrix}^\top$ be the adjoint vector in the Cartesian frame. The equilibrium adjoint dynamics are given by [130, Eq. 4.5.3]

$$\dot{\hat{\boldsymbol{\lambda}}} = -\frac{\partial \mathcal{H}}{\partial \hat{\mathbf{z}}} = 0. \qquad (12.25)$$

Without loss of generality, suppose that the Attackers seek to maximize the cost functional while the Turret seeks to minimize it. The equilibrium controls are

$$\cos \hat{\psi}_i^* = \frac{\lambda_{x_i}}{\sqrt{\lambda_{x_i}^2 + \lambda_{y_i}^2}}, \; \sin \hat{\psi}_i^* = \frac{\lambda_{y_i}}{\sqrt{\lambda_{x_i}^2 + \lambda_{y_i}^2}}, \; i = R, P \qquad (12.26)$$

$$u_T^* = -\operatorname{sign} \lambda_\beta. \qquad (12.27)$$

Because the equilibrium adjoint dynamics are 0, $\boldsymbol{\lambda}$ is constant, and thus u_T^* and $\hat{\psi}_i^*$ for $i = R, P$ are also constant. Since $\hat{\psi}_i$ are defined relative to the positive x-axis, the Attackers' trajectories are straight lines in the Cartesian plane. $\qquad\square$

Note that if $\lambda_{x_i} = \lambda_{y_i} = 0$ for $i \in \{R, P\}$ then the associated equilibrium heading $\hat{\psi}_i^*$ is not uniquely defined since it would not appear in the Hamiltonian, (12.24). This generates singular solutions, which will be addressed later.

12.4.1 *Equilibrium Turret & Runner Strategies*

Lemma 12.4 (Equilibrium Turret Strategy). *In the differential game defined by the kinematics, (12.2), cost functional, (12.16), and terminal surface, (12.17) the Turret's strategy is*

$$u_T^*(t) = k, \qquad k \in \{-1, 1\}, \forall t \in [0, t_c]. \tag{12.28}$$

Proof. The fact that k is a constant is due to Lemma 12.3. The Turret must minimize the Hamiltonian, (12.23) – in order to do so, it is clear that

$$u_T^*(t) = \arg\min_{u_T} \mathcal{H} = \text{sign}\,(\mu + \sigma_\theta).$$

Again, both μ and σ_θ are constant. The sign function ensures that $k \in \{-1, 1\}$. $\qquad\square$

Lemma 12.5 (Equilibrium Runner Strategy). *In the differential game defined by the kinematics, (12.2), cost functional, (12.16), and terminal surface, (12.17), the Runner's trajectory is a straight line perpendicular to the Turret's LOS at the time of termination.*

Proof. The Runner maximizes the Hamiltonian, (12.23), which occurs when the vector $\begin{bmatrix} \cos \psi_R & \sin \psi_R \end{bmatrix}^\top$ is parallel with the vector $\begin{bmatrix} -\lambda_{r_R} & \frac{\mu}{r_R} \end{bmatrix}^\top$. Therefore,

$$\cos \psi_R^* = \frac{-\lambda_{r_R}}{\sqrt{\lambda_{r_R}^2 + \frac{\mu^2}{r_R^2}}}, \qquad \sin \psi_R^* = \frac{\mu}{r_R \sqrt{\lambda_{r_R}^2 + \frac{\mu^2}{r_R^2}}}. \tag{12.29}$$

At terminal time, $\lambda_{r_R}(t_c) = 0$ from (12.22), which implies $\cos \psi_{R_c}^* = 0$. Thus A_R's terminal heading is $\psi_{R_c}^* \in \{\frac{\pi}{2}, -\frac{\pi}{2}\}$, and is perpendicular to T's LOS since $\theta_{R_c} = 0$. The fact that A_R's trajectory is a straight line in the Cartesian coordinate system is due to Lemma 12.3. $\qquad\square$

It remains to show in which direction (either CCW or CW) both the Turret and Runner should travel. In the present case, wherein $A_R, A_P \notin \mathcal{R}_A$, the biggest benefit for the Attacker team comes when the Runner, A_R, keeps the Turret occupied for as long as possible, thereby giving

the Penetrator, A_P, a chance to reach an advantageous position before T starts pursuing A_P in earnest.

Lemma 12.6. *The sign of the equilibrium Turret and Runner control inputs are such that*

$$\text{sign}\left(u_T^*\right) = \text{sign}\left(\sin \psi_R^*\right) = \text{sign}\left(\sin \theta_R\right). \tag{12.30}$$

That is, A_R has a component of velocity away from T, and T turns toward A_R.

Proof. There are four possibilities: i) A_R away, T towards, ii) A_R towards, T towards, iii) A_R away, T away, iv) A_R towards, T away. The cost functional (12.16) is based on the single-Attacker *Game of Angle* between T and A_P. A_R can only improve the outcome of the 1v1 game if it can cause T to implement a control other than the 1v1 equilibrium strategy (e.g., turn *away* from A_P rather than towards it). First, consider the Turret's control - if $\text{sign}\left(u_T\right) \neq \text{sign}\left(\sin \theta_R\right)$ then T is turning *away* from A_R. In order to neutralize A_R, T must go the long way around the target circle in the worst case. Thus cases iii) and iv) are excluded by inspection. It remains to determine whether A_R should head i) away from T or ii) towards. At the time A_R would be neutralized in ii), A_R would still be alive in i). The Runner can do nothing to reduce V_{A_P}, but it *can* increase V_{A_P} if it can continue to draw T away from A_P. Therefore, at the time and position of neutralization of A_R in ii) it is never worse (and generally better) for A_R to be alive, which implies A_R must run away from T. $\qquad\square$

Remark 17. The Turret strategy given by Lemma 12.4 and Lemma 12.6 corresponds to the single-Attacker circular target defense strategy from Chapter 11 (played against A_R).

Remark 18. If at any time $\text{sign}\left(\sin \theta_R\right) = \text{sign}\left(\sin \theta_P\right)$, then the Runner's heading, ψ_R, is inconsequential (i.e., ψ_R^* is not uniquely defined). The choice of direction for the Turret is to turn towards the two Attackers; by pursuing A_R, the Turret is also pursuing A_P, and thus the Runner can do nothing to help (or hinder) A_P.

12.4.2 *Equilibrium Penetrator Strategy*

The Penetrator seeks to maneuver in such a way to reach an advantageous position by the time the Runner is neutralized. By advantageous, it is meant that its terminal position maximizes the Value of the subsequent differential game which ensues once the Runner has been neutralized.

The presence of the DS in the single-Attacker game [214, 254] creates an interesting situation in this two-Attacker variant. When the state of a system lies on a DS, the equilibrium controls of one or more agents is non-unique [130]. In the case of the single-Attacker game, when $\cos \theta = -1$, there is symmetry in the system such that the T could chase A either counterclockwise (CCW) or clockwise (CW) and resulting Value of the *Game of Angle* would be the same. The consequence of the DS is that the single-Attacker Value function V_{A_P} is not smooth along the surface; thus the single-Attacker adjoint vector, σ, is undefined along the surface. Therefore, A_P's terminal heading, defined by (12.32) and (12.22) as $\psi_{P_c}^* = \tan^{-1} -\sigma_\theta/\sigma_r$, is not well-defined either. There are two cases: (1) $\cos \theta_{P_c} \neq -1$ and σ is well-defined (the regular case), and (2) $\cos \theta_{P_c} = -1$ and σ is undefined (the singular case).

Lemma 12.7 (Regular Equilibrium Penetrator Strategy). *In the differential game defined by the kinematics, (12.2), cost functional, (12.16), and terminal surface, (12.17) the Penetrator's equilibrium trajectory is a straight line that is aligned with its Game of Angle equilibrium trajectory at terminal time wherever the Game of Angle adjoints σ_r and σ_θ are defined. Moreover, A_P's control strategy is given by [214] (see, also, Chapter 11)*

$$\sin \psi_P^* = \text{sign} \left(\sin \theta_{P_c} \right) \left(\frac{v}{r_P} \right). \tag{12.31}$$

Proof. The Penetrator maximizes the Hamiltonian, (12.23), which occurs when the vector $\begin{bmatrix} \cos \psi_P & \sin \psi_P \end{bmatrix}^\top$ is parallel with the vector $\begin{bmatrix} -\lambda_{r_P} & \sigma_\theta \end{bmatrix}^\top$:

$$\cos \psi_P^* = \frac{-\lambda_{r_P}}{\sqrt{\lambda_{r_P}^2 + \frac{\sigma_\theta^2}{r_P^2}}}, \qquad \sin \psi_P^* = \frac{\sigma_\theta}{r_P \sqrt{\lambda_{r_P}^2 + \frac{\sigma_\theta^2}{r_P^2}}}. \qquad (12.32)$$

At final time, $\lambda_{r_P} = \sigma_r$ (due to (12.22)) and thus $\tan \psi_P^* = -\sigma_\theta / \sigma_r$. Thus, at final time, A_P's heading is identical to the equilibrium Attacker heading from the single-Attacker scenario (c.f. Chapter 11). Furthermore, A_P's trajectory is a straight line in the Cartesian coordinate frame due to Lemma 12.3, just as it is in the single-Attacker scenario. Therefore, A_P's regular state feedback equilibrium control is given by (12.31).

$\square$

The geometric interpretation of the following Lemma is that the Penetrator's equilibrium trajectory never crosses the $\beta + \pi$ radial. In cases where (12.31) *would* cause this, the Runner, instead, takes a shallower angle such that $\cos \theta_{P_c} = -1$.

Lemma 12.8 (Singular Penetrator Strategy)**.** *In the differential game defined by the kinematics, (12.2), cost functional, (12.16), and terminal surface, (12.17) a family of the Penetrator singular trajectories exist which terminate at $\cos \theta_{P_c} = -1$, with $r_{P_c} > 1$. These trajectories are straight lines with the following state feedback strategy*

$$\sin \psi_P^* = \frac{\chi \nu}{r_P}. \qquad (12.33)$$

where $\chi \in [-1, 1]$ and $\operatorname{sign}(\chi) = \operatorname{sign}\left(\sin \theta_{P_c}\right)$.

Proof. First, recall that the trajectories are straight lines in the Cartesian coordinate frame due to Lemma 12.3. The general form of the single-Attacker equilibrium control is given in Lemma 12.2:

$$\sin \psi_P^* = \operatorname{sign}(\sigma_\theta) \frac{\nu}{r_P}.$$

However, when $\cos \theta_P = -1$, the term $\operatorname{sign}(\sigma_\theta)$ is undefined because the Value function V_{A_P} is not differentiable on the DS. Replace the quan-

tity sign (σ_θ) with a variable χ. When $\chi = \pm 1$, the solution exactly corresponds to the limiting case of the regular equilibrium trajectories described in Lemma 12.7 where $\sin \psi_P^* = \pm \frac{\nu}{r_P}$. If $|\chi| > 1$, the approach angle to the point $(r_P, \cos \theta_P) = (r_{P_c}, -1)$ would be steeper. Backwards integrating from $(r_{P_c}, -1)$ with an angle $|\sin \psi_P| > \frac{\nu}{r_P}$ would push the state of the system into a region that is filled with regular equilibrium trajectories – see Fig. 12.3. The former trajectories would be suboptimal (nonequilibrium) compared to the latter. Therefore, it must be the case that $\chi \in [-1, 1]$. The sign of χ is governed by the sign of $\sin \theta_{P_c}$ as in the regular trajectory case. Note, this proof method is similar to the method used to solve for the simultaneous capture condition in [89] and Chapter 6. □
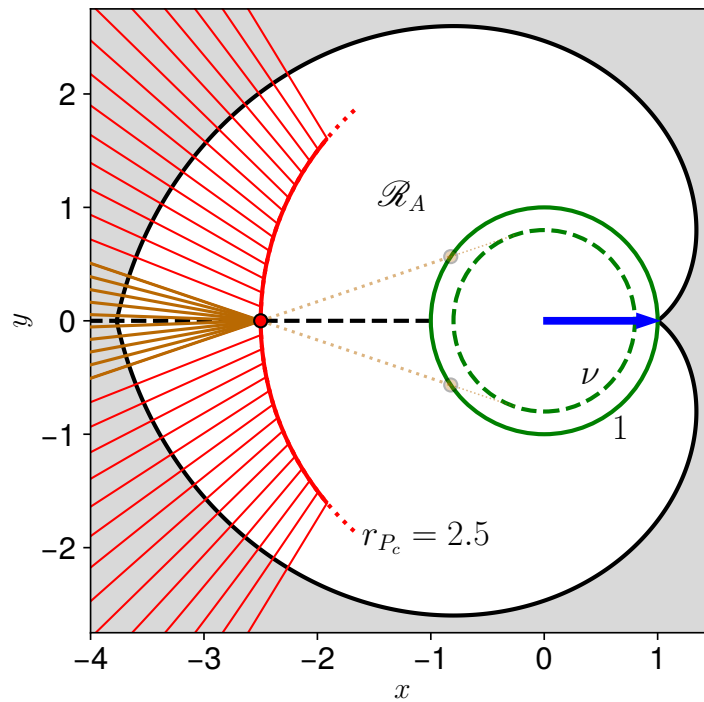
Figure 12.3: The Penetrator regular (red) and singular (dark orange) trajectories. The target circle is green; the dashed inner circle is a circle of radius ν $(= 0.8)$. Note the extension of each regular A_P trajectory are tangential to the ν circle. The position of the Turret at the time of neutralization of A_R is shown by a blue arrow. A family of trajectories is shown wherein $r_{P_c} = 2.5$. Singular A_P trajectories terminate on the dashed black DS. In the second phase of the scenario, A_P terminates at either dark orange filled circle depending on T's choice of CCW or CW.

The DS in the single-Attacker game (c.f. Chapter 11) favors the Turret. While on the DS ($\theta = \pi$) the Turret may choose to turn either CW or CCW at max turn rate and achieve the same cost in equilibrium. However, for the Attacker to achieve a payoff associated with the equilibrium it must know the Turret's choice at $t = 0$ and choose a corresponding heading (i.e., CW if T chooses CW and CCW otherwise). Without knowing $u_T(0)$, the Attacker is left to guess; a correct guess will yield the equilibrium payoff, and an incorrect guess will result in a small loss in the payoff. In the latter case, the Attacker moves *towards* T at the initial time instant and must immediately switch headings. The implication for the TRPDG is that the singularity (i.e., $\theta_{P_c} = \pi$) does not benefit the Attacker team.
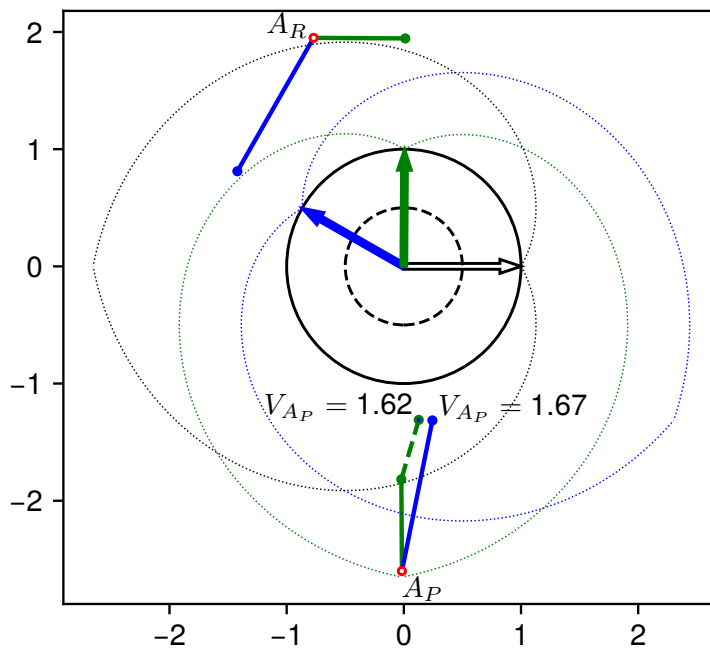
Figure 12.4: Comparison of A_R running away (blue) versus towards (green) T. The black arrow represents T's initial position and the dotted lines represent the boundary of $\mathscr{R}_A$ at the time instant associated with its color. In the green case, A_P can make maximal usage of the singularity and aim straight at the target circle. However, the resultant payoff is not as good as in the blue case wherein all three agents implement the prescribed strategies.

Consider the following example, shown in Fig. 12.4. One may be tempted to believe that the singularity could be *helpful* to the Attacker team in the sense that, under the prescribed Penetrator strategy, A_P may run directly towards the closest point on the target circle (thereby

reducing r_P as quickly as possible). In this example, A_R, contrary to Lemma 12.6 runs towards T who turns CCW (shown in green). A_P is placed directly on the line $\theta_{R_c} = \pi$; this is the most "extreme" example of A_P taking a singular path (i.e., being able to set $\psi_P = 0$). The blue trajectory corresponds to A_R running away from T and A_P taking the heading prescribed in (12.31). For the green case, $\theta_{P_c} = \pi$ but neutralization of A_R happens much sooner, whereas in the blue case $\theta_{P_c} < \pi$ but neutralization of A_R is later. The green Penetrator trajectory is continued with a dashed line for the CCW case up until T reaches the blue position. In this case, if T goes CCW the entire time, then clearly it is better for A_P to play according to the associated 1v1 strategy of aiming at the tangent to the circle of radius ν. Hence, the blue trajectories (which follow the prescribed strategies) yield the best Value for the Attackers.

12.4.3 *Full solution*

Figure 12.5 shows the state trajectory in the Cartesian coordinate frame for a regular trajectory (with $\cos\theta_{P_c} \neq -1$) and for a singular trajectory (with $\cos\theta_{P_c} = -1$). The Runner, A_R, has a trajectory which is perpendicular to the Turret's LOS at the time of termination. In the regular case, the Penetrator, A_P, has a trajectory which is aimed at the tangent of a circle of radius ν; once A_R is neutralized, A_P would continue along this course all the way to the target circle. In the singular case, the Penetrator prefers not to cross the $\cos\theta_P = -1$ radial at $t = t_c$ and therefore has taken a shallower angle to end up at $\cos\theta_{P_c} = -1$. From here, the Penetrator takes either the upper or lower trajectory depending on T's choice of rotation after neutralizing A_R (CW or CCW, respectively).

Although it wasn't explicitly stated in the problem formulation, it is required that $\cos\theta_P \neq 0$ for all $t \in [0, t_c]$ because, otherwise, the Penetrator would have been neutralized while T was *en route* to neutralize the Runner. The limiting case occurs when $\text{sign}(\sin\theta_P) = \text{sign}(\sin\theta_R)$ and $\cos\theta_P \to 0$ precisely at the moment of neutralization of A_R.
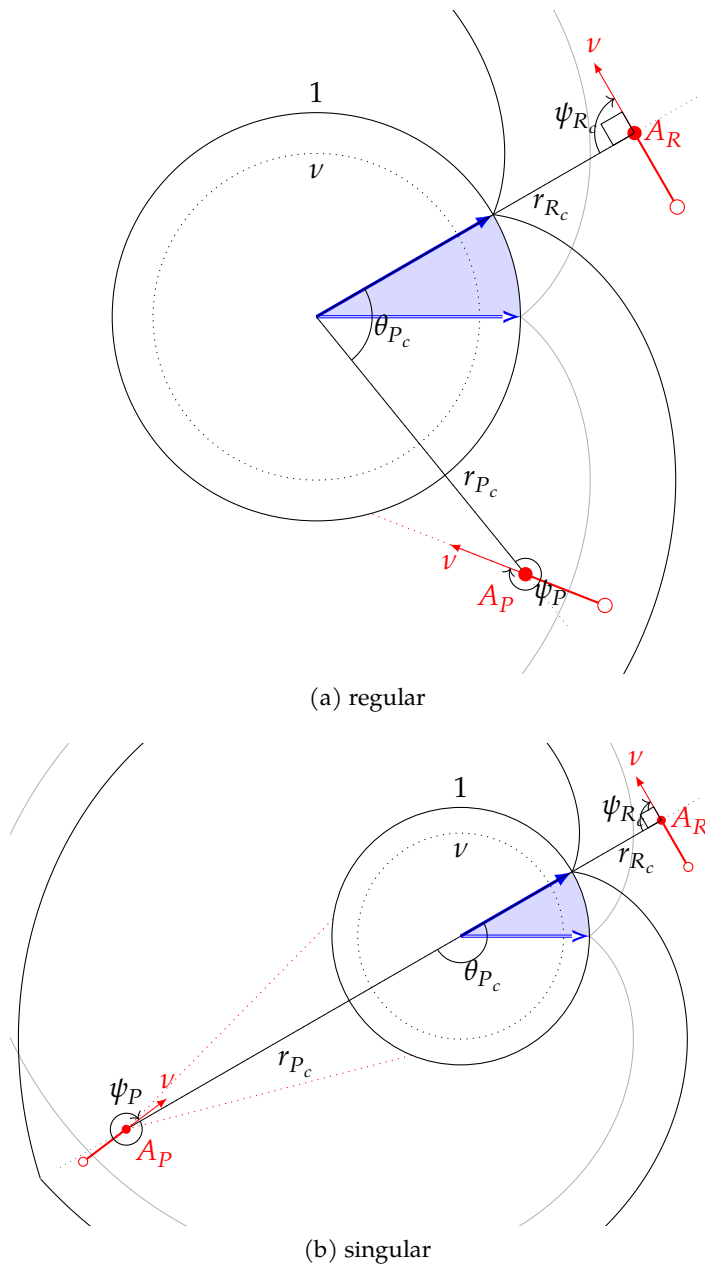
(a) regular

(b) singular

Figure 12.5: Representative solutions for the (a) regular and (b) singular cases. Initial Attacker and Turret positions are denoted by open circles and an arrow, respectively; terminal positions are filled. The boundary of $\mathscr{R}_A$ is shown at $t = 0$ (grey) and at $t = t_c$ (black).

In order for A_P to penetrate the target, it must reach $\mathscr{R}_{A_c} \equiv \mathscr{R}_A\left(\mathbf{z}(t_c)\right)$, i.e., the one-on-one Attacker win region at terminal time. The limiting case occurs when $A_{P_c} \in \partial\mathscr{R}_{A_c}$ where $\partial\mathscr{R}_{A_c}$ is the boundary of the one-on-one Attacker win region at terminal time. That is, the Penetrator is just barely able to satisfy the necessary condition to 'win' (i.e., reach the target) in the second phase of the engagement. Note that $\partial\mathscr{R}_{A_c}$ is the zero-level set of the cost functional, V_{A_P}, and thus the equilibrium Penetrator trajectories terminating at a point on $\partial\mathscr{R}_{A_c}$ are normal to the surface. The other limiting case is when A_P reaches the target exactly when A_R is neutralized.

Define $\mathscr{R}_{2A}$ as the set of states in which A_P can be guaranteed to 'win', i.e., the set of states in which $A_P \in \mathscr{R}_A$ within t_c time while avoiding premature termination. One boundary of $\partial\mathscr{R}_{2A}$ can be constructed geometrically by setting A_P on $\partial\mathscr{R}_{A_c}$ and backwards integrating the equilibrium Penetrator strategy ((12.31) for $\cos\theta_{P_c} \neq -1$, and (12.33) for $\cos\theta_{P_c} = -1$). The other boundary is obtained by setting $r_P = 1$ and backwards integrating. Care must be taken to eliminate terminal A_P positions which result in A_P paths which start and end inside the sector swept by the T's motion (which would result in premature termination.) Figure 12.6 shows a slice of $\mathscr{R}_{2A}$ for a particular initial Turret position (β) and A_R position ((r_R, θ_R)).

It's clear from Fig. 12.6 and Eqs. (12.31) and (12.33) that the solution depends on β_c (from which θ_{P_c} is measured), or equivalently, the terminal time, t_c. From Lemmas 12.4 and 12.5, along with Lemma 12.6, it is clear that A_R has a component of velocity directed away from T and terminates perpendicular to T's LOS under optimal play, while T moves in the direction of A_R at its maximum turn rate. Thus T must cover an angular sector at least $\text{mod}(|\theta_R|, 2\pi)$. For the Turret, angle traveled and time are equivalent since the Turret's turn rate and the target circle radius are both 1. Let $\gamma \geq 0$ be the amount of additional angle the Turret must cover to neutralize the Runner. Then $t_D = \text{mod}(|\theta_R|, 2\pi) + \gamma$ is time of arrival of the Turret to the candidate terminal position. The Runner's trajectory to the candidate terminal configuration covers an angular sector γ and is perpendicular to T's LOS in the terminal configuration. See Fig. 12.7 for a diagram depicting the geometry. Thus
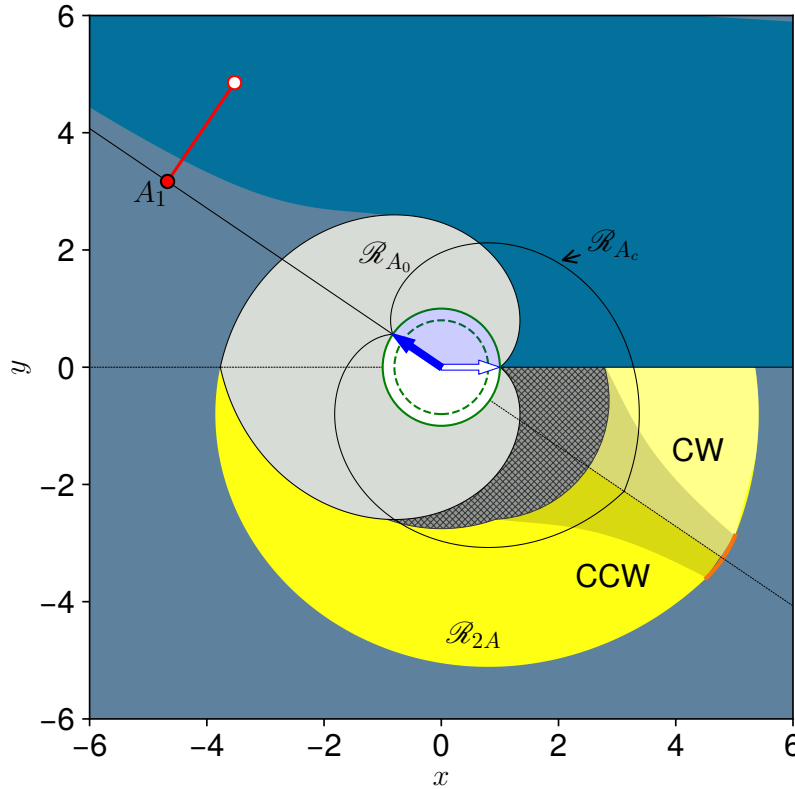
Figure 12.6: A partitioning of the state space for particular β, r_R, and θ_R. The T and A_R trajectories start at the open circles and end at the closed circles. The *Game of Kind* surface θ_{GoK} is drawn at $t = 0$ and at $t = t_c$. Note A_P positions beginning within $\mathscr{R}_{A_0}$, marked by light grey, are not considered, nor are positions in which A_P penetrates the target before t_c, marked by hatched grey. The yellow region represents $\mathscr{R}_{2A}$, the set of A_P initial conditions which in which it can be guaranteed to successfully penetrate the target. In the light shaded portion, A_P's motion has a clockwise component, otherwise it has a counter-clockwise component. The dark shaded portion is filled with singular trajectories which terminate on $\cos \theta_{P_c} = -1$. There is a segment of $\partial \mathscr{R}_{2A}$ which is a circular arc, marked by orange, which is the locus of extremal A_P singular initial conditions. Premature termination would occur for any A_P positions beginning in the bright blue region, and the faded blue region represents positions in which $\mathscr{R}_{A_c}$ cannot be reached; T is able to neutralize both Attackers in either case.

$t_A = \frac{1}{\nu} r_R \sin \gamma$ is the time of arrival of the Runner to the candidate terminal position. In the limiting case, the terminal Runner distance is $r_{R_{\min}} = 1$, which gives an upper bound for γ:

$$\gamma_{\max} = \cos^{-1}\left(\frac{1}{r_R}\right).$$

Now, define the time difference of arrival to the terminal configuration as

$$\tau(\gamma) \equiv t_D(\gamma) - t_A(\gamma)$$
$$= |\theta_R| + \gamma - \frac{1}{\nu} r_R \sin \gamma, \tag{12.34}$$

with $\gamma \in [0, \cos^{-1}(\frac{1}{r_R})]$. Clearly it would be suboptimal for the Runner to reach a point, stop, and wait for the Turret to reach that point (i.e., $\tau > 0$); similarly, if the Turret arrives before the Runner (i.e., $\tau < 0$) the Turret would have had to pass the Runner *en route*. Thus, for equilibrium, it must be the case that both agents arrive in the terminal configuration simultaneously, i.e., $\tau^* = 0$.

Lemma 12.9. *The function, $\tau(\gamma)$, (12.34), which represents the time difference of arrival of the Runner and Turret to a candidate terminal configuration, has a unique zero, γ^*, on the interval $[0, \cos^{-1}(1/r_R)]$.*

Proof. First, (12.34) is a continuous function of γ since γ and $\sin \gamma$ are both continuous. For the lower bound of τ, it must be that $\tau(0) = \mod(|\theta_R|, 2\pi)$, and thus $\tau(0) > 0$. In other words the Runner arrives first – in fact, it travels zero distance, whereas the Turret covers an an-
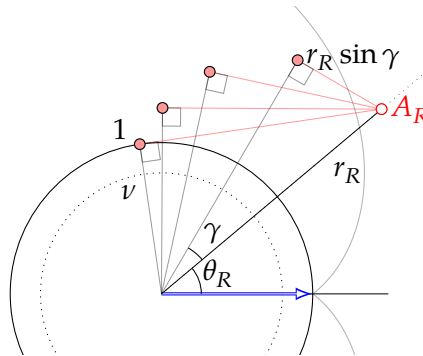

Figure 12.7: Relevant geometry for the determination of terminal time t_c. Open circles represent initial positions and the closed red circles indicate candidate terminal configurations for A_R.

gular distance of $\mathrm{mod}(|\theta_R|, 2\pi)$. For the upper bound, it will shown that $\tau(\gamma_{\max}) < 0$ by contradiction. Suppose that $\tau(\gamma_{\max}) > 0$, that is, the Runner arrives to the candidate terminal configuration before the Turret. The upper bound, $\gamma_{\max}$ is derived from the limiting case where $r_{R_c} \to 1$. This would mean the Runner was able to reach the target circle before the Turret could align with it which contradicts the assumption that $A_R \notin \mathscr{R}_A$ (which is embedded in the assumption that $\mathbf{z} \in \Omega$). Therefore, from the Intermediate Value Theorem, the function $\tau(\gamma)$ crosses zero on the interval $[0, \cos^{-1}(1/r_R)]$.

Also, $\partial\tau/\partial\gamma = 1 - r_R/\nu \cos\gamma$ which is strictly negative on the interval $[0, \cos^{-1}(1/r_R)]$ since $r_R/\nu > 1$ and $\cos\gamma > 0$ on the interval. Thus $\tau(\gamma)$ is monotonic on the interval, which implies that the zero crossing is unique. $\qquad\square$

Because of the uniqueness of γ^* in which $\tau(\gamma^*) = 0$ many standard root-finding methods are suitable for computing it. The terminal time is simply

$$t_c = \mathrm{mod}(|\theta_R|, 2\pi) + \gamma^*. \tag{12.35}$$

With the value of t_c computed, the terminal angle is

$$\beta_c = \beta + t_c \, \mathrm{sign}\left(\sin\theta_R\right).$$

From Fig. 12.6 it is clear that the effect of $\mathrm{sign}\left(\sin\theta_{P_c}\right)$ in (12.31) and (12.33) is that the Runner's motion (at least in $\mathscr{R}_A$) has a component of velocity towards the $\beta_c + \pi$ radial. The interpretation is that the Runner seeks to end up *behind* the Turret at terminal time, which is an advantageous position for the *Game of Angle*. Thus, under equilibrium play by all the agents, the terminal state is

$$\mathbf{z}_c = \begin{bmatrix} r_{R_c} \\ \theta_{R_c} \\ r_{P_c} \\ \theta_{P_c} \\ \beta_c \end{bmatrix} = \begin{bmatrix} r_R \cos\left(t_c - \mathrm{mod}(|\theta_R|, 2\pi)\right) \\ 0 \\ \sqrt{\left(\dfrac{\nu^2\chi t_c}{r_P}\right)^2 + \left(r_P - \nu t_c\sqrt{1 - \dfrac{\chi^2\nu^2}{r_P^2}}\right)^2} \\ \theta_P - \mathrm{sign}\left(\sin\theta_R\right)t_c + \sin^{-1}\left(\dfrac{\chi\nu^2 t_c}{r_P r_{P_c}}\right) \\ \beta + \mathrm{sign}\left(\sin\theta_R\right)t_c \end{bmatrix}, \tag{12.36}$$

where $\chi \in \{-1, 1\}$ for regular trajectories, $\chi \in [-1, 1]$ for singular trajectories, and

$$\text{sign}\, \chi = \text{sign}\left(\sin \theta_{P_c}\right) = \text{sign}\, \xi, \tag{12.37}$$

where $\xi \in [-\pi, \pi]$ is A_P's angle-to-go to the $\beta_c + \pi$ radial,

$$\xi = -\text{mod}\left(\theta_P - \text{sign}\left(\sin \theta_R\right) t_c,\ 2\pi\right) + \pi. \tag{12.38}$$

The trajectory is singular if A_P's regular strategy, (12.31), would cause it cross the $\beta_c + \pi$ radial, which occurs if

$$\sin^{-1}\left(\frac{v^2 t_c}{r_P \sqrt{r_P^2 + v^2 t_c^2 - 2 r_P v t_c \sqrt{1 - \frac{v^2}{r_P^2}}}}\right) > |\xi|. \tag{12.39}$$

Note the LHS of the above expression is the angular sector swept (w.r.t. the origin) by A_P's regular strategy in t_c time. If the trajectory is singular, then, by definition $\cos \theta_{P_c} = -1$. The Law of Sines gives the following relationships:

$$\frac{r_{P_c}}{\sin \psi_P} = \frac{v t_c}{\sin \xi} = \frac{r_P}{\sin\left(\pi - |\psi_P| - |\xi|\right)}.$$

The singular A_P heading is

$$\psi_P = \text{sign}\left(\xi\right)\left(\sin^{-1}\left(\frac{r_P \sin|\xi|}{v t_c}\right) - |\xi|\right), \tag{12.40}$$

and the singular terminal A_P distance is

$$r_{P_c} = \frac{v t_c \sin \psi_P}{\sin \xi}. \tag{12.41}$$

Finally, the Value function is

$$V(\mathbf{z}) = |\theta_{P_c}| - \theta_{GoK}\left(r_{P_c}\right) \tag{12.42}$$

where $\left(r_{P_c}, \theta_{P_c}\right)$ is given by (12.36) and θ_{GoK} is defined in (12.6).

The Attackers simply aim at their respective terminal point from (12.36), and the Turret rotates towards the Runner. Of course, one or

more agents could (to their detriment) deviate from the strategy which would necessitate recomputing the solution in practice. For discrete time systems, for example, it is recommended for the agent implementing its equilibrium strategy to recompute the solution at each time step.

Figure 12.8 contains an example in which the Attackers both lose when operating individually, but one is able to win when the Attackers cooperate and behave according to the solution of the TRPDG.
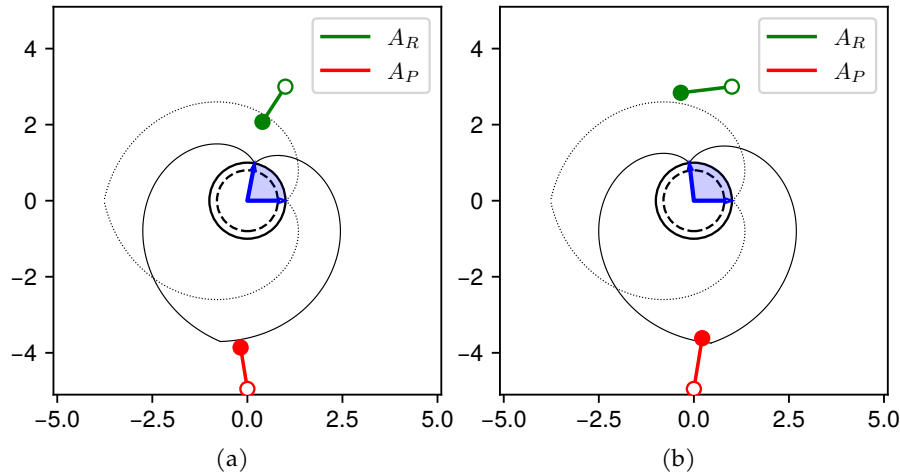
Figure 12.8: Attackers implement single-Attacker strategy, ignoring the presence of the other Attacker (a); neither Attacker wins. Attackers cooperate, implementing the TRP solution (b); A_P wins as a result.

12.5 TRPDG WITH EARLY PENETRATION

In this section the case in which the Penetrator can reach the target circle prior to the capture of the Runner is analyzed. Such is the case, e.g., in the hatched region of Fig. 12.6. It is assumed that the Turret prefers to pursue (and eventually neutralize) the Runner rather than attempt to reduce the Penetrator's angular separation at the time of penetration. If the Turret were to attempt to minimize $|\theta_P|$ it may be the case A_R could reach a more advantageous position – perhaps even inside $\mathcal{R}_A$. Therefore, the Turret's strategy is fixed to $u_T = \text{sign}(\sin \theta_R)$ as it is in the previous section. Additionally, the Runner strategy from the previous section is utilized in the following analysis.

With the Turret and Runner strategies fixed, the EP case becomes an optimal control problem for the Penetrator. Let t_f be the time instant at when the Penetrator actually penetrates the target circle. Then the cost functional is the same as in (12.4) and is restated here with some additional notation.

$$J^E = \Phi^E\left(\mathbf{z}(t), t_f\right) = |\theta_P(t_f)|, \tag{12.43}$$

where $0 \leq t_f \leq t_c$ and the superscript E denotes Early Penetration (EP). The Penetrator seeks to maximize its angular separation *at the time of penetration*. For the sake of clarity, let reaching the target circle and penetrating the target be defined as $r_P = 1$ and $r_P < 1$, respectively. The Value of the EP optimal control problem, if it exists, is defined as

$$V^E(\mathbf{z}_0) \equiv \max_{\psi_P(\cdot)} J^E. \tag{12.44}$$

An important distinction must be made at this point as to whether or not the Penetrator has anything to gain (according to (12.43)) by delaying penetration. If, for example, $|\dot{\theta}_P| < 0$ when the Penetrator has reached the target circle then it only stands to reduce its payoff by delaying penetration and thereby chooses to end the game by penetrating the target immediately upon arrival. However, delaying penetration is advantageous, for example, when the Turret's pursuit of the Runner is drawing its aim further away from the Penetrator at the time of arrival at the target (i.e., $|\dot{\theta}_P| > 0$).

12.5.1 *Regular Trajectories Ending in Immediate Penetration*

For regular trajectories ending in immediate penetration it is necessary that 1)
$$\text{sign}(\sin \theta_P) = \text{sign}(\sin \theta_R)$$

or 2)
$$\text{sign}(\sin \theta_R) \sin \theta_P < \text{sign}(\sin \theta_R)(t_c - \pi).$$

In the former case, the two Attackers are on the same side w.r.t. the Turret's look angle. Once the Penetrator reaches the target circle (i.e., $r_P = 1$) it can only be the case that $|\dot{\theta}_P| < 0$ because $\nu < 1$ and thus T has an angular velocity advantage. Case 2) corresponds to $\text{sign}(\sin \theta_{P_c}) = \text{sign}(\sin \theta_R)$, and thus the direction of A_P's motion is the same as A_R and T's (c.f. 12.4.2). Satisfaction of this condition, alone, is not sufficient. For example, it may be possible for A_P to achieve $\cos \theta_{P_f} = -1$, thereby the trajectory would be singular (to be discussed in the following subsections). The regular optimal penetrator strategy is given by (12.31) defined over $t \in \left[0, t_f\right]$.

12.5.2   *Delayed Penetration*

When the Turret is turning away from the Penetrator, it is advantageous for the Penetrator to delay penetration until the Runner's neutralization at t_c; thus $t_f = t_c$. First, begin by augmenting the problem definition by including the following path constraint

$$m(\mathbf{z}) = r_P - 1 \geq 0, \qquad \forall t \in [0, t_c] \tag{12.45}$$

which requires that A_P remain on or outside the target circle until the moment of A_R's neutralization by T. When the constraint is active, the system may remain constrained if $\dot{r}_P = 0$. Eq. (12.2) gives

$$\dot{r}_P = -\nu \cos \psi_P,$$

and thus the system will remain constrained if $\psi_P = \pm\frac{\pi}{2}$.

Note that the terminal manifold, the zero-level set of ϕ (defined in (12.17)), is only defined over $r_{P_c} \geq 1$. From (12.22) the terminal adjoint values depend on $\frac{\partial \phi}{\partial \mathbf{z}_c}$. Thus, when $r_{P_c} = 1$, the quantity $\frac{\partial \phi}{\partial r_{P_c}}$ is undefined (since the state is on edge of the bounded plane $\phi = 0$). Consequently, from (12.32), it is clear that the optimal Penetrator heading at terminal time, $\psi_{P_c}^*$, is undefined. This is another singularity which is similar to the $\cos \theta_{P_c} = -1$ singularity analyzed in Section 12.4.2.

The following subsections treat the $r_{P_c} = 1$ singularity and the constrained trajectories. The $r_{P_c} = 1$ and $\cos \theta_{P_c} = -1$ singularities are referred to as the *distance singularity* and *angle singularity*, respectively.

12.5.2.1 Distance Singularity

The following lemma provides bounds on the value of the terminal Penetrator heading for this case.

Lemma 12.10. *For EP, if $r_{P_c} = 1$ and $\cos \theta_{P_c} \neq -1$, the terminal Penetrator heading is bounded according to*

$$\text{sign}(\sin \theta_{P_c}) \sin^{-1} \nu \leq \text{sign}(\sin \theta_{P_c}) \psi_{P_c}^* \leq \text{sign}(\sin \theta_{P_c}) \frac{\pi}{2}. \quad (12.46)$$

Proof. The inclusion of $\text{sign}(\sin \theta_{P_c})$ is necessary to account for the fact that A_P seeks to aim towards the $\beta_c + \pi$ radial, which maximizes its payoff (c.f., Section 12.4.2). The upper bound of (12.46) is due to the path constraint, $r_P \geq 1$. A larger heading angle would yield $\dot{r}_{P_c} > 0$, which implies that A_P had arrived at the target circle from the inside, which clearly violates the constraint. The lower bound corresponds to the regular/unconstrained Penetrator control, (12.31). A smaller heading angle would push the state of the system into a region where regular/unconstrained trajectories exist. They, by definition have $r_{P_c} > 1$ and $|\theta_{P_c}|$ necessarily smaller. Thus coming in to $r_{P_c} = 1$ with this heading would have been suboptimal for the Penetrator. □

For any $|\psi_{P_c}| \in \left[\sin^{-1} \nu, \frac{\pi}{2} \right)$, it must be that $\dot{r}_2 < 0$ and thus the system immediately leaves the constraint in backwards time. The following Lemma gives the optimal Penetrator heading for this case.

Lemma 12.11. *For EP, if the constraint m is active only at terminal time, that is, if $r_P > 1$ for $t \in [0, t_c)$ and $r_{P_c} = 1$, and $\theta_{P_c} \neq \pi$ the optimal Penetrator heading is*

$$\sin \psi_P^* = \text{sign}\left(\sin \theta_{P_c} \right) \left(\frac{\kappa}{r_P} \right), \quad (12.47)$$

where

$$\kappa = + \sqrt{ \frac{4 r_P^2 - \left(\nu^2 t_c^2 - r_P^2 - 1 \right)^2}{4 \nu^2 t_c^2} }, \quad (12.48)$$
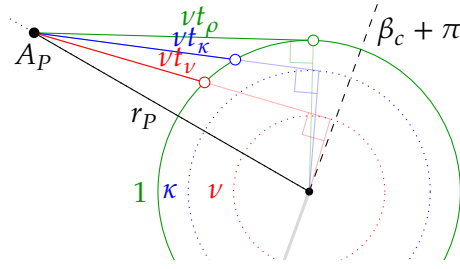
and $\nu < \kappa < 1$.

Figure 12.9: An illustration of three time instants of interest: t_ν, the time required for A_P to reach the target circle whilst aiming at the tangent to the ν circle, t_κ, the time to reach the target circle whilst aiming at the tangent to the κ circle, and t_ρ, the time to reach the target circle tangentially.

Proof. Because the trajectory is unconstrained (except for at the moment of termination), Lemma 12.3 applies, which means A_P's trajectory is a straight line in the Cartesian plane. The premise of this Lemma is that A_P ends on the target circle, thus one need only determine the line segment joining A_P's position to a point on the target circle which is of length νt_c, which is the distance A_P can cover in the time it takes T to neutralize A_R. Consider a circle of radius κ, $\nu < \kappa < 1$, centered at the origin. The distance from A_P to a tangent point on the κ circle is $\sqrt{r_P^2 - \kappa^2}$ (see Fig. 12.9). Since $\kappa < 1$, the line segment joining A_P to the tangent point passes through the target circle. The distance from this intersection to the tangent point on the κ circle is $\sqrt{1 - \kappa^2}$. Thus the time it takes for A_P to reach the target circle while aiming at a tangent point on the κ circle is

$$t_\kappa = \frac{1}{\nu} \left(\sqrt{r_P^2 - \kappa^2} - \sqrt{1 - \kappa^2} \right). \tag{12.49}$$

Eq. (12.48) is obtained by setting the above expression equal to t_c and solving for κ. Finally, (12.47) is obtained from the right-triangle geometry, since A_P's aim point is tangent to the κ circle, and, once again, $\text{sign}\left(\sin \theta_{P_c} \right)$ appears for reasons described in Lemma 12.7.   □

12.5.2.2 *Constrained Trajectories*

The following Lemma provides the optimal Penetrator heading for the case where the trajectory is constrained (or partly constrained).

Lemma 12.12. *For EP, if the constraint m activates at a time t_ρ, where $0 \leq t_\rho \leq t_c$ the optimal Penetrator heading is*

$$\sin \psi_P^* = \begin{cases} \text{sign}\left(\sin \theta_{P_c}\right)\left(\frac{1}{r_P}\right) & 0 \leq t < t_\rho, \ r_P > 1, \\ \text{sign}\left(\sin \theta_{P_c}\right) & t_\rho \leq t \leq t_c, \ r_P = 1, \end{cases} \tag{12.50}$$

where

$$t_\rho = \frac{1}{\nu}\sqrt{r_P^2 - 1}. \tag{12.51}$$

Proof. In this case, the constraint m is activated partway through the trajectory. The trajectory may be considered as two parts: an unconstrained arc, $\mathcal{U}$ (wherein $r_P > 1$), and a constrained arc, $\mathcal{C}$ ($r_P = 1$). In the transition from $\mathcal{U}$ to $\mathcal{C}$, it is necessary that there exist controls ψ_P and/or u_T that keep the system on the constraint (i.e., maintain $m = 0$). This is referred to as the tangency condition [40]. Here, the Penetrator heading, ψ_P, appears in the expression $\dot{m} = \dot{r}_P = -\nu \cos \psi_P$, and thus the system may remain constrained if the Penetrator applies a heading s.t. $\dot{m} = 0$ which is $\psi_P = \pm \frac{\pi}{2}$. Over the $\mathcal{U}$ arc, the results of Lemma 12.3 apply, which states the Penetrator trajectory is a straight line in the Cartesian plane. Thus the Penetrator must reach the target circle tangentially, in a straight line, and proceed thereafter by traveling along its perimeter until terminal time. The time at which A_P reaches the target circle tangentially, (12.51), is obtained from the relevant right-triangle geometry (see Fig. 12.9). Eq. 12.50 is then synthesized from the controls associated with the $\mathcal{U}$ and $\mathcal{C}$ arcs. The inclusion of $\text{sign}(\sin \theta_{P_c})$ is necessary for reasons discussed in Lemma 12.7. □

12.5.2.3 *Case Determination*

From Fig. 12.9 and the delayed penetration control policies in Lemmas 12.11 and 12.12, the conditions which determine the type of trajectory based on the system's current state may be established. Let t_κ and t_ρ be defined as in (12.49) and (12.51), respectively. Also, t_ν (as depicted in Fig. 12.9) can be derived:

$$t_\nu = \frac{1}{\nu}\left(\sqrt{r_P^2 - \nu^2} - \sqrt{1 - \nu^2}\right). \tag{12.52}$$

If $t_\nu > t_c$, then A_P, employing the regular optimal strategy, (12.31), could not have reached the target circle by the time of A_R's neutralization (then the optimal penetrator control is governed by Lemma 12.7). Else if $t_\nu \leq t_c < t_\tau$, there exists a straight line trajectory terminating on the target circle of length νt_c (Section 12.5.2.1, Lemma 12.11). Otherwise, $t_c \geq t_\tau$ and A_P can aim at a tangent point to the target circle, reach the target circle, and apply the constrained control until $t = t_c$ (Section 12.5.2.2, Lemma 12.12).

12.5.3 *Max Payoff Possible*

In this subsection the case in which the upper limit of the cost/payoff functional, (12.43), is realizable (i.e., $J = \pi$) is addressed. Recall from Section 12.4.2 the singularity which arises when $\cos \theta_{P_c} = -1$ – that is, the Turret is looking directly away from the Penetrator at terminal time. There, the terminal Penetrator heading $\psi_{P_c}^*$ was undefined according to the first-order necessary conditions for equilibrium. This singularity also comes into play for the two early penetrator cases discussed so far.

Consider the case where $(r_{P_c}, \cos \theta_{P_c}) = (1, -1)$ which is the corner of the bounded plane $\phi = 0$ and the intersection of the distance and angle singularities. In Section 12.4.2, the angle singularity gave rise to a symmetric cone of incoming trajectories bounded by $\psi_{P_c}^* \in \left[-\sin^{-1}\left(\frac{\nu}{r_{P_c}}\right), \sin^{-1}\left(\frac{\nu}{r_{P_c}}\right) \right]$. In Section 12.5.2.1, the distance singularity gave rise to a cone of incoming trajectories bounded by $\sin^{-1}(\nu)$ and $\frac{\pi}{2}$. The combination of the two singularities results in a cone of incoming trajectories bounded by,

$$-\frac{\pi}{2} \leq \psi_{P_c}^* \leq \frac{\pi}{2}. \tag{12.53}$$

The locus of initial A_P positions thus forms a semi-circle of radius νt_c centered on $(r_{P_c}, \cos \theta_{P_c}) = (1, -1)$. Concerning constrained trajectories with $\psi_{P_c} = \pm \pi$, the optimal trajectories emanate in both directions in backwards time. Figure 12.10b shows the locus of initial Penetrator positions which terminate on the corner point exactly at $t_f = t_c$.
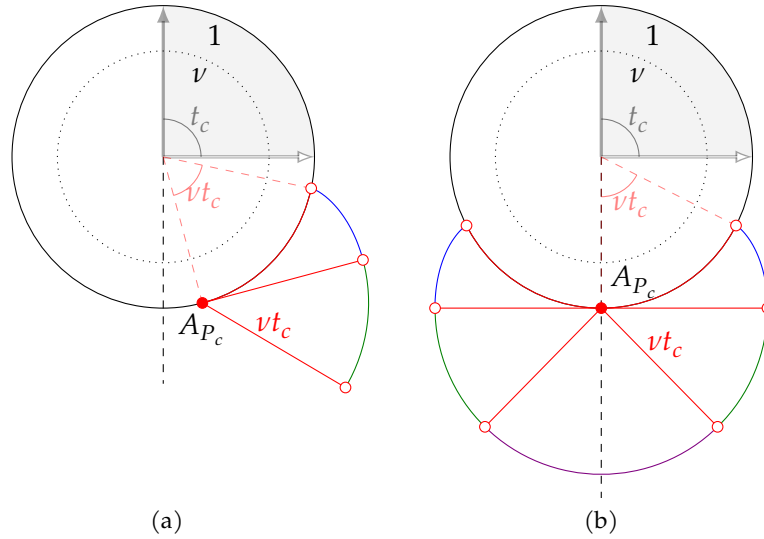
(a)   (b)

Figure 12.10: Delayed penetration case (a) away from the angle singularity
and (b) at the angle singularity. The curve shows the locus of
initial A_P positions which terminate at the point shown. Blue
sections are involutes, green sections are circular sectors corre-
sponding to the distance singularity, and the purple section in
(b) is a circular sector corresponding to the angle singularity.

Remark 19. For initial A_P positions which lie *inside* the locus shown in
Fig.12.10b the optimal Penetrator heading ψ_P^* is non-unique.

The non-uniqueness of A_P's control in this region is due to the fact
that the cost functional, (12.43), is upper-bounded by π. Termination at
the corner point $((r_P, \cos\theta_P) = (1, -1))$ achieves the upper-bound for
A_P's payoff; thus any trajectory which reaches the corner point within
t_c time is an optimal trajectory for A_P.

12.5.4   *Full Solution*

To summarize the solution of the Early Penetration optimal control
problem within the TRPDG, the regions for which each particular Pene-
trator control is optimal are constructed. The green region in Fig. 12.11
corresponds to the immediate penetration case (Section 12.5.1). For
initial A_P positions in the yellow region, the Penetrator aims clock-
wise and penetrates exactly at the moment A_R is neutralized at $t = t_c$
(see Section 12.5.2). Lastly, the red region represents all the positions
from which A_P can reach $(r_P, \cos\theta_P) = (1, -1)$. The "Max Value" re-

gion arises from the fact that A_P may penetrate the target at any time $t_f \in [0, t_c]$, and thus the distance and angle singularities, as shown in Fig. 12.10b, must be considered *for all possible* t_f. The red region in Fig. 12.11 is the union of all such regions. Inside the red region, the Penetrator may be able to reach $(1, -1)$ for a continuum of different t_f.
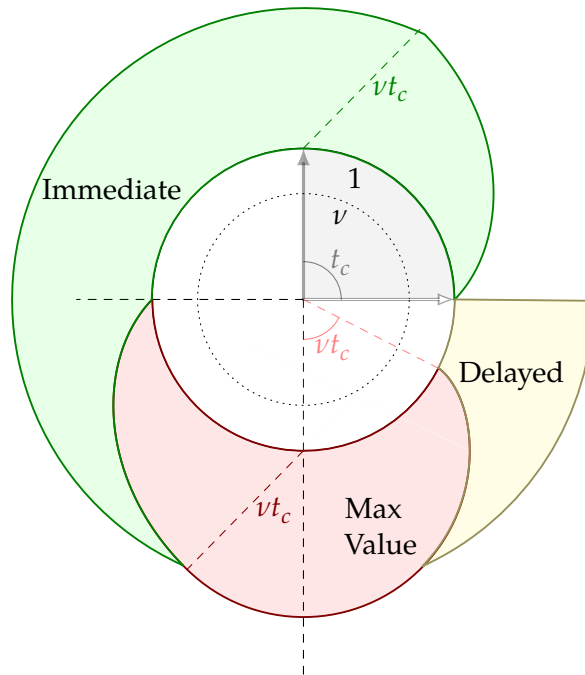
Figure 12.11: Solution of the Early Penetration optimal control problem for $t_c = \frac{\pi}{2}$.

12.6 ROLE SELECTION

In the previous sections the roles of A_R and A_P were considered to be fixed, *a priori*, to Runner and Penetrator, respectively. The Turret adhered to this convention by blindly pursuing the Runner, regardless of the position of the Penetrator. Moreover, the formulation was restricted such that the roles could not switch. Now, consider the more likely scenario in which the roles of the Attackers are not specified; thus the Turret gets to choose which Attacker to pursue (continuously throughout the game). Let the Attackers be specified generally as A_1 and A_2.

In order to specify the equilibrium "status" of the TRPDG policies in the context of this more general version of the problem, recall the definitions of the equilibrium concepts GSE, FNE, and subgame perfection (or time consistency) from Chapter 8, Section 8.3.1.

For the context of the TRPDG with role selection, consider a new action space for the agents. The Turret must only decide to turn CCW or CW at $t = 0$, which effectively determines which Attacker will be the Penetrator and which Attacker will be the Runner. From there, all 3 agents proceed according to their TRPDG equilibrium controls associated with this assignment (according to Section 12.4). Let the Value associated with the two assignments be defined

$$V_{1,2}\left(A_1, A_2\right) \equiv V\left(\begin{bmatrix} r_1 & \theta_1 & r_2 & \theta_2 & \beta \end{bmatrix}^\top\right) \tag{12.54}$$

$$V_{2,1}\left(A_2, A_1\right) \equiv V\left(\begin{bmatrix} r_2 & \theta_2 & r_1 & \theta_1 & \beta \end{bmatrix}^\top\right), \tag{12.55}$$

where V is defined in (12.18). Thus, e.g., $V_{2,1}$ is the Value of the TRPDG with A_2 assigned to Runner and A_1 assigned to Penetrator.

Lemma 12.13. *The TRPDG strategies given in Section 12.4 corresponding to* $\min\left(V_{1,2}, V_{2,1}\right)$ *constitute a GSE with the Turret as the leader and Attackers as the follower. The associated Value of the Stackelberg Game is* $V_S = \min\left(V_{1,2}, V_{2,1}\right)$, *where the subscript S denotes "Stackelberg".*

Proof. By construction, the strategies satisfy the first-order necessary conditions for equilibrium. Now it remains to show that the Value, V_S, is given by the minimum of the TRPDG Value associated with each assignment. The Value of the TRPDG, (12.42), is mathematically premised on the fact that the only way to terminate the game is by neutralizing the Runner (i.e., the terminal surface is $\theta_R = 0$, see (12.17)). Thus the formulation ignores the possibility of neutralization of the Penetrator (i.e., by driving $\cos\theta_P \to 1$). The TRPDG Value functions (without role selection) satisfy the saddle-point equilibrium property

$$J(u_T^*, \psi_i, \psi_j) \leq V_{i,j} \leq J(u_T, \psi_i^*, \psi_j^*), \qquad i, j \in \{1, 2\} \tag{12.56}$$

So, in the general scenario in which T can choose which Attacker to neutralize, only T's side (the left side) of the saddle-point property

holds – that is, T can do no worse than the TRPDG Value associated with each direction. Therefore, T is free to choose the smaller of the two. The Attackers' best response is to respond accordingly, assigning the Runner and Penetrator roles as dictated by T's choice. Thus T pursues A_1 if $V_{1,2} < V_{2,1}$, otherwise, it is better to pursue A_2 first. $\qquad\square$

It appears that the GSE is also the FNE everywhere in the state space. However, a rigorous proof of such for is left future work.
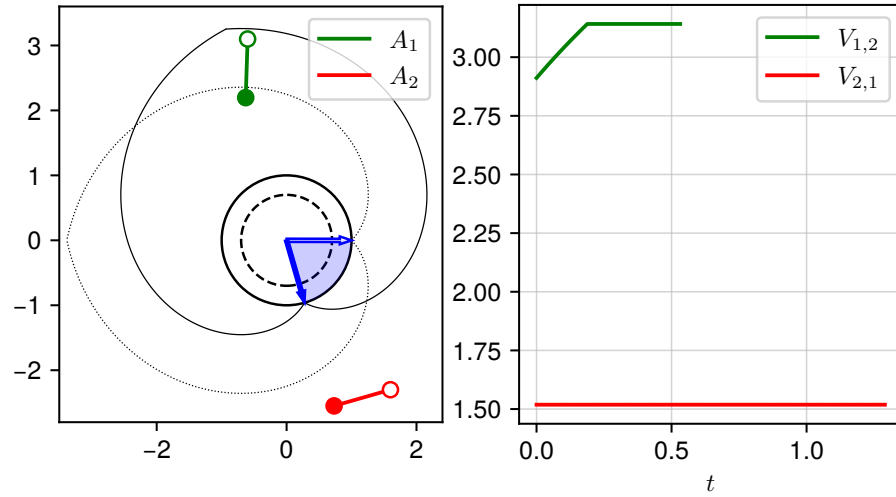
12.7   SIMULATIONS

In this section, some simulation results are presented wherein the prescribed strategies are implemented in a discrete-time fashion. In particular, the agents' control signal is held constant over a fixed time interval (which is denoted as Δt) and all agents' controls are updated synchronously as a function of the current state only. Neither agent has access to its opponents current control action (unlike in the Stackelberg version of the game).

The assignment of roles for the Attackers is not specified. Instead, the "equilibrium" action involves comparing $V_{1,2}$ and $V_{2,1}$ (at the current time) and implementing the TRPDG controls associated with the lesser Value (as described in the previous section). The Value of the "wrong" sequence (i.e., $\max\left(V_{1,2}, V_{2,1}\right)$) is of the utmost importance in demonstrating whether or not a GSE trajectory is subgame perfect (and thus also FNE). Thus, in Fig. 12.12, both values, $V_{1,2}$ and $V_{2,1}$, are computed along the entire trajectory.

As shown in Fig. 12.12a, when all 3 agents implement the GSE strategy (corresponding to A_2 being Runner and A_1 being Penetrator) the Value $V_{2,1}$ is constant and remains the smaller of the two Values throughout the playout of the game. This trajectory is subgame perfect, and thus the GSE and FNE are equivalent for all points along this trajectory.

Figure 12.12b demonstrates what happens when T deviates by pursuing A_1 first. Without knowing T's current or future control actions, the Attackers proceed with implementing the equilibrium strategy (treating A_2 as the Runner). Around $t = 0.5$, the TRPDG Values of the two

(a) Equilibrium – all agents implement controls associated with min $\left(V_{1,2}, V_{2,1}\right)$.
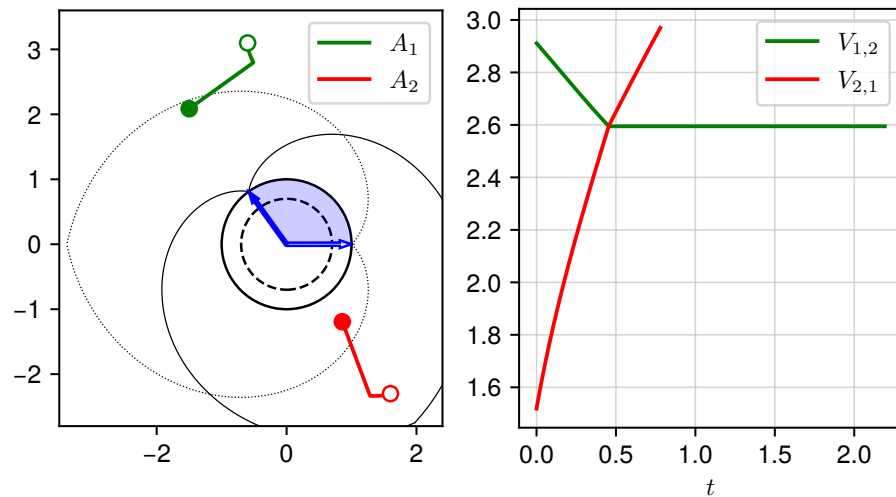
(b) Attackers implement equilibrium strategy, and the Turret chooses to pursue A_1 first.

Figure 12.12: Simulation results; $\nu = 0.7$, $\Delta t = 1e-3$. Once the Attacker designated as Runner crosses into $\mathscr{R}_A$, the associated Value of the TRPDG does not exist, which is the reason for the reason for, e.g., $V_{1,2}$ stopping early in (a).

assignments cross, and immediately afterwards the Attackers switch to A_1 being Runner and A_2 being Penetrator. The Attackers, as the maximizers, receive a significant gain over the equilibrium Value *without knowing what T will do or whether it will suddenly switch*. This is a symptom of the saddle-point property of the equilibrium. However, this example, alone, is far from sufficient to prove that the GSE and FNE are equivalent everywhere.

12.8 CONCLUSION

In this chapter, the two-Attacker, single-Turret circular target guarding problem was introduced. The focus was on a region of the state space in which neither Attacker can guarantee to reach the target, individually. The case where one Attacker can guarantee to reach the target when the Turret pursues its fellow Attacker was considered. Within this case, a differential game was posed and solved which terminates when the Runner is neutralized, and an optimal control problem for when the Penetrator can reach the target before the Runner is neutralized was posed and solved. Most of the analysis was done under the assumption that Attackers' roles were predefined and that no switching could occur. This assumption was later lifted, and it was shown how the solution already obtained can be used to determine the "best" roles for the Attackers and that this corresponds to the GSE. The next chapter, rather than adding more agents, considers the twist of multiple choices for the Attacker.

"Sometimes, people only seem determined upon one course because the have been offered no other options." – Sazed

— Brandon Sanderson, Mistborn Trilogy

ENGAGE OR RETREAT WITH A KINETIC TURRET

13.1 INTRODUCTION

In this chapter, the Attacker is given the option to engage with the kinetic turret (which ends with the Attacker reaching the circle centered on the Defender) or to retreat to a safe zone (i.e., the TEoR with kinetic turret,). Research Objectives 8 and 9, which pertain to multi-objective scenarios and discrete events, respectively, are partly addressed by the work in this chapter. The latter comes into play in the sense that the moment at which the Defender's LOS aligns with the Attacker serves as a discrete event which is paramount to the determination of the equilibrium behaviors. Here, the Attacker is not neutralized when this event occurs, but rather, the Attacker may begin accruing damage once this alignment has occurred. This material is based on the paper [249], which is in review.

Section 13.2 provides the problem formulation and solution methodology. The derivations for engagement with lock-on and retreat with lock-on are contained in Sections 13.3.2 and 13.3.3, respectively. Sections 13.4.1 and 13.4.2 cover engagement without lock-on and retreat without lock-on, respectively. Section 13.5 discusses how to obtain the state space partitioning and provides examples. The chapter is concluded in Section 13.6.

Engagement Zone
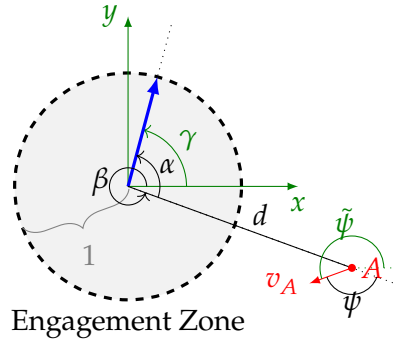
Retreat Zone

Figure 13.1: Lock-Evade, Engage or Retreat Scenario with relative and Cartesian state representations.

13.2   PROBLEM FORMULATION AND SOLUTION METHODOLOGY

Begin by defining the state in a relative coordinate system; $\mathbf{z} \equiv [d, \alpha, \beta] \in \mathbb{R}^3$, which is comprised of the distance from A to D, d, the look angle of D w.r.t. A, α, and the azimuth of A's position relative to D w.r.t. the positive x-axis, β. Figure 13.1 shows a diagram of the TEoRK scenario. Some of the subsequent analysis is eased by utilizing the Cartesian coordinate system. Let $\tilde{\mathbf{z}} \equiv [x, y, \gamma]$ be the state of the system expressed in the Cartesian frame, where x and y are the coordinates of A and γ is D's look angle w.r.t. the positive x-axis. D is positioned at the origin of the Cartesian coordinate system. The transformation between the two state representations is

$$\begin{bmatrix} x \\ y \\ \gamma \end{bmatrix} = \begin{bmatrix} d \cos \beta \\ d \sin \beta \\ \beta + \alpha \end{bmatrix}. \tag{13.1}$$

A has a fixed speed (normalized to 1) and controls its instantaneous heading, $\psi \in \mathbb{R}$ (i.e. simple motion, or single integrator dynamics). D is stationary and has control over its turn rate, $\omega \in [-\rho, \rho]$ where $\rho > 1$ is the maximum turn rate. Positive ω corresponds to D turning counterclockwise. D has an additional control variable $w \in [0, \bar{w}]$, which appears in A's cost functional. It represents the amount of turret firepower to apply; the turret must be aimed directly at A for it to

be effective. Let A and D's control vectors be defined, respectively, as $\mathbf{u}_A \equiv [\psi]$ and $\mathbf{u}_D \equiv [\omega, w]$. In the relative coordinate system, the kinematics are

$$f(\mathbf{z}, \mathbf{u}_A, \mathbf{u}_D) = \dot{\mathbf{z}} = \begin{bmatrix} \dot{d} \\ \dot{\alpha} \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} \cos\psi \\ \omega - \frac{1}{d}\sin\psi \\ \frac{1}{d}\sin\psi \end{bmatrix}. \tag{13.2}$$

The kinematics, expressed in the Cartesian coordinate system, are

$$\tilde{f}(\tilde{\mathbf{z}}, \mathbf{u}_A, \mathbf{u}_D) = \dot{\tilde{\mathbf{z}}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\gamma} \end{bmatrix} = \begin{bmatrix} \cos\tilde{\psi} \\ \sin\tilde{\psi} \\ \omega \end{bmatrix}, \tag{13.3}$$

where $\tilde{\psi} \equiv \beta + \psi$ is A's heading w.r.t. the positive x-axis. The agents are assumed to have full state information.

Ultimately, A chooses between two endings for the overall TEoRK scenario: engagement or retreat. In the former, A moves towards D and ultimately collides with it. Let the terminal surface for engagement be defined as

$$\mathcal{E} \equiv \{\mathbf{z} \mid d = 1\} \tag{13.4}$$

Alternatively, A bypasses D and maneuvers towards a retreat zone, which is specified *a priori*. Here, the retreat zone is the surface $y = y_R$, and thus the terminal surface for retreat is defined as

$$\mathcal{R} \equiv \{\mathbf{z} \mid d\sin\beta = y_R\} \tag{13.5}$$

The region of admissible initial conditions is defined as

$$\Omega \equiv \{\mathbf{z} \mid d\sin\beta > y_R, \, d > 1\}. \tag{13.6}$$

Within both the engage and retreat cases, D may or may not achieve a lock on A. Lock occurs when D's look angle is aligned with A's position (i.e., when $\cos\alpha = 1$). Note that because $\rho > 1$ and $d > 1$, D has an angular velocity advantage over all of Ω; once a lock is achieved, D has sufficient control authority to keep $\cos\alpha = 1$. Thus there are four cases: engagement wherein D achieves a lock on A at some point (Locked Engagement, LE), retreat with a lock (Locked Retreat, LR),

engagement without a lock (Unlocked Engagement, UE), retreat without a lock (Unlocked Retreat, UR).

The lock function is defined as

$$L(\mathbf{z}) = \begin{cases} 1 & \text{if } \cos \alpha = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{13.7}$$

The cost functionals for A and D are defined, respectively, as

$$J_A\left(\mathbf{z}_0; \mathbf{u}_A(\cdot), \mathbf{u}_D(\cdot)\right) = \Psi_A\left(\mathbf{z}_f\right) + \tag{13.8}$$
$$\int_{t_0}^{t_f} L\left(\mathbf{z}(t)\right)\left(w(t) + c\right) \mathrm{d}t$$
$$J_D\left(\mathbf{z}_0; \mathbf{u}_A(\cdot), \mathbf{u}_D(\cdot)\right) = \Psi_D\left(\mathbf{z}_f\right), \tag{13.9}$$

where $c > 0$ is a constant which penalizes A for time spent locked-on. The terminal cost functions are defined as

$$\Psi_A\left(\mathbf{z}_f\right) = \begin{cases} 0 & \mathbf{z}_f \in \mathcal{E} \\ c_A & \mathbf{z}_f \in \mathcal{R} \end{cases} \tag{13.10}$$

$$\Psi_D\left(\mathbf{z}_f\right) = \begin{cases} c_D & \mathbf{z}_f \in \mathcal{E} \\ 0 & \mathbf{z}_f \in \mathcal{R} \end{cases}, \tag{13.11}$$

where $c_A > 0$ is a constant penalty given to A for retreating instead of engaging, and $c_D > 0$ is a constant penalty given to D if it is destroyed. A and D simultaneously seek to minimize their respective cost functionals, giving rise to a nonzero-sum differential game,

$$J_A^*\left(\mathbf{z}_0; \mathbf{u}_D(\cdot)\right) = \min_{\mathbf{u}_A(\cdot)} J_A\left(\mathbf{z}_0; \mathbf{u}_A(\cdot), \mathbf{u}_D(\cdot)\right) \tag{13.12}$$

$$J_D^*\left(\mathbf{z}_0; \mathbf{u}_A(\cdot)\right) = \min_{\mathbf{u}_D(\cdot)} J_D\left(\mathbf{z}_0; \mathbf{u}_A(\cdot), \mathbf{u}_D(\cdot)\right). \tag{13.13}$$

D strictly prefers retreat and therefore seeks to make retreat as attractive as possible for A by making engagement as costly as possible. Note that the integral cost in (13.8) is nonzero only while D has a lock on A. While A can evade D's turret, the integral cost is zero.

For the cases in which D is unable to achieve a lock on A, $L = 0$ for all $t \in \left[t_0, t_f \right]$ and thus (13.8) simplifies to $J_A \left(\mathbf{z}_0; \mathbf{u}_A (\cdot), \mathbf{u}_D (\cdot) \right) = \Psi_A \left(\mathbf{z}_f \right)$. Therefore A's cost is constant: either 0, if $\mathbf{z}_f \in \mathcal{E}$ or c_A, if $\mathbf{z}_f \in \mathcal{R}$. As long as A can guarantee arrival at $\mathcal{E}$ or $\mathcal{R}$ with $\cos \alpha \neq 1$ for all $t \in \left[t_0, t_f \right]$ the optimal actions for A and D are not uniquely defined. The construction of trajectories for A which can guarantee satisfaction of this constraint is the subject of Section 13.4.

As stated previously, the goal of this chapter is to determine the equilibrium actions for both A and D consisting of the choice of whether to engage or retreat as well as the instantaneous heading angle and turning direction, respectively, throughout the scenario. Ultimately, the solution consists of a partitioning over the state space corresponding to the four different ways the scenario can terminate. Within a particular region, the end-game of the scenario is known and the necessary conditions for equilibrium yield the equilibrium control inputs. The boundaries between the regions are generally found by considering the most limiting case for a particular type of termination and backwards-integrating the associated equilibrium kinematics until the specified initial condition. For example, boundary between Locked and Unlocked Retreat corresponds to the limiting case of LR in which lock-on occurs at the moment A reaches the Retreat Zone. Figure 13.2 shows an example solution along with all of the pertinent features, as described in Table 13.1. Note the presence of several DSs and other curves for which there are multiple equally-optimal choices for direction or even termination.

13.3 LOCKED SCENARIOS

In this section, the cases in which D is able to achieve a lock on A are solved. In order for this to occur, D must be able to drive $\cos \alpha$ to 1 before A is able to reach the retreat surface, $\mathcal{R}$, or reach the engagement surface, $\mathcal{E}$.

In the following two subsections, the cases wherein D achieves a lock on A, Locked Engagement and Locked Retreat, are analyzed individ-
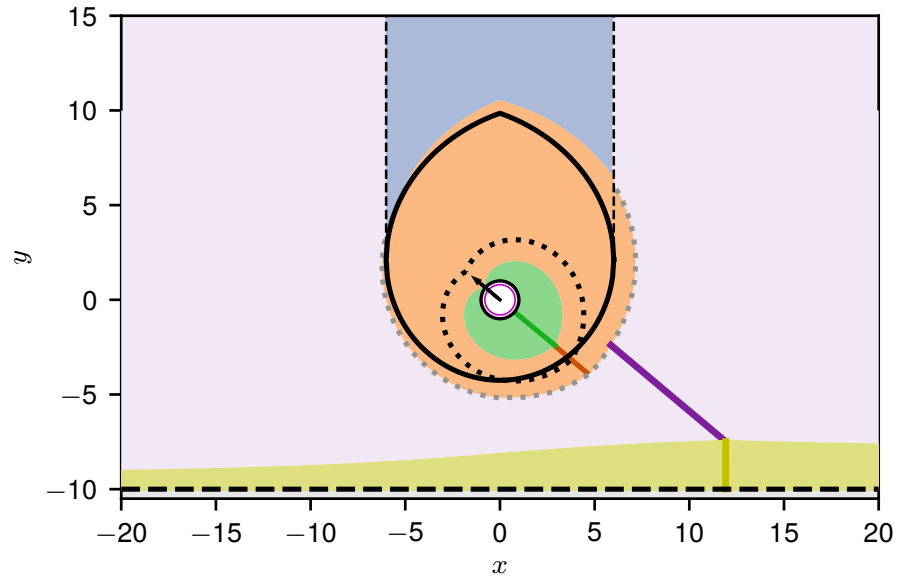
Figure 13.2: A partitioning of the state space for a particular initial turret look angle, $\gamma_0 = \frac{7}{9}\pi$. Table 13.1 contains the legend information along with mathematical notation for each feature.

Table 13.1: Solution legend

Legend	Symbol	Description
– – –	$\mathscr{R}$	Retreat Surface
——	$V_E = V_R$	Simple Engage or Retreat surface
——	$1/\rho$	Surface of equal max angular velocity
▢	LR_1	Locked Retreat ending in $\mathscr{R}_{R_1}$
▢	LR_2	Locked Retreat ending in $\mathscr{R}_{R_2}$
▢	LE	Locked Engagement
▢	UR	Unlocked Retreat
▢	UE	Unlocked Engagement
····	$LE = LR$	Engagement and Retreat equally optimal
——	$\mathscr{D}_{LR_1}$	Locked Retreat DS
——	$\mathscr{D}_{LE}$	Locked Engagement DS
——	$\mathscr{D}_{UR}$	Unlocked Retreat DS
——	$\mathscr{D}_{UE}$	Unlocked Engagement DS
····	$LE = c_A$	LE Value equal to retreat penalty

ually. Here, the general solution approach is described which will be specialized for each of these cases. The LE and LR cases will be solved by splitting the game into a pre-lock and post-lock segment, solving each segment individually, and stitching the solutions together. A similar approach is taken in Chapter 12 wherein a 2-Attacker turret defense scenario is broken up into a segment prior to and after the termination of the first Attacker. See also [217] where a partial information scenario is solved as individual phases (or segments). In all of these cases, the agents employ controls which put them in the most advantageous configuration for subsequent segments.

For the cases in which D achieves a lock on A, the first time at which $\cos \alpha = 1$ is defined as $t_l \in (t_0, t_f]$. In general, the subscript l refers to conditions evaluated at $t = t_l$. Thus the lock function $L = 0$ for $t \in [t_0, t_l)$ and $L = 1$ for $t \in [t_l, t_f]$. The pre-lock segment ends at t_l when $\cos \alpha = 1$; at this point, neither agent has accrued any cost (c.f. (13.8) and (13.9)) For the pre-lock segment, a terminal-valued zero-sum differential game [130] is defined with a terminal value, $\Phi(\mathbf{z}_l)$, based on the resulting Value of the post-lock segment initialized at $\mathbf{z}_l \equiv \mathbf{z}(t_l)$,

$$\Phi_l(\mathbf{z}_l) = V_{EoR}(\mathbf{z}_l) \tag{13.14}$$

with A as the minimizer and D as the maximizer, along with the terminal boundary condition

$$\phi_l(\mathbf{z}_l, t_l) = \alpha_l = 0. \tag{13.15}$$

The Value function for the post-lock segment, V_{EoR}, corresponds to the canonical engage or retreat game [90] and will be defined precisely in Section 13.3.1 (c.f. (13.37)). If an equilibrium exists, the Value function of the pre-lock differential game satisfies

$$V_l(\mathbf{z}) = \min_{\mathbf{u}_A(\cdot)} \max_{\mathbf{u}_D(\cdot)} \Phi_l(\mathbf{z}_l, t_l). \tag{13.16}$$

Because the Value of the pre-lock segment depends on the Value of the post-lock segment, (13.14)–(13.16) may be thought of as a one-step dynamic programming problem.

The solution methodology proceeds with the formation of the Hamiltonian

$$\mathcal{H} = \dot{\mathbf{z}}^\top \boldsymbol{\lambda} = \lambda_d \cos\psi + \lambda_\alpha \left(\omega - \frac{1}{d}\sin\psi \right) + \lambda_\beta \frac{1}{d}\sin\psi, \qquad (13.17)$$

where $\boldsymbol{\lambda} \equiv \begin{bmatrix} \lambda_d & \lambda_\alpha & \lambda_\beta \end{bmatrix}^\top$ is a vector of adjoint variables. The adjoint dynamics are given by [40] $\dot{\boldsymbol{\lambda}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{z}}$, which simplifies to

$$\dot{\lambda}_d = \left(\lambda_\beta - \lambda_\alpha \right) \frac{1}{d^2}\sin\psi, \;\; \dot{\lambda}_\alpha = 0, \;\; \dot{\lambda}_\beta = 0, \qquad (13.18)$$

thus λ_α and λ_β are constant. At $t = t_l$, the adjoint vector is [40]

$$\boldsymbol{\lambda}^\top (t_l) = \frac{\partial \Phi_l}{\partial \mathbf{z}_l} + \nu \frac{\partial \phi_l}{\partial \mathbf{z}_l}, \qquad (13.19)$$

where ν is another adjoint variable. The value of the Hamiltonian at the time of lock is [40]

$$\mathcal{H} (t_l) = -\frac{\partial \Phi_l}{\partial t_l} - \nu \frac{\partial \phi_l}{\partial t_l} \qquad (13.20)$$

When the functions Φ_l and ϕ_l do not depend on terminal time $\mathcal{H}_l = 0$, and since the system kinematics are time-autonomous, $\mathcal{H}(t) = 0$ for all $t \in [t_0, t_l]$.

A and D's equilibrium control strategies minimize and maximize the Hamiltonian, respectively:

$$\psi^* = \arg\min_{\psi} \mathcal{H}, \qquad \omega^* = \arg\max_{\omega} \mathcal{H},$$

which becomes

$$\cos\psi^* = \frac{-\lambda_d}{\sqrt{\lambda_d^2 + \frac{\left(\lambda_\alpha - \lambda_\beta\right)^2}{d^2}}}, \;\; \sin\psi^* = \frac{\lambda_\alpha - \lambda_\beta}{d\sqrt{\lambda_d^2 + \frac{\left(\lambda_\alpha - \lambda_\beta\right)^2}{d^2}}} \qquad (13.21)$$

$$\omega^* = \rho \cdot \operatorname{sign}\lambda_\alpha. \qquad (13.22)$$

The value of the adjoint variable ν is generally determined by evaluating the Hamiltonian, (13.17), at lock time and substituting in (13.19)–(13.22).

In each case, sign λ_α is not fully determined by the above first order necessary conditions for equilibrium. Consequently, there arises a DS (c.f. [20, 130]) which partitions the region of admissible initial conditions into a region where D either moves CCW or CW. The region of admissible initial conditions may be constructed via backwards integration of the equilibrium kinematics ((13.21) and (13.22) substituted into (13.2)) from the terminal manifold (or some particular limiting manifold, depending on the case). The following Lemma greatly simplifies the process of constructing these regions.

Lemma 13.1. *The Attacker's equilibrium trajectory for the pre-lock segment of the TEoRK game, $t \in [0, t_l]$, is a straight line in the Cartesian frame.*

Proof. Consider the Cartesian version of the Hamiltonian,

$$\mathscr{H}^{\sim} = \dot{\tilde{\mathbf{z}}}^\top \tilde{\boldsymbol{\lambda}} = \lambda_x \cos \tilde{\psi} + \lambda_y \sin \tilde{\psi} + \lambda_\gamma \omega, \qquad (13.23)$$

where $\tilde{\boldsymbol{\lambda}} \equiv \begin{bmatrix} \lambda_x & \lambda_y & \lambda_\gamma \end{bmatrix}^\top$ is a vector of Cartesian adjoint variables. As in the relative coordinate system, the equilibrium adjoint dynamics are given by [40] $\dot{\tilde{\boldsymbol{\lambda}}} = -\frac{\partial \mathscr{H}^{\sim}}{\partial \tilde{\mathbf{z}}}$, which simplifies to $\dot{\tilde{\boldsymbol{\lambda}}} = \mathbf{0}$ since the state components, $\tilde{\mathbf{z}}$ do not appear in (13.23). Again, the equilibrium Attacker strategy minimizes the Cartesian Hamiltonian, $\tilde{\psi}^* = \arg\min_\psi \mathscr{H}^{\sim}$:

$$\cos \tilde{\psi}^* = \frac{-\lambda_x}{\sqrt{\lambda_x^2 + \lambda_y^2}}, \quad \sin \tilde{\psi}^* = \frac{-\lambda_y}{\sqrt{\lambda_x^2 + \lambda_y^2}} \qquad (13.24)$$

Since $\dot{\lambda}_x = \dot{\lambda}_y = 0$ it must also be that $\dot{\tilde{\psi}}^* = 0$ and therefore the Cartesian Attacker heading, $\tilde{\psi}^*$, is constant. A similar style proof appears in Chapter 12. $\qquad \square$

Thus, A's initial position can be obtained using its terminal Cartesian heading $\tilde{\psi}_l^* = \psi_l^* + \beta_l$ for a particular t_l. This result is used repeatedly throughout the remainder to obtain boundaries which partition the space of initial A positions into regions where, e.g., retreat or engagement is optimal. For a limiting terminal configuration (e.g., $\cos \alpha = 1$, $d = 1$), the necessary conditions for equilibrium yield $\tilde{\psi}_l^*$ and ω^* and the length of A's trajectory comes from the time for D to turn back to its starting position.

13.3.1   *The Post-Lock Engage or Retreat Game*

This section of the paper addresses the definition and solution of the post-lock segment of the TEoRK game, which is needed to solve the pre-lock segment (as dictated by (13.14)). The Engage or Retreat Game with an Attacker moving with simple motion and a Defender whose control appears directly in the Attacker's integral cost was formulated and solved in [84]. This corresponds to D having a lock on A over the entire game, i.e., $L = 1$ for all $t \in [t_0, t_f]$.

When A chooses engagement, D incurs its maximum cost and has "lost" in some sense. So, generally, D seeks to make engagement as costly for A as possible to make retreat a more attractive option for A. Thus the agents play a zero-sum differential game with J_A as a cost functional; the associated Value function of the differential Game of Engagement (GoE) is defined as

$$V_E\left(\mathbf{z}_0\right) = \min_{\mathbf{u}_A(\cdot)} \max_{\mathbf{u}_D(\cdot)} J_A\left(\mathbf{z}_0; \mathbf{u}_A(\cdot), \mathbf{u}_D(\cdot)\right), \qquad (13.25)$$

with the constraint $\phi\left(\mathbf{z}_f\right) = d_f - 1 = 0$. The solution of the game is comprised of the following expression for the Value function [84]

$$V_E\left(\mathbf{z}\right) = \left(\bar{w} + c\right)\left(d - 1\right) \qquad (13.26)$$

along with the state-feedback equilibrium strategies [84]

$$\psi^*\left(\mathbf{z}\right) = \pi, \qquad w^*\left(\mathbf{z}\right) = \bar{w}, \qquad (13.27)$$

which corresponds to A aiming directly at the D and D applying its maximum defense.

For retreat, D wants to encourage A to continue retreating and thus seeks to minimize A's cost. The Value function for the associated optimal control problem is defined as

$$V_R\left(\mathbf{z}_0\right) = \min_{\mathbf{u}_A(\cdot), \mathbf{u}_D(\cdot)} J_A\left(\mathbf{z}_0; \mathbf{u}_A(\cdot), \mathbf{u}_D(\cdot)\right), \qquad (13.28)$$

with the constraint $\phi\left(\mathbf{z}_f\right) = d_f \sin \beta_f - y_R = 0$ which corresponds to $y_f = y_R$. An additional constraint is necessary in order to ensure the validity of the OCR trajectory [84, 248]:

$$V_E\left(\mathbf{z}(t)\right) - V_R\left(\mathbf{z}(t)\right) \geq 0 \ \ \forall t \in \left[t_0, t_f\right], \qquad (13.29)$$

which essentially requires that engagement must be as costly or more costly than retreating along the entire retreat trajectory. The manifold $V_E(\mathbf{z}) = V_R(\mathbf{z})$ partitions Ω into a region where engagement is optimal and a region where retreat is optimal. The latter may be further partitioned depending on whether or not A must maneuver around the engagement region in order to retreat (which occurs when (13.29) is activated). Let the partitioning of Ω be defined as

$$\mathscr{R}_E = \{\mathbf{z} \mid V_E(\mathbf{z}) < V_R(\mathbf{z})\} \qquad (13.30)$$

$$\mathscr{R}_{R_1} = \{\mathbf{z} \notin \mathscr{R}_E \mid |x| > x_2 \text{ or } y < y_2\} \qquad (13.31)$$

$$\mathscr{R}_{R_2} = \Omega \ \left(\mathscr{R}_E \cup \mathscr{R}_{R_1}\right) \qquad (13.32)$$

where

$$x_2 = \frac{\bar{w} + c + c_A - c y_R}{\sqrt{(\bar{w} + c)^2 - c^2}}, \ y_2 = \frac{c\left(\bar{w} + c + c_A - c y_R\right)}{(\bar{w} + c)^2 - c^2}. \qquad (13.33)$$

Figure 13.3 shows the partitioning of the state space for the parameter settings used throughout the remainder of the paper. In $\mathscr{R}_{R_1}$, the solution is comprised of the Value function expression [90]

$$V_{R_1} \equiv V_R\left(\mathbf{z}\right) = c\left(d \sin \beta - y_R\right) + c_A, \ \mathbf{z} \in \mathscr{R}_{R_1} \qquad (13.34)$$

along with the state-feedback equilibrium strategies [90]

$$\psi^*\left(\mathbf{z}\right) = \frac{3\pi}{2} - \beta, \qquad w^*\left(\mathbf{z}\right) = 0, \qquad (13.35)$$

which corresponds to A running straight down and D holding fire. In $\mathscr{R}_{R_2}$, the solution is comprised of the Value function expression

$$V_{R_2} \equiv V_R\left(\mathbf{z}\right) = c t_f^* + c_A, \ \mathbf{z} \in \mathscr{R}_{R_2} \qquad (13.36)$$
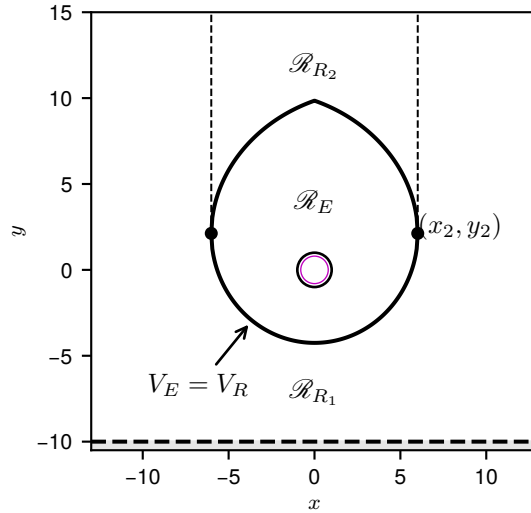
Figure 13.3: Engage or retreat regions. $c = 0.5$, $c_A = 2$, $\rho = 0.8$, $\bar{w} = 1$, and $y_R = -10$.

where t_f^* is determined numerically; the details are contained in [90].

Finally, the overall Value function for the post-lock segment of the TEoRK game is defined as

$$V_{EoR}\left(\mathbf{z}\right) \equiv \begin{cases} V_E\left(\mathbf{z}\right) & \mathbf{z} \in \mathscr{R}_E \\[2ex] V_{R_1}\left(\mathbf{z}\right) & \mathbf{z} \in \mathscr{R}_{R_1} \\[2ex] V_{R_2}\left(\mathbf{z}\right) & \mathbf{z} \in \mathscr{R}_{R_2} \end{cases} \tag{13.37}$$

13.3.2 *Locked Engagement*

In this section, the case in which the TEoRK game terminates in engagement is analyzed. Again, the post-lock segment corresponds to the classical Engage or Retreat Game discussed in the previous section. Thus in order for engagement to be optimal Attacker behavior, it is necessary for the state of the system at the time of lock-on to be in the engage region. Define the region for which lock-on occurs and engagement is optimal in the TEoRK game as

$$LE \equiv \left\{\mathbf{z} \mid \mathbf{z}_l^* \in \mathscr{R}_E\right\}. \tag{13.38}$$

Note the region LE is conditioned on $\mathbf{z}_l^*$ – the (equilibrium) state of the system when D locks onto A, which is the terminal state for the pre-lock segment. Therefore, the first order necessary conditions for equilibrium, derived below, will be utilized to construct the region LE via backwards integration.

By construction $\mathbf{z}_l \in \mathscr{R}_E$ and thus the terminal cost function is obtained by substituting (13.26) into (13.14)

$$\Phi_l\left(\mathbf{z}_l, t_l\right) = V_E\left(\mathbf{z}_l\right) = \left(\bar{w} + c\right)\left(d_l - 1\right), \tag{13.39}$$

with the terminal boundary condition given in (13.15). The terminal adjoint values are obtained by substituting (13.39) and (13.15) into (13.19)

$$\lambda_{d_l} = \bar{w} + c, \quad \lambda_\alpha = \nu, \quad \lambda_\beta = 0 \tag{13.40}$$

Evaluating (13.21) and (13.22) at $t = t_l$ and substituting in (13.40) yields the following pre-lock terminal equilibrium controls

$$\cos \psi_l^* = \frac{-\left(\bar{w} + c\right)}{\sqrt{\left(\bar{w} + c\right)^2 + \frac{\nu^2}{d_l^2}}}, \quad \sin \psi_l^* = \frac{\nu}{d_l \sqrt{\left(\bar{w} + c\right)^2 + \frac{\nu^2}{d_l^2}}} \tag{13.41}$$

$$\omega^* = \rho \operatorname{sign} \nu. \tag{13.42}$$

Note ω^* is constant over $t \in [0, t_l]$ because ν is constant. Evaluating (13.17) at $t = t_l$, substituting in the terminal adjoints (13.40) and terminal equilibrium controls (13.41), (13.42), and setting equal to 0 allows ν to be obtained:

$$\nu = \frac{\pm\left(\bar{w} + c\right) d_l}{\sqrt{\rho^2 d_l^2 - 1}} \tag{13.43}$$

Lemma 13.2. *The terminal equilibrium controls for Locked Engagement are given by*

$$\cos \psi_l^* = \frac{-\sqrt{\rho^2 d_l^2 - 1}}{\rho d_l}, \quad \sin \psi_l^* = \frac{-\operatorname{sign}\left(\sin \alpha_0\right)}{\rho d_l} \tag{13.44}$$

$$\omega^* = -\rho \operatorname{sign}\left(\sin \alpha_0\right), \tag{13.45}$$

Proof. Begin by proving that sign $\nu = -\text{sign}(\sin \alpha_0)$. D wishes to maximize (13.39) and thus seeks to drive $\alpha \to 0$ with maximum d. Meanwhile, A minimizes (13.39) and wishes to terminate as close to D as possible. From (13.21) $\cos \psi^* < 0$ and thus d is monotonically decreasing over $t \in [0, t_l]$. Therefore, D must drive $\alpha \to 0$ as quickly as possible, which corresponds to turning at its maximum angular rate towards A: $\omega^* = -\rho \, \text{sign}(\sin \alpha_0)$. Equating this expression to (13.42) yields sign $\nu = -\text{sign}(\sin \alpha_0)$. Substituting this along with (13.43) into (13.41) and (13.42) yields (13.44) and (13.45). □

Corollary 13.1. *In the region LE there exists a DS wherein the Defender and Attacker may either turn CCW or CW and achieve the same Value in the TEoRK game, and is defined by*

$$\mathcal{D}_{LE} \equiv \{ \mathbf{z} \mid \mathbf{z} \in LE, \; \cos \alpha = -1 \}. \tag{13.46}$$

Proof. The DS $\mathcal{D}_{LE}$ arises due to the symmetry in the problem geometry and the equilibrium controls. When $\cos \alpha = -1$ it must be that $\sin \alpha = 0$ and thus the direction of the equilibrium controls, (13.44) and (13.45) is undefined. The pre-lock segment terminates at a particular d_l with $\alpha_l = 0$. Thus, the initial state wherein $\cos \alpha = -1$ can be reached via backwards integration with CCW or CW motion. These two trajectories have the same Value since V_E is only a function of d. This proof is similar to one used in [2]; the interested reader is referred therein for further detail. □

Figure 13.4 shows an example pair of trajectories emanating from the DS $\mathcal{D}_{LE}$. When the state of the system begins on $\mathcal{D}_{LE}$, the Defender may choose either CCW or CW without suffering any penalty w.r.t. the Value of the game. The Attacker's control becomes uniquely defined the instant the state of the system departs from $\mathcal{D}_{LE}$.

13.3.3 *Locked Retreat*

There are two regions, $\mathscr{R}_{R_1}$ and $\mathscr{R}_{R_2}$, in which A would choose to retreat. Trajectories beginning in $\mathscr{R}_{R_1}$ are unconstrained, while those be-
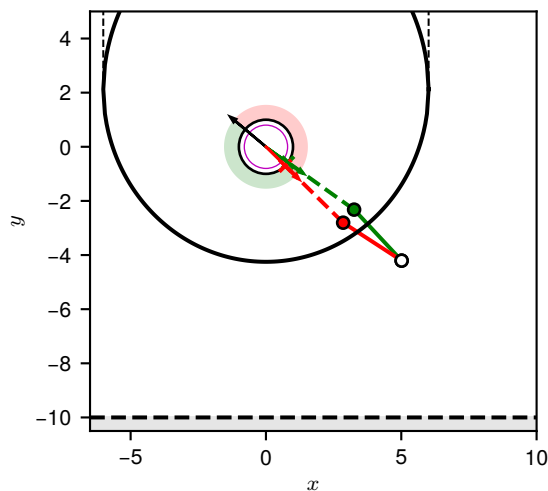
Figure 13.4: Example DS trajectories emanating from $\mathcal{D}_{LE}$ corresponding to D choosing CCW (green) or CW (red). The open circle is the initial A position, closed circles are the A position when lock-on occurs. The post-lock trajectories are shown in dashed lines and an × marks A's position at t_f. Initially, D's turret is aimed along the black vector.

ginning in $\mathcal{R}_{R_2}$ are constrained. The solutions will be discussed in Sections 13.3.4 and 13.3.5, respectively.

13.3.4   *LR ending in $\mathcal{R}_{R_1}$*

Similar to the previous section, define a region where retreat is optimal in the TEoRK game, lock occurs, and the retreat trajectory is unconstrained (i.e., (13.29) remains inactive):

$$LR_1 \equiv \left\{ \mathbf{z} \mid \mathbf{z}_l^* \in \mathcal{R}_{R_1} \right\}. \tag{13.47}$$

Substituting (13.34) into (13.14) yields the terminal cost function

$$\Phi_l\left(\mathbf{z}_l, t_l\right) = V_{R_1}\left(\mathbf{z}_l\right) = c\left(d_l \sin \beta_l - y_R\right) + c_A, \tag{13.48}$$

with the terminal boundary condition given in (13.15).

Lemma 13.3. *The terminal equilibrium controls for Locked Retreat ending in* $\mathscr{R}_{R_1}$ *are given by*

$$\cos \psi_l^* = \frac{-c \sin \beta_l}{\chi}, \quad \sin \psi_l^* = \frac{-c \cos \beta_l + \frac{\nu}{d_l}}{\chi} \tag{13.49}$$

$$\omega^* = \rho \operatorname{sign} \nu, \tag{13.50}$$

where

$$\nu = -\frac{c d_l \cos \beta_l}{\rho^2 d_l^2 - 1} \pm c d_l \sqrt{\frac{c d_l \cos^2 \beta_l}{\left(\rho^2 d_l^2 - 1\right)^2} + \frac{1}{\rho^2 d_l^2 - 1}}, \tag{13.51}$$

$$\chi = \sqrt{c^2 \sin^2 \beta_l + \left(c \cos \beta_l - \frac{\nu}{d_l}\right)^2}. \tag{13.52}$$

Proof. Substituting (13.48) and (13.15) into (13.19) yields the terminal adjoint values

$$\lambda_{d_l} = c \sin \beta_l, \quad \lambda_\alpha = \nu, \quad \lambda_\beta = c d_l \cos \beta_l. \tag{13.53}$$

Substituting (13.53) into (13.21) and (13.22) at $t = t_l$ yields the terminal pre-lock equilibrium controls, (13.49) and (13.50). The pre-lock terminal equilibrium control expressions are then substituted into the Hamiltonian, (13.17), and evaluated at $t = t_l$:

$$\mathscr{H}^*(t_l) = -\sqrt{c^2 \sin^2 \beta_l + \left(c \cos \beta_l - \frac{\nu}{d_l}\right)^2} + \rho |\nu| = 0, \tag{13.54}$$

where the last equality follows from the discussion after (13.20). Then, (13.51) is obtained by solving (13.54) for ν. $\qquad \square$

Corollary 13.2. *In the region* LR_1 *there exists a* DS, $\mathcal{D}_{LR_1}$, *wherein the Defender and Attacker may either turn CCW or CW and achieve the same Value in the* TEoRK *game.*

Proof. The logic is similar to Corollary 13.1, however, the polar symmetry from the LE case is lost in the LR case since A is ultimately heading for the retreat zone. Hence there is no analytic expression for the surface $\mathcal{D}_{LR_1}$, and as a result, the surface is computed numerically. $\qquad \square$

The DS $\mathcal{D}_{LR_1}$ partitions LR_1 into two regions wherein D has either CCW motion or CW motion (which corresponds to the two sign possibilities for ν in (13.51)). Computation of $\mathcal{D}_{LR_1}$ is thus useful for determining the sign of ν for particular terminal conditions. This computation is accomplished via coupled backwards shooting process adapted from Chapter 14 and demonstrated in Section 13.4.2 wherein pairs of trajectories are obtained which have equal Value and integrate back to the same initial condition. Figure 13.5 gives an example pair of trajectories which are initialized on a point along $\mathcal{D}_{LR_1}$. Like the Locked Engagement DS, $\mathcal{D}_{LE}$, A can respond with its corresponding equilibrium control the instant the state leaves $\mathcal{D}_{LR_1}$.
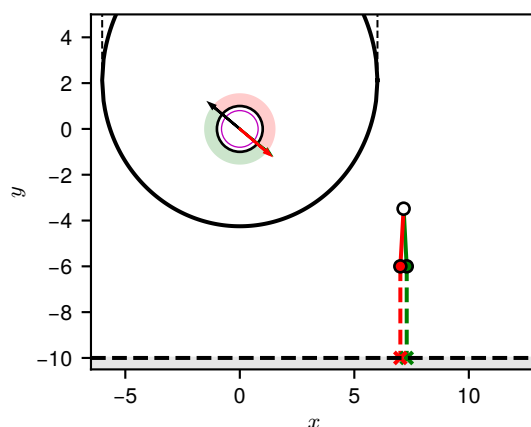

Figure 13.5: Representative trajectories emanating from the DS $\mathcal{D}_{LR_1}$. The CCW (red) and CW (green) trajectories have the same Value, which, in this case, means lock-on occurs at the same y coordinate.

Special care must be taken when backwards integrating as the state of the system may enter a region where Locked Engagement becomes optimal. Figure 13.6 shows an example trajectory for LR and LE, $\mathbf{z}_{LR}$ and $\mathbf{z}_{LE}$, respectively, which have the same Value. The manifold for which the Value of LE and LR are identical is akin to the $V_E = V_R$ manifold shown in Fig. 13.3 except that it is a 2D manifold in the 3D state space.
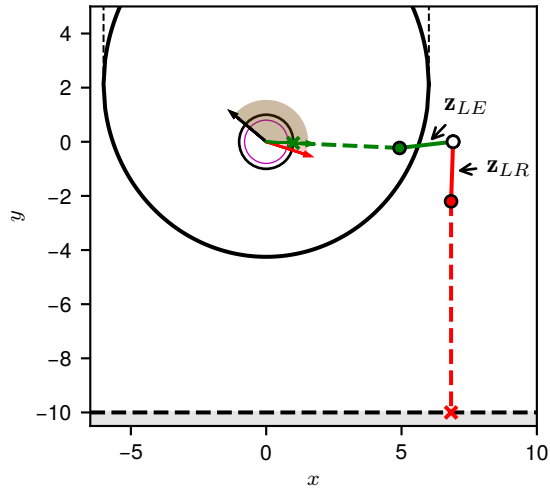
Figure 13.6: Example trajectories for LE (green) and LR (red) with the same Value.

13.3.5   *LR ending in $\mathscr{R}_{R_2}$*

As in previous sections, define a region where retreat is optimal in the TEoRK game, lock occurs, and the retreat trajectory is constrained (i.e., (13.29) is activated):

$$LR_2 \equiv \left\{ \mathbf{z} \mid \mathbf{z}_l^* \in \mathscr{R}_{R_2} \right\}. \tag{13.55}$$

For convenience, the analysis is done in the Cartesian frame for this case. Substituting (13.36) into (13.14) yields the terminal cost function

$$\Phi_l\left(\tilde{\mathbf{z}}_l, t_l\right) = V_{R_2}\left(\tilde{\mathbf{z}}_l\right) = ct_f^* + c_A. \tag{13.56}$$

In the Cartesian frame, the terminal manifold (13.15) becomes

$$\phi\left(\tilde{\mathbf{z}}_l, t_l\right) = \gamma_1 - \arctan\left(\frac{y_l}{x_l}\right) = 0. \tag{13.57}$$

The solution for this case is complicated by the lack of an analytical expression for t_f^*, which is the optimal time for A to reach the retreat zone while satisfying (13.29). From [90], the OCR trajectory for A is comprised of a straight segment which is tangent to the $V_R = V_E$ manifold, a curved segment which rides along the manifold, and a straight segment departing from (x_2, y_2) terminating at (x_2, y_R). Let $\tilde{\psi}_R^*$ be the

optimal retreat heading for A at $t = t_l$ which is tangent to the $V_R = V_E$ manifold. Then, specializing (13.19) the pre-lock terminal adjoint values are

$$\lambda(t_l) = \frac{\partial V_{R_2}}{\partial \tilde{\mathbf{z}}} + \nu \frac{\partial \phi_l}{\partial \tilde{\mathbf{z}}} \tag{13.58}$$

$$\implies \lambda_{x_l} = -c \cos \tilde{\psi}_R^* + \nu \frac{y_l}{x_l^2 + y_l^2} \tag{13.59}$$

$$\lambda_{y_l} = -c \sin \tilde{\psi}_R^* - \nu \frac{x_l}{x_l^2 + y_l^2} \tag{13.60}$$

$$\lambda_{\gamma_l} = \nu. \tag{13.61}$$

The constrained retreat Value partial derivatives $\frac{\partial V_{R_2}}{\partial \tilde{\mathbf{z}}}$ are derived as follows. Let $\sigma \equiv \frac{\partial V_{R_2}}{\partial \tilde{\mathbf{z}}} = \begin{bmatrix} \sigma_x & \sigma_y & \sigma_\gamma \end{bmatrix}^\top$. The focus is on the post-lock segment wherein lock-on has occurred, and thus γ no longer has any bearing on the optimality since D has sufficient control authority to maintain the lock-on (i.e., keep $\cos \alpha = 1$). Therefore, $\sigma_\gamma = 0$. Thus the Hamiltonian for the retreat case in the Cartesian frame is

$$\mathcal{H}_R = \sigma_x \cos \tilde{\psi}_R + \sigma_y \sin \tilde{\psi}_R \tag{13.62}$$

Here, the terminal cost function is simply $\Phi_R \left(\tilde{\mathbf{z}}_f, t_f \right) = c t_f + c_A$, and the terminal manifold is $\phi_R \left(\tilde{\mathbf{z}}_f, t_f \right) = y_f - y_R = 0$ [90]. The value of $\mathcal{H}_R$ at the post-lock termination is [40]

$$\mathcal{H}_R(t_f) = -\frac{\partial \Phi_R}{\partial t_f} - \mu \frac{\partial \phi_R}{\partial t_f} = -c. \tag{13.63}$$

The Attacker must minimize $\mathcal{H}_R$, and thus

$$\cos \tilde{\psi}_R^* = \frac{-\sigma_x}{\sqrt{\sigma_x^2 + \sigma_y^2}}, \quad \sin \tilde{\psi}_R^* = \frac{-\sigma_y}{\sqrt{\sigma_x^2 + \sigma_y^2}}. \tag{13.64}$$

Evaluating (13.64) at $t = t_f$ and substituting into (13.62) gives

$$\mathcal{H}_R(t_f) = -\sqrt{\sigma_x^2 + \sigma_y^2} = -c. \tag{13.65}$$

Substituting (13.65) into (13.64) yields the desired expressions:

$$\sigma_x = -c \cos \tilde{\psi}_R^*, \quad \sigma_y = -c \sin \tilde{\psi}_R^*. \tag{13.66}$$

Lemma 13.4. *The terminal equilibrium controls for Locked Retreat ending in $\mathscr{R}_{R_2}$ are given by*

$$\cos \tilde\psi_l^* = \frac{-\lambda_{x_l}}{\sqrt{\lambda_{x_l}^2 + \lambda_{y_l}^2}}, \quad \sin \tilde\psi_l^* = \frac{-\lambda_{y_l}}{\sqrt{\lambda_{x_l}^2 + \lambda_{y_l}^2}} \tag{13.67}$$

$$\omega^* = \rho \operatorname{sign} \nu \tag{13.68}$$

where $\lambda_{x_l}, \lambda_{y_l}, \lambda_{\gamma_l}$ are given in (13.59)–(13.61), $\tilde\psi_R^*$ is computed numerically, and

$$\nu = \frac{-b \pm \sqrt{b^2 - 4ac^2}}{2a} \tag{13.69}$$

where

$$a = \left(\frac{1}{x_l^2 + y_l^2} - \rho^2 \right) \tag{13.70}$$

$$b = 2c \left(\frac{y_l \cos \tilde\psi_R^* - x_l \sin \tilde\psi_R^*}{x_l^2 + y_l^2} \right). \tag{13.71}$$

Proof. Considering, again, the Cartesian Hamiltonian, (13.23), and the fact that A minimizes this Hamiltonian, (13.67) is directly obtained (c.f. Lemma 13.1; the vector $\begin{bmatrix} \cos \tilde\psi^* & \sin \tilde\psi^* \end{bmatrix}^\top$ must be anti-parallel with the vector $\begin{bmatrix} \lambda_x & \lambda_y \end{bmatrix}^\top$). Similarly, ω^* maximizes $\mathscr{H}^\sim$, and (13.68) is obtained. Substituting (13.67) and (13.68) into (13.23) and evaluating at $t = t_l$ yields

$$\mathscr{H}_l^* = -\sqrt{\lambda_{x_l}^2 + \lambda_{y_l}^2} + \rho|\nu| = 0 \tag{13.72}$$

where, again, the latter equality is discussed after (13.20). Substituting in the terminal adjoint values (13.59)– (13.61) gives

$$-\sqrt{\left(c \cos \tilde\psi_R + \nu \frac{y_l}{x_l^2 + y_l^2} \right)^2 + \left(c \sin \tilde\psi_R - \nu \frac{x_l}{x_l^2 + y_l^2} \right)^2} \\ + \rho|\nu| = 0, \tag{13.73}$$

which simplifies to the quadratic form:

$$\underbrace{\left(\frac{1}{x_l^2 + y_l^2} - \rho^2 \right)}_{a} \nu^2 + \underbrace{2c \left(\frac{y_l \cos \tilde\psi_R - x_l \sin \tilde\psi_R}{x_l^2 + y_l^2} \right)}_{b} \nu + c^2 = 0, \tag{13.74}$$

where a and b have been defined as the coefficients of the ν terms, corresponding to (13.70) and (13.71). Thus the solution for ν is given by the quadratic formula, as shown in (13.69). Note that in the state space of interest, $d = \sqrt{x^2 + y^2} > \frac{1}{\rho}$ and thus $a < 0$. So the term $4ac^2$ is negative and thus $\sqrt{b^2 - 4ac^2} > b$. Therefore, the two solutions for ν have opposite signs. $\qquad\square$
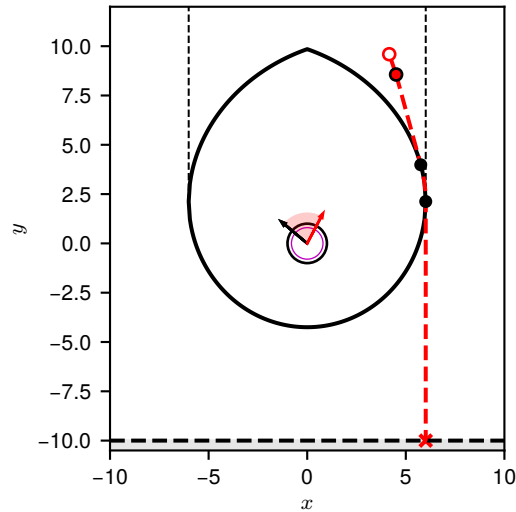

Figure 13.7: Example trajectory for LR ending in $\mathscr{R}_{R_2}$. The initial turret look angle is shown by the black arrow. A starts at the open circle, gets locked-onto at the closed red circle, enters the constrained arc at the upper black circle, leaves the $V_E = V_R$ manifold at the lower black circle, and reaches the retreat zone at the ×.

13.4   UNLOCKED SCENARIOS

In this section, regions are obtained for which A can be guaranteed to reach the engagement surface, $\mathcal{E}$, or the retreat surface, $\mathcal{R}$, without D being able to achieve a lock-on. According to the *actual* cost functional, (13.8), any trajectory ending on the engagement surface for which lock-on does not occur (and thus $L = 0 \ \forall t \in [0, t_f]$) results in an Attacker cost of 0. Likewise, any trajectory ending on the retreat surface for which lock-on does not occur results in an Attacker cost of c_A. Therefore, the Attacker strategy which can achieve one or other outcomes is not necessarily unique. Consider, for example, an initial A position which is very close to the retreat surface, and D is looking away; there

could be an infinite number of ways for A to reach the retreat surface prior to lock-on. Thus, within each of these two regions, an auxiliary differential game is formulated and solved in order to obtain an admissible Attacker trajectory.

In both cases, the terminal cost functional considered is the terminal separation angle:

$$\Phi_U \left(\mathbf{z}_f \right) = |\alpha_f|, \tag{13.75}$$

where the subscript f denotes conditions at terminal time, t_f, and the subscript U denotes "unlocked". The limiting case for which $\alpha_f = 0$ yields the boundary of the UE and UR regions. The auxiliary Value function for the unlocked scenarios is defined as,

$$V_U(\mathbf{z}) = \max_{\mathbf{u}_A(\cdot)} \min_{\mathbf{u}_D(\cdot)} \Phi_U(\mathbf{z}_f). \tag{13.76}$$

Note that, unlike in the locked formulation (c.f. (13.16)), A is the maximizer and D is the minimizer. By maximizing $|\alpha_f|$ the Attacker is trying to guarantee the largest "margin" between itself and D at terminal time. Alternatively, one may consider the terminal time to be the cost/payoff as was done in Chapter 11.

13.4.1 *Unlocked Engagement*

For UE, the Attacker terminates on $\mathcal{E}$ wherein $d_f = 1$. Define a region where engagement is optimal in the TEoRK game and lock-on does not occur:

$$UE \equiv \left\{ \mathbf{z} \mid \mathbf{z}_f^* \in \mathcal{E}, \ L(\mathbf{z}(t)) = 0 \ \forall t \in \left[0, t_f \right] \right\} \tag{13.77}$$

The auxiliary differential game ((13.2) with (13.75) and (13.76)) is identical to the *Game of Angle* in Chapter 11 with a scaling in time (i.e., $\frac{1}{\rho}$ is the equivalent speed ratio).

Lemma 13.5 (UE Region [254, 255])**.** *The region for which A can reach $\mathcal{E}$ before D can achieve lock-on is given by*

$$UE = \{ \mathbf{z} \mid |\alpha| > \alpha_{UE} \}, \tag{13.78}$$

where

$$\alpha_{UE} = \sqrt{\rho^2 d^2 - 1} + \sin^{-1}\left(\frac{1}{\rho d}\right) - \sqrt{\rho^2 - 1} - \sin^{-1}\left(\frac{1}{\rho}\right) \quad (13.79)$$

Lemma 13.6 (UE Auxiliary Equilibrium Policies [254, 255]). *The equi-librium policies for the UE auxiliary differential game are*

$$\cos\psi_{UE}^* = \sqrt{1 - \frac{1}{\rho^2 d^2}}, \ \sin\psi_{UE}^* = -\operatorname{sign}\alpha\left(\frac{1}{\rho d}\right) \quad (13.80)$$

$$\omega_{UE}^* = -\rho\operatorname{sign}\alpha. \quad (13.81)$$

For UE, the Attacker terminates on the target circle and thus the β state component has no effect on the optimality in the auxiliary dif-ferential game. As was shown in Chapter 11, there exists a dispersal surface due to the symmetry:

$$\mathcal{D}_{UE} \equiv \{\mathbf{z} \in UE \mid \cos\alpha = -1\}. \quad (13.82)$$

From (13.46) it is clear that the locked engagement dispersal surface, $\mathcal{D}_{LE}$, is coincident with $\mathcal{D}_{UE}$ at the boundary of UE and is, in fact, a continuation thereof.

13.4.2 *Unlocked Retreat*

For UR, the Attacker terminates on $\mathcal{R}$ wherein $y_f = y_R$. As before, define a region where retreat is optimal in the TEoRK game and lock-on does not occur:

$$UR \equiv \left\{\mathbf{z} \mid \mathbf{z}_f^* \in \mathcal{R}, \ L(\mathbf{z}(t)) = 0 \ \forall t \in \left[0, t_f\right]\right\} \quad (13.83)$$

The associated terminal boundary condition is thus

$$\phi_{UR}\left(\mathbf{z}_f\right) = d_f \sin\beta_f - y_R = 0. \quad (13.84)$$

From the transversality condition [40], the adjoint values at terminal time are

$$\lambda^\top(t_f) = \frac{\partial \Phi_U}{\partial \mathbf{z}_f} + \nu \frac{\partial \phi_{UR}}{\partial \mathbf{z}_f} \tag{13.85}$$

$$\implies \lambda_{d_f} = \nu \sin \beta_f \tag{13.86}$$

$$\lambda_\alpha = \text{sign} \, \alpha_f \tag{13.87}$$

$$\lambda_\beta = \nu d_f \cos \beta_f. \tag{13.88}$$

Again, the Hamiltonian for this auxiliary UR game is constant since the dynamics are autonomous. Moreover,

$$\mathscr{H}(t_f) = -\frac{\partial \Phi_U}{\partial t_f} - \nu \frac{\partial \phi_{UR}}{\partial t_f} = 0, \tag{13.89}$$

and therefore, $\mathscr{H}(t) = 0 \; \forall t$.

Lemma 13.7. *The terminal equilibrium controls for the UR auxiliary differential game given by* (13.2), (13.75), (13.76), *and* (13.84) *are*

$$\cos \psi_f^* = \frac{\nu \sin \beta_f}{\zeta}, \quad \sin \psi_f^* = \frac{\nu \cos \beta_f - \frac{1}{d_f} \text{sign} \, \alpha_f}{\zeta} \tag{13.90}$$

$$\omega^* = -\rho \, \text{sign} \, \alpha_f, \tag{13.91}$$

where

$$\zeta = \sqrt{\nu^2 \sin \beta_f^2 + \left(\nu \cos \beta_f - \frac{1}{d_f} \text{sign} \, \alpha_f \right)^2}, \tag{13.92}$$

and

$$\nu = \frac{1}{d_f} \left(\text{sign} \, \alpha_f \cdot \cos \beta_f - \sqrt{\cos^2 \beta_f - 1 + d_f^2 \rho^2} \right) \tag{13.93}$$

Proof. Evaluating (13.17) at time t_f and substituting in the terminal adjoint values (13.86)–(13.88) yields

$$\mathscr{H}_f = \nu \sin \beta_f \cos \psi_f + \text{sign} \, \alpha_f \left(\omega_f - \frac{1}{d_f} \sin \psi_f \right) \tag{13.94}$$

$$+ \nu \cos \beta_f \sin \psi_f$$

$$= \left[\nu \sin \beta_f \quad \left(\nu \cos \beta_f - \frac{1}{d_f} \text{sign} \, \alpha_f \right) \right] \tag{13.95}$$

$$\left[\cos \psi_f \quad \sin \psi_f \right]^\top + \omega_f \, \text{sign} \, \alpha_f$$

Then (13.90) and (13.91) are obtained by maximizing and minimizing (13.95), respectively. Clearly, for ψ^*, the vector $\begin{bmatrix} \cos \psi_f^* & \sin \psi_f^* \end{bmatrix}$ must be parallel to the vector $\begin{bmatrix} \nu \sin \beta_f & \left(\nu \cos \beta_f - \frac{1}{d_f} \operatorname{sign} \alpha_f \right) \end{bmatrix}$. Substituting (13.89), (13.90), and (13.91) into (13.95) gives

$$\mathcal{H}_f^* = \sqrt{\nu^2 \sin^2 \beta_f + \left(\nu \cos \beta_f - \frac{1}{d_f} \operatorname{sign} \alpha_f \right)^2} - \rho = 0. \quad (13.96)$$

Solving for ν gives

$$\nu = \frac{1}{d_f} \left(\operatorname{sign} \alpha_f \cos \beta_f \pm \sqrt{\cos^2 \beta_f - 1 + d_f^2 \rho^2} \right). \quad (13.97)$$

Within the game space, Ω, the distance $d > 1$, and it was assumed that $\rho > 1$. Therefore, $\rho^2 d_f^2 > 1$, so the argument in the radical is always positive. Moreover, the radical term is larger in magnitude than the $\cos \beta_f$ term and thus dominates the sign of ν. Because $y_R \ll -1$ it must be that $d_f > 0$, that is, A has a component of velocity away from D. Also, from (13.2) and (13.90) it must be that $\dot{d}_f = \cos \psi_f^* \propto \nu \sin \beta_f$. The term $\sin \beta_f$ is strictly negative (again, since $y_R \ll -1$) which means that ν must also be strictly negative. Therefore, the $\pm$ sign in (13.97) can be replaced with a minus sign, since the positive version would make ν positive. $\square$

Remark 20. The Attacker control, (13.90), produces motion to the left or right in the Cartesian frame for positive or negative α_f (CW or CCW Defender motion), respectively.

Corollary 13.3. *In the region UR there exists a DS, denoted $\mathcal{D}_{UR}$, wherein the Defender may either turn CCW or CW and achieve the same Value in the auxiliary game.*

Proof. This DS is present for reasons similar to Corollaries 13.1 and 13.2, but is more similar to the latter in that it does not have a closed form analytic expression. $\square$

When considering a general initial state within UR, one does not necessarily know the conditions at final time – in particular, sign α_f is unknown, which determines whether D turns CCW or CW. Thus the DS

partitions UR into a region for which $\alpha_f > 0$ and $\alpha_f < 0$ (corresponding to D turning CW or CCW, respectively). It must be the case that the antipodal point of D's look-angle on the line $y = y_R$ is one of the endpoints of $\mathcal{D}_{UR}$:

$$\mathbf{z}_{end} \equiv \begin{bmatrix} |\frac{y_R}{\cos\gamma}| & \pi & \pi + \gamma \end{bmatrix} \in \partial\mathcal{D}_{UR} \tag{13.98}$$

where ∂ denotes the boundary of a set (or surface, in this case). At this point, the Value of the auxiliary game is $|\alpha_f| = \pi$, and both the CW and CCW solution trivially yield this Value since this point is also on the terminal surface $\mathcal{R}$. The other endpoint of $\mathcal{D}_{UR}$ corresponds to $|\alpha_f| = 0$, which also lies on ∂UR – the boundary of the Unlocked Retreat region. Thus, the DS $\mathcal{D}_{UR}$ can be parameterized by the Value of the auxiliary game, $|\alpha_f|$, for which the two equi-Valued trajectories naturally arise because of the absolute value (i.e., there is a positive and negative solution).

These facts are important as they allow the construction of $\mathcal{D}_{UR}$, given an initial look-angle γ_0, by employing concepts from numerical continuation (or homotopy) [3]. Let $\left\{\alpha_{f_k}\right\}_{k=0}^{N}$ be an ordered, nonincreasing set of α_f values where $\alpha_{f_0} = \pi$, $\alpha_{f_N} = 0$, and $N \in \mathbb{N}^+$. For each α_{f_k}, a coupled shooting problem is solved to find a pair of of trajectories which integrate back to the same initial conditions, $\begin{bmatrix} x_D & y_D & \gamma_0 \end{bmatrix}_k$. However, because of Lemmas 13.1 and 13.7, given a terminal state on $\mathcal{R}$ the associated initial state for a given γ_0 can be obtained analytically (without the need for numerical backwards integration as is generally the case for shooting). The DS, for a particular γ_0, is then approximated by

$$\mathcal{D}_{UR}(\gamma_0) \approx \{(x_D, y_D)\}_{k=0}^{N} \tag{13.99}$$

For a particular α_{f_k}, the following constraint satisfaction problem is solved:

$$x_r, \ x_l \ \text{s.t.}$$

$$\begin{bmatrix} x_{\mathcal{D}} & y_{\mathcal{D}} \end{bmatrix}_l^{\mathsf{T}} = \begin{bmatrix} x_l & y_R \end{bmatrix}^{\mathsf{T}} - \frac{\Delta\gamma_l}{\rho} \begin{bmatrix} \cos\psi^* & \sin\psi^* \end{bmatrix}_l^{\mathsf{T}}$$

$$\begin{bmatrix} x_{\mathcal{D}} & y_{\mathcal{D}} \end{bmatrix}_r^{\mathsf{T}} = \begin{bmatrix} x_r & y_R \end{bmatrix}^{\mathsf{T}} - \frac{\Delta\gamma_r}{\rho} \begin{bmatrix} \cos\psi^* & \sin\psi^* \end{bmatrix}_r^{\mathsf{T}} \qquad (13.100)$$

$$\begin{bmatrix} x_{\mathcal{D}} & y_{\mathcal{D}} \end{bmatrix}_r^{\mathsf{T}} = \begin{bmatrix} x_{\mathcal{D}} & y_{\mathcal{D}} \end{bmatrix}_l^{\mathsf{T}}$$

$$x_r > x_l$$

where subscripts r and l correspond to the right and left sides (in the Cartesian frame) of the DS (c.f. the preceding Remark). The quantities on the RHS are given by

$$\Delta\gamma_l = \gamma_0 + 2\pi - \operatorname{atan2}(y_R, \ x_l) - \alpha_{f_k} \qquad (13.101)$$

$$\Delta\gamma_r = -\gamma_0 + \operatorname{atan2}(y_R, \ x_r) - \alpha_{f_k} \qquad (13.102)$$

$$\nu_l = \frac{1}{d_l}\left(\frac{x_l}{d_l} - \sqrt{\frac{x_l^2}{d_l^2} - 1 + d_l^2\rho^2}\right) \qquad (13.103)$$

$$\nu_r = \frac{1}{d_r}\left(-\frac{x_r}{d_r} - \sqrt{\frac{x_r^2}{d_r^2} - 1 + d_r^2\rho^2}\right) \qquad (13.104)$$

$$\psi_l^* = \operatorname{atan2}\left(\frac{1}{d_l}(\nu_l x_l - 1), \ \frac{\nu_l y_R}{d_l}\right) \qquad (13.105)$$

$$\psi_r^* = \operatorname{atan2}\left(\frac{1}{d_r}(\nu_r x_r + 1), \ \frac{\nu_r y_R}{d_r}\right) \qquad (13.106)$$

The other important surface is the (upper) boundary of the Unlocked Retreat region, ∂UR (the lower is simply $y = y_R$). This surface is obtained by setting $\alpha_f = 0$ and integrating back from $\mathcal{R}$. It can be thought of as being computed in two parts: for $x_f \in \left(-\infty, x_{r_{max}}\right]$ set $\operatorname{sign}\alpha_f = -1$ to compute ν, and for $x_f \in \left[x_{l_{min}}, \infty\right)$ set $\operatorname{sign}\alpha_f = 1$. Figure 13.8 shows the Unlocked Retreat DS $\mathcal{D}_{UR}$, along with the boundary of the Unlocked Retreat region ∂UR.
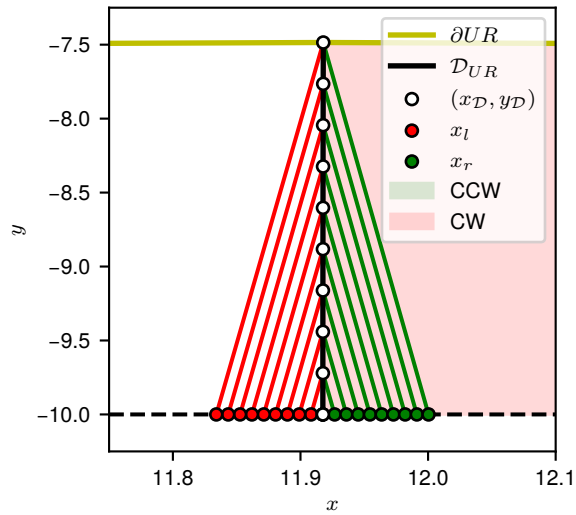
(a) Unlocked Retreat DS showing the result of solving (13.100) for each α_{f_k} for $N = 10$
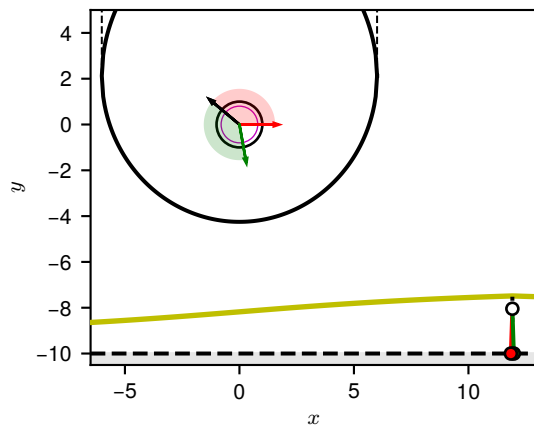
(b) Representative trajectories emanating from $\mathcal{D}_{UR}$ along with the associated Defender sweep angles

Figure 13.8: Unlocked Retreat DS

13.5 FULL SOLUTION

In this section the results of the previous sections are utilized to develop the solution over the whole state space. In particular, the state space is partitioned into 5 distinct regions in which a particular A behavior and termination is optimal – LE, LR ending in $\mathscr{R}_{R_1}$ or $\mathscr{R}_{R_2}$, UE, and UR. Along the boundaries of these regions, A may have two or more choices which yield the same Value, J_A^*, in the TEoRK game. Because the state space Ω is in 3D, the regions are 3D as well, and the boundaries are 2D manifolds in 3D space. In order to visualize the partitioning in a more meaningful way, the initial look angle of D's turret is fixed. Then the regions correspond to initial A positions (in 2D) and the boundaries become 1D. For a particular initial turret look angle, γ_0, the 1D boundaries are obtained by identifying the appropriate terminal manifold(s) and extending A's position backwards in time using the associated terminal equilibrium heading until $\gamma = \gamma_0$.

For Locked Engagement there are two critical configurations to consider: 1) lock-on occurs exactly when A reaches D ($d_l = 1$) and 2) lock-on occurs exactly when $V_E = V_R$. The first is critical in the sense that, for LE, it is required that lock-on actually occur; the limiting case is when it occurs just before the overall TEoRK game terminates. The boundary obtained via backwards integration from $d_l = 1$ divides the LE and UE regions – that is, if A is initialized *inside* this boundary, A can reach D and avoid being locked-on. The terminal manifold $V_E = V_R$ is critical because for engagement to be optimal, A must be in $\mathscr{R}_E$ by the time lock-on occurs; the limiting case is A just barely reaches $\mathscr{R}_E$ when $\cos\alpha \to 1$. Backwards integration from this terminal manifold divides the LE and LR regions.

For Locked Retreat, a distinction is made between LR ending in $\mathscr{R}_{R_1}$ or $\mathscr{R}_{R_2}$. LR ending in $\mathscr{R}_{R_1}$ has one critical terminal manifold: lock-on occurs exactly when A reaches the retreat zone ($y_l = y_R$). The associated boundary divides LR ending in $\mathscr{R}_{R_1}$ from UR – that is, if A is initialized *below* this boundary, A can reach y_R and avoid being locked-on. LR ending in $\mathscr{R}_{R_2}$ has one critical terminal manifold in which lock-on occurs exactly when A reaches the boundary between $\mathscr{R}_{R_1}$ and $\mathscr{R}_{R_2}$

$(|x_l| = x_2, y_l > y_2)$. The associated boundary merely distinguishes between LR ending in $\mathscr{R}_{R_2}$ versus $\mathscr{R}_{R_1}$.

The terminal manifolds identified above do not all have the same cost. Depending on the problem parameters $(\rho, c, \bar{w})$, points on the boundary between LE and LR must be obtained by integrating back from equi-Valued terminal manifolds (instead of the critical terminal manifolds identified above). The process is similar to the process described in Section 13.4.2 for computing the Unlocked Retreat DS. Figures 13.2 and 13.9 each show an example partitioning for a particular γ_0.
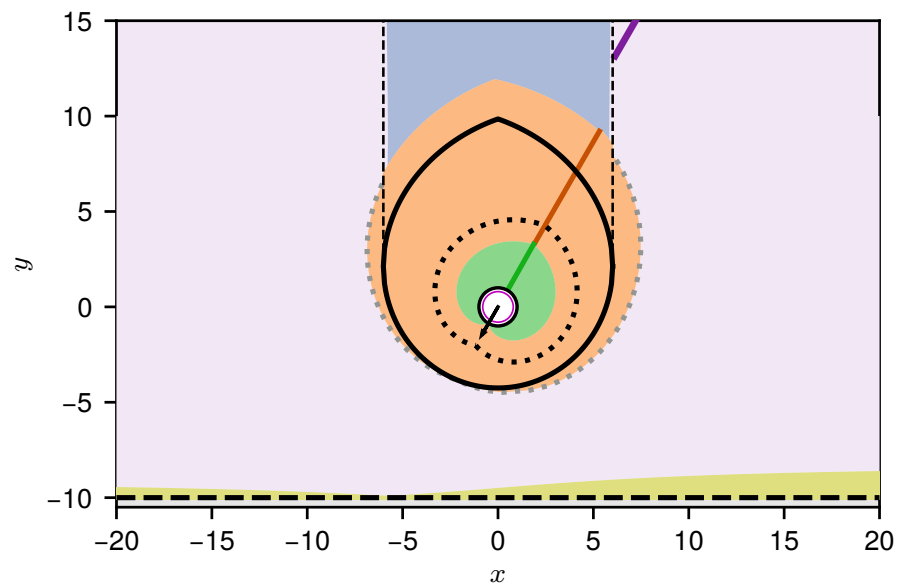
Figure 13.9: A partitioning of the state space for a particular initial turret look angle, $\gamma_0 = \frac{12}{9}\pi$. The legend is in Tab. 13.1.

Along with these examples, Fig. 13.10 is included in order to demonstrate the effect of the 3 parameters, ρ, c, and P, on the state space partitioning. As shown in Figs. 13.10a and 13.10e the nondimensional turret maximum turn rate, ρ, affects the *size* of the UE, LE, and UR regions; increasing ρ decreases their size and *vice versa*. This is because D is able to drive $\cos\alpha \to 1$ quicker which diminishes what A can accomplish in the pre-lock segment; in the limit as $\rho \to \infty$, the unlocked regions disappear and the scenario is identical to the original engage or retreat game in [90]. Increasing the time penalty, c, mainly results in stretching the $V_E = V_R$ surface, and thereby the LE region, vertically (see

Figs. 13.10b and 13.10f). Higher c means that time is of the essence and the total time cost begins to outweigh the damage that D can inflict, so engagement is more advantageous for the portion of the state space which is closer to D than to $\mathcal{R}$ (i.e., the space above D). Finally, increasing the retreat penalty, P, results in a larger LE region while essentially maintaining its shape. Again, this is because larger P means engagement is generally more favorable, however, the effect is consistent over the state space.

(a) $\rho = \frac{1}{0.5}$ (b) $c = 0.2$ (c) $P = 1$

(d) reproduction of Fig. 13.2

(e) $\rho = \frac{1}{0.95}$ (f) $c = 1$ (g) $P = 5$
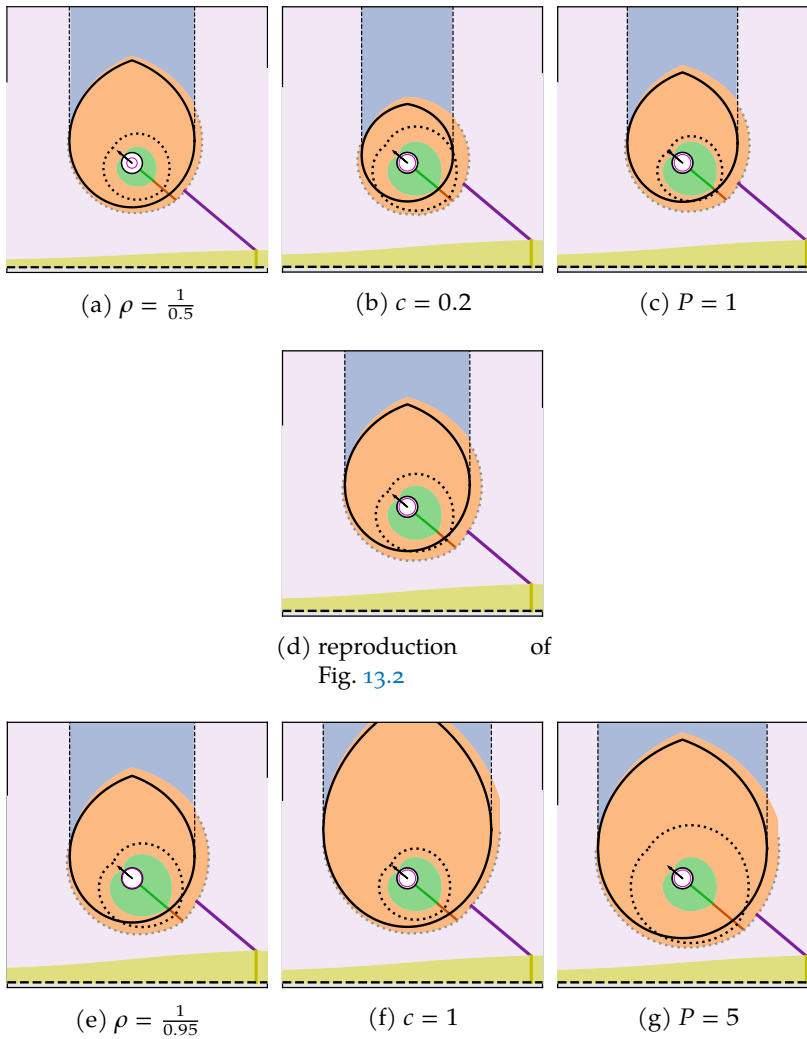
Figure 13.10: State space partitioning for varying sets of parameters. Unless indicated, the parameters match Fig. 13.2: $\rho = \frac{1}{0.8}, c = 0.5, P = 2$. The legend is in Tab. 13.1.

13.6   CONCLUSION

In this chapter the canonical engage or retreat game was extended by considering the Defender to have a turn-constrained turret. The cases in which the Defender could aim its turret onto the Attacker's position prior to the Attacker reaching either the engagement or retreat terminal surfaces were solved. For Locked Engagement and Locked Retreat, the equilibrium control strategies for the Attacker and Defender were derived as a function of the terminal state. Using backwards integration, these strategies were used to construct a partitioning of the state space whose regions correspond to the various terminal cases. Additionally, two auxiliary differential games were posed to derive unique Attacker controls for the cases in which the Defender could not aim its turret onto the Attacker prior to the latter reaching a terminal surface. The need for additional cost/payoff criteria (in this case, terminal look angle) arose because the original cost/payoff is the same for all such "unlocked" trajectories. The qualitative influence of each of the 3 parameters on the solution was demonstrated and discussed.

SINGLE ATTACKER AND A WIDE-BEAM TURRET

14.1 INTRODUCTION

Understanding singular surfaces within differential games and their relationship to the game's parameters is important (c.f. Section 2.4.5). The focus in this chapter is on numerical techniques for computing singular surfaces, for which there is a dearth of literature. As mentioned previously, there are three singular surfaces are present in the wide-beam version of the TDDG: (1) the Turret DS (TDS), (2) the Turret Universal Surface (TUS), and (3) the Attacker DS (ADS) [2]. These names describe the type of the surface as well as which agent has control authority on the surface. In the case of the DSs, the agent with control authority must make a choice between two (or more) optimal actions. Here, the Value function, which gives the saddle-point equilibrium cost of the game for a particular initial condition, is not differentiable [20]. The Value function is also not differentiable on the Turret Universal Surface; and in this case, the Turret's optimal control is not defined. Fortunately, both the Turret DS and the TUS can be handled analytically, as was done in [2]. Neither of these surfaces are affected by the parameter settings as they arise due to inherent symmetries in the game. The focus, in this chapter, is on the Attacker DS which does not have an analytical expression and thus must be treated numerically [see 191, §2.5.1]. In particular, it is shown that the parameters' settings affect not only the shape of the surface but also whether or not the surface is even present (thereby partly addressing Research Objectives 2 and 3). The numerical process developed to accomplish this task is enabled by collapsing the five natural parameters into two composite parameters. This material is based upon [247].

The chapter is organized as follows. Section 14.2 contains the problem formulation and details the conversion from the natural parameter

space to the collapsed parameter space. Section 14.3 formally describes DSs and includes necessary and sufficient conditions which are used in the following section. Section 14.4 details two different numerical approaches used to characterize the Attacker DS. Section 14.5 concludes the paper.
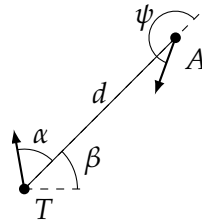
14.2 KINEMATICS AND SCALING

Figure 14.1: Coordinate system

Figure 14.1 shows the coordinate system used throughout the remainder of the paper. Note this is a relative coordinate system. As presented in [2], the coordinate system cannot be reduced further than three. Note, the 'global rotation' of the system, β, is necessary for the conversion of the relative states back to the natural states (i.e. Cartesian coordinates of both agents), but not necessary for computing the equilibrium trajectories of the other states. The motivation to reduce the parameter set is similar to the motivation for reducing the state dimension. Different parameter sets may yield quite different player behaviors across the state space. Characterizing these different regions of optimal play over the space of possible parameter settings is an important part of solving the game (in a full sense). Doing so for large (> 3) numbers of parameters may be difficult or intractable. Fortunately, it is often possible to reduce the number of parameters via various scalings and non-dimensional quantities.

For inspiration, consider the famous Buckingham π Theorem, which is based upon unit compatibility, and is used (particularly within the aerospace community) to derive physical laws centered around non-dimensional numbers (e.g., Reynolds number, Mach number, etc.). Without an understanding of Reynolds number, $Re = \frac{\rho u L}{\mu}$, one may unnec-

essarily repeat experiments with different length scales (L) and speeds (u), but constant Re.

Such a collapsing of physical quantities can also be seen in the HCDG wherein it is clear that the *ratio* of capture radius to Chauffeur turn radius is pertinent to the different solution regions. [164] characterized the entire parameter space, dividing it into 20 distinct regions with different characteristics.

Here, the collapsing of the physical parameters is facilitated by scaling time and distance. For notational convenience, the dimensional variables are barred; the non-barred variables denote the scaled quantities. The TDDG kinematics were expressed in [2] originally and are given here with a small correction and a simplification, along with the optimal costate kinematics:

$$
\begin{aligned}
\dot{\bar{d}} &= v_A \cos\bar{\psi}, & \bar{\psi} &\in [0, 2\pi] \\
\dot{\bar{\alpha}} &= \bar{\omega} - v_A \frac{1}{\bar{d}} \sin\bar{\psi}, & \bar{\omega} &\in [-\Omega, \Omega] \\
\dot{\bar{\beta}} &= v_A \frac{1}{\bar{d}} \sin\bar{\psi} \\
\dot{\lambda}_{\bar{d}} &= -\lambda_{\bar{\alpha}} v_A \frac{1}{\bar{d}^2} \sin\bar{\psi} \\
\dot{\lambda}_{\bar{\alpha}} &= -c_1 \frac{1}{2} \sin\bar{\alpha} \\
\dot{\lambda}_{\bar{\beta}} &= 0.
\end{aligned}
\tag{14.1}
$$

where $\bar{d}(\bar{t}_f) = d_c$ and $\bar{d}(\bar{t}) > d_c$, $\forall \bar{t} < \bar{t}_f$ – that is, the game terminates when the attacker reaches the capture distance d_c. From (14.1) and the boundary conditions, the four parameters are apparent: attacker speed v_A, maximum turret slew rate Ω, capture distance d_c, and cost parameter c_1. The fifth parameter c_2 appears in the cost function [2],

$$
\bar{J} = \int_{\bar{t}_0}^{\bar{t}_f} \left(c_1 \frac{1}{2} \left(1 + \cos(\bar{\alpha}(\bar{t}))\right) + c_2 \right) d\bar{t}
\tag{14.2}
$$

as well as in the equilibrium terminal value of λ_d [2],

$$
\lambda_{\bar{d}}(\bar{t}_f) = -\frac{c_1 \frac{1}{2}(1 + \cos(\bar{\alpha}(\bar{t}_f))) + c_2}{v_A}.
\tag{14.3}
$$

Fact 1. The cost in (14.2) is already dimensionless and altering its value with an additive offset or multiplicative scaling does not change the equilibrium strategies.

The presence of the parameters in (14.1)–(14.3) suggests the need to consider each of the equations when performing any type of scaling. The methodology employed here is described more fully by [140]. First, consider the following scaling of distance and time,

$$d = \frac{\bar{d}}{d_c}, \qquad t = \frac{\bar{t}}{t_c} \tag{14.4}$$

where d and t are the scaled/non-dimensional versions of distance and time, d_c is, as before, the capture distance, and t_c is a time constant representing a characteristic time associated with the problem. Let t_c be the time it would take the Attacker to travel a distance d_c in a straight line,

$$t_c = \frac{d_c}{v_A} \tag{14.5}$$

Note that the independent variable, time, has been scaled, and all of the states are now essentially dimensionless since α and β are angles. Scaling in time affects all of the states in (14.1), but scaling in distance requires special care in obtaining $\dot{d}$. The closing velocity can now be expressed in terms of non-dimensional distance and time,

$$\dot{\bar{d}} = \frac{\mathrm{d}\bar{d}}{\mathrm{d}\bar{t}} = \frac{\mathrm{d}(dd_c)}{\mathrm{d}(tt_c)} = \frac{d_c}{t_c}\frac{\mathrm{d}d}{\mathrm{d}t} = \frac{d_c}{t_c}\dot{d}$$
$$\implies \dot{d} = t_c\frac{1}{d_c}\dot{\bar{d}} \tag{14.6}$$

For α and β states,

$$\dot{\bar{\theta}} = \frac{\mathrm{d}\bar{\theta}}{\mathrm{d}\bar{t}} = \frac{\mathrm{d}\bar{\theta}}{\mathrm{d}(tt_c)} = \frac{1}{t_c}\frac{\mathrm{d}\theta}{\mathrm{d}t}$$
$$\implies \dot{\theta} = t_c\dot{\bar{\theta}}, \qquad \theta = \alpha, \beta \tag{14.7}$$

as a form for the dynamics. The relationship between the scaled (non-barred) versions of the angles and their natural counterparts is simply,

$$\theta(t) = \bar{\theta}(\bar{t}), \qquad \theta = \alpha, \beta \tag{14.8}$$

Substituting (14.1), (14.4), and (14.5) into (14.6) and (14.7) yields the following scaled state kinematics,

$$
\begin{aligned}
\dot{d} &= t_c \frac{1}{d_c} v_A \cos \psi = \cos \psi, & \psi &\in [0, 2\pi], \\
\dot{\alpha} &= t_c \left(\bar{\omega} - \frac{v_A}{\bar{d}} \sin \psi \right) = \omega - \frac{1}{d} \sin \psi, & \omega &\in [-\rho, \rho], & (14.9) \\
\dot{\beta} &= t_c v_A \frac{1}{\bar{d}} \sin \psi = \frac{1}{d} \sin \psi,
\end{aligned}
$$

where,

$$
\rho = \frac{\Omega d_c}{v_A}. \tag{14.10}
$$

Concerning the c_1 and c_2 parameters, let $c_1 = 1$ without loss of generality (see Fact 1), and $c_2 = c$. The "natural" cost $\bar{J}$ may be recovered by $\bar{J} = c_1 J$. Define the set of admissible initial conditions as

$$
\mathbf{X} := \{ \mathbf{x} = (d, \alpha, \beta) \mid d \geq 1 \}. \tag{14.11}
$$

Now, the scaled Hamiltonian is written, using new, non-barred versions of the costates, as,

$$
\begin{aligned}
\mathscr{H} = {}& \lambda_d \cos \psi + \lambda_\alpha \left(\omega - \frac{1}{d} \sin \psi \right) + \\
& \lambda_\beta \frac{1}{d} \sin \psi - \frac{1}{2}(1 + \cos \alpha) - c
\end{aligned} \tag{14.12}
$$

The non-barred costate kinematics are then obtained by taking the partial of $\mathscr{H}$ w.r.t. each (scaled) state,

$$
\begin{aligned}
\dot{\lambda}_d &= -\frac{\partial \mathscr{H}}{\partial d} = (\lambda_\beta - \lambda_\alpha) \frac{1}{d^2} \sin \psi, \\
\dot{\lambda}_\alpha &= -\frac{\partial \mathscr{H}}{\partial \alpha} = -\frac{1}{2} \sin \alpha, & (14.13) \\
\dot{\lambda}_\beta &= -\frac{\partial \mathscr{H}}{\partial \beta} = 0.
\end{aligned}
$$

To obtain the saddle point control strategies the Hamiltonian is minimized and maximized w.r.t. ω and ψ, respectively. Incidentally, the

expressions to follow are identical to those given by [2] except with scaled/non-barred versions of the variables in place.

$$\omega^* = \arg\min_{\omega} \mathscr{H} = -\rho \, \mathrm{sign} \, \lambda_\alpha \tag{14.14}$$

$$\psi^* = \arg\max_{\psi} \mathscr{H}$$

$$\implies \cos \psi^* = \frac{\lambda_d}{\sigma_A}, \qquad \sin \psi^* = \frac{-\lambda_\alpha}{d\sigma_A}, \tag{14.15}$$

where,

$$\sigma_A = \sqrt{\lambda_d^2 + \left(\frac{\lambda_\alpha}{d}\right)^2}. \tag{14.16}$$

Another consequence of the scaling in distance is that the terminal condition is,

$$\Gamma(\mathbf{x}) := d - 1, \tag{14.17}$$

where $\mathbf{x} := (d, \alpha, \beta)^\top$. Termination occurs when the state enters the terminal surface,

$$\mathscr{C} := \{\mathbf{x} \mid \Gamma(\mathbf{x}) = 0\}. \tag{14.18}$$

Using the Lagrange multiplier ν, the adjoined terminal Value function is written,

$$\Phi(\mathbf{x}_f) = \nu\Gamma(\mathbf{x}_f) = \nu\,(d - 1). \tag{14.19}$$

Just as was done by [2], differentiating (14.19) w.r.t. the states yields the terminal costate variables,

$$\begin{aligned}
\lambda_d(t_f) &= \frac{\partial \Phi(\mathbf{x}_f)}{\partial d} = \nu, \\
\lambda_\alpha(t_f) &= \frac{\partial \Phi(\mathbf{x}_f)}{\partial \alpha} = 0, \\
\lambda_\beta(t_f) &= \frac{\partial \Phi(\mathbf{x}_f)}{\partial \beta} = 0.
\end{aligned} \tag{14.20}$$

Because the game is independent of time, the Hamiltonian, (14.12), must be equal to zero at all time. Substituting the equilibrium con-

trol strategies, (14.14) and (14.15), and terminal costate values, (14.20), into (14.12) and setting time to t_f gives

$$\mathcal{H}^* \big|_{t=t_f} = |\nu| - \frac{1}{2}(1 + \cos \alpha_f) - c = 0. \qquad (14.21)$$

Solving for ν gives

$$|\nu| = \frac{1}{2}(1 + \cos \alpha_f) + c. \qquad (14.22)$$

All together, (14.9), (14.13)–(14.16), (14.20), and (14.22) fully express the regular solutions (i.e., the non-singular saddle point equilibrium dynamics) of the TDDG in the scaled space. Of primary importance is the fact that the five natural parameters (v_A, d_c, Ω, c_1, and c_2) have been replaced with two: ρ and c. In a general sense, ρ represents the control authority of the Turret w.r.t. the Attacker: increasing ρ favors the Turret (in terms of Value, given some initial condition), whereas decreasing ρ favors the Attacker. The cost parameter, c, functions in much the same way as in the original (natural) representation of the game – it is merely a weight affecting the relative importance of time versus the Turret-induced cost to the Attacker. For small settings of c, the Attacker will prefer to avoid the Turret's gaze even if it takes longer to reach termination (and vice versa for large settings of c). [2] considered one particular set of natural parameters throughout: $d_c = 1$, $v_A = 1$, $\Omega = 0.05$, $c_1 = 1$, and $c_2 = 0.01$. This set corresponds to $\rho = 0.05$ and $c = 0.01$, which will henceforth be referred to as the canonical parameters for the TDDG. The process of filling the state space with equilibrium trajectories is mostly standard for the TDDG: backwards integration from the terminal set and from points along the Turret Universal Surface. However, the ADS requires special care as it does not have an analytical expression.

14.3   GENERAL DISPERSAL SURFACE CHARACTERISTICS

In this section, the characteristics of a DS are formally stated and criteria for use in the computation of the ADS is established. First, define the Value function as

$$
\begin{aligned}
V(\mathbf{x}(t)) &= \min_{\psi} \max_{\omega} J(\mathbf{x}(t); \psi(t), \omega(t)) \\
&= \max_{\omega} \min_{\psi} J(\mathbf{x}(t); \psi(t), \omega(t)) \\
&= \max_{\omega} \min_{\psi} \int_{t}^{t_f} \left(\frac{1}{2} (1 + \cos \alpha(\tau)) + c \right) d\tau.
\end{aligned}
\tag{14.23}
$$

The DS is characterized by one (or both) of the agents' equilibrium actions being non-unique [130]. That is, the agent whose equilibrium control is non-unique may choose which equilibrium action to take, leaving the Value of the game unaffected. Let $\mathbf{X}_{DS}$ denote the set of points on the DS. Let this be a DS in which the Attacker's equilibrium action is non-unique, but the Turret's equilibrium action is uniquely defined. Let ψ_A and ψ_B represent two different headings the Attacker can choose at particular state $\mathbf{x}'$.

Criterion 1. The condition

$$
V(\mathbf{x}'(t)) = J(\mathbf{x}'(t); \psi_A(t), \omega^*(t)) = J(\mathbf{x}'(t); \psi_B(t), \omega^*(t))
\tag{14.24}
$$

where $\omega^* = \arg\max_{\omega} J(\mathbf{x}'(t); \psi(t), \omega(t))$ is necessary for $\mathbf{x}' \in \mathbf{X}_{DS}$ to hold.

That is, the DS is characterized by two equilibrium trajectories intersecting at the same state with the same Value. Because the system dynamics are autonomous (i.e. not time-dependent), it is not necessary for this intersection to occur at the same time.

14.4   ATTACKER DISPERSAL SURFACE

In this section, the focus is on the Attacker DS in the TDDG. At a conceptual level, the ADS arises due to the interplay between the $(1 + \cos \alpha)$ and the c terms of the cost functional's integrand. On the ADS, the

Attacker has a choice between taking a quicker route in which the $(1 + \cos\alpha)$ term's contribution to the cost is higher, and a more round-about route in which the c term's contribution to the cost is higher. These paths are referred to as the direct (D) and indirect (I) paths, respectively. Based on the analysis in the previous section, two numerical procedures are developed to characterize the ADS. The first procedure computes the extent of the state space that is covered by equilibrium trajectories emanating from (in forward time) the ADS. This procedure provides the primary evidence for the presence or absence of the ADS, depending on the parameters. The second procedure computes the surface itself and is included here to confirm the results of the first procedure. Note that because of the lack of an analytical expression for the ADS both procedures are necessary to fill the state space with equilibrium trajectories, which is synonymous with solving the TDDG for particular parameter settings.

Before proceeding into the discussion of the procedures themselves, some background is included on the other singular surfaces and their interactions with the ADS. Ref. [2] showed that the Turret DS is defined by

$$\mathbf{X}_{TDS} := \{\mathbf{x} \mid \alpha = \pi\}, \tag{14.25}$$

which is the configuration in which the Turret is looking directly away from the Attacker. In this symmetrical configuration, the Turret has the choice of turning clockwise or counter-clockwise at its maximum turn rate: $\omega^* = \pm\rho$ – both choices are optimal. Upon the Turret making a choice, the state of the system immediately departs the TDS and both agents' equilibrium controls are uniquely defined. Ref. [2] also showed that the Turret Universal Surface is defined by

$$\mathbf{X}_{TUS} := \{\mathbf{x} \mid \alpha = 0\}. \tag{14.26}$$

On the TUS, the Turret is looking directly at the Attacker; the Turret's control is $\omega^* = 0$ and the Attacker heads directly towards the Turret, $\psi^* = \pi$. The agents remain in this configuration for the remainder of the game – neither gains an advantage by deviating from this "locked on" configuration.

Note that the TUS does not interact with the ADS, but its tributaries must also be computed in order to fill the state space with equilibrium trajectories. The TDS is particularly important in the present context because it coincides with the ADS.

14.4.1   *ADS Envelope Computation*

In this section, a procedure is presented for computing the envelope of the ADS. Because of the symmetrical nature of the TDDG's solution (c.f [2]), one may, without loss of generality, consider the part of the state space in which $0 \leq \alpha \leq \pi$. For the remainder of the paper, the solution of the TDDG (14.9), (14.13)–(14.16), (14.20) and (14.22) is used repeatedly to backwards integrate from the terminal surface (14.18). This backwards integration is carried out until the trajectory reaches the TDS, where $\alpha = \pi$ since the trajectories can go no further and the goal is to compute the maximal envelope of the ADS. Define the following mapping between points on the terminal surface and points along the trajectory obtained by backwards integration of the solution:

$$B : \alpha_f, t_f \rightarrow \mathbf{x}_0,$$
$$\text{s.t. } (14.9), (14.13)–(14.16), (14.20) \text{ and } (14.22), \tag{14.27}$$

where $\mathbf{x}_0 = \mathbf{x}(0)$. Making use of (14.24), define the envelope of the ADS as the pair

$$(\alpha_{f_L}, \alpha_{f_U}) \text{ s.t. } \begin{cases} d_{0_L} = d_{0_U}, \\ \alpha_{0_L} = \alpha_{0_U} = \pi, \\ \alpha_{f_L} < \alpha_{f_U}, \\ V(\mathbf{x}_{0_L}) = V(\mathbf{x}_{0_U}) \end{cases} \tag{14.28}$$

where $\mathbf{x}_{0_L} = B(\alpha_{f_L}, t_{f_L}) = \begin{bmatrix} d_{0_L} & \alpha_{0_L} & \beta_{0_L} \end{bmatrix}^\top$, etc. In general, t_{f_L} and t_{f_U} are not known *a priori*. However, this is of little consequence since one can simply integrate trajectories backwards in time until the condition $\alpha = \pi$ is met. Indeed, the trajectories ought not go any further, due to

the TDS. The purpose of defining and computing (14.28) is twofold. First, for $\alpha_{f_L} \leq \alpha_f \leq \alpha_{f_U}$ the trajectories terminate (in retrograde time) on the ADS while for $\alpha_f < \alpha_{f_L}$ and $\alpha_f > \alpha_{f_U}$ the trajectories terminate on the TUS and do not interact with the ADS. Second, when $\alpha_{f_L} = \alpha_{f_U}$, i.e., there is no solution to (14.28) the ADS is not present. The following is stated, without proof:

Proposition 14.1. *For the Attacker* DS *to exist in the solution of the* TDDG, (14.28) *must be satisfied.*

That is, if the ADS exists in the solution, the ADS's envelope $(\alpha_{f_L}, \alpha_{f_U})$ must be computable.

14.4.1.1 *Procedure*

Now, a process is described by which the ADS envelope, (14.28), is computed. Figures 14.2 and 14.3 contain the results for this proce-
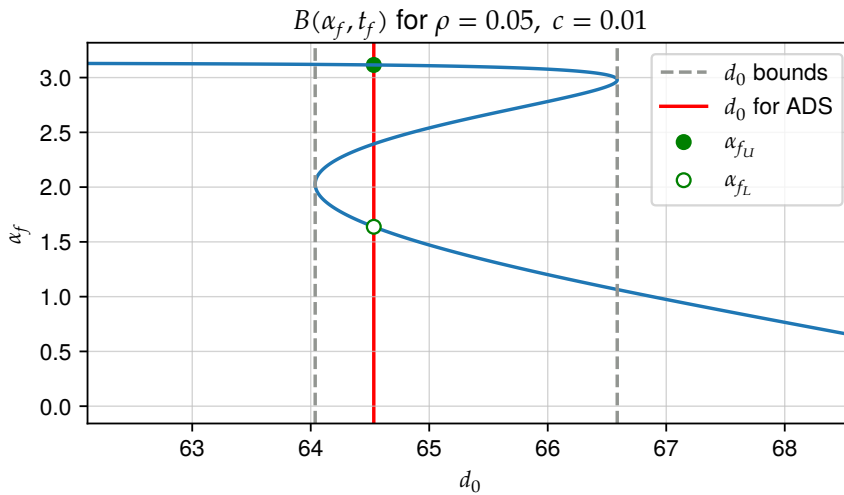
Figure 14.2: Mapping of α_f to d_0 along the TDS indicating the region of d_0 (dashed gray lines) that may satisfy the ADS envelope conditions and the d_0 and $(\alpha_{f_L}, \alpha_{f_U})$ for which the Value of the L and U trajectories are equal. Initially 4,000 samples were used in the α_f sweep, and then 1,000 samples were used in the d_0 sweep.

dure which are illustrative for describing the procedure itself. First, the terminal surface, (14.18), is swept along a grid of α_f, $0 < \alpha_f < \pi$. The mapping B is computed for each α_f by backwards integration until $\alpha = \pi$, and the corresponding d_0 and V values are recorded. Figure 14.2 shows the d_0 values along the horizontal axis associated with
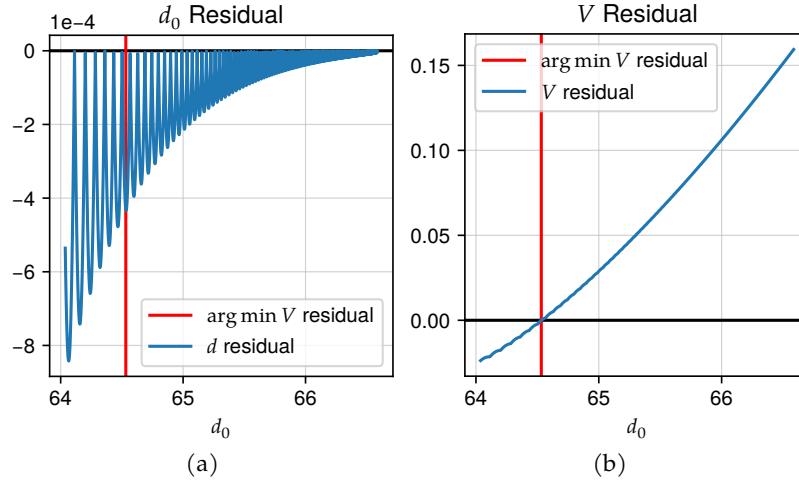
Figure 14.3: Residuals between the L and U trajectories corresponding to Fig. 14.2.

the α_f values. From (14.28), it is required that there are two different α_f with the same d_0. Thus the gray dashed lines in the figure denote the bounds for which this condition can be satisfied. All of the d_0 within this range satisfy the first three conditions in (14.28). Satisfaction of the last condition, $V(\mathbf{x}_{0_L}) = V(\mathbf{x}_{0_U})$, is the distinguishing criteria by which the ADS envelope is determined. Then, a line of constant d_0 (i.e., the red line in the figure) is swept within the bounds. Because the process started with a grid of α_f, the spacing of points in the d_0 axis is non-uniform. A new grid over the d_0 is used and the backwards "S" curve is interpolated to get approximate candidate values for α_{f_L} and α_{f_U}. The mapping B is recomputed for the approximate candidate α_f values and the initial distance residual $d_{0_L} - d_{0_U}$ (Fig. 14.3a) and Value residual $V(\mathbf{x}_{0_L}) - V(\mathbf{x}_{0_U})$ (Fig. 14.3b) are recorded. Note the d_0 residual is larger for smaller d_0 because the upper part of the curve is very flat here – a small error in α_{f_U} produces a large deviation in d_{0_U}. Finally, when the number of d_0 samples is sufficiently large, the Value residual in Fig. 14.3b is fairly smooth and monotonic with a unique zero crossing (marked with the red line). This value for d_0, and the corresponding $(\alpha_{f_L}, \alpha_{f_U})$ satisfies (14.28) and thus defines the envelope of the ADS. Figure 14.4 shows the ADS envelope trajectories in the state space along with the terminal surface, the TDS, and the TUS.
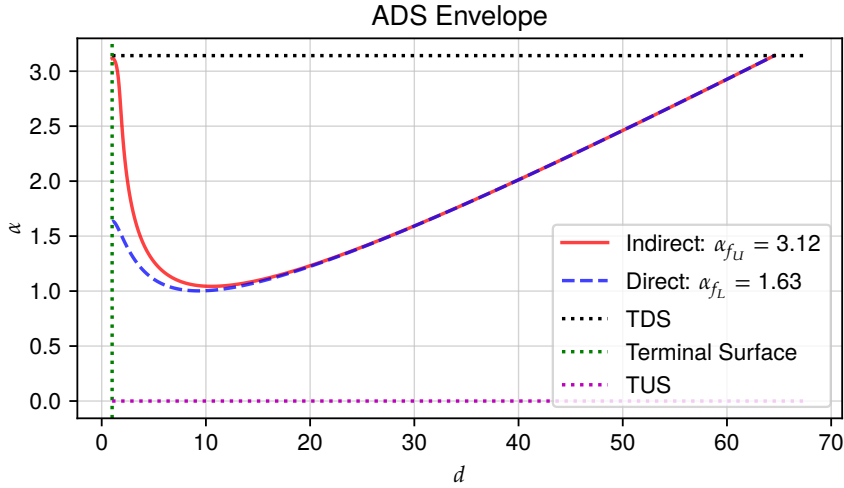
Figure 14.4: Depiction of the ADS envelope in the state space for $\rho = 0.05$ and $c = 0.01$. Equilibrium trajectories emanating from the ADS fill the space between the red and blue curves (i.e. the upper and lower extent of the ADS envelope).

14.4.1.2   *Remarks*

Note that the ADS envelope conditions in (14.28) can be determined by setting up an appropriate nonlinear program (NLP). However, it can be difficult to enforce the condition $\alpha_{f_L} < \alpha_{f_U}$ as the optimizer tends to end up in the situation where $\alpha_{f_L} = \alpha_{f_U}$ and dwells there. Also, the second stage of the procedure, in which d_0 is swept, may be replaced by a binary search process due to the monotonicity of the Value residual w.r.t. d_0. This drastically reduces the number of backwards shots needed in the second stage. For example, to obtain a Value residual $< 1 \times 10^{-12}$ only 37 backwards shots are needed for the example in Fig. 14.2. In this way, the Value residual tolerance can be directly specified, and the d_0 residual is directly affected by the number of samples in the initial α_f grid. Figure 14.5 shows the result of using binary search to find the zero crossing of the Value residual.

The trajectories generated within the range $\alpha_{f_L} < \alpha_f < \alpha_{f_U}$ do not actually reach back to the TDS, since this would require passing through the ADS. Therefore, the middle section of the backwards "S" curve in Fig. 14.2 is meaningless. Algorithm 3 summarizes the procedure used to compute the ADS envelope $\left(\alpha_{f_L}, \alpha_{f_U}\right)$.
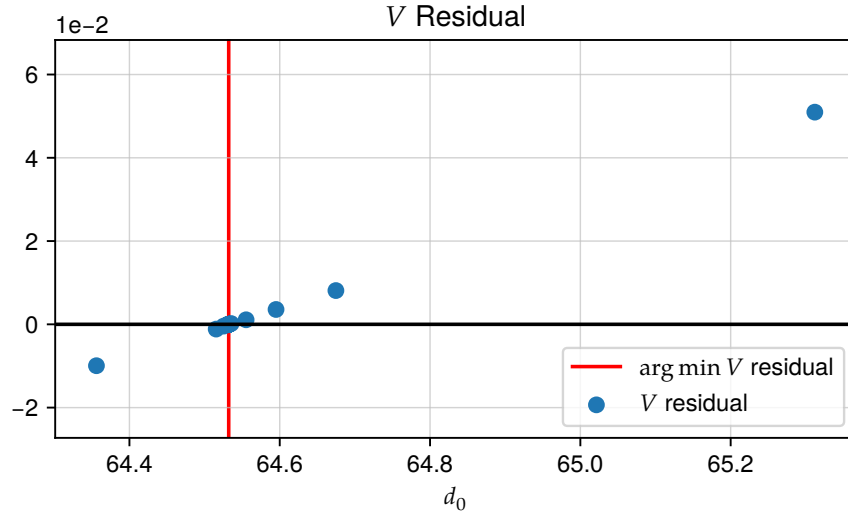
Figure 14.5: Value residual for the binary search algorithm.

Algorithm 3 ADS Envelope (with binary search)

$d_0 \leftarrow$ empty vector
for α_{f_i} **in** α_f grid **do**
    $d_{0_i} \leftarrow B(\alpha_{f_i}, t_f)$ $\triangleright$ backwards integrate

$d_{0_{\min}}, d_{0_{\max}} \leftarrow$ local min & max of d_0
$V_{\text{resid}} \leftarrow \infty$
while $V_{\text{resid}} > \varepsilon$ **do** $\triangleright$ binary search
    $d_{\text{mid}} \leftarrow (d_{0_{\min}} + d_{0_{\max}})/2$
    $\alpha_{f_L}, \alpha_{f_U} = B^{-1}(d_{\text{mid}})$ $\triangleright$ see Fig. 14.2
    $\mathbf{x}_{0_L} \leftarrow B(\alpha_{f_L}, t_f)$ $\triangleright$ backwards integrate
    $\mathbf{x}_{0_U} \leftarrow B(\alpha_{f_U}, t_f)$ $\triangleright$ backwards integrate
    $V_{\text{resid}} \leftarrow V(\mathbf{x}_{0_L}) - V(\mathbf{x}_{0_U})$
    if $V_{\text{resid}} > 0$ **then**
        $d_{0_{\max}} \leftarrow d_{\text{mid}}$
    else
        $d_{0_{\min}} \leftarrow d_{\text{mid}}$

14.4.2 *Different Parameters*

In this section, the utility of the previously described numerical proce-
dure is demonstrated in the determination of the regions of the param-
eter space in which the ADS is present. As Proposition 14.1 suggests,
the inability for the procedure to find $(\alpha_{f_L}, \alpha_{f_U})$ that satisfies (14.28) in-
dicates that there is no ADS in the solution to the TDDG. Figure 14.6
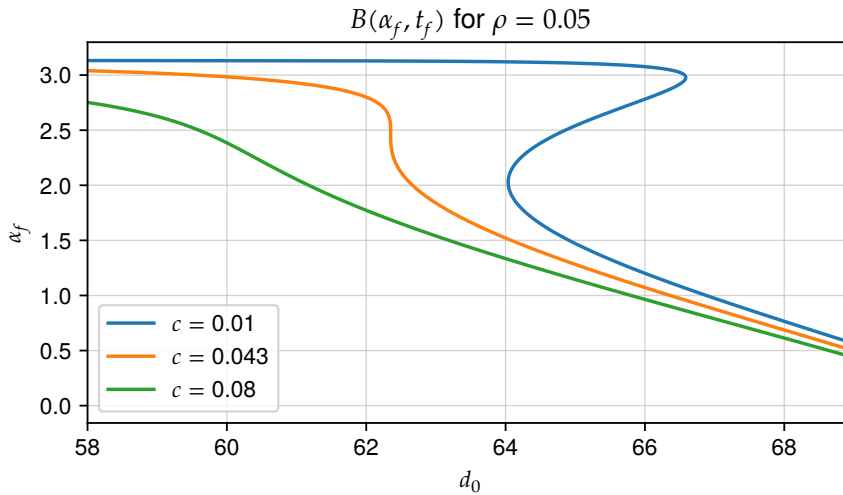
Figure 14.6: Mapping of α_f to d_0 for different settings of the cost parameter c.

shows the results of the backwards integration process (i.e. the back-
wards "S" curves) for different settings of the cost parameter c for a
particular Turret effectiveness setting, ρ. For $c = 0.043$, the curve flat-
tens out to the point that for each d_0 there is a unique α_f. Thus, the con-
ditions for the ADS envelope (14.28), in particular the requirements
that $\alpha_{f_L} < \alpha_{f_U}$ and $d_{0_L} = d_{0_U}$, cannot be met. It is precisely at this set-
ting of c, i.e., where there exists a point on the curve s.t. $dd_0/d\alpha_f = 0$,
that the ADS disappears from the solution of the TDDG. As suggested
by the curve in Fig. 14.6 corresponding to $c = 0.08$, the ADS is absent
for all $c > 0.043$, in this case. Repeating the process across a range of ρ
yields the curve in Fig. 14.7.
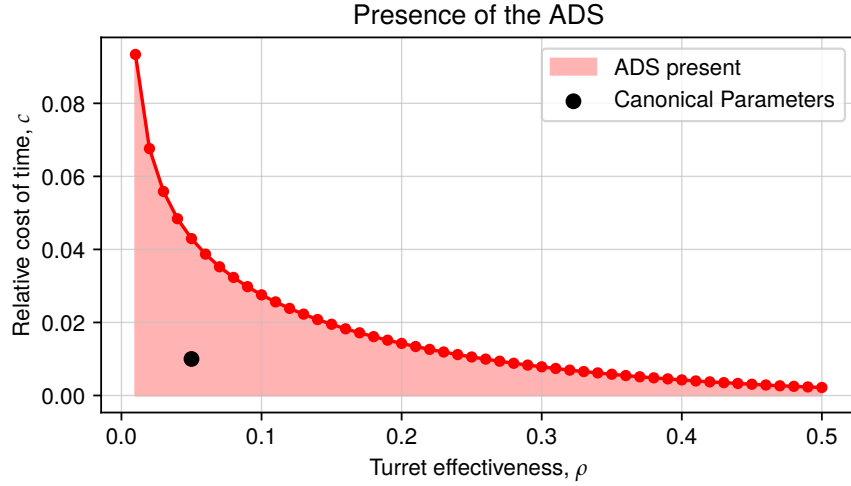
Figure 14.7: Presence of the ADS in the solution of the TDDG over the parameter space.

14.4.3 *ADS Termination Point Computation*

In order to gain a better understanding of the ADS and how it is affected by the parameter settings, a procedure is developed for computing the surface itself. It has been observed that one of the endpoints of the ADS (if it exists) lies on the TDS. Ref. [2] showed that, for the canonical parameter settings, the other endpoint of the ADS is *not* coincident with any other surface. The approach for computing points on the ADS is based upon solving for a pair of trajectories satisfying

$$
(\alpha_{f_l}, \alpha_{f_u}) \text{ s.t. } \begin{cases} d_{0_l} = d_{0_u} = d_0, \\[1mm] \alpha_{0_l} = \alpha_{0_u} = \alpha_0, \\[1mm] \alpha_{f_l} < \alpha_{f_u}, \\[1mm] V(\mathbf{x}_{0_l}) = V(\mathbf{x}_{0_u}) \end{cases} \tag{14.29}
$$

where d_0 is the specified distance associated with the point on the ADS that is being solved for. The condition (14.29) is similar to (14.28) but with a free α_0. Again, the t_f associated with these trajectories is not known *a priori*. However, using integrator callbacks (c.f. [197]), one can integrate backwards until the specified distance d_0 is reached. To

solve for $(\alpha_{f_l}, \alpha_{f_u})$, an NLP solver is employed which seeks to minimize the following residual

$$r = \left\| \begin{matrix} V(\mathbf{x}_{0_l}) - V(\mathbf{x}_{0_u}) \\ \alpha_{0_l} - \alpha_{0_u} \end{matrix} \right\|$$

To generate a series of points on the ADS, (14.29) is solved along a grid of d_0. First, some appropriate d_0 is chosen and the procedure continues in decreasing order of d_0 along this grid. A homotopy approach is used in which the NLP solver is given the $(\alpha_{f_l}, \alpha_{f_u})$ of the previous d_0 as an initial guess. Figure 14.8 contains the results of this procedure
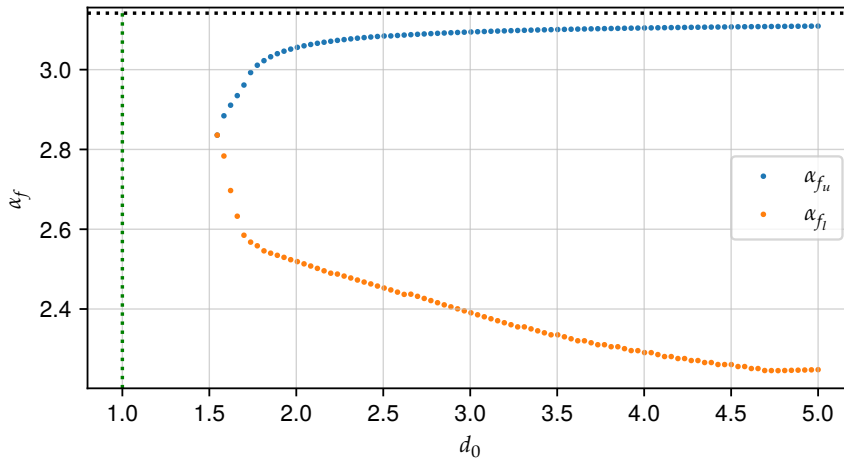
Figure 14.8: Solutions to (14.29) over a d_0 grid for $c = 0.01$.

for a relative time cost setting $c = 0.01$. Note on the left side of the plot the two curves quickly converge s.t. $\alpha_{f_l} \not< \alpha_{f_u}$. Thus for $d_0 \leq 1.5$ the condition $\alpha_{f_l} < \alpha_{f_u}$ in (14.29) cannot be satisfied. The leftmost d_0 point corresponds to the non-TDS endpoint of the ADS. For each solution in Fig. 14.8, the α_0 to which the upper and lower trajectories integrate back to is recorded. The points (d_0, α_0) represent the ADS itself. The procedure is repeated for different settings of c up until $c \approx 0.043$ whereupon, based on Fig. 14.7, it is expected that the ADS will disappear. Figure 14.9 shows the Attacker DSs for $\rho = 0.05$ and several different c. As c is increased from 0.01 to ≈ 0.43, the endpoint of the ADS recedes away from the terminal surface and towards the TDS. These results corroborate earlier statements about the disappearance of the ADS from the solution of the TDDG for high settings of c. Algorithm 4
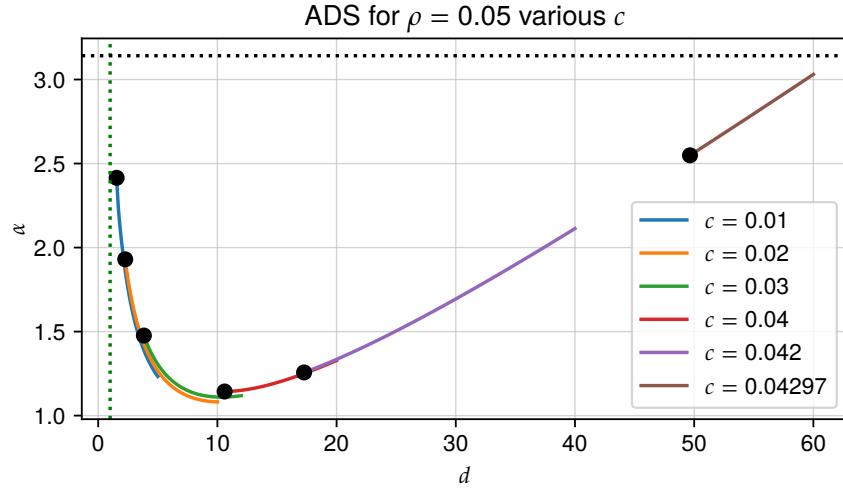
Figure 14.9: ADS for $\rho = 0.05$ and various c showing the recession of the ADS into the TDS as $c \to 0.043$.

summarizes the process for computing the ADS once for a particular setting of c.

Algorithm 4 ADS Computation

$\mathbf{X}_{ADS} \leftarrow \emptyset$
$i \leftarrow$ LENGTH $(d_0 \text{ grid})$
$\alpha_{f_l} \leftarrow 0$ and $\alpha_{f_u} \leftarrow \pi$
$guess \leftarrow \left(\alpha_{f_l}, \alpha_{f_u}\right)$
while $\alpha_{f_u} - \alpha_{f_l} > \varepsilon$ and $i \geq 0$ **do**
$\quad d_0 \leftarrow d_0 \text{ grid } [i]$
$\quad \alpha_{f_l}, \alpha_{f_u}, \alpha_0 \leftarrow$ solution to (14.29) with $guess$ $\qquad\qquad \triangleright$ NLP
$\quad$ **if** $\alpha_{f_u} - \alpha_{f_l} > \varepsilon$ **then**
$\quad\quad \mathbf{X}_{ADS} \leftarrow \mathbf{X}_{ADS} \cup \{(d_0, \alpha_0)\}$
$\quad\quad guess \leftarrow \left(\alpha_{f_l}, \alpha_{f_u}\right)$
$\quad i \leftarrow i - 1$

14.5 CONCLUSION

In this chapter the TDDG – a game with three states, only two of which affect the solution – has been analyzed. Despite the simplicity of its dynamics, the full solution of the game is difficult to obtain due to the presence of a non-analytical singular surface: the Attacker DS. Numerical methods are required to characterize the Attacker DS as well as to compute the surface itself. Furthermore, the original description of the game contains five natural parameters, which makes understand-

ing the impact of parameters on the solution (particularly the Attacker DS) difficult. The five natural parameters have been collapsed into two composite parameters: the Turret effectiveness, ρ, and the relative time cost c. The scaling of the kinematics facilitated the exploration of the parameter space. Two numerical approaches, based on the general definition of a dispersal surface, were developed to characterize the Attacker DS. These approaches exposed an interesting feature of the game's solution: the Attacker DS is only present over a portion of the parameter space. In the following, this wide-beam turret defense scenario will be considered in the context of engage or retreat; the numerical procedures developed in here will be useful in computing the retreat trajectories.

ENGAGE OR RETREAT WITH A WIDE-BEAM TURRET

15.1 INTRODUCTION

In this chapter, the TEoR game is revisited for the case of a wide-beam Turret (TEoRW), and particular focus is given to the case of OCR. As mentioned in the previous chapter, the solution of the wide-beam TDDG (referred to henceforth as the *Game of Engagement*, or GoE) can only be obtained numerically. Thus, in order to impose the path constraint needed to ensure engagement doesn't become optimal, the GoE solution must be interpolated along the OCR trajectory. The main contributions of this chapter are (1) a numerical interpolation scheme of the TDDG's solution (2) derivation of the first order necessary conditions for optimality for the OCR problem, (3) specification of the Boundary Value Problem (BVP) for when the constraint becomes active along the trajectory, and (4) a process for solving the BVP. The first of these supports Research Objective 1 (approximations to improve computability), while the latter directly address Research Objective 7 (solution of optimal control problems embedded within games). This material is based on [248].

Section 15.2 contains a formal description of the OCR problem. Section 15.3 contains the derivation of the first order necessary conditions for optimality, specification of the BVP, and special considerations for the Turret's turn control. Section 15.4 contains the simulation results for a particular initial condition, and Section 15.5 concludes the chapter.

15.2   PROBLEM DESCRIPTION

Given a stationary Turret with bounded turn rate and a mobile At-
tacker, moving with simple motion, a trajectory is sought which moves
the system state to the retreat surface, while minimizing cost, subject to
a path constraint. The coordinate system is depicted in Fig. 14.1 (in the
previous chapter) and the Turret's position is defined to be the origin
of the (x, y)-plane. The dynamics of the system are a modified version
of the dynamics presented in [2] wherein the natural parameters are
consolidated into relative turret effectiveness, ρ, and relative time cost,

$$f(\mathbf{x}, \mathbf{u}, t) = \dot{\mathbf{x}} = \begin{bmatrix} \dot{d} \\ \dot{\alpha} \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} \cos \psi \\ \omega - \frac{1}{d} \sin \psi \\ \frac{1}{d} \sin \psi \end{bmatrix}, \ \omega \in [-\rho, \rho], \qquad (15.1)$$

where the Attacker controls ψ, its instantaneous heading angle relative
to the LOS from the Turret, and the Turret controls ω, its turn-rate (as
in the previous chapter).

Define the following boundary conditions,

$$\phi(\mathbf{x}(t_0), \mathbf{x}(t_f)) = \begin{bmatrix} d(t_0) \cos \beta(t_0) + 5 \\ d(t_0) \sin \beta(t_0) - 20 \\ \alpha(t_0) - \beta(t_0) + \frac{\pi}{2} \\ d(t_f) \sin \beta(t_f) - y_R \end{bmatrix} = 0 \qquad (15.2)$$

with $t_0 = 0$ and t_f free. Note that in the natural (x, y)-plane, the first two
constraints can be expressed as $x(0) = -5$ and $y(0) = 20$, respectively,
while the last can be expressed $y(t_f) = y_R$. Thus, the retreat region, in
this case is $y \leq y_R$. Here, $y_R = -20$. The initial α constraint stipulates
that the Turret's initial global look angle is $\frac{\pi}{2}$ (i.e. along the positive y
axis).

The cost functional is also slightly modified from that of [2],

$$J = \int_{t_0}^{t_f} \left[\theta \left(\frac{1}{2} (1 + \cos \alpha) \right) + c \right] dt + \Phi \left(\mathbf{x}(t_f), t_f \right), \qquad (15.3)$$

where $\theta \in [0,1]$ is an additional Turret control representing some kind of rate of fire. The terminal Value function,

$$\Phi\left(\mathbf{x}(t_f), t_f\right) = P, \tag{15.4}$$

where $P > 0$ is a constant, penalizes the Attacker for retreating instead of engaging. The results herein are based on setting $\rho = 0.05$, $c = 0.01$, and $P = 8$. The settings for ρ and c match the settings used in [2]. For the optimal constrained retreat, the Turret and Attacker cooperate to minimize the cost in (15.3). Thus the Value function is defined as,

$$V_R(\mathbf{x}(t)) = \min_{\omega, \theta} \min_{\psi} J = \min_{\omega, \theta, \psi} J. \tag{15.5}$$

Similarly, the optimal controls are defined as,

$$u^* = (\omega^*(t), \theta^*(t), \psi^*(t)) = \underset{\omega(t), \theta(t), \psi(t)}{\arg\min} \; J. \tag{15.6}$$

The path constraint is defined as,

$$g(\mathbf{x}) = V_E(\mathbf{x}(t)) - V_R(\mathbf{x}(t)) \geq 0, \; \forall t \in \left[t_0, t_f\right], \tag{15.7}$$

where V_E is the Value function associated with the GoE. Let L be the integrand of the cost functional and suppose an additional state component, $l(t)$, is appended to $\mathbf{x}$ where,

$$\dot{l}(t) = -L, \; l(t_f) = 0 \tag{15.8}$$

which represents the remaining integral cost-to-go. Then the retreat Value function may be written,

$$\begin{aligned}
V_R(\mathbf{x}(t)) &= P + \int_t^{t_f} L \, dt \\
&= P - \int_t^{t_f} \dot{l}(t) \, dt \\
&= P - l(t_f) + l(t) \\
&= P + l(t)
\end{aligned} \tag{15.9}$$

In the case of backwards shooting, $l(t)$ may be easily computed, making the computation of $V_R(\mathbf{x})$ trivial.

Due to the path constraint, (15.7), it is necessary to solve the GoE so that $V_E(\mathbf{x})$ may be known for every point in the state space. An analytical solution of the GoE, however, is not available and a numerical solution must suffice. For a particular $\mathbf{x}$, V_E may be obtained by solving the associated boundary value problem via indirect backwards shooting. That approach, however, is infeasible for the purposes of the present work because the path constraint will need to be evaluated at *every* point along the trajectory. Instead, indirect backward shooting is used to fill the state space with equilibrium trajectories which are then sampled to generate a large set of data,

$$D = \left\{ \left(d_i, \alpha_i, \sigma_{d_i}, \sigma_{\alpha_i}, V_{E_i}(d_i, \alpha_i) \right) \right\}, \tag{15.10}$$

where each element is a tuple comprised of the state, GoE adjoints, and Value associated with starting in this state (for the GoE). The GoE adjoints are used in indirect optimal control analysis in Section 15.3. Fig. 15.1 shows $V_E(d, \alpha)$ (i.e. the last column of D) for each of the sam-
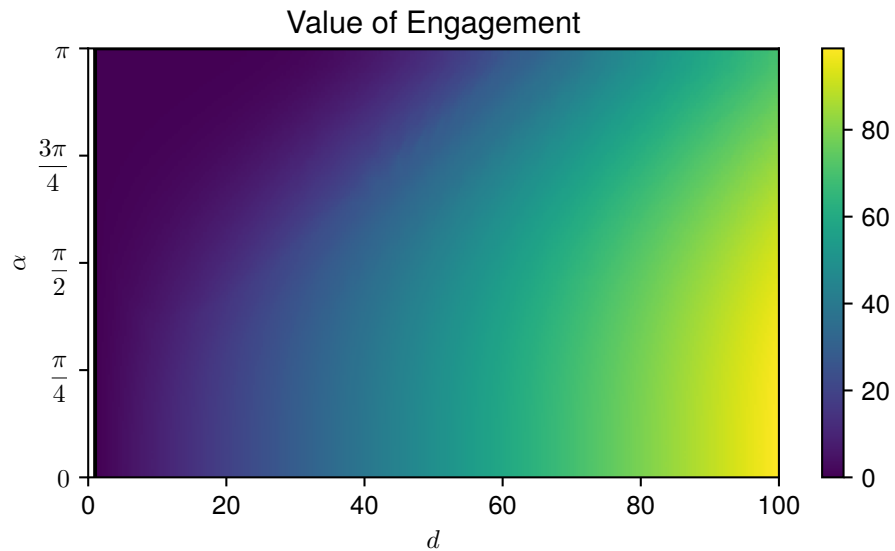
Figure 15.1: Value of engagement, $V_E(\mathbf{x})$

ple points.

The quantities $V_E(\mathbf{x})$ (and σ_d and σ_α) are computed via a k-nearest neighbor (kNN) search with $k = 3$ and take a distance weighted av-

erage of the neighbors. The kNN search is made viable through the usage of the efficient *NearestNeighbors* package for the Julia programming language [41].

The difficulty in sampling the saddle-point equilibrium trajectories for the GoE lies in handling the singular surfaces. As mentioned previously, there are three singular surfaces in the GoE [2]: the Turret's Universal Surface (at $\alpha = 0$), the Turret's Dispersal Surface (at $\alpha = \pi$), and the Attacker's Dispersal Surface (ADS), which cannot be described analytically. To produce Fig. 15.1, the state space ($0 \leq \alpha \leq \pi$ and $1 \leq d \leq 100$) can be divided into 5 distinct regions: 1) trajectories with $\alpha_f \neq 0$ above the envelope of the ADS, 2) trajectories with $\alpha_f \neq 0$ below the envelope of the ADS, 3) "direct" trajectories emanating from the ADS, 4) "indirect" paths emanating from the ADS, and 5) trajectories in which $\alpha_f = 0$ (trajectories ending on the Turret's Universal Surface).

With all of the above definitions in place, the problem definition is formally stated:

$$\min_{\omega(t), \theta(t), \psi(t)} J$$

$$\text{s.t. } \phi(\mathbf{x}(t_0), \mathbf{x}(t_f)) = 0, \qquad (15.11)$$

$$g(\mathbf{x}(t)) \geq 0 \ \ \forall t \in \left[0, t_f\right]$$

15.3 METHODOLOGY

15.3.1 *Optimality Conditions*

First, the first order optimality conditions for problem (15.11) are developed. The trajectory constraint, g, is transformed to a control constraint, h, by differentiating with respect to time,

$$\begin{aligned} h(\mathbf{x}) = \frac{\mathrm{d}g}{\mathrm{d}t} &= \frac{\mathrm{d}V_E}{\mathrm{d}t} - \frac{\mathrm{d}l}{\mathrm{d}t} \\ &= \frac{\partial V_E}{\partial \mathbf{x}}\dot{\mathbf{x}} + L \\ &= \sigma\dot{\mathbf{x}} + L \end{aligned} \qquad (15.12)$$

where $\sigma = \begin{bmatrix} \sigma_d & \sigma_\alpha & \sigma_\beta \end{bmatrix}^\top$ are the adjoint variables associated with the GoE. When the path constraint is active, $g(\mathbf{x}) = 0$, and in order to remain on the constraint $h(\mathbf{x}) = 0$ as well. Thus when transitioning from an unconstrained arc to a constrained arc (or vice versa) $h(\mathbf{x})$ must be zero; this is referred to as the tangency condition. The Hamiltonian may be written,

$$H = \lambda \dot{\mathbf{x}} + \mu h + L, \tag{15.13}$$

where μ is the adjoint variable associated with the path constraint derivative, h, and

$$\mu(t) = \begin{cases} 0 & g(\mathbf{x}(t)) > 0 \\ > 0 & g(\mathbf{x}(t)) = 0. \end{cases} \tag{15.14}$$

Substituting (15.12) into (15.13) gives,

$$H = \lambda \dot{\mathbf{x}} + \mu(\sigma \dot{\mathbf{x}} + L) + L$$
$$= (\lambda + \mu\sigma)\,\dot{\mathbf{x}} + (1 + \mu)\,L. \tag{15.15}$$

Note that $\sigma_\beta = 0$ for all time [2]. The optimal adjoint dynamics are obtained by differentiating (15.15) w.r.t. each state component and substituting the state dynamics, (15.1), and $\sigma_\beta = 0$,

$$\dot{\lambda}_d = -\frac{\partial H}{\partial d} = \left(\lambda_\beta - \lambda_\alpha - \mu\sigma_\alpha\right)\frac{1}{d^2}\sin\psi \tag{15.16}$$

$$\dot{\lambda}_\alpha = -\frac{\partial H}{\partial \alpha} = \theta\left(1 + \mu\right)\frac{1}{2}\sin\alpha \tag{15.17}$$

$$\dot{\lambda}_\beta = -\frac{\partial H}{\partial \beta} = 0. \tag{15.18}$$

The fact that $\dot{\lambda}_\beta = 0$ comes from the fact that $\frac{\partial \dot{\mathbf{x}}}{\partial \beta} = 0$ and $\frac{\partial L}{\partial \beta} = 0$ and implies that $\lambda_\beta(t)$ is constant.

The optimizing controls are obtained via PMP.

$$\theta^* = \arg\min_\theta H$$
$$= \arg\min_\theta (\lambda \dot{\mathbf{x}} + \mu\sigma) + (1 + \mu)\left[\theta\frac{1}{2}\left(1 + \cos\alpha\right) + c\right]$$

Since $\mu \geq 0$, the term $(1 + \mu)$ must be positive, which implies,

$$\theta^*(t) = 0. \tag{15.19}$$

It is clear that the Turret should have its rate-of-fire set to zero for optimal constrained retreat. As a result,

$$\dot{\lambda}_\alpha = 0. \tag{15.20}$$

For the Turret's optimal turn control,

$$\omega^* = \arg\min_\omega H = \arg\min_\omega (\lambda + \mu\sigma)\,\dot{x} + (1 + \mu)L$$

$$= \arg\min_\omega (\lambda_\alpha + \mu\sigma_\alpha)\left(\omega - \frac{1}{d}\sin\psi\right),$$

which implies,

$$\omega^* = -\rho\,\mathrm{sign}\left(\lambda_\alpha + \mu\sigma_\alpha\right). \tag{15.21}$$

The Attacker's optimal heading control is given by,

$$\psi^* = \arg\min_\psi H = \arg\min_\psi (\lambda + \mu\sigma)\dot{x} + (1 + \mu)L$$

$$= \arg\min_\psi (\lambda + \mu\sigma)$$

$$= \arg\min_\psi (\lambda_d + \mu\sigma_d)\cos\psi$$

$$+ (\lambda_\alpha + \mu\sigma_\alpha)\left(\omega - \frac{1}{d}\sin\psi\right)$$

$$+ \left(\lambda_\beta + \mu\sigma_\beta\right)\frac{1}{d}\sin\psi.$$

Thus ψ^* is determined by setting the vector $\begin{bmatrix} \cos\psi^* & \sin\psi^* \end{bmatrix}^\top$ to be antiparallel to the vector $\begin{bmatrix} \frac{\lambda_d + \mu\sigma_d}{\zeta} & \frac{\lambda_\beta - \mu\sigma_\alpha}{d\zeta} \end{bmatrix}^\top$:

$$\cos\psi^* = \frac{-\lambda_d - \mu\sigma_d}{\zeta} \tag{15.22}$$

$$\sin\psi^* = \frac{-\lambda_\beta + \mu\sigma_\alpha}{d\zeta}, \tag{15.23}$$

where,

$$\zeta = \sqrt{(\lambda_d + \mu\sigma_d)^2 + \frac{1}{d^2}\left(\lambda_\beta - \mu\sigma_\alpha\right)^2}.$$

In order to proceed, let $\mu(t_f) = 0$, that is, the system is unconstrained at final time. This is not strictly necessary, however, it will end up being the case for the particular parameter settings used in Section 15.3. Because components of the final state $\mathbf{x}(t_f)$ are free, the transversality condition yields [40],

$$
\begin{aligned}
\boldsymbol{\lambda}^\top(t_f) &= \frac{\partial \Phi}{\partial \mathbf{x}(t_f)} + \nu \frac{\partial \phi}{\partial \mathbf{x}(t_f)} \\
&= 0 + \nu \left[-\frac{y_R}{d(t_f)^2} \quad 0 \quad -\cos\beta \right]
\end{aligned}
$$

and thus,

$$
\lambda_d(t_f) = -\nu \frac{y_R}{d(t_f)^2} \tag{15.24}
$$

$$
\lambda_\alpha(t_f) = 0 \tag{15.25}
$$

$$
\lambda_\beta(t_f) = -\nu \cos\beta. \tag{15.26}
$$

The cosine of the optimal terminal Attacker heading is obtained by substituting (15.24) and (15.26) into (15.22) with $\mu(t_f) = 0$

$$
\begin{aligned}
\cos\psi^*(t_f) &= \frac{\nu y_R / d(t_f)^2}{\sqrt{\nu^2 \frac{y_R^2}{d(t_f)^2} + \frac{1}{d(t_f)^2} \nu^2 \cos\beta(t_f)^2}} \\
&\propto y_R \operatorname{sign}(\nu) \\
&\propto -\operatorname{sign}(\nu).
\end{aligned}
$$

The last expression is due to the fact that $y_R < 0$. If $\nu > 0$, then,

$$
\dot{d}(t_f) = \cos\psi(t_f) \propto -\operatorname{sign}(\nu) = -1, \tag{15.27}
$$

which states that distance at final time is decreasing, but for optimal constrained retreat, as long as the Turret is not placed directly on the retreat surface, the distance at final time must be increasing. Therefore, it must be the case that $\nu < 0$. Note, also, that because $\theta^* = 0$ it must be

that $L = c$. With knowledge of the sign of ν, the cosine of the optimal Attacker's heading angle at final time can be further simplified,

$$\cos \psi^*(t_f) = \frac{\text{sign}(\nu) y_R / d(t_f)^2}{\sqrt{\frac{y_R^2}{d(t_f)^4} + \frac{1}{d(t_f)^2} \cos \beta(t_f)^2}}$$

$$= \frac{-y_R}{d(t_f) \sqrt{\frac{y_R^2}{d(t_f)^2} + \cos \beta(t_f)^2}}.$$

Define,

$$\chi = \sqrt{\frac{y_R^2}{d(t_f)^2} + \cos \beta(t_f)^2}.$$

Then the optimal terminal Attacker heading angle can be written,

$$\cos \psi^*(t_f) = \frac{-y_R}{d(t_f)\chi}, \ \sin \psi^*(t_f) = \frac{-\cos \beta(t_f)}{\chi}. \tag{15.28}$$

Based on the necessary conditions for optimality [40], the Hamiltonian at final time is given by

$$H^*(t_f) = -\frac{\partial \Phi}{\partial t_f} - \nu \frac{\partial \phi}{\partial t_f} = 0 \tag{15.29}$$

Substituting (15.1) and (15.24)–(15.26) into the Hamiltonian, (15.13), at final time gives

$$H(t_f) = -\nu \frac{y_R}{d(t_f)^2} \cos \psi(t_f) \tag{15.30}$$

$$- \nu \cos \beta(t_f) \frac{1}{d(t_f)} \sin \psi(t_f) + L$$

Substituting in (15.28) and (15.29) with $\mu = 0$ and $L = c$ into (15.30) yields

$$H^*(t_f) = 0 = \nu \frac{y_R}{d(t_f)^2} \cdot \frac{y_R}{\chi d(t_f)}$$

$$+ \nu \cos \beta(t_f) \frac{1}{d(t_f)} \cdot \frac{\cos \beta(t_f)}{\chi} + c. \tag{15.31}$$

Solving for ν:

$$\nu = \frac{-c\chi}{\frac{y_R^2}{d(t_f)^3} + \frac{\cos^2 \beta(t_f)}{d(t_f)}}. \tag{15.32}$$

Because the retreat surface is a straight line in the (x, y)-plane, selecting either $d(t_f)$ or $\beta(t_f)$ determines the other. Therefore, the ν may be computed readily given a choice in one of these variables, and thus the terminal adjoint values, $\lambda(t_f)$, may be computed as well.

In order to compute μ when the path constraint is active (i.e. $g(\mathbf{x}) = 0$), the optimal state dynamics $\dot{\mathbf{x}}^*$ (found by substituting in the optimal adjoint variables and optimal controls) may be substituted into (15.12), which can be then solved for the μ in which $h(\mathbf{x}) = 0$. Alternatively, consider that in order to keep on the constraint $h(\mathbf{x}) = 0$ and similarly $\dot{h}(\mathbf{x}) = 0$:

$$\dot{h} = \frac{\partial h}{\partial \mathbf{x}} \dot{\mathbf{x}} + \frac{\partial h}{\partial \lambda} \dot{\lambda} + \frac{\partial h}{\partial \mu} \dot{\mu} = 0$$

Also, when the trajectory enters or exits a constrained arc $\mu = 0$. Thus μ can be appended to $\mathbf{x}$ over the constrained portion of the trajectory with $\mu(t_1) = 0$ and,

$$\dot{\mu} = \frac{-\frac{\partial h}{\partial \mathbf{x}} \dot{\mathbf{x}} - \frac{\partial h}{\partial \lambda} \dot{\lambda}}{\frac{\partial h}{\partial \mu}},$$

where t_1 is the time at which the trajectory enters a constrained arc (in backwards time). Let t_2 be the time (in backwards time) at which the trajectory leaves the constrained arc and enters an unconstrained arc. In order for the tangency condition to be met, it must be that $h(\mathbf{x}(t_2)) = 0$. As a result, the adjoint variables are subject to an additional internal boundary condition [40, 84],

$$\lambda^\top(t_2^-) = \lambda^\top(t_2^+) + \pi \frac{\partial h}{\partial \mathbf{x}}, \tag{15.33}$$

where π is an additional adjoint variable.

Generally, one would use (15.33) to solve for π by substituting into the Hamiltonian, (15.13), evaluated at $t = t_2^-$. However, computing $\frac{\partial h}{\partial \mathbf{x}}$ is nontrivial as it is subject to the numerical inaccuracies introduced by sampling σ (the GoE adjoints). A different approach is taken here which avoids the computation of $\frac{\partial h}{\partial \mathbf{x}}$. Since $\dot{\mathbf{x}}$, σ_d, and σ_α do not depend on β, $\frac{\partial h}{\partial \beta} = 0$. Thus, from (15.33), it must be that $\lambda_\beta(t_2^-) = \lambda_\beta(t_2^+)$ and thus there are two unknowns: $\lambda_d(t_2^-)$ and $\lambda_\alpha(t_2^-)$. At $t = t_2^-$ it must be that $h(\mathbf{x}(t_2^-)) = 0$ due to the tangency condition [40]. It is also the

case that $h\left(\mathbf{x}(t_2^+)\right) = 0$ since the system is constrained after t_2. Expanding (15.12) yields,

$$h(\mathbf{x}) = \sigma_d \cos \psi^* + \sigma_\alpha \left(\omega^* - \frac{1}{d} \sin \psi^* \right) + c, \tag{15.34}$$

since $\sigma_\beta = 0$. The quantities σ_d and σ_α are continuous at t_2 and c is constant, thus, from (15.34) it must be true that $\psi^*(t_2^-) = \psi^*(t_2^+)$. Eqs. (15.22) and (15.23) can be manipulated to solve for $\lambda_d(t_2^-)$, noting that $\mu(t_2^-) = 0$

$$\lambda_d(t_2^-) = \frac{\lambda_\beta(t_2^-)\left(\lambda_d(t_2^+) + \mu(t_2^+)\sigma_d\right)}{\left(\lambda_\beta(t_2^+) - \mu(t_2^+)\sigma_\alpha\right)}. \tag{15.35}$$

Because the dynamics, (15.1), are autonomous and $H(t_f) = 0$ the Hamiltonian must be zero at all times including at time t_2^-. Evaluating (15.13) at t_2^- and substituting the value of (15.35) in along with $\mu = 0$ gives

$$
\begin{aligned}
H^*(t_2^-) = 0 &= \boldsymbol{\lambda}\dot{\mathbf{x}} + \mu h + L \tag{15.36} \\
&= \lambda_d(t_2^-)\dot{d}(t_2^-) + \lambda_\alpha(t_2^-)\dot{\alpha}(t_2^-) + \lambda_\beta\dot{\beta}(t_2^-) + 0 + c \\
&= \lambda_d(t_2^-) \cos \psi^* + \lambda_\alpha(t_2^-)\left(\omega^* - \frac{1}{d}\sin\psi^*\right) \\
&\qquad + \lambda_\beta(t_2^-)\frac{1}{d}\sin\psi^* + c.
\end{aligned}
$$

Substituting in (15.22) and (15.23), rearranging, and solving for $\lambda_\alpha(t_2^-)$ gives

$$\lambda_\alpha(t_2^-) = \frac{\sqrt{\lambda_d(t_2^-)^2 + \frac{1}{d^2}\lambda_\beta(t_2^-)^2} - c}{\omega^* + \frac{\lambda_\beta}{d^2\xi}} \tag{15.37}$$

Note that ω^* depends on the sign of λ_α (c.f. (15.21)), so its sign may be assumed *a priori* and then (15.37) must be checked for consistency.

15.3.2 *Boundary Value Problem*

For backwards shooting, the BVP consists of choosing a $\beta(t_f)$ (which also determines λ_β, $d(t_f)$, and $\lambda_d(t_f)$), a value for t_f, and a value for $\alpha(t_f)$ and then integrating backwards in time from t_f to t_0. At t_0 the

state values may be substituted into the first three components of ϕ in (15.2) to yield a residual (the fourth component is 0 by construction/selection of the terminal state). Formally,

$$\beta_f^*, \alpha_f^*, t_f^* = \underset{\beta_f, \alpha_f, t_f}{\arg\min} \|\phi\|$$

$$\text{s.t. } g(\mathbf{x}(t)) \geq 0 \; \forall t \in \left[t_0, t_f\right], \tag{15.38}$$

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

$$\text{Eqs. (15.17)–(15.18), (15.19), (15.21), (15.22), (15.23).}$$

In practice, dealing with the path constraint forces one to assume a sequence of arcs, e.g., UCU, where U denotes an unconstrained arc and C denotes a constrained arc (e.g., see [50]). Then, the times at which the system switches from an arc of one type to another must be solved for as well. Due to the similarity of this instantiation of constrained optimal retreat to that of the example in [84], it is assumed that UCU is indeed the proper sequence to consider, and that the boundary conditions at $t = 0$ are such that the path constraint will indeed be activated. That is, the trajectory will begin unconstrained, transition into a constrained arc, and finally end with another unconstrained arc. In backwards time, let t_1 be the time instant of the first switch, from U to C, and let t_2 be the time instant of the second switch, from C to U. Note, also, that in order to transition from U to C, it must also be the case that $h(\mathbf{x}(t_1)) = 0$ (i.e., the tangency condition must be satisfied). Then the BVP in (15.38) may be reposed as

$$\beta_f^*, \alpha_f^*, t_f^*, t_1^*, \mathbf{x}_1^*, t_2^* = \underset{\beta_f, \alpha_f, t_f, t_1, \mathbf{x}_1, t_2}{\arg\min} \|\phi\|$$

$$\text{s.t. } t_2 \leq t_1,$$

$$g(\mathbf{x}(t_1)) = 0, \tag{15.39}$$

$$h(\mathbf{x}(t_1)) = 0,$$

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

$$\text{Eqs. (15.17)–(15.18), (15.19), (15.21), (15.22), (15.23).}$$

At t_1, when the constraint first becomes active the integration proceeds backwards in time, using the constrained version of the optimal state,

adjoint, and control dynamics wherein $\mu \neq 0$. Then, at t_2, the jump condition (15.33) is used to update λ_d and λ_α and then the integration proceeds with $\mu = 0$. The satisfaction of $g(\mathbf{x}(t)) = 0$, $\forall t \in [0, t_f]$, $t \neq t_1$ is guaranteed by way of the optimality conditions derived in this section. In terms of the implementation, a multiple shooting scheme [226] is employed wherein the first shot comprises the first U arc (in backwards time) from t_f to t_1 and the second shot comprises the remaining C and U arcs. Thus the BVP in (15.39) is augmented to include the stitching constraints at t_1.

15.3.3 *Turret Turn Control*

If the path constraint does not become active, the trajectory is comprised of a single U arc. Consider the case where the constraint becomes active at some point along the trajectory and later becomes inactive; this corresponds to the UCU sequence of arcs.

Case 1: (Terminal unconstrained arc, $t \in [t_1, t_f]$.) This case corresponds to the final U arc in the UCU sequence, or the singular U arc in a completely unconstrained trajectory. Here, $\mu = 0$, and $\lambda_\alpha = 0$ due to (15.17) and (15.25). According to (15.21) the optimal Turret control, ω^* is undefined. Once the trajectory is on the fina unconstrained arc, the Turret's control has no bearing on the outcome. In this time interval, any $\omega_u \in [-\rho, \rho]$ is trivially optimal [84]. As a convention, it is assumed that $\omega_u = -\rho \operatorname{sign}(\sigma_\alpha)$ in this case, which corresponds to the GoE control [2].

Case 2: (Constrained arc, $t \in [t_2, t_1)$.) This case corresponds to the C arc wherein the path constraint is active. Here, $\mu \neq 0$, since $\mu(t_1) = 0$ and $\mu > 0$ when the constraint is active. As in the previous case, $\lambda_\alpha = 0$. Therefore, ω^* is generally well defined according to (15.21) except in the following (sub)cases.

Case 2.1: (Turret locked on, $\alpha = 0$.) If $\alpha = 0$, the Turret's gaze is lined up exactly with the Attacker's position. This configuration corresponds to the Turret's Universal Surface wherein the GoE adjoint $\sigma_\alpha = 0$ [2].

Since $\lambda_\alpha = \sigma_\alpha = 0$, the Turret's optimal control is undefined. In this case, ω^* is given by (15.41) (see Proposition 15.1).

Case 2.2: (Turret looking away, $\alpha = \pi$.) If $\alpha = \pi$, the Turret is looking directly away from the Attacker's position. This configuration corresponds to the Turret's Dispersal Surface wherein the GoE adjoint σ_α is undefined [2]. Thus, the Turret's optimal control, ω^* is undefined as well. In the GoE, the Turret has the authority to choose to turn either clockwise or counterclockwise to remain on an optimal (equilibrium) trajectory. Here, however, the choices are not equivalent because the Attacker's heading control is different than in the GoE. In this case, ω is set to

$$\omega_{DS} = \begin{cases} \rho & \text{if } \cos\beta \geq 0, \\ -\rho & \text{otherwise,} \end{cases} \tag{15.40}$$

which states that if the Attacker is to the right of the Turret, the Turret should turn counterclockwise at full rate. In this configuration, turning counterclockwise results in a larger $\dot\alpha$, bringing the Attacker closer to the Turret's view for the remainder of the constrained arc than if the Turret had turned clockwise (see, e.g., Fig. 14.1).

Case 3: (Initial unconstrained arc, $t \in [t_0, t_2]$.) This case corresponds to the initial U arc in the UCU sequence. Here $\lambda_\alpha \neq 0$ in general, due to (15.37). Since $\mu = 0$ when the constraint is inactive ω^* is well-defined. Moreover, because $\dot\lambda_\alpha = 0$, the Turret always turns in one particular direction during this part of the trajectory.

Proposition 15.1. *When the constraint is active ($g(\mathbf{x}) = 0$) and $\alpha = 0$, the Turret's optimal control is given by*

$$\omega_{US} = -\operatorname{sign}(\lambda_\beta) \cdot \min\left(\left|\frac{\lambda_\beta}{d^2\sqrt{(\lambda_d + \mu\sigma_d)^2 + \frac{1}{d^2}\lambda_\beta^2}}\right|, \rho\right) \tag{15.41}$$

Proof. In the GoE, the Turret's singular control along the Universal Surface is $\omega = 0$, which is the control associated with keeping $\alpha = \sigma_\alpha = 0$. However, in the optimal constraint retreat scenario, a different singu-

lar control may be required to keep $\alpha = 0$. Note that $\sigma_\alpha = 0$ on the Universal Surface. Thus,

$$\dot{\alpha} = \omega - \frac{1}{d} \sin \psi$$

$$= \omega + \frac{\lambda_\beta - \mu \sigma_\alpha}{d^2 \sqrt{(\lambda_d + \mu \sigma_d)^2 + \frac{1}{d^2}(\lambda_\beta - \mu \sigma_\alpha)^2}}$$

$$= \omega + \frac{\lambda_\beta}{d^2 \sqrt{(\lambda_d + \mu \sigma_d)^2 + \frac{1}{d^2}\lambda_\beta^2}},$$

and thus to keep $\alpha = 0$ the Turret should select ω such that $\dot{\alpha} = 0$. However, it is entirely possible that the Turret does not have enough control authority to keep $\dot{\alpha} = 0$ and thus control saturation must be considered – hence the min with ρ in (15.41).

Suppose that the Turret were to disregard (15.41) in the case where the first argument of the min is less than ρ. The Turret's gaze would actually turn *past* the Attacker thereby making $\alpha \neq 0$ and $\sigma_\alpha \neq 0$ and the non-singular turn control, (15.21), would come into play. Eq. (15.21) would dictate that the Turret reverse direction in order to drive $\alpha \to 0$, and the dilemma would begin again. Thus the Turret's turn control would chatter/modulate in such a way as to emulate the behavior captured in (15.41). □

15.4 SOLUTIONS AND RESULTS

The results in this section pertain to the following initial conditions and an assumed terminal Turret look-angle

$$\mathbf{x}_0^\top = \begin{bmatrix} 20.61 & -0.2450 & 1.816 \end{bmatrix}, \qquad \alpha_f = -0.6736. \qquad (15.42)$$

These values correspond to the Attacker beginning at $(-5, 20)$ in the xy-plane and the Turret looking along the y-axis, initially. In order to ensure feasible transition from U to C a sweep of β_f is performed for the assumed α_f. The trajectories in the sweep are integrated starting at $t = t_f$ and proceed (backwards) until the first time at which either $g(\mathbf{x}) = 0$ or $d(t) \sin \beta(t) = 20$.

The purpose of the sweep is to find a value for β_f which, when integrated until $g(\mathbf{x}) = 0$ results in the tangency condition to be satisfied: $h(\mathbf{x}) = 0$. The β_f associated with this point is then used as an initial guess for a Nonlinear Program (NLP) to fine-tune β_f to drive $h(\mathbf{x}) \to 0$ when the constraint becomes active ($g(\mathbf{x}) = 0$). The solution of the NLP also yields a guess for $t_1, d_1, \alpha_1, \beta_1$, and λ_{d_1}. A guess for t_2 (or $\Delta t = t_1 - t_2$, $\Delta t > 0$) is the last piece needed to specify a guess for the BVP, (15.39).

Figure 15.2 shows the result of the BVP solution procedure. The red portion of the Attacker trajectory indicates where the path constraint is active. The Turret's initial look angle is shown by the black arrow; the first blue arrow counterclockwise is the Turret's look angle when the constraint becomes active; the red arrow is the look angle when the trajectory leaves the constraint; the last blue arrow indicates the final Turret look angle. The open black circle indicates the initial boundary conditions at $(x, y) = (-5, 20)$. Note that the Turret is constantly turning counterclockwise in this case. Incidentally, the best trajectory found by the NLP corresponds to the initial guess. The reasons for this may be that the solver was not allowed to run long enough, the integrator was encumbered with too many callbacks (with too high of precision), the solver had trouble finding or moving to $\mathbf{x}(t_2)$ that satisfy $g(\mathbf{x}(t_1)) = 0$ and $h(\mathbf{x}(t_1)) = 0$, or the solver algorithm (constrained optimization by linear approximation, or COBYLA) has difficulties with this particular problem.

Figure 15.3 shows the state, adjoint, and control trajectories associated with the indirect solution. Figure 15.4 compares the Value of Engagement and Value of Retreat along the trajectory; $g(\mathbf{x})$ is the difference between the upper and lower curves, and the curves are coincident along the red constrained arc. The validity of the trajectory is guaranteed by construction: it satisfies all of the necessary conditions for optimality in (15.39) and, from Fig. 15.4, it is clear that $V_E \geq V_R$ along the entire trajectory. In particular, we've shown the existence of and computed optimal retreat control strategies, $\psi^*(t; \mathbf{x})$ and $\omega^*(t; \mathbf{x})$ which is sufficient to state this is a valid solution according to [84, Definition 1 & Theorem 1].
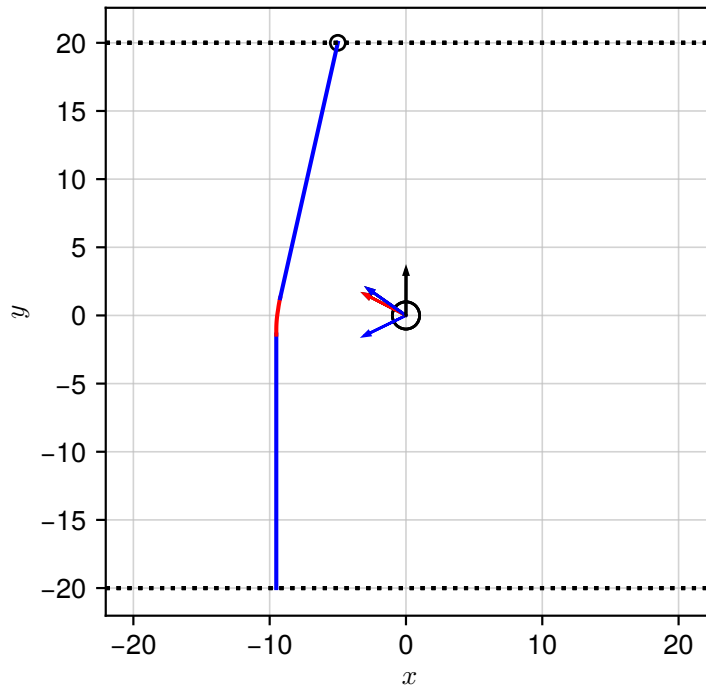
Figure 15.2: Indirect backwards multiple shooting solution trajectory.
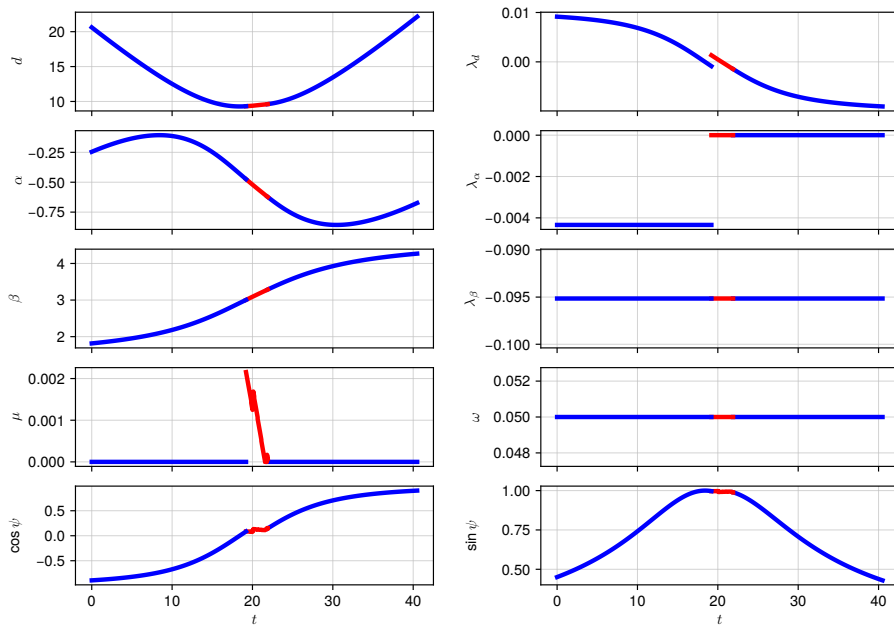
Figure 15.3: Indirect backwards multiple shooting state, adjoint, and control trajectories.
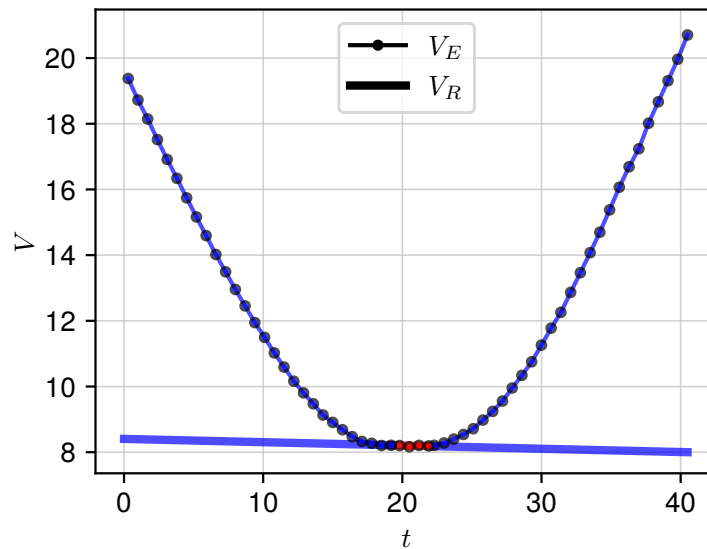
Figure 15.4: Comparison of Value of Engagement and Value of Retreat along
the trajectory.

15.5 CONCLUSION

In this chapter, an engage or retreat scenario was applied to a mobile
Attacker against a stationary Turret with bounded turn rate. Since the
solution of the Game of Engagement is known, the focus was on solv-
ing the Optimal Constrained Retreat problem. The path constraint im-
posed by the Game of Engagement makes the OCR problem both in-
teresting and challenging. The first-order necessary conditions for op-
timality were developed which yielded a system of differential equa-
tions describing the optimal dynamics. Then, a general BVP was speci-
fied with known initial conditions and solved it for a particular point
in the state space using backwards (multiple) shooting.

Part IV

CONCLUSION

16

SUMMARY OF THE CONTRIBUTIONS

Overall, the goal of all of this document has been to argue that what *can* be done rigorously and analytically *should* be done so. In addition, it has been shown that the usefulness of solutions to "small" problems transcends the problem itself. The work followed along two main problem domains within the study of adversarial confrontations: pursuit-evasion and turret defense. With the exception of the Turret, which was modeled as a stationary agent with bounded turn rate and infinite range, the agents were modeled as having control over their instantaneous heading. One may argue that such models are too simplistic to be of any practical use, however, at the very least, these types of solution may provide excellent initial guesses for the optimization of higher fidelity models or serve as starting points for agents who learn and refine their strategies for their particular environment.

One thing that cannot be overstated is the power of the equilibrium strategies in providing performance guarantees, without any regard to the adversary's actions, to the one who implements them. Often, the hope of obtaining the equilibrium is abandoned far too early in the process in favor of utilizing solution methods which are easier to implement (but which yield sub-optimal performance and no insight into the problem). It should be clear that there is still much that can be done in the way of obtaining equilibrium strategies for interesting and relevant scenarios. Not only this, but the approaches demonstrated herein provide additional understanding of the entire state space in the way of win regions or regions corresponding to different end-games.

In particular, the technical contents build up the two problem areas from the smallest (in terms of number of agents) and most restrictive (in terms of the assumptions about each agent's strategy) instances into larger, more complex scenarios. Thus an overarching process has been developed and demonstrated which sits atop traditional game-

theoretic and optimal control methods. Following are several ways in which the "small" solutions have been utilized in larger problems. Often, within the larger problem, the solution degenerates to the smaller solution for a large portion of the state space (as was evident in the two-Pursuer optimal evasion problem, and the two-Turret problem). This may also be described as decomposition, but, here, there is assurance that the decomposition is correct (in the sense that the saddle-point equilibrium property still holds) – it is often the case that decomposing a problem is done out of necessity, at the cost of performance. Next, the small solutions have been used to provide constraints in order to ensure the validity of a particular choice of intention in multi-objective scenarios (as demonstrated in the chapters on engage or retreat). Lastly, the Value function associated with the solution of a small problem was used as the *cost functional* for a larger scenario (in the two-Attacker turret defense game). All of these usages can find application in other adversarial problems. Along the way, many challenges were identified and addressed – even some of the more nuanced challenges, like the difficulty in finding a state-feedback Nash equilibrium for the three-Pursuer game of $\min \max$ capture time, have resulted in solutions and techniques with wide applicability.

FUTURE WORK

This chapter expounds on a some directions for further research in light of the developments of this work. First, a list of direct extensions is given for the specific scenarios considered. The chapter concludes with some less direct extensions which contain significant challenges and warrant, perhaps, a more prolonged research effort.

17.1 DIRECT EXTENSIONS

The extensions in this list follow along in the order of appearance of the technical chapter to which they pertain.

PURE PURSUIT AND CONSTANT EVADER HEADING   Concerning the work on pursuit curves, the case of escape from persistent surveillance is a rich area for further development, especially when one must consider sequences of entries and exits of the Pursuers' effector range. For example, a particular heading may cause the Evader to spend less time in the initial Pursuers' range but allow a subsequent Pursuer to accrue more surveillance time. A more significant extension for the escape scenario is to allow the agents total freedom over their control (i.e., treat it as a differential game). The simpler case, wherein the Evader's course is given and one wishes to obtain the optimal control for the (single) Pursuer which maximizes time-of-surveillance has recently been solved in [264].

PURE PURSUIT AND OPTIMAL EVADER   Recall, for the optimal evasion against two Pursuers employing PP, that the solution, when dual capture is optimal, must be computed numerically. For onboard implementation in a feedback control sense, since the entire state space was filled with optimal trajectories, one may sample the state space to

generate a lookup table for the Evader's instantaneous heading. The method developed to interpolate the wide-beam TDDG could certainly be adapted for this purpose. There are two research directions for furthering and utilizing these results: (1) comparison (and simulation) against the optimal game policies, wherein the Pursuers seek to minimize the Evader's life, and (2) using the Value function from this problem to make task assignments in larger scenarios involving many agents.

M PURSUER GAME OF TIME   There is still much interest in obtaining the Feedback Nash Equilibrium strategies for this differential game, however, the approach presented in Chapter 8 to obtain *robust* strategies has application in many other games whose solutions are not known. In particular, the bounds described by robust policies may be useful for making assignments (as in who ought to pursue who) in a multi-Pursuer multi-Evader scenario.

M PURSUER BORDER DEFENSE   There are several possible extensions to this work, including multiple Evaders, a heterogeneous Pursuer team, Evader with escorts, static border defenses, etc. In particular, for multi-Evaders, one may expect to see certain Evaders performing a sacrificial role as seen in the two-Attacker TDDG. Again, the geometric policies developed have application to larger, multi-Pursuer, multi-Evader border defense scenarios wherein decomposition, team forming, and target assignment may feature prominently. Some work along these lines has already been completed [100]. As regards a more general border geometry, [60] contains some results.

KINETIC TURRET DEFENSE DIFFERENTIAL GAME   For the two-Attacker variant, it remains to formulate and solve a *Game of Degree* in the portion of the state space in which the Turret can guarantee neutralization of both Attackers (the cost functional may be a function of the two distances). Additionally, it would be beneficial to rigorously prove that the Global Stackelberg Equilibrium and State Feedback Nash Equilib-

rium are coincident across the state space. This problem can be made even more interesting with the introduction of additional Attackers.

KINETIC TURRET ENGAGE OR RETREAT   Possible extensions to this work may include considering multiple defensive agents as in [91, 221]. Some initial results, for non-turret defensive agents, are given in [91].

WIDE-BEAM TURRET DEFENSE DIFFERENTIAL GAME   In general, analysis of non-analytic singular surfaces, even "benign" types such as the dispersal surface, is difficult. The numerical approaches developed in Chapter 15 make extensive use of the unique aspects of the game and are built upon existing general purpose algorithms such as numerical integration, root finding, binary search, nonlinear programming, etc. An interesting research direction would be to generalize some of these concepts to better address non-analytic singular surfaces in other games.

WIDE-BEAM TURRET ENGAGE OR RETREAT   There are still some difficulties in applying the optimality conditions to find an optimal trajectory, in general. The three singular surfaces that appear in the Game of Engagement impute some of their complexities and subtleties onto the Optimal Constrained Retreat, particularly when the Value function constraint is active. Specification of the Boundary Value Problem relied on having to assume a sequence of constrained and unconstrained arcs. Some care was taken to choose parameters and initial boundary conditions to yield a trajectory in which the path constraint would become active. The process of computing an initial guess to satisfy the tangency condition when the constraint becomes active is nontrivial. Finally, the barrier surface, which partitions the state space into a region where engagement is optimal and a region where retreat is optimal, could be computed.

17.2   RESEARCH OUTLOOK

As mentioned previously, the developments pertaining to analyzing multi-agent differential games, multi-objective scenarios, and applying the solutions of smaller games as costs in larger games are ready to be applied to novel scenarios. Although each new scenario may present its own nuanced challenges, these approaches should yield some interesting results. Then there is the problem of sensing limitations – in this work, full state information has been given to the agents. For more real-world scenarios involving vehicles with noisy observations, the simple solution is to implement a suitable filter which passes its filtered observations to the policies developed herein. Even more challenging, however, is the case where no information about the opponent's state is known. These problems are related to the field of study known as adversarial search. Some initial progress for engagements similar to those studied here is contained in [217]. Next, there is the problem of limited information about an opponent's capability (i.e., uncertainty in the parameters of the differential game). In all of this work, it has been assumed that the agents know all of the system parameters, though it is clearly more realistic if these were unknown or at least uncertain. Again, a simple solution is to utilize some kind of onboard estimator – if, for example, an opponent's speed is unknown, one may simply measure its speed and update its model accordingly. Whatever current best estimate is available may then be used to compute the appropriate policy. However, this approach may be ignoring the possibility of one or other agent disguising its capability. Finally, obtaining equilibrium (or near-equilibrium) policies for models with more complex dynamics is desired. For this, a lifted representation of the model may be beneficial (e.g., [276]).

Part V

APPENDIX

SUBSTITUTE
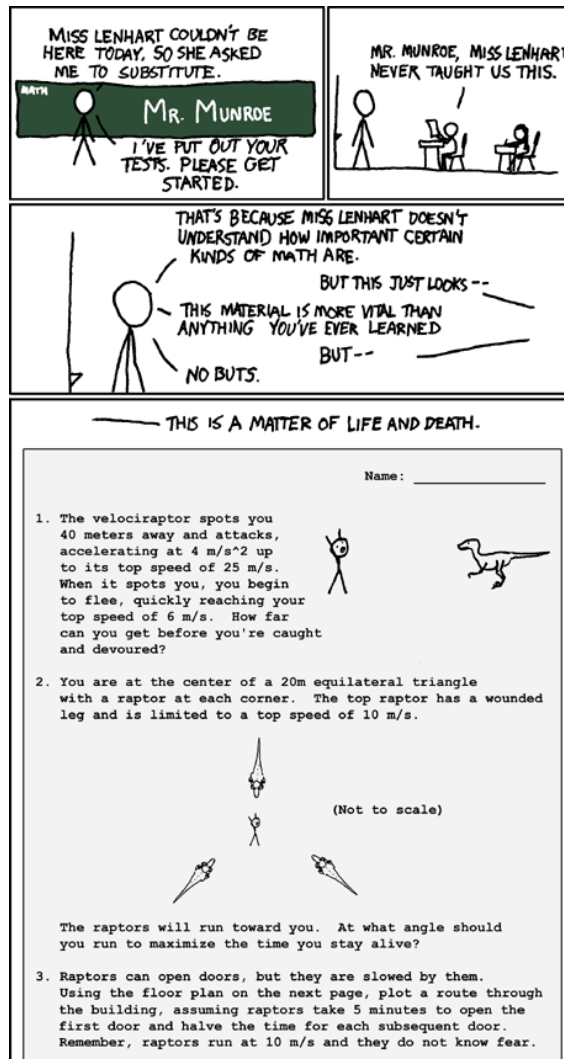
Figure A.1: XKCD comic showing an even more challenging variant of the three-Pursuer, single-Evader differential game of min max capture time [169]. The proposed solutions are based on an assumption that the Pursuers know the Evader's current and future control inputs.

2 PURSUERS AND N EVADERS

The contents of this appendix are based upon the paper [244].

B.1 INTRODUCTION

Cooperation is essential to the survival and success of many different species of animals. Remarkable feats are made possible through cooperative behaviors which emerge from relatively simple processes at the individual level. For example, consider the nest building, foraging, and decision making abilities that insect 'societies' are capable of [4], while at the same time only possessing a modest amount of processing power. Similarly, cooperation among autonomous vehicles (or more abstractly, agents) may enable new types of tasks to be accomplished or better performance on existing tasks. Particularly interesting is the use of cooperation within a group in adversarial scenarios *between* groups or species. In nature, these scenarios include collective defense such as meerkat mobbing [116], distributed nest defense [176], and musk oxen who press together with their horns facing outward [123, 235]; cooperative predation such as the yellowsaddle goatfish [225], wolves [230], and dolphins [111]; and cooperative sensing such as predator inspection performed by guppies [51].

Biological systems have evolved such innovative ways of cooperating that they often serve as the inspiration for optimization algorithms [235], robotic control algorithms [78, 227, 265], and methods for designing controllers (e.g. via evolutionary algorithms) [8, 78, 122]. Regarding the design of controllers, the difficulty often lies in the dimensionality of either the state space or the control space. Genetic algorithms, for example, tend to settle in local optima when the chromosome encoding the controller is large. One approach to reducing the size of control space is to constrain the controller at the outset based

on heuristics or expert knowledge (c.f. [8, 229]). This approach finds its roots in Connell's ideas of minimalist robotics [47], wherein he discussed how a controller comprised of a few simple behaviors could mostly account for the seemingly complex behavior of a snail. Of course, the principle of simplicity in robotic control was heavily inspired by Braitenberg Vehicles whose two sensors are directly connected to its two motors; the different ways of connecting sensors to motors results in fundamentally different, but understandable behaviors [35].

The focus, in this appendix, is on a particular type of predator-prey interaction, though aspects of some of the other natural cooperative behaviors were influential in the problem setup and technical approach. An analogous relationship is that between a pursuer and an evader which is prevalent in controls and optimization literature (see, e.g., [130]). Often, differential game theory is employed to obtain saddle-point equilibrium strategies for the pursuer and the evader depending on the particular cost functional being considered. This type of analysis is typically only possible for systems with simple dynamics and/or few numbers of agents. For example, the two-pursuer one-evader game of min/max capture time was solved in [107] wherein the two pursuers employ a "pincer" maneuver to reduce the capture time of the evader w.r.t. either of their individual capture times. Similarly, Breakwell et al. [37] solved the one-pursuer two-evader game of min/max capture time (of the second evader captured). Other cost functionals have been explored in the one-against-two game which drastically changes the equilibrium control strategies [85].

A similar style of pursuit-evasion (or predator-prey) game with similar (single-integrator) dynamics is considered here, but with much larger numbers of prey. Also, while the mentioned literature pits two against one (or vice versa), a two-against-many scenario is explored. The goal of the two predators is to maximize the collective number of prey captured within a specified time horizon. They do so without communicating or directly coordinating with their fellow teammate, whereas teammates in a differential game context are essentially treated as a single entity. Thus the proposed control strategy is inherently decentralized, which more closely represents a biological system.

In order to focus on the cooperation of the predators, in particular, the prey control strategy is fixed to a Boids-like controller (c.f. [200]). As a result, the predators can learn to exploit the underlying prey flocking behavior by affecting the *shape* of the prey distribution. The predator control is based on a nearest-neighbor approach based on [8], though the weighted nearest-neighbor is considered here. Each predator measures the (meta) state of the system and implements the baseline behavior associated with the nearest anchor point.

These approaches and concepts are applicable to behavior generation for non-player characters (NPCs) as well as military training simulations. In particular, a neural network-based controller was evolved using a genetic algorithm for a real-time strategy (RTS)-like game in [122]. To contrast, the controller structure employed here is not a neural network and the focus is on movement and positioning (as opposed to game elements like resource gathering). A neural network controller generated via genetic algorithm was also employed in [204] where the emphasis was on so-called "multi-modal" behaviors. One of the games considered was a fight or flight game where at any given time the player is executing one or other particular behavior; this concept of a multi-modal behavior or controller is quite related to this appendix. Reference [62] addresses generation of NPC behaviors but from the angle of planning (e.g., via meta-heuristic search) which occurs on a much larger timescale than the type of control focused on in this appendix. Finally, the application of artificial intelligence and game concepts to military training simulations is demonstrated in [238] on task allocation and planning problems with an emphasis on "explainability".

The contributions of this appendix are listed as follows: (i) specification of baseline cooperative predator strategies; (ii) specification of many-against-many metastates; (iii) an anchor point control architecture based on weighted nearest anchor point; and (iv) a decentralized predator control strategy which outperforms a baseline strategy. Section B.2 contains the problem formulation. Sections B.3 and B.4 describe the prey model and predator model, respectively. The controller structure is also specified in Section B.4. Section B.5 specifies the evo-

lutionary algorithm along with all of the genetic operators particular to the control structure employed. Simulation results are contained in Section B.6, and Section B.7 concludes the paper.

B.2   PROBLEM FORMULATION

The environment in which the scenario takes place is a flat plane with no boundaries or obstacles. Let the predator positions be given as $\mathbf{P}_1 = (x_{\mathbf{P}_1}, y_{\mathbf{P}_1})$, $\mathbf{P}_2 = (x_{\mathbf{P}_2}, y_{\mathbf{P}_2}) \in \mathbb{R}^2$. Similarly, the prey positions are denoted by $\mathbf{E}_j = (x_{\mathbf{E}_j}, y_{\mathbf{E}_j}) \in \mathbb{R}^2$ for $j \in 1, \dots, M$, and $M \in \mathbb{Z}$. Let the full system state be denoted as $\mathbf{x} = \begin{bmatrix} \mathbf{P}_1^\top & \mathbf{P}_2^\top & \mathbf{E}_1^\top & \dots & \mathbf{E}_M^\top \end{bmatrix}^\top$. In general, it is assumed that the maximum predator and prey speeds, v_P and v_E, are such that $v_P > v_E$, and that agent speeds are homogeneous within each respective group. All the agents move with simple motion, i.e.,

$$\dot{\mathbf{P}}_i = v_P \begin{bmatrix} \cos \psi_i \\ \sin \psi_i \end{bmatrix}, \quad \dot{\mathbf{E}}_j = \bar{v}_E \begin{bmatrix} \cos \phi_j \\ \sin \phi_j \end{bmatrix},$$

where $\psi_i, \phi_j \in [0, 2\pi]$ are the instantaneous heading angle of predator i and prey j, respectively, and $\bar{v}_E \in [0, v_E]$. Note that the predator is assumed to always move with its maximum speed whereas the prey is allowed to slow down (subject to the prey model discussed in detail in Section B.3). For the purposes of numerical simulation, the scenario takes place in discretized time and the agents' headings are computed/updated simultaneously. Thus the discrete time step $t_k = k\Delta t \in [0, T]$, where T is the time horizon of the simulation and is an integer multiple of Δt, and $k \in \{1, 2, \dots, T/\Delta t\}$. Throughout the remainder of the paper, the current positions of the agents are generally referred to without explicitly stating its dependence on time (e.g. $\mathbf{E}_j$ rather than $\mathbf{E}_j(t_k)$).

Capture is said to occur when a predator comes within a distance d_c of a prey agent. An indicator function $\mathbb{I}_{\text{cap}}(i,j,t_k)$ is used to represent predator i ($i \in \{1,2\}$) capturing prey j in time step t_k:

$$\mathbb{I}_{\text{cap}}(i,j,t_k) = \begin{cases} 1 & \text{if } \|\mathbf{P}_i - \mathbf{E}_j\| \leq d_c, \\ 0 & \text{otherwise}, \end{cases} \tag{B.1}$$

where $\|\cdot\|$ is the Euclidean norm. Once a prey agent has been captured, it is effectively removed from the scenario and the capturing predator is free to move on to other targets immediately thereafter.

The predators have a shared goal of maximizing the number of prey captured over a time horizon of T simulation seconds,

$$U\left(\psi_1\left(\mathbf{x}(t)\right), \psi_2\left(\mathbf{x}(t)\right)\right) = \sum_{i=1}^{2} \sum_{j=1}^{M} \sum_{t_k} \mathbb{I}_{\text{cap}}(i,j,t_k), \tag{B.2}$$

where $\psi_i\left(\mathbf{x}(t)\right)$ is the state-feedback control law of predator i. In practice, (B.2) must be modified to exclude double-counted captures (i.e. when prey j is within d_c of both predators). Because the utility is shared, this scenario is not an example of by-product mutualism, wherein cooperation arises from selfish acts, as in [225]. With the prey behavior fixed, the aim of this study is to design a state-feedback controller for the predators.

B.3 PREY MODEL

The behavior of the prey agents is governed by a Boids-like model (c.f. [200]); this section describes the details of the model. In order to mimic the flocking behavior of animals, the Boids model introduces inertia into the agents' motion and models various influences as virtual forces. The virtual forces governing the behavior of the prey are alignment, cohesion, and separation [200] as well as avoidance (of nearby predators). Another augmentation to the original Boids model is the inclusion of a finite sensing radius, r, for the prey. That is, a prey agent only knows, or takes into consideration, the positions of predators and

prey within r distance of its current position. Here, the prey agents are treated as if they have unit mass.

Three sets are used in formulating the virtual forces. The first is the set of prey agents within the sensing radius r,

$$\mathcal{R}_j := \left\{ j' \mid j' \neq j, \|\mathbf{E}_j - \mathbf{E}_{j'}\| \leq r \right\}. \tag{B.3}$$

Next is the set of prey agents within a desired minimum separation distance s,

$$\mathcal{S}_j := \left\{ j' \mid j' \neq j, \|\mathbf{E}_j - \mathbf{E}_{j'}\| \leq s \right\}. \tag{B.4}$$

Last is the set of predators within the sensing radius r,

$$\mathcal{P}_j := \left\{ i \mid \|\mathbf{E}_j - \mathbf{P}_i\| \leq r \right\}. \tag{B.5}$$

The current centroid of prey positions and velocity vectors within distance r of prey j are respectively given by

$$\bar{\mathbf{E}}_j = \frac{1}{|\mathcal{R}_j|} \sum_{j' \in \mathcal{R}_j} \mathbf{E}_{j'}, \quad \dot{\bar{\mathbf{E}}}_j = \frac{1}{|\mathcal{R}_j|} \sum_{j' \in \mathcal{R}_j} \dot{\mathbf{E}}_{j'}. \tag{B.6}$$

Alignment is based on the principle that the prey naturally seek to align their velocity with the overall direction of travel of the flock,

$$\mathbf{F}_{\mathrm{ali}_j} = \frac{\dot{\bar{\mathbf{E}}}_j}{\|\dot{\bar{\mathbf{E}}}_j\|}. \tag{B.7}$$

Cohesion provides some influence for prey to gravitate towards the centroid of prey positions which prevents the flock from splitting apart,

$$\mathbf{F}_{\mathrm{coh}_j} = \frac{\bar{\mathbf{E}}_j - \mathbf{E}_j}{\|\bar{\mathbf{E}}_j - \mathbf{E}_j\|}. \tag{B.8}$$

Because the agents represent some physical entity, a force designed to maintain a desired minimum separation distance, s, is necessary,

$$\mathbf{F}_{\mathrm{sep}_j} = \frac{1}{|\mathcal{S}_j|} \sum_{j' \in \mathcal{S}_j} \frac{\mathbf{E}_j - \mathbf{E}_{j'}}{\|\mathbf{E}_j - \mathbf{E}_{j'}\|^2} \tag{B.9}$$

In this case, the contribution due to each neighboring prey is distance-weighted to prioritize nearer violators of the desired minimum separation distance. Finally, avoidance forces the prey to flee from predators within the sensing radius r,

$$\mathbf{F}_{\text{avo}_j} = \frac{1}{|\mathcal{P}_j|} \sum_{i \in \mathcal{P}_j} \frac{\mathbf{E}_j - \mathbf{P}_i}{\|\mathbf{E}_j - \mathbf{P}_i\|^2}. \tag{B.10}$$

Again, closer predators provide more force.

With all of the necessary forces defined, the (candidate) velocity update law of the prey is given as,

$$\dot{\mathbf{E}}_j'(t_k) = \dot{\mathbf{E}}_j(t_{k-1}) + \Delta t \, (w_{\text{ali}} \mathbf{F}_{\text{ali}} \tag{B.11}$$
$$+ w_{\text{coh}} \mathbf{F}_{\text{coh}} + w_{\text{sep}} \mathbf{F}_{\text{sep}} + w_{\text{avo}} \mathbf{F}_{\text{avo}}),$$

where the w's are weights corresponding to the level at which the prey are influenced by each force. Then the candidate velocity is passed through a saturation function to ensure it does not exceed the maximum prey speed,

$$\dot{\mathbf{E}}_j = \begin{cases} \dot{\mathbf{E}}_j' & \text{if } \|\dot{\mathbf{E}}_j'\| \leq v_E, \\ \dot{\mathbf{E}}_j' \frac{v_E}{\|\dot{\mathbf{E}}_j'\|} & \text{otherwise.} \end{cases} \tag{B.12}$$

B.4 PREDATOR MODEL

The overall predator control approach is based on the idea of specifying simple atomic behaviors and then letting the predator decide which behavior to implement in each time step. Unlike the prey, the predators have full access to the state, $\mathbf{x}$ (the positions of each agent). Each predator selects its own atomic behavior to implement – that is, they do not collaborate on which behavior to select. Again, this approach is decentralized, in contrast to more explicit cooperative maneuvers in studies like [229] which require communication.

The first atomic behavior is the `pursue` behavior wherein the predator, $\mathbf{P}_i$, aims directly at the nearest prey agent. The index of the nearest prey agent is given by

$$j^* = \arg\min_j \|\mathbf{P}_i - \mathbf{E}_j\|, \tag{B.13}$$

and the associated heading angle is

$$\psi_{\mathrm{pur}} = \mathrm{atan2}\left(y_{\mathbf{E}_{j^*}} - y_{\mathbf{P}_i}, x_{\mathbf{E}_{j^*}} - x_{\mathbf{P}_i}\right), \tag{B.14}$$

where atan2 is the four-quadrant inverse tangent function. In the one-pursuer, one-evader differential game of $\min\max$ capture time with simple motion this is, in fact, the equilibrium strategy for the pursuer (c.f. [130]). Note, this may not be optimal for the predator since the prey is not necessarily implementing its equilibrium strategy. Nonetheless, it is robust to any prey strategy. It is also useful here as the one-on-one pursuit-evasion game may be considered to be a subproblem to the overall scenario.

Next, the `converge` and `diverge` behaviors are based on mixing pure convergence (i.e. aiming directly towards the predator centroid) or divergence (i.e. aiming directly away from the predator centroid) with aiming at the prey centroid. Inclusion of `diverge` is partly based on [138], wherein the author utilized an intra-predator repulsive force in controlling a group of pursuers pursuing a single evader in a decentralized fashion. Let M' be the number of prey currently living; the angles from the predator to the prey centroid and predator centroid are given as

$$\psi_{\bar{\mathbf{E}}} = \mathrm{atan2}\left(\left(\sum_j \frac{y_{\mathbf{E}_j}}{M'}\right) - y_{\mathbf{P}_{i'}}, \left(\sum_j \frac{x_{\mathbf{E}_j}}{M'}\right) - x_{\mathbf{P}_i}\right), \tag{B.15}$$

$$\psi_{\bar{\mathbf{P}}} = \mathrm{atan2}\left(\left(\sum_{i'} \frac{y_{\mathbf{E}_{i'}}}{N}\right) - y_{\mathbf{P}_{i'}}, \left(\sum_{i'} \frac{x_{\mathbf{E}_{i'}}}{N}\right) - x_{\mathbf{P}_i}\right), \tag{B.16}$$

respectively. Then the `converge` and `diverge` behaviors are governed by

$$\psi_{\text{con}} = \text{atan2} \left(m \sin \psi_{\bar{\mathbf{P}}} + \sin \psi_{\bar{\mathbf{E}}}, m \cos \psi_{\bar{\mathbf{P}}} + \cos \psi_{\bar{\mathbf{E}}} \right), \tag{B.17}$$

$$\psi_{\text{div}} = \text{atan2} \left(-m \sin \psi_{\bar{\mathbf{P}}} + \sin \psi_{\bar{\mathbf{E}}}, -m \cos \psi_{\bar{\mathbf{P}}} + \cos \psi_{\bar{\mathbf{E}}} \right), \tag{B.18}$$

where m is the mixing weight; the larger m the more the behaviors approach pure convergence or divergence.

The last two behaviors, `drive` and `flank`, both make use of an angle, ψ_{den}, which represents the angle from the predator to the highest distance-weighted density of prey. A Kernel Density Estimator (KDE) is used to compute an estimate of ψ_{den}. First, the angles from the predator to each living prey are computed as

$$\psi_{i,j} = \text{atan2} \left(y_{\mathbf{E}_j} - y_{\mathbf{P}_i}, x_{\mathbf{E}_j} - x_{\mathbf{P}_i} \right). \tag{B.19}$$

Then the distribution of $\psi_{i,j}$ is smoothed via the following estimator:

$$\hat{f}_h(\psi) = \frac{1}{hM'} \sum_j \frac{1}{\|\mathbf{P}_i - \mathbf{E}_j\|} K \left(\frac{\psi - \psi_{i,j}}{h} \right), \tag{B.20}$$

where $h > 0$ is the bandwidth of the estimator and K is the kernel. For this study, the bandwidth is set using Silverman's Rule [222], and the standard Gaussian distribution is used for the kernel. Finally, the angle from the predator to the highest distance-weighted density of prey is defined as

$$\psi_{\text{den}} = \arg\max_{\psi} \hat{f}_h(\psi). \tag{B.21}$$

The `drive` behavior is governed by

$$\psi_{\text{dri}} = \psi_{\text{den}}, \tag{B.22}$$

thus the predator aims always in the direction of maximum distance-weighted prey density. The main purpose of including this behavior is to avoid the potential pitfall of wasting time pursuing a single prey agent away from an advantageous cluster of prey.

Influence over the distribution or shape of the prey flock is the main purpose of the `flank` behavior. In general, it is better for the predators when the prey are highly concentrated; however, once a predator approaches, the flock will evade and disperse. Thus the `flank` behavior is designed to aim the predator in a direction tangential to the flock. Let $i' = \arg\min_i \|\mathbf{P}_i - \mathbf{P}_{i'}\|$ be the index of the predator closest to the ith predator.

$$\psi_{\text{fla}} = \psi_{\text{den}} - \frac{\pi}{2} \cdot \text{sign}\left(\begin{bmatrix} \cos\psi_{\text{den}} \\ \sin\psi_{\text{den}} \end{bmatrix} \times (\mathbf{P}_{i'} - \mathbf{P})\right) \tag{B.23}$$

The last term ensures that the tangential direction is one which points away from the nearest predator.

(a) `pursue`   (b) `converge`   (c) `diverge`

(d) `drive`   (e) `flank`
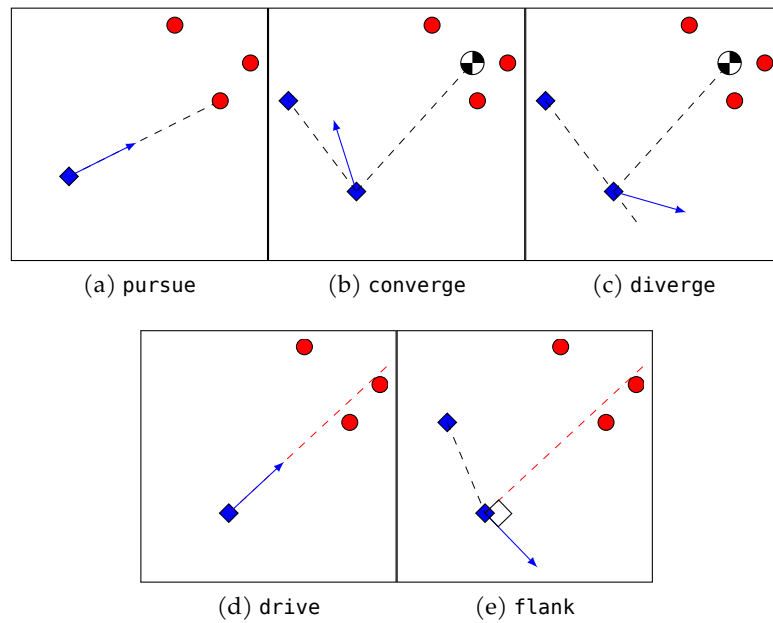
Figure B.1: Demonstration of each of the five predator behaviors. The red dashed line indicates the direction of maximum distance-weighted density of prey, ψ_{den}.

Figure B.1 shows a pictorial representation of each of the predator behaviors. Although `pursue` and `drive` (and, similarly `diverge` and `flank`) appear quite similar, the presence of additional prey can effect more obvious distinctions.

B.4.2 *Meta State*

The overall state of the system at any given time step is essentially the positions of all of the predators and all of the prey agents who have not yet been captured. One may consider including another state for the prey which specifies whether a particular agent is alive or dead, but ultimately the positions of dead prey agents ought not have any bearing on the predator's decisions. Aside from the variability in the size of the state space, the main issue is that even with a few living prey agents the number of dimensions is quite large. In an effort to avoid the curse of dimensionality, and to be able to function in the presence of many prey agents, it is prudent to collapse the positional state information into a smaller (meta) state space.

Let the centroid of *all* currently living prey be $\bar{\mathbf{E}} = \frac{1}{M'} \sum_j \mathbf{E}_j$. For the ith predator, define the meta state as

$$\hat{\mathbf{x}}_i(t_k) = \begin{bmatrix} \min_j \|\mathbf{P}_i - \mathbf{E}_j\| \\ \|\mathbf{P}_1 - \bar{\mathbf{E}}\| \\ \vdots \\ \|\mathbf{P}_N - \bar{\mathbf{E}}\| \\ \|\mathbf{P}_i - \mathbf{P}_{\neq i}\| \\ \vdots \end{bmatrix}, \tag{B.24}$$

where N is the number of predators. The meta state is comprised of the distance to the nearest prey, the distances from each predator to the prey centroid, and the distances to each other predator; $\hat{\mathbf{x}}_i \in \mathbb{R}^{2N}$. For $N = 2$ predators the size of the meta state space is 4, regardless of the number of prey in the simulation.

B.4.3 *Controller Structure*

In order to limit the complexity of this initial study, it is assumed that the predators to be homogeneous – that is, they share the same controller, ψ, which is a function of the meta state. Thus let $\psi_i(\mathbf{x}) := \psi(\hat{\mathbf{x}}_i)$ for $i \in \{1, 2\}$. In order to evolve the predators' feedback controller, it must be parameterized in such a way that will allow the application of

crossing and mutation operations to it. For the research, this is accomplished by using an anchor point method [8].

Anchor points $a := [\hat{x}, w, u]$ are comprised of a position (in the meta state space) $\hat{x}$, a weight w, and a control u. In this case, the control is selected as one of the atomic behaviors described in Section B.4.1:

$$u \in \left\{ \psi_{\text{pur}}, \psi_{\text{con}}, \psi_{\text{div}}, \psi_{\text{dri}}, \psi_{\text{fla}} \right\}. \tag{B.25}$$

A set of anchor points $S = [a_1, a_2, \ldots a_N]$ can be used to parameterize a feedback controller by using the anchor points as the basis for a weighted nearest neighbor switching controller.

$$[\hat{x}^*, w^*, u^*] = \underset{[\hat{x}_i, w_i, u_i] \in S}{\arg \min} \; w_i \|\hat{x} - \hat{x}_i\| \tag{B.26}$$

Figure B.2 shows a 2D example of a feedback controller parameterized
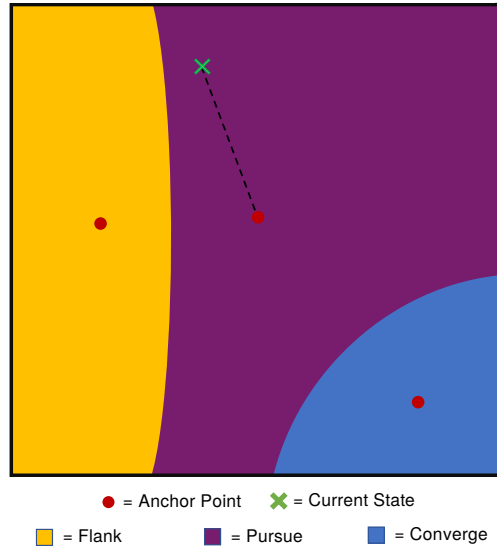
Figure B.2: Example of an anchor point feedback controller in a fictitious 2-dimensional meta state space.

by a set of 3 anchor points. At any given state $\hat{x}$, the control can be computed by finding the weighted nearest anchor point, a^*, by (B.26) and implementing the associated control, u^*. The higher the weight associated with a particular anchor point the further it will appear, in this case.

The evolutionary algorithm (EA) used in this research follows a canonical EA format. Generation G_0 begins with an initial population of P candidate controllers, each of which is comprised of $n_l \sim \text{Uniform}(1, n)$ anchor points that are randomly placed in the 4-dimensional meta state space with a randomly assigned behavior and a random weight, $w \sim \text{Uniform}(0, 1)$. This initial population is evaluated, assigned a fitness, and entered as the first generation of the evolutionary loop. In each generation, the population is crossed, mutated, and evaluated using the methods outlined in this section. Once a predetermined number of generations is completed, the EA ends and the candidate controller with the highest fitness is considered the best evolved controller.

B.5.1   *Crossing*

Crossing begins by selecting two unique parent controllers from the current generation. Each parent has an equal probability of being selected. The child controller is created by using a uniform crossover technique in which each child anchor is randomly selected from the corresponding anchors of the two parents. This process is repeated until P children are created. The new set of children controllers are then added into the current population to create a combined population of size $2P$.

B.5.2   *Mutation*

After the crossing is complete, the combined population (comprised of both parents and children) is passed through a mutation operation. There are two distinct types of mutation that can occur; *major* and *minor* mutations. The probability for these mutations to occur to a given anchor point is given by their respective mutation rates: μ_M for major mutation and μ_m for minor mutation.

Major mutations randomly change the value of the control associated with a given anchor point. For example, if an anchor point has a control value of `pursue`, a major mutation can change it to a control value of `drive`. This can have a drastic effect on the performance of the controller and therefor is assigned a relatively small probability of happening: $\mu_M = 1\%$.

Minor mutations shift the position of a given anchor point a small amount in a random direction. Given an anchor point $\hat{\mathbf{x}}$ with weight w, the mutated version $\hat{\mathbf{x}}'$, w' can be computed as

$$\hat{\mathbf{x}}' = \hat{\mathbf{x}} + \mathbf{R}, \qquad w' = w + \text{Uniform}(0, 1)$$

where $\mathbf{R} \in \mathbb{R}^4$ is a vector of uniform random numbers in the range 0 to 1. Minor mutations are designed to slightly modify the boundaries between the different regions of control and thus will usually have a relatively small effect on the controllers performance. Minor mutations are applied to anchor points with a probability of $\mu_m = 10\%$

B.5.3 *Fitness Evaluation*

After the crossing and mutation operations have been completed, the combined population is then evaluated and each candidate controller is assigned a fitness. The fitness is defined as the average utility of 9 different simulations (all with the same settings) with various initial conditions. Figure B.3 shows all 9 of the sets of initial conditions. These same 9 configurations are used to assess the fitness in every generation. All of the configurations begin with the prey concentrated in a circular ball. In the first column of Fig. B.3, the prey start with zero velocity; the second column has the prey all moving in the same direction at max velocity; and the third column has the prey moving with random velocities. For the predators, the first row of Fig. B.3 starts the predators relatively close together; the second row places the two predators nearly opposite one another w.r.t. the prey and at similar distances; and the third row starts the predators nearly opposite one another but with one of the predators closer to the prey. The purpose of the differ-
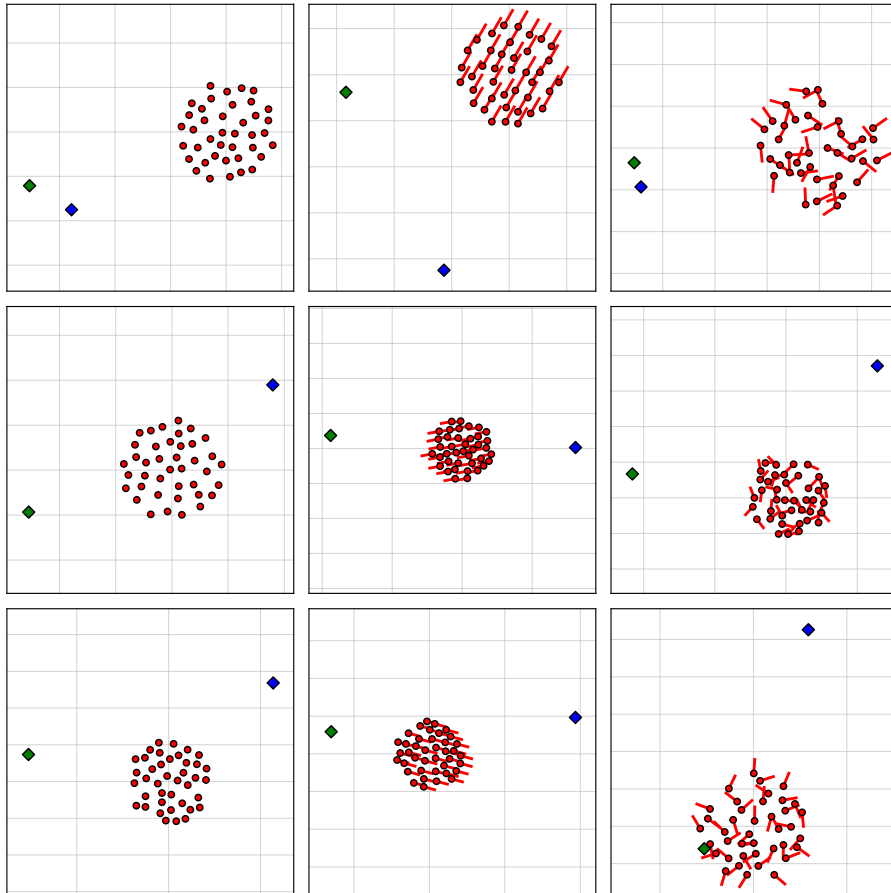
Figure B.3: Suite of initial conditions to simulate to determine fitness of an individual controller.

ent initial conditions is to expose the controller to a variety of scenarios in an effort to avoid over-fitting to a particular configuration. Once all the agents are assigned a fitness, the top 30% of the combined population are selected along with $0.4P$ individuals chosen from the bottom 70% to pass on to the next generation.

Table B.1 summarizes the settings used for the EA simulation in the following section. Note the odd value for population size is due to the number of cores available on the computer in which the EA was run. In each generation, the fitness evaluation is performed in parallel (one individual per core).

Table B.1: EA Parameter Settings

Parameter	Value	Description
n	10	maximum number of anchor points
μ_m	0.1	minor mutation probability
μ_M	0.01	major mutation probability
G_max	100	number of generations
P	190	population size

B.6   RESULTS

This section contains the results of the EA as well as a Monte Carlo simulation comparing the best evolved controller to two baseline controllers across many different initial conditions not previously seen by the EA. Table B.2 contains the simulation parameters used in all of the experiments. Note that the EA-learned controller was trained for these particular settings of prey virtual force weights. Thus the EA-learned controller is specifically tuned to this particular prey behavior. The cohesion weight, w_coh, in particular, is quite high in comparison to the other forces. This is to encourage the prey to cluster together more since the interest is in examining how predators should approach/maneuver around clusters of prey. For most cases, T is not large enough for the predators to capture all of the prey.

Figure B.4 shows the evolution of the best and average utility (over the whole population) for 100 generations. As indicated by the consistent gap between the best and average fitnesses at each generation,

Table B.2: Simulation Parameter Settings

Parameter	Value	Description
M	40	number of prey
N	2	number of predators
Δt	0.01	timestep
T	15	final time
v_E	0.1	max evader speed
v_P	0.2	max pursuer speed
r	0.6	prey sensing range
s	0.05	desired separation
d_c	0.01	capture distance
w_{coh}	7	weight for cohesion force
w_{ali}	0.1	weight for alignment force
w_{sep}	1	weight for separation force
w_{avo}	2	weight for avoiding the predator

some diversity within the population is maintained throughout the evolution. By the end of the 100 generations, the best evolved controller's utility is 20% higher than the baseline controller, which always uses the `pursue` behavior (aim at the nearest prey).
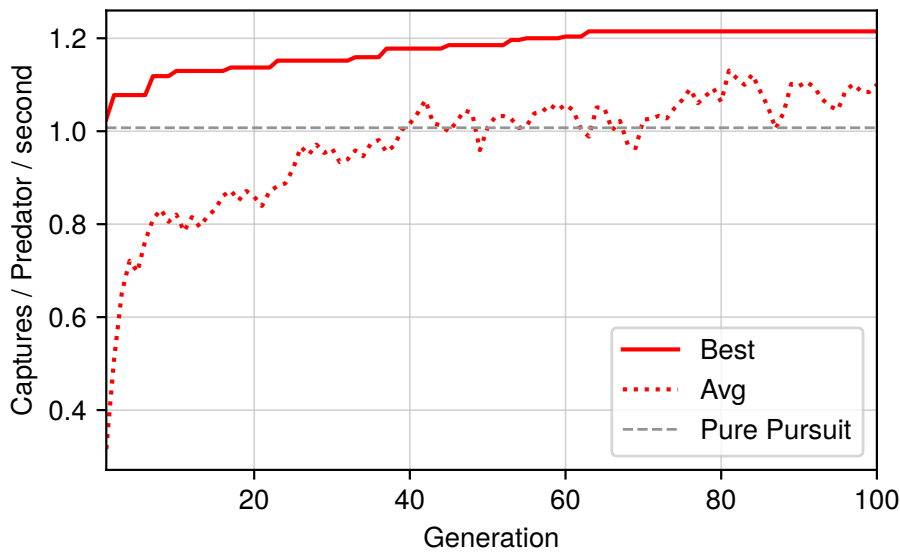
Figure B.4: Evolutionary algorithm results - best and average normalized fitness for each generation. The gray line is the average performance of the baseline Pure Pursuit controller on the test suite.

Figure B.5 contains an animation of the best evolved controller starting from a particular initial condition (seen during evolution). For most of the simulation, the predators utilize the `pursue` behavior; Table B.3 summarizes the amount of time spent using each of the behaviors. Interestingly, the best evolved controller's anchor points are comprised
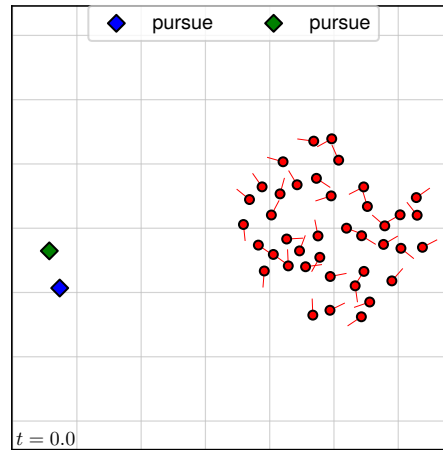
Figure B.5: Simulation of the best controller evolved after 100 generations on one of nine initial conditions in which it was tested. Also available at avonmoll.github.io/files/pred_prey.gif

Table B.3: Summary of behaviors used in simulation shown in Fig. B.5

Behavior	P_1 Percentage of Time	P_2 Percentage of Time
pursue	99.6	93.4
converge	0	0
diverge	0	0
drive	0	0.13
flank	0.4	6.47

only of the pursue, drive, and flank behaviors. Neither converge nor diverge are even present in the controller. It appears that these two behaviors were mostly "evolved out" of the population (or at least moved to a remote area of the meta state space, thereby limiting its activation). Because individual predators are capable of capturing individual prey agents alone, the converge behavior is almost never necessary. As far as diverge goes, the flank behavior appears to be a slightly more useful means of creating separation between the predators when they come too close. Although drive and flank are used, they are used sparingly, suggesting they are useful only in specific circumstances. Nonetheless, the results in Table B.3 suggest that behaviors that are active for a small amount of time can have a large influence on the overall utility.

Because the controllers in the EA saw the same 9 initial conditions in each generation, over fitting is a concern. It is possible that the controllers learned only how to handle these 9 initial conditions and may generalize poorly for other scenarios. In order to corroborate the per-

formance gain of the evolved controller over the baseline (as shown
in Fig. B.4) a Monte Carlo experiment is run over 1000 different (pre-
viously unseen) initial conditions. Figure B.6 contains the results of
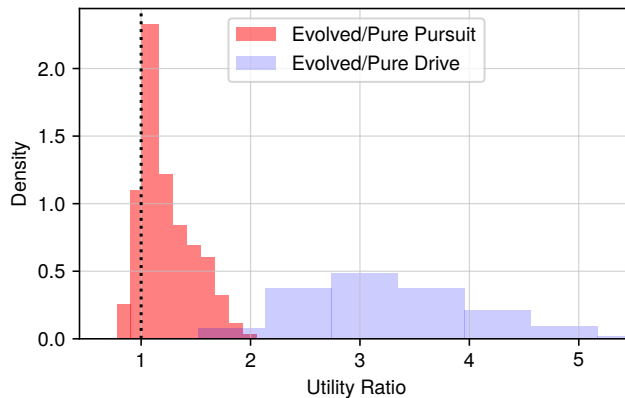
Figure B.6: Monte Carlo results for 1000 simulations - histograms of the util-
ity ratio of the best evolved controller compared to two baseline
controllers (always `pursue` and always `drive`)

.

the Monte Carlo experiment. Compared to the Pure Drive controller
(i.e. always use `drive`), the evolved controller typically captures three
times the number of prey. The utility gain over the Pure Pursuit con-
troller is more modest – the peak of the histogram occurs at 1, mean-
ing the two controllers have the same utility. Pure Pursuit outperforms
the evolved controller in less than 14% of the experiments, and the loss
is never more than 23%. The evolved controller performs 100% better
than Pure Pursuit for several cases. Pure Pursuit's biggest drawback
is that it's possible for the predators to end up quite close to one an-
other; thereafter the two predators make the same moves, always pur-
suing the same prey, which is clearly a waste. Although the evolved
controller only saw 9 initial conditions during evolution it was able to
generalize reasonably well, generally matching or outperforming the
baseline controllers.

B.7   CONCLUSION

A predator-prey scenario was considered in which the prey's behav-
ior is governed by a Boids-like flocking model and the predators co-
operate to maximize the number of captures within the time horizon.

One of the purposes of the paper was to demonstrate the efficacy of controllers based on a small set of pre-specified behaviors for a many-on-many type of adversarial engagement. Moreover, an evolutionary algorithm was employed to determine the appropriate base behavior to implement as a function of a meta state space. The meta state space used here was independent of the number of prey and thus the predator controller is scalable. In essence, the meta state space is just a feature space; it is possible that other features or functions of features could be useful in determining an appropriate behavior for this particular scenario. The notion of weighting was introduced onto the existing anchor point control architecture in order to have finer control over the partitioning of the meta state space. Based on the simulation results, the optimal (or approximately optimal) behavior for the two predator case is to aim at the nearest prey while avoiding getting too close to the other predator. The evolved controller was shown to have this type of behavior, generally, and thus it performed much better than always aiming at the nearest prey, in most cases.

Analytical solutions to many-on-many pursuit evasion problems like this one do not exist. However, approximately optimal solutions are often intuitive. The approach used in this study yielded an overall strategy that deconflicted the predators when necessary and otherwise employed a control that's optimal under several assumptions. It is possible, perhaps, to design a controller which plans a sequence of targets to pursue that is more optimal than always aiming at the nearest. However, that approach would be much more computationally complex; deconfliction would also require either explicit communication or heuristics to avoid wasting resources. Thus the advantages of this approach are that it is scalable, decentralized, and simple.

The goal was to show emergent cooperative predator behaviors along similar lines as those described in the referenced literature. The cooperation that was evolved in this study is intuitive, but the predators do not spend much time "herding" the prey or otherwise limiting prey dispersion. It is possible that, based on this particular definition of the utility and the settings of the prey model parameters, this type of behavior is unnecessary. Future research efforts may explore the simulation

parameter space, consider heterogeneous predators (i.e. each preda-
tor has its own anchor points, and, perhaps, role), specifying different
utility functionals (for example, requiring two predators to capture,
or explicitly including prey dispersion in the utility), and, of course,
studying larger scenarios with more predators.

[1] Behçet Açikmeşe and Scott R. Ploen. "Convex Programming Approach to Powered Descent Guidance for Mars Landing". en. In: *Journal of Guidance, Control, and Dynamics* 30 (5 Sept. 2007), pp. 1353–1366. DOI: 10.2514/1.27553.

[2] Z. Akilan and Z. Fuchs. "Zero-sum turret defense differential game with singular surfaces". In: *2017 IEEE Conference on Control Technology and Applications (CCTA)*. 2017 IEEE Conference on Control Technology and Applications (CCTA). Aug. 1, 2017, pp. 2041–2048. DOI: 10.1109/CCTA.2017.8062754.

[3] Eugene L Allgower and Kurt Georg. *Numerical continuation methods: an introduction*. Vol. 13. Springer Science & Business Media, 2012.

[4] C. Anderson, G. Theraulaz, and J.-L. Deneubourg. "Self-assemblages in insect societies". In: *Insectes Sociaux* 49 (2 May 2002), pp. 99–110. DOI: 10.1007/s00040-002-8286-y.

[5] Roger Anderson. "Defender-Assisted Evasion and Pursuit Maneuvers". MA thesis. Air Force Institute of Technology, Mar. 1, 2018.

[6] Ross P. Anderson and Dejan Milutinović. "On the Construction of Minimum-Time Tours for a Dubins Vehicle in the Presence of Uncertainties". In: *Journal of Dynamic Systems, Measurement, and Control* 137 (3 Mar. 2015). DOI: 10.1115/1.4028552.

[7] Pavlos Androulakakis. "Analysis of Evolutionary Algorithms in the Control of Path Planning Problems". PhD thesis. Wright State University, 2018.

[8] Pavlos Androulakakis and Zachariah E. Fuchs. "Evolutionary design of engagement strategies for turn-constrained agents". In: *2017 IEEE Congress on Evolutionary Computation (CEC)*. June 2017, pp. 2354–2363. DOI: 10.1109/CEC.2017.7969590.

[9] Pavlos Androulakakis, Zachariah E. Fuchs, and Jason E. Shroyer. "Evolutionary Design of an Open-Loop Turn Circle Intercept Controller". In: *2018 IEEE Congress on Evolutionary Computation (CEC)* (Rio de Janeiro). IEEE, July 2018. DOI: 10.1109/cec.2018.8477726.

[10] Timothy J. Arnett, Kelly Cohen, Brandon Cook, and David Casbeer. "A Supervisory Neural Network and Fuzzy Inference System Approach to an Unmanned Aerial Vehicle Tail-Chase Scenario". In: *AIAA Scitech 2019 Forum* (San Diego, California). American Institute of Aeronautics and Astronautics, Jan. 2019. DOI: 10.2514/6.2019-1455.

[11] F. Aurenhammer and H. Edelsbrunner. "An optimal algorithm for constructing the weighted voronoi diagram in the plane". In: *Pattern Recognition* 17 (2 Jan. 1, 1984), pp. 251–257. DOI: 10.1016/0031-3203(84)90064-5.

[12] Allan M. Avila, Maria Fonoberova, Joao P. Hespanha, Igor Mezic, Daniel Clymer, Jonathan Goldstein, Marco A. Pravia, and Daniel Javorsek. "Game Balancing using Koopman-based Learning". In: *2021 American Control Conference (ACC)* (New Orleans, LA, USA). IEEE, May 2021. DOI: 10.23919/acc50511.2021.9483027.

[13] David Avis, Gabriel D. Rosenberg, Rahul Savani, and Bernhard von Stengel. "Enumeration of Nash equilibria for two-player games". en. In: *Economic Theory* 42 (1 Jan. 1, 2010), pp. 9–37. DOI: 10.1007/s00199-009-0449-x.

[14] Mostafa D. Awheda and Howard M. Schwartz. "A Residual Gradient Fuzzy Reinforcement Learning Algorithm for Differential Games". en. In: *International Journal of Fuzzy Systems* 19 (4 Aug. 1, 2017), pp. 1058–1076. DOI: 10.1007/s40815-016-0284-8.

[15] Shivam Bajaj, Eric Torng, and Shaunak D. Bopardikar. "Competitive Perimeter Defense on a Line". In: *2021 American Control Conference (ACC)* (New Orleans, LA, USA). IEEE, May 2021. DOI: 10.23919/acc50511.2021.9483308.

[16] Efstathios Bakolas and Panagiotis Tsiotras. "Relay pursuit of a maneuvering target using dynamic Voronoi diagrams". In: *Automatica* 48 (9 Sept. 1, 2012), pp. 2213–2220. DOI: `10.1016/j.automatica.2012.06.003`.

[17] D. Barton. "Universal Equations for Radar Target Detection". en. In: *IEEE Transactions on Aerospace and Electronic Systems* 41 (3 July 2005), pp. 1049–1052. DOI: `10.1109/taes.2005.1541448`.

[18] J. C. Barton and C. J. Eliezer. "On pursuit curves". In: *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics* 41 (3 Jan. 2000), pp. 358–371. DOI: `10.1017/S0334270000011292`.

[19] Tamer Başar and Geert Jan Olsder. "Chapter 8: Pursuit-Evasion Games". In: Dynamic Noncooperative Game Theory 160 (Jan. 1, 1982), pp. 344–398.

[20] Tamer Başar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory*. 2nd. Vol. 160. Mathematics in Science and Engineering. Elsevier, Jan. 1, 1982.

[21] Richard Bellman. *Dynamic programming*. Princeton, N.J: Univ. Pr, 1972. ISBN: 0-691-07951-X.

[22] Richard Bellman. *Eye Of The Hurricane*. World Scientific, 1984.

[23] Carl T. Bergstrom and Michael Lachmann. "Alarm calls as costly signals of antipredator vigilance: the watchful babbler game". en. In: *Animal Behaviour* 61 (3 Mar. 2001), pp. 535–543. DOI: `10.1006/anbe.2000.1636`.

[24] Pierre Bernhard. "Linear pursuit-evasion games and the isotropic rocket". PhD thesis. 1971.

[25] Pierre Bernhard. "Singular surfaces in differential games an introduction". en. In: Lecture Notes in Control and Information Sciences (Jan. 1, 1977), pp. 1–33.

[26] Pierre Bernhard. "Isaacs, Breakwell, and Their Sons". In: 8th ISDG International Symposium on the Theory and Applications of Differential Games. The Netherlands, Jan. 1, 1998.

[27] Pierre Bernhard. "Pursuit-evasion games and zero-sum two-person differential games". In: *Encyclopaedia of Systems and Control*. Ed. by J. Bailleul and T. Samad. Springer, Jan. 1, 2014, pp. 1103–1109.

[28] Josh Bertram and Peng Wei. "An Efficient Algorithm for Multiple-Pursuer-Multiple-Evader Pursuit/Evasion Game". In: *AIAA Scitech 2021 Forum* (VIRTUAL EVENT). American Institute of Aeronautics and Astronautics, Jan. 2021. DOI: 10.2514/6.2021-1862.

[29] Hossein Bolandi, Mohammad Rezaei, Reza Mohsenipour, Hossein Nemati, and Seed Majid Smailzadeh. "Attitude control of a quadrotor with optimized pid controller". In: (Jan. 1, 2013).

[30] G. Boole. *A Treatise on Differential Equations*. A Treatise on Differential Equations 1. Macmillan, 1872.

[31] Carl de Boor and Blair Swartz. "Collocation at Gaussian Points". en. In: *SIAM Journal on Numerical Analysis* 10 (4 Sept. 1973), pp. 582–606. DOI: 10.1137/0710052.

[32] Shaunak D. Bopardikar, Francesco Bullo, and João P. Hespanha. "A cooperative Homicidal Chauffeur game". In: *Automatica* 45 (7 July 1, 2009), pp. 1771–1777. DOI: 10.1016/j.automatica.2009.03.014.

[33] P. Bouguer. "Lignes de Poursuite". In: *Mémoires de l'Académie Royale des Sciences* (1732).

[34] John P Boyd. *Chebyshev and Fourier spectral methods*. Dover, 2000.

[35] Valentino Braitenberg. *Vehicles: Experiments in synthetic psychology*. MIT press, 1986.

[36] J. V. Breakwell and P. Bernhard. "A simple game with a singular focal line". en. In: *Journal of Optimization Theory and Applications* 64 (2 Feb. 1990), pp. 419–428. DOI: 10.1007/bf00939457.

[37] J. V. Breakwell and P. Hagedorn. "Point capture of two evaders in succession". en. In: *Journal of Optimization Theory and Applications* 27 (1 Jan. 1, 1979), pp. 89–97. DOI: 10.1007/BF00933327.

[38] M. H. Breitner. "The Genesis of Differential Games in Light of Isaacs' Contributions". en. In: *Journal of Optimization Theory and Applications* 124 (3 Mar. 2005), pp. 523–559. DOI: 10.1007/s10957-004-1173-0.

[39] Kevin Q. Brown. "Geometric Transforms for Fast Geometric Algorithms". In: (Dec. 1, 1979).

[40] Arthur Earl Bryson and Yu-Chi Ho. *Applied Optimal Control: Optimization, Estimation and Control*. CRC Press, 1975. DOI: 10.1201/9781315137667.

[41] Kristoffer Carlsson. *NearestNeighbors.jl*. Accessed 2020-01-14. 2020.

[42] Elliot Cartee, Lexiao Lai, Qianli Song, and Alexander Vladimirsky. "Time-Dependent Surveillance-Evasion Games". In: *2019 IEEE 58th Conference on Decision and Control (CDC)* (Nice, France). IEEE, Dec. 2019. DOI: 10.1109/cdc40024.2019.9029329.

[43] Marco Casini and Andrea Garulli. "On the advantage of centralized strategies in the three-pursuer single-evader game". en. In: *Systems & Control Letters* 160 (Feb. 2022), p. 105122. DOI: 10.1016/j.sysconle.2021.105122.

[44] M. Chen, Z. Zhou, and C. J. Tomlin. "A path defense approach to the multiplayer reach-avoid game". In: *53rd IEEE Conference on Decision and Control*. 53rd IEEE Conference on Decision and Control. Dec. 1, 2014, pp. 2420–2426. DOI: 10.1109/CDC.2014.7039758.

[45] M. Chen, Z. Zhou, and C. J. Tomlin. "Multiplayer Reach-Avoid Games via Pairwise Outcomes". In: *IEEE Transactions on Automatic Control* 62 (3 Mar. 1, 2017), pp. 1451–1457. DOI: 10.1109/TAC.2016.2577619.

[46] Warren A Cheung. "Constrained pursuit-evasion problems in the plane". University of British Columbia, Jan. 1, 2005.

[47] Jonathan H Connell. *Minimalist mobile robotics*. Vol. 5. Elsevier, 2012. DOI: 10.1016/c2009-0-27873-1.

[48]   Bruce A. Conway and Mauro Pontani. "Optimal Interception of Evasive Missile Warheads: Numerical Solution of the Differential Game". en. In: *Journal of Guidance, Control, and Dynamics* 31 (4 July 2008), pp. 1111–1122. DOI: 10.2514/1.30893.

[49]   John B. Conway. *A Course in Functional Analysis*. en. Graduate Texts in Mathematics. New York: Springer-Verlag, Jan. 1, 1985.

[50]   Olivier Cots, Joseph Gergaud, and Damien Goubinat. "Direct and indirect methods in optimal control with state constraints and the climbing trajectory of an aircraft". In: *Optimal Control Applications and Methods* 39 (1 Jan. 2018), pp. 281–301. DOI: 10.1002/oca.2347.

[51]   D P Croft, R James, P O R Thomas, C Hathaway, D Mawdsley, K N Laland, and J Krause. "Social structure and co-operative interactions in a wild population of guppies (Poecilia reticulata)". English. In: *Behavioral Ecology and Sociobiology* 59.5 (Mar. 2006), pp. 644–650. ISSN: 0340-5443. DOI: 10.1007/s00265-005-0091-y.

[52]   J. B. Cruz. "Survey of Nash and Stackelberg Equilibrium Strategies in Dynamic Games". In: *Annals of Economic and Social Measurement* 4 (2 Jan. 1, 1975).

[53]   J.B. Cruz, M.A. Simaan, A. Gacic, None Huihui Jiang, B. Letellier, None Ming Li, and None Yong Liu. "Game-theoretic modeling and control of a military air operation". In: *IEEE Transactions on Aerospace and Electronic Systems* 37 (4 2001), pp. 1393–1405. DOI: 10.1109/7.976974.

[54]   Brigitte d'Andrea-Novel, Michel Fliess, Cedric Join, Hugues Mounier, and Bruno Steux. "A mathematical explanation via "intelligent" PID controllers of the strange ubiquity of PIDs". In: *Automation (MED 2010)* (Marrakech, Morocco). IEEE, June 2010. DOI: 10.1109/med.2010.5547700.

[55]   Christopher Darby and Anil Rao. "A State Approximation-Based Mesh Refinement Algorithm for Solving Optimal Control Problems Using Pseudospectral Methods". In: *AIAA Guidance, Nav-*

igation, and Control Conference (Chicago, Illinois). American Institute of Aeronautics and Astronautics, Aug. 2009. ISBN: 9781600869785. DOI: `10.2514/6.2009-5791`.

[56] Marc Deisenroth and Carl E. Rasmussen. "PILCO: A model-based and data-efficient approach to policy search". In: *Proceedings of the 28th International Conference on machine learning* (*ICML-11*). Jan. 1, 2011, pp. 465–472.

[57] Marc Peter Deisenroth, Carl Edward Rasmussen, and Jan Peters. "Gaussian process dynamic programming". In: *Neurocomputing*. Advances in Machine Learning and Computational Intelligence 72 (7 Mar. 1, 2009), pp. 1508–1524. DOI: `10.1016/j.neucom.2008.12.019`.

[58] Adam Dobrin. "A review of properties and variations of Voronoi diagrams". In: *Whitman College* (Jan. 1, 2005).

[59] Engelbert J. Dockner, Steffen Jorgensen, Ngo Van Long, and Gerhard Sorger. *Differential Games in Economics and Management Science by Engelbert J. Dockner*. en. Nov. 1, 2000.

[60] Michael Dorothy, Dipankar Maity, Daigo Shishika, and Alexander Von Moll. "One Apollonius Circle is Enough for Many Pursuit-Evasion Games". In: *Automatica* (Nov. 17, 2021). Submitted for Review. eprint: `2111.09205`.

[61] Michael R. Dorothy. "Neuroinspired control strategies with applications to flapping flight". PhD thesis. 2016.

[62] John Drake, Alla Safonova, and Maxim Likhachev. "Towards adaptability of demonstration-based training of npc behavior". In: *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*. 2017.

[63] Matthew G. Earl and Raffaello D'Andrea. "A decomposition approach to multi-vehicle cooperative control". In: *Robotics and Autonomous Systems* 55 (4 Apr. 30, 2007), pp. 276–291. DOI: `10.1016/j.robot.2006.11.002`.

[64]   C. J. Eliezer and J. C. Barton. "Pursuit Curves". In: *Bulletin of the Institute of Mathematics and its Applications* 28 (1992), pp. 182–184.

[65]   C. J. Eliezer and J. C. Barton. "Pursuit Curves II". In: *Bulletin of the Institute of Mathematics and its Applications* 31 (1995), pp. 139–141.

[66]   R. Emery-Montemerlo, G. Gordon, J. Schneider, and S. Thrun. "Approximate solutions for partially observable stochastic games with common payoffs". In: *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, 2004. AAMAS 2004.* 2004, pp. 136–143.

[67]   Nicholas Ernest, Kelly Cohen, Elad Kivelevitch, Corey Schumacher, and David Casbeer. "Genetic Fuzzy Trees and their Application Towards Autonomous Training and Control of a Squadron of Unmanned Combat Aerial Vehicles". In: *Unmanned Systems* 03 (03 May 13, 2015), pp. 185–204. DOI: 10.1142/S2301385015500120.

[68]   Ioannis Exarchos, Panagiotis Tsiotras, and Meir Pachter. "On the Suicidal Pedestrian Differential Game". en. In: *Dynamic Games and Applications* 5 (3 Sept. 2015), pp. 297–317. DOI: 10.1007/s13235-014-0130-2.

[69]   Ioannis Exarchos, Panagiotis Tsiotras, and Meir Pachter. "UAV Collision Avoidance based on the Solution of the Suicidal Pedestrian Differential Game". In: *AIAA Guidance, Navigation, and Control Conference* (San Diego, California, USA). American Institute of Aeronautics and Astronautics, Jan. 2016. DOI: 10.2514/6.2016-2100.

[70]   Fariba Fahroo and I. Ross. "Pseudospectral Methods for Infinite-Horizon Nonlinear Optimal Control Problems". In: *AIAA Guidance, Navigation, and Control Conference and Exhibit* (San Francisco, California). American Institute of Aeronautics and Astronautics, Aug. 2005. ISBN: 9781624100567. DOI: 10.2514/6.2005-6076.

[71] Mariam Faied and Anouck Girard. "Game formulation of multiteam target assignment and suppression mission". In: *IEEE Transactions on Aerospace and Electronic Systems* 50 (2 Apr. 2014), pp. 1234–1248. DOI: `10.1109/taes.2014.110676`.

[72] M. Falcone. "Numerical methods for differential games based on partial differential equations". In: *International Game Theory Review* 08 (02 June 1, 2006), pp. 231–272. DOI: `10.1142/S0219198906000886`.

[73] C. Fershtman. "Identification of classes of differential games for which the open loop is a degenerate feedback Nash equilibrium". en. In: *Journal of Optimization Theory and Applications* 55 (2 Nov. 1, 1987), pp. 217–231. DOI: `10.1007/BF00939082`.

[74] A. Festa and R. B. Vinter. "A decomposition technique for pursuit evasion games with many pursuers". In: *52nd IEEE Conference on Decision and Control*. 52nd IEEE Conference on Decision and Control. Dec. 1, 2013, pp. 5797–5802. DOI: `10.1109/CDC.2013.6760803`.

[75] Adriano Festa and Richard B. Vinter. "Decomposition of Differential Games with Multiple Targets". en. In: *Journal of Optimization Theory and Applications* 169 (3 June 1, 2016), pp. 848–875. DOI: `10.1007/s10957-016-0908-z`.

[76] Aleksej Fedorovič Filippov. *Differential Equations with Discontinuous Righthand Sides: Control Systems*. Vol. 18. Springer Science & Business Media, Jan. 1, 2013.

[77] J. F. Fisac and S. S. Sastry. "The pursuit-evasion-defense differential game in dynamic constrained environments". In: *2015 54th IEEE Conference on Decision and Control (CDC)*. 2015 54th IEEE Conference on Decision and Control (CDC). Dec. 1, 2015, pp. 4549–4556. DOI: `10.1109/CDC.2015.7402930`.

[78] D. Floreano and F. Mondada. "Automatic Creation of an Autonomous Agent: Genetic Evolution of a Neural Network Driven Robot". en. In: Third International Conference on Simulation

of Adaptive Behavior (SAB'94). The MIT Press, Jan. 1, 1994, pp. 421–430. ISBN: 9780262531221. DOI: 10.3929/ethz-a-010111549.

[79]   Steven Fortune. "A sweepline algorithm for Voronoi diagrams". en. In: *Algorithmica* 2 (1-4 Nov. 1, 1987), p. 153. DOI: 10.1007/BF01840357.

[80]   Han Fu and Hugh H. Liu. "On the Strategies of Defending a Target against Multiple Intruders". In: *AIAA Scitech 2021 Forum* (VIRTUAL EVENT). American Institute of Aeronautics and Astronautics, Jan. 2021. DOI: 10.2514/6.2021-1861.

[81]   Han Fu and Hugh H.-T. Liu. "Guarding a Territory Against an Intelligent Intruder: Strategy Design and Experimental Verification". In: *IEEE/ASME Transactions on Mechatronics* 25 (4 Aug. 2020), pp. 1765–1772. DOI: 10.1109/tmech.2020.2996901.

[82]   Maozhong Fu, Yixiong Zhang, Risheng Wu, Zhenmiao Deng, Yunjian Zhang, and Xiangyu Xiong. "Fast Range and Motion Parameters Estimation for Maneuvering Targets Using Time-Reversal Process". In: *IEEE Transactions on Aerospace and Electronic Systems* 55 (6 Dec. 2019), pp. 3190–3206. DOI: 10.1109/taes.2019.2901586.

[83]   Z. E. Fuchs and P. P. Khargonekar. "Encouraging attacker retreat through defender cooperation". In: *2011 50th IEEE Conference on Decision and Control and European Control Conference*. 2011 50th IEEE Conference on Decision and Control and European Control Conference. Dec. 1, 2011, pp. 235–242. DOI: 10.1109/CDC.2011.6161096.

[84]   Z. E. Fuchs and P. P. Khargonekar. "Generalized Engage or Retreat Differential Game With Escort Regions". In: *IEEE Transactions on Automatic Control* 62 (2 Feb. 1, 2017), pp. 668–681. DOI: 10.1109/TAC.2016.2562921.

[85]   Z. E. Fuchs, P. P. Khargonekar, and J. Evers. "Cooperative defense within a single-pursuer, two-evader pursuit evasion differential game". In: *49th IEEE Conference on Decision and Control*

(*CDC*). 49th IEEE Conference on Decision and Control (CDC). Dec. 1, 2010, pp. 3091–3097. DOI: 10.1109/CDC.2010.5717894.

[86] Z. E. Fuchs and J. Metcalf. "Equilibrium radar-target interactions in an ATR scenario: A differential game". In: *2018 IEEE Radar Conference (RadarConf18)*. 2018 IEEE Radar Conference (RadarConf18). Apr. 1, 2018, pp. 1228–1233. DOI: 10.1109/RADAR.2018.8378738.

[87] Zachariah E Fuchs. "Cooperative Control Strategies and Deception in Adversiarial Systems". en. University of Florida, Jan. 1, 2012.

[88] Zachariah E. Fuchs, David W. Casbeer, and Eloy Garcia. "Singular analysis of a multi-agent, turn-constrained, defensive game". In: *2016 American Control Conference (ACC)* (Boston, MA, USA). IEEE, July 2016. DOI: 10.1109/acc.2016.7526097.

[89] Zachariah E. Fuchs, Eloy Garcia, and David W. Casbeer. "Two-Pursuer, One-Evader Pursuit Evasion Differential Game". In: *2018 IEEE National Aerospace and Electronics Conference (NAECON)*. 2018 IEEE National Aerospace and Electronics Conference (NAECON). Dayton, OH: IEEE, July 23, 2018, pp. 456–464. DOI: 10.1109/NAECON.2018.8556827.

[90] Zachariah E. Fuchs and Pramod P. Khargonekar. "An engage or retreat differential game with an escort region". In: *2014 IEEE 53rd Annual Conference on Decision and Control (CDC)* (Los Angeles, CA, USA). IEEE, Dec. 2014. ISBN: 9781467360906. DOI: 10.1109/cdc.2014.7040058.

[91] Zachariah E. Fuchs, Alexander Von Moll, and David Casbeer. "Engage or Retreat Differential Game with N-Targets and Distributed Defensive Assets". In: *Conference on Controls Technology and Applications*. Conference on Controls Technology and Applications. San Diego, CA, Aug. 31, 2021. DOI: 10.1109/CCTA48906.2021.9658922.

[92] Runze Gan, Bashar I. Ahmad, and Simon J. Godsill. "Lévy State-Space Models for Tracking and Intent Prediction of Highly Ma-

neuverable Objects". In: *IEEE Transactions on Aerospace and Electronic Systems* 57 (4 Aug. 2021), pp. 2021–2038. DOI: 10.1109/taes.2021.3088430.

[93] Manan Gandhi. "Trajectory Optimization Algorithm Studies". In: *arXiv preprint arXiv:1506.00731* (Jan. 1, 2015).

[94] Manan Gandhi and Evangelos Theodorou. "A Comparison between Trajectory Optimization Methods: Differential Dynamic Programming and Pseudospectral Optimal Control". In: *AIAA Guidance, Navigation, and Control Conference*. Jan. 1, 2016, p. 0385.

[95] Sergey A. Ganebny, Sergey S. Kumkov, Stéphane Le Ménec, and Valerii S. Patsko. "Model Problem in a Line with Two Pursuers and One Evader". en. In: *Dynamic Games and Applications* 2 (2 June 1, 2012), pp. 228–257. DOI: 10.1007/s13235-012-0041-z.

[96] E. Garcia, D. W. Casbeer, and M. Pachter. "Design and Analysis of State-Feedback Optimal Strategies for the Differential Game of Active Defense". In: *IEEE Transactions on Automatic Control* (Jan. 1, 2018), pp. 1–1. DOI: 10.1109/TAC.2018.2828088.

[97] E. Garcia, D. W. Casbeer, K. Pham, and M. Pachter. "Cooperative aircraft defense from an attacking missile". In: *53rd IEEE Conference on Decision and Control*. 53rd IEEE Conference on Decision and Control. Dec. 1, 2014, pp. 2926–2931. DOI: 10.1109/CDC.2014.7039839.

[98] Eloy Garcia. "Cooperative Target protection from a superior Attacker". en. In: *Automatica* 131 (Sept. 2021), p. 109696. DOI: 10.1016/j.automatica.2021.109696.

[99] Eloy Garcia, David Casbeer, Khanh D. Pham, and Meir Pachter. "Cooperative Aircraft Defense from an Attacking Missile using Proportional Navigation". In: *AIAA Guidance, Navigation, and Control Conference*. AIAA Guidance, Navigation, and Control Conference. AIAA SciTech Forum. American Institute of Aeronautics and Astronautics, Jan. 2, 2015.

[100] Eloy Garcia, David W Casbeer, Alexander Von Moll, and Meir Pachter. "Multiple Pursuer Multiple Evader Differential Games". In: *Transactions on Automatic Control* (May 1, 2020). DOI: 10 . 1109/TAC.2020.3003840.

[101] Eloy Garcia, David W. Casbeer, Zachariah E. Fuchs, and Meir Pachter. "Cooperative Missile Guidance for Active Defense of Air Vehicles". In: *IEEE Transactions on Aerospace and Electronic Systems* 54 (2 Apr. 2018), pp. 706–721. DOI: 10 . 1109 / taes . 2017.2764269.

[102] Eloy Garcia, David W. Casbeer, and Meir Pachter. "The Capture-the-Flag Differential Game". In: *2018 IEEE Conference on Decision and Control (CDC)* (Miami Beach, FL). IEEE, Dec. 2018. ISBN: 9781538613955. DOI: 10.1109/cdc.2018.8619026.

[103] Eloy Garcia, David W. Casbeer, and Meir Pachter. "Optimal Strategies of the Differential Game in a Circular Region". In: *IEEE Control Systems Letters* 4 (2 Apr. 2020), pp. 492–497. DOI: 10.1109/lcsys.2019.2963173.

[104] Eloy Garcia, David W. Casbeer, and Meir Pachter. "The Complete Differential Game of Active Target Defense". en. In: *Journal of Optimization Theory and Applications* (Feb. 2021). DOI: 10 . 1007/s10957-021-01816-z.

[105] Eloy Garcia, David W. Casbeer, Dzung Tran, and Meir Pachter. "A Differential Game Approach for Beyond Visual Range Tactics". In: *2021 American Control Conference (ACC)* (New Orleans, LA, USA). IEEE, May 2021. DOI: 10 . 23919 / acc50511 . 2021 . 9482650.

[106] Eloy Garcia, David W. Casbeer, Alexander Von Moll, and Meir Pachter. "Cooperative Two-Pursuer One-Evader Blocking Differential Game". In: American Control Conference. July 10, 2019. DOI: 10.23919/ACC.2019.8814294.

[107] Eloy Garcia, Zachariah E. Fuchs, Dejan Milutinović, David W. Casbeer, and Meir Pachter. "A Geometric Approach for the Cooperative Two-Pursuer One-Evader Differential Game". In:

IFAC-PapersOnLine. 20th IFAC World Congress 50 (1 July 1, 2017), pp. 15209–15214. DOI: 10.1016/j.ifacol.2017.08.2366.

[108]   Eloy Garcia, Dzung M. Tran, David Casbeer, Dejan Milutinović, and Meir Pachter. "Beyond Visual Range Tactics". In: *AIAA Scitech 2021 Forum* (VIRTUAL EVENT). American Institute of Aeronautics and Astronautics, Jan. 2021. DOI: 10.2514/6.2021-1229.

[109]   Eloy Garcia, Alexander Von Moll, David Casbeer, and Meir Pachter. "Strategies for Defending a Coastline Against Multiple Attackers". In: *IEEE Conference on Decision and Control*. IEEE Conference on Decision and Control. Nice, France, Dec. 31, 2019. DOI: 10.1109/CDC40024.2019.9029340.

[110]   Andrew Gard. "The Wild Goose Chase Problem". en. In: *The American Mathematical Monthly* 125 (7 Aug. 2018), pp. 602–611. DOI: 10.1080/00029890.2018.1465785.

[111]   Stefanie K. Gazda, Richard C. Connor, Robert K. Edgar, and Frank Cox. "A division of labour with role specialization in group–hunting bottlenose dolphins (Tursiops truncatus) off Cedar Key, Florida". en. In: *Proceedings of the Royal Society B: Biological Sciences* 272 (1559 Jan. 2005), pp. 135–140. DOI: 10.1098/rspb.2004.2937.

[112]   W. M. Getz and M. Pachter. "Capturability in a two-target 'game of two cars'". en. In: *Journal of Guidance and Control* 4 (1 Jan. 1981), pp. 15–21. DOI: 10.2514/3.19715.

[113]   Debasish Ghose. "An Introduction to Proportional Navigation". In: (Mar. 1, 2012).

[114]   Debasish Ghose. "Missile Guidance Laws". In: (Mar. 1, 2012).

[115]   Robert H. Goddard. *A Method of Reaching Extreme Altitudes*. Vol. 71. Princeton, New Jersey: Smithsonian Institution, 1919.

[116]   Beke Graw and Marta B. Manser. "The function of mobbing in cooperative meerkats". In: *Animal Behaviour* 74 (3 Sept. 2007), pp. 507–517. DOI: 10.1016/j.anbehav.2006.11.021.

[117] Yue Guan, Dipankar Maity, Christopher M. Kroninger, and Panagiotis Tsiotras. "Bounded-Rational Pursuit-Evasion Games". In: *2021 American Control Conference (ACC)* (New Orleans, LA, USA). IEEE, May 2021. DOI: 10.23919/acc50511.2021.9483152.

[118] Otomar Hájek. "Toward a general theory of pursuit and evasion". In: Springer-Verlag, 1977, pp. 143–152. ISBN: 354008407X. DOI: 10.1007/bfb0009068.

[119] David Hambling. "AI outguns a human fighter pilot". en. In: *New Scientist* 247 (3297 Aug. 2020), p. 12. DOI: 10.1016/s0262-4079(20)31477-9.

[120] Nicholas Hanlon, Eloy Garcia, David Casbeer, and Meir Pachter. "AFSIM Implementation and Simulation of the Active Target Defense Differential Game". In: *2018 AIAA Guidance, Navigation, and Control Conference* (Kissimmee, Florida). American Institute of Aeronautics and Astronautics, Jan. 2018. DOI: 10.2514/6.2018-1582.

[121] A. Hashemi, D. W. Casbeer, and D. Milutinović. "Scalable value approximation for multiple target tail-chase with collision avoidance". In: *2016 IEEE 55th Conference on Decision and Control (CDC)*. 2016 IEEE 55th Conference on Decision and Control (CDC). Dec. 1, 2016, pp. 2543–2548. DOI: 10.1109/CDC.2016.7798645.

[122] Thomas Haynes and Sandip Sen. "Evolving behavioral strategies in predators and prey". In: *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 1996, pp. 113–126. DOI: 10.1007/3-540-60923-7_22.

[123] Douglas C. Heard. "The Effect of Wolf Predation and Snow Cover on Musk-Ox Group Size". en. In: *The American Naturalist* 139 (1 Jan. 1992), pp. 190–204. DOI: 10.1086/285320.

[124] João Hespanha. "Linear Systems Theory". In: (Dec. 2018). DOI: 10.23943/9781400890088.

[125] Y. Ho, A. Bryson, and S. Baron. "Differential games and optimal pursuit-evasion strategies". In: *IEEE Transactions on Auto-*

matic Control 10 (4 Oct. 1, 1965), pp. 385–389. DOI: 10.1109/TAC.1965.1098197.

[126]   Kazuhiro Horie and Bruce A. Conway. "Optimal Fighter Pursuit-Evasion Maneuvers Found Via Two-Sided Optimization". en. In: *Journal of Guidance, Control, and Dynamics* 29 (1 Jan. 2006), pp. 105–112. DOI: 10.2514/1.3960.

[127]   Taylor A. Howell, Brian E. Jackson, and Zachary Manchester. "ALTRO: A Fast Solver for Constrained Trajectory Optimization". In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Macau, China). IEEE, Nov. 2019. ISBN: 9781728140049. DOI: 10.1109/iros40897.2019.8967788.

[128]   Haomiao Huang, Wei Zhang, Jerry Ding, Dusan M Stipanović, and Claire J Tomlin. "Guaranteed Decentralized Pursuit-Evasion in the Plane with Multiple Pursuers". In: *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*. IEEE, Jan. 1, 2011, pp. 4835–4840.

[129]   Rufus Isaacs. "Games of Pursuit". In: *RAND report* (Jan. 1, 1951).

[130]   Rufus Isaacs. *Differential Games: A Mathematical Theory with Applications to Optimization, Control and Warfare*. Wiley, New York, Jan. 1, 1965. ISBN: 9780486406824.

[131]   Rufus Isaacs. "Differential games: Their scope, nature, and future". In: *Journal of Optimization Theory and Applications* 3 (5 May 1, 1969), pp. 283–295. DOI: 10.1007/BF00931368.

[132]   Steven G. Johnson. *The NLopt nonlinear-optimazation package*. Accessed 2019-8-01. 2019.

[133]   Atsushi Kamimura and Toru Ohira. *Group Chase and Escape*. Springer Singapore, 2019. ISBN: 9789811517303. DOI: 10.1007/978-981-15-1731-0_2.

[134]   Matthew Kelly. "An Introduction to Trajectory Optimization: How to Do Your Own Direct Collocation". en. In: *SIAM Review* 59 (4 Jan. 2017), pp. 849–904. DOI: 10.1137/16m1062569.

[135] Pramod P. Khargonekar and Munther A. Dahleh. "Advancing systems and control research in the era of ML and AI". In: *Annual Reviews in Control* 45 (Jan. 1, 2018), pp. 1–4. DOI: `10.1016/j.arcontrol.2018.04.001`.

[136] Donald Kirk. *Optimal control theory; an introduction*. Englewood Cliffs, N.J: Prentice-Hall, 1970. ISBN: 0486434842.

[137] J. Ko, D. J. Klein, D. Fox, and D. Haehnel. "Gaussian Processes and Reinforcement Learning for Identification and Control of an Autonomous Blimp". In: *Proceedings 2007 IEEE International Conference on Robotics and Automation*. Proceedings 2007 IEEE International Conference on Robotics and Automation. Apr. 1, 2007, pp. 742–747. DOI: `10.1109/ROBOT.2007.363075`.

[138] Richard E Korf. "A simple solution to pursuit games". In: *Working Papers of The 11th International Workshop on DAI*. 1992, pp. 183–194.

[139] Sergey S. Kumkov, Stéphane Le Ménec, and Valerii S. Patsko. "Zero-Sum Pursuit-Evasion Differential Games with Many Objects: Survey of Publications". en. In: *Dynamic Games and Applications* 7 (4 Dec. 1, 2017), pp. 609–633. DOI: `10.1007/s13235-016-0209-z`.

[140] Hans Petter Langtangen and Geir K. Pedersen. *Scaling of Differential Equations*. en. Oslo: Springer-Verlag, June 20, 2016.

[141] Daniel T. Larsson, Georgios Kotsalis, and Panagiotis Tsiotras. "Nash and Correlated Equilibria for Pursuit-Evasion Games Under Lack of Common Knowledge". In: *2018 IEEE Conference on Decision and Control (CDC)*. Dec. 2018, pp. 3579–3584. DOI: `10.1109/CDC.2018.8619168`.

[142] Yoonjae Lee and Efstathios Bakolas. "Optimal Strategies for Guarding a Compact and Convex Target Area: A Differential Game Approach". In: (Apr. 2021). eprint: `2104.00717v1`.

[143] Dario Leone. *That time North Korea fired SA-2 SAMs at SR-71 Blackbird*. NEWS article. Apr. 30, 2018.

[144] J. Lewin. *Differential Games: Theory and Methods for Solving Game Problems with Singular Surfaces*. Springer-Verlag London Limited, Jan. 1, 1994.

[145] J. Lewin and J. V. Breakwell. "The surveillance-evasion game of degree". en. In: *Journal of Optimization Theory and Applications* 16 (3-4 Aug. 1975), pp. 339–353. DOI: 10.1007/bf01262940.

[146] J. Lewin and G. J. Olsder. "Conic surveillance evasion". en. In: *Journal of Optimization Theory and Applications* 27 (1 Jan. 1979), pp. 107–125. DOI: 10.1007/bf00933329.

[147] D. Li and J. B. Cruz. "Defending an Asset: A Linear Quadratic Game Approach". In: *IEEE Transactions on Aerospace and Electronic Systems* 47 (2 Apr. 1, 2011), pp. 1026–1044. DOI: 10.1109/TAES.2011.5751240.

[148] Dongxu Li, J. B. Cruz, Genshe Chen, Chiman Kwan, and Mou-Hsiung Chang. "A Hierarchical Approach To Multi-Player Pursuit-Evasion Differential Games". In: *Proceedings of the 44th IEEE Conference on Decision and Control*. Proceedings of the 44th IEEE Conference on Decision and Control. Dec. 1, 2005, pp. 5674–5679. DOI: 10.1109/CDC.2005.1583067.

[149] Ch Lin. "Intelligent control theory in guidance and control system design: An overview". In: *Proceedings of the National Science Council, ROC, 2000* 24 (2000), pp. 15–30.

[150] S. Y. Liu, Z. Zhou, C. Tomlin, and K. Hedrick. "Evasion as a team against a faster pursuer". In: *2013 American Control Conference*. 2013 American Control Conference. June 1, 2013, pp. 5368–5373. DOI: 10.1109/ACC.2013.6580676.

[151] Ping Lu. "What Is Guidance?" en. In: *Journal of Guidance, Control, and Dynamics* 44 (7 July 2021), pp. 1237–1238. DOI: 10.2514/1.g006191.

[152] Venkata Ramana Makkapati, Wei Sun, and Panagiotis Tsiotras. "Optimal Evading Strategies for Two-Pursuer/One-Evader Problems". In: *Journal of Guidance, Control, and Dynamics* 41 (4 Jan. 1, 2018), pp. 851–862. DOI: 10.2514/1.G003070.

[153]   Venkata Ramana Makkapati and Panagiotis Tsiotras. "Optimal Evading Strategies and Task Allocation in Multi-player Pursuit–Evasion Problems". en. In: *Dynamic Games and Applications* (June 30, 2019). DOI: 10.1007/s13235-019-00319-x.

[154]   G. Manyam Satyanarayana, David Casbeer, Alexander Von Moll, and Zachariah Fuchs. "Optimal Dubins Paths to Intercept a Moving Target on a Circle". In: American Control Conference. July 10, 2019. DOI: 10.23919/ACC.2019.8814913.

[155]   Satyanarayana G. Manyam, David Casbeer, Alexander Von Moll, and Eloy Garcia. "Coordinating Defender Path Planning for Optimal Target-Attacker-Defender Game". In: *AIAA Infotech*. AIAA Infotech. San Diego, CA: AIAA, Jan. 11, 2019. DOI: 10.2514/6.2019-0388.

[156]   Satyanarayana Gupta Manyam, David Casbeer, Alexander Von Moll, and Zachariah Fuchs. "Shortest Dubins Path to a Circle". In: *AIAA SciTech*. AIAA SciTech. San Diego, CA: AIAA, Jan. 7, 2019. DOI: 10.2514/6.2019-0919.

[157]   Kostas Margellos and John Lygeros. "Hamilton-Jacobi Formulation for Reach–Avoid Differential Games". In: *IEEE Transactions on Automatic Control* 56 (8 Aug. 2011), pp. 1849–1861. DOI: 10.1109/tac.2011.2105730.

[158]   T.G. McGee and J.K. Hedrick. "Guaranteed strategies to search for mobile evaders in the plane". In: *2006 American Control Conference* (Minneapolis, MN, USA). IEEE, June 2006. ISBN: 1424402093. DOI: 10.1109/acc.2006.1656651.

[159]   Arik Melikyan and Pierre Bernhard. "Geometry of Optimal Paths around Focal Singular Surfaces in Differential Games". In: *Applied Mathematics and Optimization* 52.1 (June 2005), pp. 23–37. ISSN: 1432-0606. DOI: 10.1007/s00245-004-0816-8.

[160]   Arik A. Melikyan. "The method of characteristics for constructing singular paths and manifolds in optimal control and differential games". In: Springer-Verlag, 1991, pp. 81–90. ISBN: 3540537872. DOI: 10.1007/bfb0040229.

[161] Arik A. Melikyan. "Singular Paths in Differential Games with Simple Motion". en. In: Annals of the International Society of Dynamic Games (Jan. 1, 1994), pp. 125–135.

[162] A. W. Merz. "The game of two identical cars". en. In: *Journal of Optimization Theory and Applications* 9 (5 May 1972), pp. 324–343. DOI: 10.1007/bf00932932.

[163] A. W. Merz. "To pursue or to evade -That is the question". en. In: *Journal of Guidance, Control, and Dynamics* 8 (2 Mar. 1985), pp. 161–166. DOI: 10.2514/3.19954.

[164] Antony W. Merz. "The Homicidal Chauffeur -A Differential Game". PhD thesis. Stanford, Jan. 1, 1971.

[165] Dejan Milutinović, David W. Casbeer, Alexander Von Moll, Meir Pachter, and Eloy Garcia. "Rate of Loss Characterization that Resolves the Dilemma of the Wall Pursuit Game Solution". In: *Transactions on Automatic Control* (Dec. 27, 2021). DOI: 10.1109/TAC.2021.3137786.

[166] I.M. Mitchell, A.M. Bayen, and C.J. Tomlin. "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games". In: *IEEE Transactions on Automatic Control* 50 (7 July 2005), pp. 947–957. DOI: 10.1109/tac.2005.851439.

[167] Dov Monderer and Lloyd S. Shapley. "Potential Games". en. In: *Games and Economic Behavior* 14 (1 May 1996), pp. 124–143. DOI: 10.1006/game.1996.0044.

[168] Alexey A. Munishkin, Dejan Milutinović, and David W. Casbeer. "Min-max time efficient inspection of ground vehicles by a UAV team". en. In: *Robotics and Autonomous Systems* 125 (Mar. 2020), p. 103370. DOI: 10.1016/j.robot.2019.103370.

[169] Randall Munroe. *Substitute*. July 31, 2006. URL: https://xkcd.com/135/.

[170] Keiko Nagami and Mac Schwager. "HJB-RL: Initializing Reinforcement Learning with Optimal Control Policies Applied to Autonomous Drone Racing". In: *Robotics: Science and Systems*

2021. Robotics: Science and Systems Foundation, July 2021. DOI: 10.15607/rss.2021.xvii.062.

[171]   Paul J. Nahin. *Chases and Escapes*. Princeton University Press, Dec. 2012. ISBN: 9781400842063. DOI: 10.1515/9781400842063.

[172]   John Nash. "Non-Cooperative Games". In: *The Annals of Mathematics* 54 (2 Sept. 1951), p. 286. DOI: 10.2307/1969529.

[173]   Nhan Nguyen and Stephen A. Jacklin. "Neural net adaptive flight control stability, verification and validation challenges, and future research". In: *Workshop on Applications of neural networks in high assurance systems-International Joint Conference on Neural networks, Orlando, Florida*. Jan. 1, 2007, pp. 12–17.

[174]   A. O'Dwyer. "PI and PID controller tuning rules: an overview and personal perspective". In: *IET Irish Signals and Systems Conference (ISSC 2006)* (Dublin, Ireland). IEE, 2006. DOI: 10.1049/cp:20060431.

[175]   Ernest J. Ohlmeyer and Craig A. Phillips. "Generalized Vector Explicit Guidance". en. In: *Journal of Guidance, Control, and Dynamics* 29 (2 Mar. 2006), pp. 261–268. DOI: 10.2514/1.14956.

[176]   Robert Olendorf, Thomas Getty, and Kim Scribner. "Cooperative nest defence in red–winged blackbirds: reciprocal altruism, kinship or by–product mutualism?" In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271 (1535 Jan. 2004), pp. 177–182. DOI: 10.1098/rspb.2003.2586.

[177]   G. J. Olsder and J. V. Breakwell. "Role determination in an aerial dogfight". en. In: *International Journal of Game Theory* 3 (1 Mar. 1974), pp. 47–66. DOI: 10.1007/bf01766218.

[178]   Geert Jan Olsder. "Information Structures in Differential Games". In: *Differential Games and Control Theory II*. Ed. by Emilio O. Roxin, Pan-Tai Liu, and Robert L. Sternberg. Marcel Dekker, 1977, pp. 99–135.

[179]   Geert Jan Olsder. "On observation costs and information structures in stochastic differential games". In: Springer-Verlag, 1977, pp. 172–185. ISBN: 354008407X. DOI: 10.1007/bfb0009070.

[180]   Martin J Osborne. *An introduction to game theory*. Vol. 3. 3. Oxford university press New York, 2004. ISBN: 9780195128963.

[181]   Dave W. Oyler, Pierre T. Kabamba, and Anouck R. Girard. "Pursuit-evasion games in the presence of obstacles". In: *Automatica* 65 (Supplement C Mar. 1, 2016), pp. 1–11. DOI: `10.1016/j.automatica.2015.11.018`.

[182]   David Oyler. "Contributions To Pursuit-Evasion Game Theory." University of Michigan, Jan. 1, 2016.

[183]   Meir Pachter. "Isaacs' Two-on-One Pursuit Evasion Game". In: *18th International Symposium on Dynamic Games and Applications (ISDG)*. 18th International Symposium on Dynamic Games and Applications. Grenoble, France, 2018.

[184]   Meir Pachter, Eloy Garcia, and David W. Casbeer. "Active target defense differential game". In: *2014 52nd Annual Allerton Conference on Communication, Control, and Computing*. 2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton) (Monticello, IL, USA). IEEE, Sept. 2014. ISBN: 9781479980093. DOI: `10.1109/allerton.2014.7028434`.

[185]   Meir Pachter, Eloy Garcia, and David W. Casbeer. "Toward a Solution of the Active Target Defense Differential Game". en. In: *Dynamic Games and Applications* (Mar. 16, 2018), pp. 1–52. DOI: `10.1007/s13235-018-0250-1`.

[186]   Meir Pachter, Alexander Von Moll, Eloy Garcia, David Casbeer, and Dejan Milutinović. "Singular Trajectories in the Two Pursuer One Evader Differential Game". In: *2019 International Conference on Unmanned Aircraft Systems*. 2019 International Conference on Unmanned Aircraft Systems. Atlanta, GA, June 15, 2019. DOI: `10.1109/ICUAS.2019.8798244`.

[187]   Meir Pachter, Alexander Von Moll, Eloy Garcia, David Casbeer, and Dejan Milutinović. "Two-on-One Pursuit". In: *Journal of Guidance, Control, and Dynamics* 42.7 (July 8, 2019). DOI: `10.2514/1.G004068`.

[188]   Meir Pachter and Patrick Wasz. "On a Two Cutters and Fugitive Ship Differential Game". In: *IEEE Control Systems Letters* 3 (4 Oct. 2019), pp. 913–917. DOI: `10.1109/lcsys.2019.2919418`.

[189]   Yunpeng Pan and Evangelos Theodorou. "Probabilistic differential dynamic programming". In: *Advances in Neural Information Processing Systems*. Jan. 1, 2014, pp. 1907–1915.

[190]   A. G. Pashkov and S. D. Terekhov. "A differential game of approach with two pursuers and one evader". en. In: *Journal of Optimization Theory and Applications* 55 (2 Nov. 1987), pp. 303–311. DOI: `10.1007/bf00939087`.

[191]   Valerii Patsko, Sergey Kumkov, and Varvara Turova. "Pursuit-Evasion Games". In: Springer International Publishing, 2018, pp. 1–87. ISBN: 9783319273358. DOI: `10.1007/978-3-319-27335-8_30-2`.

[192]   A. Pierson, Z. Wang, and M. Schwager. "Intercepting Rogue Robots: An Algorithm for Capturing Multiple Evaders With Multiple Pursuers". In: *IEEE Robotics and Automation Letters* 2 (2 Apr. 1, 2017), pp. 530–537. DOI: `10.1109/LRA.2016.2645516`.

[193]   Lev Semenovich Pontryagin, V G Boltyanskii, R V Gamkrelidze, and E F Mishchenko. *The mathematical theory of optimal processes*. New York, NY: Wiley, 1962.

[194]   Adrian P. Pope, Jaime S. Ide, Daria Micovic, Henry Diaz, David Rosenbluth, Lee Ritholtz, Jason C. Twedt, Thayne T. Walker, Kevin Alcedo, and Daniel Javorsek. "Hierarchical Reinforcement Learning for Air-to-Air Combat". In: (May 2021). eprint: `2105.00990v2`.

[195]   M. J. D. Powell. "A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation". In: Springer Netherlands, 1994, pp. 51–67. ISBN: 9789048143580. DOI: `10.1007/978-94-015-8330-5_4`.

[196]   Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. de. John Wiley & Sons, Aug. 28, 2014.

[197] Christopher Rackauckas and Qing Nie. "DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia". en. In: *Journal of Open Research Software* 5 (May 2017). DOI: 10.5334/jors.151.

[198] Anil V. Rao. "A Survey of Numerical Methods in Optimal Control". In: vol. 135. 2010, pp. 497–528.

[199] J. F. Reinganum. "A class of differential games for which the closed-loop and open-loop Nash equilibria coincide". en. In: *Journal of Optimization Theory and Applications* 36 (2 Feb. 1, 1982), pp. 253–262. DOI: 10.1007/BF00933832.

[200] Craig W. Reynolds. "Flocks, herds and schools: A distributed behavioral model". In: *the 14th annual conference* (Not Known). ACM Press, 1987. ISBN: 0897912276. DOI: 10.1145/37401.37406.

[201] S. J. Rubio. "On Coincidence of Feedback Nash Equilibria and Stackelberg Equilibria in Economic Applications of Differential Games". en. In: *Journal of Optimization Theory and Applications* 128 (1 Jan. 1, 2006), pp. 203–220. DOI: 10.1007/s10957-005-7565-y.

[202] Stuart Russell and Peter Norvig. *Artificial Intelligence -A Modern Approach*. en. Pearson Education, 2010.

[203] Thomas C. Schelling. *The Strategy of Conflict*. en. Harvard University Press, Jan. 1, 1980.

[204] Jacob Schrum and Risto Miikkulainen. "Evolving multi-modal behavior in NPCs". In: *2009 IEEE Symposium on Computational Intelligence and Games (CIG)* (Milano, Italy). IEEE, Sept. 2009. ISBN: 9781424448142. DOI: 10.1109/cig.2009.5286459.

[205] Howard M. Schwartz. "Differential Games". en. In: (Jan. 1, 2014), pp. 144–199.

[206] Iris Rubi Seaman, Jan-Willem van de Meent, and David Wingate. "Nested Reasoning About Autonomous Agents Using Probabilistic Programs". In: (Dec. 2018). eprint: 1812.01569v2.

[207]   William A Searcy and Stephen Nowicki. *The evolution of animal communication: reliability and deception in signaling systems*. Princeton University Press, 2005.

[208]   Eduardo Sebastian and Eduardo Montijano. "A Multi-robot Cooperative Control Strategy for Non-linear Entrapment Problems". In: *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)* (Zaragoza, Spain). IEEE, Sept. 2019. ISBN: 9781728103037. DOI: 10.1109/etfa.2019.8869099.

[209]   Jhanani Selvakumar and Efstathios Bakolas. "Feedback Strategies for a Reach-Avoid Game With a Single Evader and Multiple Pursuers". In: *IEEE transactions on cybernetics* (2019).

[210]   Jhanani Selvakumar and Efstathios Bakolas. *Min-Max Q-Learning for Multi-Player Pursuit-Evasion Games*. 2020. eprint: 2003.03727.

[211]   Robert L Shaw. *Fighter Combat -Tactics and maneuvering*. 1985.

[212]   J. Shinar, M. Medinah, and M. Biton. "Singular surfaces in a linear pursuit-evasion game with elliptical vectograms". en. In: *Journal of Optimization Theory and Applications* 43 (3 July 1984), pp. 431–456. DOI: 10.1007/bf00934465.

[213]   Josef Shinar, Valery Y. Glizer, and Vladimir Turetsky. "A Pursuit-evasion Game with Hybrid Pursuer Dynamics". en. In: *European Journal of Control* 15 (6 Jan. 2009), pp. 665–684. DOI: 10.3166/ejc.15.665-684.

[214]   Daigo Shishika and Vijay Kumar. "Local-game Decomposition for Multiplayer Perimeter-defense Problem". In: *2018 IEEE Conference on Decision and Control (CDC)* (Miami Beach, FL). IEEE, Dec. 2018. ISBN: 9781538613955. DOI: 10.1109/cdc.2018.8618879.

[215]   Daigo Shishika and Vijay Kumar. "Perimeter-defense Game on Arbitrary Convex Shapes". In: *arXiv* (Sept. 9, 2019).

[216]   Daigo Shishika and Vijay Kumar. "A Review of Multi Agent Perimeter Defense Games". In: Springer International Publishing, Dec. 2020, pp. 472–485. DOI: 10.1007/978-3-030-64793-3_26.

[217] Daigo Shishika, Dipankar Maity, and Michael Dorothy. "Partial Information Target Defense Game". In: *International Conference on Robotics and Automation*. International Conference on Robotics and Automation. Xi'an, China, May 30, 2021. DOI: `10.1109/ICRA48506.2021.9561995`.

[218] Daigo Shishika, James Paulos, Michael R. Dorothy, M. Ani Hsieh, and Vijay Kumar. "Team Composition for Perimeter Defense with Patrollers and Defenders". In: *2019 IEEE 58th Conference on Decision and Control (CDC)* (Nice, France). IEEE, Dec. 2019. ISBN: 9781728113982. DOI: `10.1109/cdc40024.2019.9030082`.

[219] Daigo Shishika, James Paulos, and Vijay Kumar. "Cooperative Team Strategies for Multi-player Perimeter-Defense Games". In: *Robotics and Automation Letters* (Feb. 10, 2020), pp. 2738–2745. DOI: `10.1109/LRA.2020.2972818`.

[220] Neryahu A Shneydor. *Missile guidance and pursuit: kinematics, dynamics and control*. Elsevier, 1998.

[221] Bikash Shrestha and Zachariah E. Fuchs. "Engage or Retreat Differential Game with Two Targets". In: *NAECON 2019 - IEEE National Aerospace and Electronics Conference* (Dayton, OH, USA). IEEE, July 2019. ISBN: 9781728114163. DOI: `10.1109/naecon46414.2019.9057945`.

[222] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1986. ISBN: 9780412246203.

[223] Emmanuel Sin, Murat Arcak, Douglas Philbrick, and Peter Seiler. "Iterative Best Response for Multi-Body Asset-Guarding Games". In: (Nov. 2020). eprint: `2011.01893v1`.

[224] P. Souères and J. -D. Boissonnat. "Optimal trajectories for non-holonomic mobile robots". en. In: 229 (Jan. 1, 1998). Ed. by J. -P. Laumond, pp. 93–170.

[225] Marc Steinegger, Dominique G. Roche, and Redouan Bshary. "Simple decision rules underlie collaborative hunting in yellow saddle goatfish". en. In: *Proceedings of the Royal Society B:*

Biological Sciences 285 (1871 Jan. 2017), p. 20172488. DOI: 10 . 1098/rspb.2017.2488.

[226] Josef Stoer and Roland Bulirsch. *Introduction to numerical analysis*. Vol. 12. Springer Science & Business Media, 2013.

[227] Laura Strickland, Kaitlin Baudier, Kenneth Bowers, Theodore P. Pavlic, and Charles Pippin. "Bio-inspired Role Allocation of Heterogeneous Teams in a Site Defense Task". In: *Distributed Autonomous Robotic Systems*. Ed. by Nikolaus Correll, Mac Schwager, and Michael Otte. Springer International Publishing, 2019, pp. 139–151. ISBN: 978-3-030-05816-6.

[228] Laura G. Strickland, Charles E. Pippin, and Matthew Gombolay. "Learning to Steer Swarm-vs.-swarm Engagements". In: *AIAA Scitech 2021 Forum* (VIRTUAL EVENT). American Institute of Aeronautics and Astronautics, Jan. 2021. DOI: 10.2514/6.2021-0165.

[229] Laura G. Strickland, Eric G. Squires, Michael A. Day, and Charles E. Pippin. "On Coordination in Multiple Aerial Engagement". In: *2019 International Conference on Unmanned Aircraft Systems (ICUAS)* (Atlanta, GA, USA). IEEE, June 2019. ISBN: 9781728103334. DOI: 10.1109/icuas.2019.8798042.

[230] John O. Sullivan. "Variability in the Wolf, a Group Hunter". In: Elsevier, 1978, pp. 31–40. ISBN: 9780123192509. DOI: 10.1016/b978-0-12-319250-9.50009-8.

[231] Wei Sun and Panagiotis Tsiotras. "Sequential pursuit of multiple targets under external disturbances via Zermelo–Voronoi diagrams". In: *Automatica* 81 (Supplement C July 1, 2017), pp. 253–260. DOI: 10.1016/j.automatica.2017.03.015.

[232] Wei Sun, Panogiotis Tsiotras, Tapovan Lolla, Deepak N Subramani, and Pierre FJ Lermusiaux. "Multiple-Pursuer/One-Evader Pursuit--Evasion Game in Dynamic Flowfields". In: *Journal of guidance, control, and dynamics* (Jan. 1, 2017).

[233]   Brian A. Swanson, Zachariah E. Fuchs, and Jason Shroyer. "Singular Solutions of the Optimal Pursuer Response to Fixed Control Strategies of a Dual-Evader Defensive Team". In: *2021 American Control Conference (ACC)* (New Orleans, LA, USA). IEEE, May 2021. DOI: `10.23919/acc50511.2021.9483204`.

[234]   Ken Tan, Zhenwei Wang, Hua Li, Shuang Yang, Zongwen Hu, Gerald Kastberger, and Benjamin P. Oldroyd. "An 'I see you' prey–predator signal between the Asian honeybee, Apis cerana, and the hornet, Vespa velutina". en. In: *Animal Behaviour* 83 (4 Apr. 2012), pp. 879–882. DOI: `10.1016/j.anbehav.2011.12.031`.

[235]   Surafel Luleseged Tilahun and Hong Choon Ong. "Prey-Predator Algorithm: A New Metaheuristic Algorithm for Optimization Problems". In: *International Journal of Information Technology & Decision Making* 14 (06 Nov. 2015), pp. 1331–1352. DOI: `10.1142/S021962201450031X`.

[236]   Emanuel Todorov and Weiwei Li. "A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems". In: *Proceedings of the 2005, American Control Conference, 2005.* IEEE, Jan. 1, 2005, pp. 300–306.

[237]   U.S. Air Force. *U.S. Air Force 2030 Science and Technology Strategy*. Tech. rep. Apr. 1, 2019.

[238]   Michael Van Lent, William Fisher, and Michael Mancuso. "An explainable artificial intelligence system for small-unit tactical behavior". In: *IAAI Emerging Applications*. AAAI, 2004.

[239]   Petr Vaňa and Jan Faigl. "Optimal Solution of the Generalized Dubins Interval Problem". en. In: (Jan. 1, 2018), p. 11.

[240]   Thomas L. Vincent. "Collision avoidance at sea". In: Springer-Verlag, 1977, pp. 205–221. ISBN: 354008407X. DOI: `10.1007/bfb0009072`.

[241]   J. Vlassenbroeck and R. Van Dooren. "A Chebyshev technique for solving nonlinear optimal control problems". In: *IEEE Trans-*

actions on Automatic Control 33 (4 Apr. 1988), pp. 333–340. DOI: 10.1109/9.192187.

[242] Jacques Vlassenbroeck. "A chebyshev polynomial method for optimal control with state constraints". en. In: *Automatica* 24 (4 July 1988), pp. 499–506. DOI: 10.1016/0005-1098(88)90094-5.

[243] Alexander Von Moll. "Machine Learning Applications in Complex Control Systems". Special Project AE8900. Georgia Institute of Technology, Dec. 1, 2015.

[244] Alexander Von Moll, Pavlos Androulakakis, Zachariah Fuchs, and Dieter Vanderelst. "Evolutionary Design of Cooperative Predation Strategies". In: *Conference on Games*. Conference on Games. © 2020 IEEE. Reprinted, with permission. Osaka, Japan: IEEE, Aug. 24, 2020. DOI: 10.1109/CoG47356.2020.9231945.

[245] Alexander Von Moll, David Casbeer, Eloy Garcia, Dejan Milutinović, and Meir Pachter. "The Multi-Pursuer Single-Evader Game: A Geometric Approach". In: *Journal of Intelligent and Robotic Systems* 96 (2 Jan. 2, 2019). © 2019 Springer Nature. Reprinted, with permission, pp. 193–207. DOI: 10.1007/s10846-018-0963-9.

[246] Alexander Von Moll, David W. Casbeer, Eloy Garcia, and Dejan Milutinović. "Pursuit-evasion of an Evader by Multiple Pursuers". In: *2018 International Conference on Unmanned Aircraft Systems (ICUAS)*. 2018 International Conference on Unmanned Aircraft Systems (ICUAS). © 2018 IEEE. Reprinted, with permission. Dallas, TX: IEEE, June 1, 2018, pp. 133–142. DOI: 10.1109/ICUAS.2018.8453470.

[247] Alexander Von Moll and Zachariah Fuchs. "Attacker Dispersal Surface in the Turret Defense Differential Game". In: *IFAC World Congress on Automatic Control*. IFAC World Congress on Automatic Control. Vol. 53. © 2020 The Authors. Previously published in IFAC-PapersOnline. Berlin, Germany, July 31, 2020, pp. 15659–15666. DOI: 10.1016/j.ifacol.2020.12.2549.

[248]   Alexander Von Moll and Zachariah Fuchs. "Optimal Constrained Retreat within the Turret Defense Differential Game". In: *Conference on Control Technology and Applications*. Conference on Control Technology and Applications. © 2020 IEEE. Reprinted, with permission. Montreal, Canada, Aug. 24, 2020. DOI: 10 . 1109/CCTA41146.2020.9206388.

[249]   Alexander Von Moll and Zachariah Fuchs. "A Lock-Evade, Engage or Retreat Game". In: *Transactions on Aerospace & Electronic Systems* (Oct. 1, 2022). In Preparation.

[250]   Alexander Von Moll, Zachariah Fuchs, and Meir Pachter. "Optimal Evasion Against Dual Pure Pursuit". In: *American Control Conference*. American Control Conference. © 2020 IEEE. Reprinted, with permission. Denver, CO: IEEE, July 5, 2020. DOI: 10.23919/ACC45564.2020.9147776.

[251]   Alexander Von Moll, Eloy Garcia, David Casbeer, Suresh Manickam, and Sufal Chandra Swar. "Multiple Pursuer Single Evader Border Defense Differential Game". In: *Journal of Aerospace Information Systems* (Feb. 1, 2020). Reprinted, with permission. DOI: 10.2514/1.I010740.

[252]   Alexander Von Moll, Meir Pachter, and Zachariah Fuchs. "Pure Pursuit". In: *Dynamic Games and Applications* (2022). In Preparation.

[253]   Alexander Von Moll, Meir Pachter, Eloy Garcia, David Casbeer, and Dejan Milutinović. "Robust Policies for a Multiple Pursuer Single Evader Differential Game". In: *Dynamic Games and Applications* (10 May 4, 2019). © 2019 Springer Nature. Reprinted, with permission, pp. 202–221. DOI: 10 . 1007 / s13235 - 019 - 00313-3.

[254]   Alexander Von Moll, Meir Pachter, Daigo Shishika, and Zachariah Fuchs. "Guarding a Circular Target By Patrolling its Perimeter". In: *Conference on Decision and Control*. Conference on Decision and Control. © 2020 IEEE. Reprinted, with permission. Jeju Island, South Korea: IEEE, Dec. 11, 2020. DOI: 10 . 1109 / CDC42340.2020.9304018.

[255]   Alexander Von Moll, Meir Pachter, Daigo Shishika, and Zachariah Fuchs. "Circular Target Defense Differential Games". In: *Transactions on Automatic Control* (June 30, 2021). Submitted for Review.

[256]   Alexander Von Moll, Daigo Shishika, Zachariah Fuchs, and Michael Dorothy. "The Turret-Runner-Penetrator Differential Game with Role Selection". In: *Transactions on Aerospace & Electronic Systems* (Feb. 26, 2022). Submitted for Review.

[257]   Draguna Vrabie, Kyriakos Vamvoudakis, and Frank Lewis. *Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles*. Institution of Engineering and Technology, Jan. 2012. ISBN: 9781849194891. DOI: 10.1049/pbce081e.

[258]   Isaac Weintraub, Eloy Garcia, David Casbeer, and Meir Pachter. "An Optimal Aircraft Defense Strategy for the Active Target Defense Scenario". In: *AIAA Guidance, Navigation, and Control Conference* (Grapevine, Texas). American Institute of Aeronautics and Astronautics, Jan. 2017. ISBN: 9781624104503. DOI: 10.2514/6.2017-1917.

[259]   Isaac Weintraub, Eloy Garcia, and Meir Pachter. "Optimal guidance strategy for the defense of a non-manoeuvrable target in 3-dimensions". en. In: *IET Control Theory & Applications* 14 (11 July 2020), pp. 1531–1538. DOI: 10.1049/iet-cta.2019.0541.

[260]   Isaac Weintraub, Alexander Von Moll, Eloy Garcia, and Meir Pachter. "Maximum Observation of a Target by a Slower Observer in Three Dimensions". In: *Journal of Guidance, Control, and Navigation* (Dec. 15, 2020). DOI: 10.2514/1.G005619.

[261]   Isaac E. Weintraub, Richard G. Cobb, William Baker, and Meir Pachter. "Direct Methods Comparison for the Active Target Defense Scenario". In: *AIAA Scitech 2020 Forum* (Orlando, FL). American Institute of Aeronautics and Astronautics, Jan. 2020. ISBN: 9781624105951. DOI: 10.2514/6.2020-0612.

[262]   Isaac E. Weintraub, Meir Pachter, and Eloy Garcia. "An Introduction to Pursuit-evasion Differential Games". In: *2020 Amer-*

ican Control Conference (ACC) (Denver, CO, USA). IEEE, July 2020. ISBN: 9781538682661. DOI: `10.23919/acc45564.2020.9147205`.

[263]   Isaac E. Weintraub, Alexander Von Moll, Christian Carrizales, Nicholas Hanlon, and Zachariah Fuchs. "An Optimal Engagement Zone Avoidance Scenario in 2-D". In: *Scitech*. Scitech. Accepted. San Diego: AIAA, Jan. 3, 2022.

[264]   Isaac E. Weintraub, Alexander Von Moll, Eloy Garcia, David Casbeer, Zachary J. L. Demers, and Meir Pachter. "Maximum Observation of a Faster Non-Maneuvering Target by a Slower Observer". In: *American Control Conference*. American Control Conference. Denver, CO, July 5, 2020. DOI: `10.23919/ACC45564.2020.9147340`.

[265]   Alfredo Weitzenfeld, Alberto Vallesa, and Horacio Flores. "A Biologically-Inspired Wolf Pack Multiple Robot Hunting Model". In: *2006 IEEE 3rd Latin American Robotics Symposium* (Santiago, Chile). IEEE, Oct. 2006. ISBN: 142440536X. DOI: `10.1109/lars.2006.334327`.

[266]   Artur Wolek, Eugene M. Cliff, and Craig A. Woolsey. "Time-Optimal Path Planning for a Kinematic Car with Variable Speed". In: *Journal of Guidance, Control, and Dynamics* 39 (10 Jan. 1, 2016), pp. 2374–2390. DOI: `10.2514/1.G001317`.

[267]   Rui Yan, Zongying Shi, and Yisheng Zhong. "Defense game in a circular region". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)* (Melbourne, Australia). IEEE, Dec. 2017. DOI: `10.1109/cdc.2017.8264502`.

[268]   Rui Yan, Zongying Shi, and Yisheng Zhong. "Reach-Avoid Games With Two Defenders and One Attacker: An Analytical Approach". In: *IEEE Transactions on Cybernetics* 49 (3 Mar. 2019), pp. 1035–1046. DOI: `10.1109/tcyb.2018.2794769`.

[269]   Rui Yan, Zongying Shi, and Yisheng Zhong. "Task Assignment for Multiplayer Reach–Avoid Games in Convex Domains via

Analytical Barriers". In: *IEEE Transactions on Robotics* 36 (1 Feb. 2020), pp. 107–124. DOI: 10.1109/tro.2019.2935345.

[270] Rui Yan, Zongying Shi, and Yisheng Zhong. "Cooperative strategies for two-evader-one-pursuer reach-avoid differential games". en. In: *International Journal of Systems Science* (Jan. 2021), pp. 1–19. DOI: 10.1080/00207721.2021.1872116.

[271] Rui Yan, Zongying Shi, and Yisheng Zhong. "Optimal strategies for the lifeline differential game with limited lifetime". en. In: *International Journal of Control* 94 (8 Aug. 2021), pp. 2238–2251. DOI: 10.1080/00207179.2019.1698770.

[272] Y. Yavin. "Pursuit-evasion differential games with deception or interrupted observation". en. In: *Computers & Mathematics with Applications* 13 (1-3 1987), pp. 191–203. DOI: 10.1016/0898-1221(87)90104-0.

[273] Zou Yuan, Liu Teng, Sun Fengchun, and Huei Peng. "Comparative study of dynamic programming and Pontryagin's minimum principle on energy management for a parallel hybrid electric vehicle". In: *Energies* 6 (4 Jan. 1, 2013), pp. 2305–2318.

[274] Noah D. Zepp, Han Fu, and Hugh H. Liu. "Autonomous Strategic Defense: An Adaptive Clustering Approach to Capture Order Optimization". In: *AIAA SCITECH 2022 Forum* (San Diego, CA & Virtual). American Institute of Aeronautics and Astronautics, Jan. 2022. DOI: 10.2514/6.2022-2215.

[275] Kemin Zhou and John Comstock Doyle. *Essentials of robust control*. Vol. 104. Prentice hall Upper Saddle River, NJ, 1998.

[276] Vrushabh Zinage and Efstathios Bakolas. "Koopman Operator Based Modeling and Control of Rigid Body Motion Represented by Dual Quaternions". In: (Oct. 2021). eprint: 2110.04967v1.

PUBLICATIONS

JOURNAL

[J1] Alexander Von Moll and Zachariah Fuchs. "A Lock-Evade, Engage or Retreat Game". In: *Transactions on Aerospace & Electronic Systems* (2022). In Preparation.

[J2] Alexander Von Moll, Meir Pachter, and Zachariah Fuchs. "Pure Pursuit". In: *Dynamic Games and Applications* (2022). In Preparation.

[J3] Alexander Von Moll, Daigo Shishika, Zachariah Fuchs, and Michael Dorothy. "The Turret-Runner-Penetrator Differential Game with Role Selection". In: *Transactions on Aerospace & Electronic Systems* (Feb. 26, 2022). Submitted for Review.

[J4] Michael Dorothy, Dipankar Maity, Daigo Shishika, and Alexander Von Moll. "One Apollonius Circle is Enough for Many Pursuit-Evasion Games". In: *Automatica* (Nov. 17, 2021). Submitted for Review. eprint: 2111.09205.

[J5] Dejan Milutinović, David W. Casbeer, Alexander Von Moll, Meir Pachter, and Eloy Garcia. "Rate of Loss Characterization that Resolves the Dilemma of the Wall Pursuit Game Solution". In: *Transactions on Automatic Control* (Dec. 27, 2021). DOI: 10.1109/TAC.2021.3137786.

[J6] Alexander Von Moll, Meir Pachter, Daigo Shishika, and Zachariah Fuchs. "Circular Target Defense Differential Games". In: *Transactions on Automatic Control* (June 30, 2021). Submitted for Review.

[J7] Eloy Garcia, David W Casbeer, Alexander Von Moll, and Meir Pachter. "Multiple Pursuer Multiple Evader Differential Games". In: *Transactions on Automatic Control* (May 1, 2020). DOI: 10.1109/TAC.2020.3003840.

[J8]   Satyanarayana G. Manyam, David W. Casbeer, Alexander Von
        Moll, and Zachariah Fuchs. "Curvature Constrained Paths to
        Intercept a Target Moving on a Circle". In: *Transactions on Au-
        tomation Science and Engineering* (Dec. 31, 2020). Submitted for
        Review.

[J9]   Meir Pachter, Alexander Von Moll, Eloy Garcia, David Cas-
        beer, and Dejan Milutinović. "Cooperative Pursuit by Multi-
        ple Pursuers of a Single Evader". In: *Journal of Aerospace Infor-
        mation Systems* (Feb. 1, 2020). DOI: 10.2514/1.I010739.

[J10]   John L. Salmon, Landon C. Willey, David Casbeer, Eloy Garcia,
        and Alexander Von Moll. "Single Pursuer Multiple-Cooperative
        Evaders in the Border Defense Differential Game". In: *Journal
        of Aerospace Information Systems* (Mar. 19, 2020). DOI: 10.2514/
        1.I010766.

[J11]   Alexander Von Moll, Eloy Garcia, David Casbeer, Suresh Man-
        ickam, and Sufal Chandra Swar. "Multiple Pursuer Single Evader
        Border Defense Differential Game". In: *Journal of Aerospace In-
        formation Systems* (Feb. 1, 2020). DOI: 10.2514/1.I010740.

[J12]   Isaac Weintraub, Alexander Von Moll, Eloy Garcia, and Meir
        Pachter. "Maximum Observation of a Target by a Slower Ob-
        server in Three Dimensions". In: *Journal of Guidance, Control,
        and Navigation* (Dec. 15, 2020). DOI: 10.2514/1.G005619.

[J13]   Meir Pachter, Alexander Von Moll, Eloy Garcia, David Cas-
        beer, and Dejan Milutinović. "Two-on-One Pursuit". In: *Jour-
        nal of Guidance, Control, and Dynamics* 42.7 (July 8, 2019). DOI:
        10.2514/1.G004068.

[J14]   Alexander Von Moll, David Casbeer, Eloy Garcia, Dejan Mi-
        lutinović, and Meir Pachter. "The Multi-Pursuer Single-Evader
        Game: A Geometric Approach". In: *Journal of Intelligent and
        Robotic Systems* 96 (2 Jan. 2, 2019), pp. 193–207. DOI: 10.1007/
        s10846-018-0963-9.

[J15]   Alexander Von Moll, Meir Pachter, Eloy Garcia, David Cas-
        beer, and Dejan Milutinović. "Robust Policies for a Multiple

Pursuer Single Evader Differential Game". In: *Dynamic Games and Applications* (10 May 4, 2019), pp. 202–221. DOI: 10.1007/s13235-019-00313-3.

CONFERENCE

[C1] Shivam Bajaj, Eric Torng, Shaunak D. Bopardikar, Alexander Von Moll, Isaac Weintraub, Eloy Garcia, and David W. Casbeer. "Competitive Perimeter Defense of Conical Environments". In: *American Control Conference*. American Control Conference. Submitted for Review. June 8, 2022.

[C2] Isaac E. Weintraub, Alexander Von Moll, Christian Carrizales, Nicholas Hanlon, and Zachariah Fuchs. "An Optimal Engagement Zone Avoidance Scenario in 2-D". In: *Scitech*. Scitech. Accepted. San Diego: AIAA, Jan. 3, 2022.

[C3] Zachariah E. Fuchs, Alexander Von Moll, and David Casbeer. "Engage or Retreat Differential Game with N-Targets and Distributed Defensive Assets". In: *Conference on Controls Technology and Applications*. Conference on Controls Technology and Applications. San Diego, CA, Aug. 31, 2021. DOI: 10.1109/CCTA48906.2021.9658922.

[C4] Eloy Garcia, David W. Casbeer, Alexander Von Moll, and Meir Pachter. "The Cooperative Blocking Differential Game". In: IEEE, Dec. 31, 2021. DOI: 10.1109/CCTA48906.2021.9658654.

[C5] Alexander Von Moll and Zachariah Fuchs. "Turret Lock-on in an Engage or Retreat Game". In: *American Control Conference*. American Control Conference. New Orleans: IEEE, May 28, 2021. DOI: 10.23919/ACC50511.2021.9483106.

[C6] Alexander Von Moll, Daigo Shishika, Zachariah Fuchs, and Michael Dorothy. "The Turret-Runner-Penetrator Differential Game". In: *American Control Conference*. American Control Conference. New Orleans, May 26, 2021. DOI: 10.23919/ACC50511.2021.9483094.

[C7] Isaac Weintraub, Alexander Von Moll, David Casbeer, Eloy Garcia, and Meir Pachter. "Engagement Zone Defense of a Non-Maneuvering Evader". In: *Conference on Control Technology and Applications*. Conference on Control Technology and Applications. San Diego, CA, Aug. 31, 2021. DOI: 10.1109/CCTA48906.2021.9659042.

[C8] Eloy Garcia, David W. Casbeer, Alexander Von Moll, and Meir Pachter. "Pride of Lions and Man Differential Game". In: *Conference on Decision and Control*. Conference on Decision and Control. Jeju Island, South Korea, Dec. 8, 2020. DOI: 10.1109/CDC42340.2020.9304060.

[C9] Alexander Von Moll, Pavlos Androulakakis, Zachariah Fuchs, and Dieter Vanderelst. "Evolutionary Design of Cooperative Predation Strategies". In: *Conference on Games*. Conference on Games. Osaka, Japan: IEEE, Aug. 24, 2020. DOI: 10.1109/CoG47356.2020.9231945.

[C10] Alexander Von Moll and Zachariah Fuchs. "Attacker Dispersal Surface in the Turret Defense Differential Game". In: *IFAC World Congress on Automatic Control*. IFAC World Congress on Automatic Control. Vol. 53. Berlin, Germany, July 31, 2020, pp. 15659–15666. DOI: 10.1016/j.ifacol.2020.12.2549.

[C11] Alexander Von Moll and Zachariah Fuchs. "Optimal Constrained Retreat within the Turret Defense Differential Game". In: *Conference on Control Technology and Applications*. Conference on Control Technology and Applications. Montreal, Canada, Aug. 24, 2020. DOI: 10.1109/CCTA41146.2020.9206388.

[C12] Alexander Von Moll, Zachariah Fuchs, and Meir Pachter. "Optimal Evasion Against Dual Pure Pursuit". In: *American Control Conference*. American Control Conference. Denver, CO: IEEE, July 5, 2020. DOI: 10.23919/ACC45564.2020.9147776.

[C13] Alexander Von Moll, Meir Pachter, Daigo Shishika, and Zachariah Fuchs. "Guarding a Circular Target By Patrolling its Perimeter". In: *Conference on Decision and Control*. Conference on Deci-

sion and Control. Jeju Island, South Korea: IEEE, Dec. 11, 2020. DOI: 10.1109/CDC42340.2020.9304018.

[C14]   Isaac E. Weintraub, Alexander Von Moll, Eloy Garcia, David Casbeer, Zachary J. L. Demers, and Meir Pachter. "Maximum Observation of a Faster Non-Maneuvering Target by a Slower Observer". In: *American Control Conference*. American Control Conference. Denver, CO, July 5, 2020. DOI: 10.23919/ACC45564.2020.9147340.

[C15]   Eloy Garcia, David W. Casbeer, Alexander Von Moll, and Meir Pachter. "Cooperative Two-Pursuer One-Evader Blocking Differential Game". In: American Control Conference. July 10, 2019. DOI: 10.23919/ACC.2019.8814294.

[C16]   Eloy Garcia, Alexander Von Moll, David Casbeer, and Meir Pachter. "Strategies for Defending a Coastline Against Multiple Attackers". In: *IEEE Conference on Decision and Control*. IEEE Conference on Decision and Control. Nice, France, Dec. 31, 2019. DOI: 10.1109/CDC40024.2019.9029340.

[C17]   G. Manyam Satyanarayana, David Casbeer, Alexander Von Moll, and Zachariah Fuchs. "Optimal Dubins Paths to Intercept a Moving Target on a Circle". In: American Control Conference. July 10, 2019. DOI: 10.23919/ACC.2019.8814913.

[C18]   Satyanarayana G. Manyam, David Casbeer, Alexander Von Moll, and Eloy Garcia. "Coordinating Defender Path Planning for Optimal Target-Attacker-Defender Game". In: *AIAA Infotech*. AIAA Infotech. San Diego, CA: AIAA, Jan. 11, 2019. DOI: 10.2514/6.2019-0388.

[C19]   Satyanarayana Gupta Manyam, David Casbeer, Alexander Von Moll, and Zachariah Fuchs. "Shortest Dubins Path to a Circle". In: *AIAA SciTech*. AIAA SciTech. San Diego, CA: AIAA, Jan. 7, 2019. DOI: 10.2514/6.2019-0919.

[C20]   Meir Pachter, Alexander Von Moll, Eloy Garcia, David Casbeer, and Dejan Milutinović. "Singular Trajectories in the Two Pursuer One Evader Differential Game". In: *2019 International*

Conference on Unmanned Aircraft Systems. 2019 International Conference on Unmanned Aircraft Systems. Atlanta, GA, June 15, 2019. DOI: 10.1109/ICUAS.2019.8798244.

[C21]   Alexander Von Moll, Eloy Garcia, David Casbeer, Suresh Manickam, and Sufal Chandra Swar. "Multiple Pursuer Single Evader Border Defense Differential Game". In: *AIAA SciTech*. AIAA SciTech. San Diego, CA: AIAA, Jan. 11, 2019. DOI: 10.2514/6.2019-1162.

[C22]   Richard J. Black, Behzad Moslehi, W Price, Vahid Sotoudeh, Michael Osterman, Diganta Das, Alireza Behbahani, Alexander Von Moll, and Kenneth Semega. "Integrated Fiber-Optic Sensor Network System Reliability Modeling and Analysis for Aerospace Applications". In: *2018 AIAA Information Systems-AIAA Infotech @ Aerospace*. AIAA SciTech. AIAA SciTech Forum. American Institute of Aeronautics and Astronautics, Jan. 7, 2018. DOI: 10.2514/6.2018-0714.

[C23]   Krishna Kalyanam, Satyanarayana Manyam, Alexander Von Moll, David Casbeer, and Meir Pachter. "Scalable and Exact MILP Methods for UAV Persistent Visitation Problem". In: Conference on Control Technology and Applications. IEEE, Aug. 1, 2018, pp. 337–342. DOI: 10.1109/CCTA.2018.8511587.

[C24]   Behzad Moslehi, W Price, Richard J. Black, Ming Han, Alireza Behbahani, Alexander Von Moll, and Kenneth Semega. "High-Bandwidth Fiber-Optic Pressure Sensors for High-Temperature Aerospace Applications". In: *2018 AIAA Information Systems-AIAA Infotech @ Aerospace*. AIAA SciTech. AIAA SciTech Forum. American Institute of Aeronautics and Astronautics, Jan. 7, 2018. DOI: 10.2514/6.2018-0715.

[C25]   Alexander Von Moll, David W. Casbeer, Eloy Garcia, and Dejan Milutinović. "Pursuit-evasion of an Evader by Multiple Pursuers". In: *2018 International Conference on Unmanned Aircraft Systems (ICUAS)*. 2018 International Conference on Unmanned Aircraft Systems (ICUAS). Dallas, TX: IEEE, June 1, 2018, pp. 133–142. DOI: 10.1109/ICUAS.2018.8453470.

[C26]   Alexander Von Moll, Krishna Kalyanam, David Casbeer, and
        Satyanarayana G. Manyam. "Genetic Algorithm Approach for
        UAV Persistent Visitation Problem". In: Dynamic Systems and
        Control Conference. Atlanta, GA: ASME, Oct. 1, 2018. DOI: 10.
        1115/DSCC2018-8950.

[C27]   Alireza R. Behbahani, Alex Von Moll, Robert Zeller, and James
        Ordo. "Aircraft Integration Challenges and Opportunities for
        Distributed Intelligent Control, Power, Thermal Management,
        Diagnostic and Prognostic Systems". In: *2014 SAE Aerospace
        Systems and Technology Conference*. 2014 SAE Aerospace Systems
        and Technology Conference. SAE International, Sept. 1, 2014.
        DOI: 10.4271/2014-01-2161.

[C28]   Alexander Von Moll, Alireza R. Behbahani, Gustave C. Fralick,
        John D. Wrbanek, and Gary W. Hunter. "A Review of Exhaust
        Gas Temperature Sensing Techniques for Modern Turbine En-
        gine Controls". In: *50th AIAA/ASME/SAE/ASEE Joint Propul-
        sion Conference*. AIAA Propulsion and Energy Forum. Amer-
        ican Institute of Aeronautics and Astronautics, July 25, 2014.
        DOI: 10.2514/6.2014-3977.

[C29]   Alexander Von Moll, Ken Semega, Alireza Behbahani, and John
        Hoying. "Recent Progress, Challenges, and Future Develop-
        ment Needs of Thermally/Energy Efficient Fuel Actuator Pump-
        ing Systems for Military Gas Turbine Engine Applications".
        In: *JANNAF Interagency Propulsion Committee 34th Airbreathing
        Propulsion*. JANNAF Interagency Propulsion Committee, Jan. 1,
        2014.

[C30]   Alexander Von Moll and Alireza R. Behbahani. "Comparison
        of Communication Architectures and Network Topologies for
        Distributed Propulsion Controls". In: *59th IIS 2013*. ISA, Jan. 1,
        2013.