Ballistic Missile Defense Technical Studies Series

MISSILE DEFENSE FOR THE 21ST CENTURY

Gregory H. Canavan

The Heritage Foundation

Published by
The Heritage Foundation
214 Massachusetts Avenue, NE
Washington, DC 20002–4999
800-544-4843
heritage.org

ISBN 0-89195-261-6

Cover design by Mark Hurlburt
Layout by Michelle Fulton Smith
Charts and Graphs by Russell Martin

Table of Contents

Preface

This study is part of a series of technical reports commissioned by The Heritage Foundation to examine programmatic issues related to ballistic missile defense. Since the Bush Administration officially declared in June 2002 that the United States had withdrawn from the 1972 Anti-Ballistic Missile Treaty with the Soviet Union, the focus has turned to the best way to build and operate a system that dramatically lessens America's vulnerability to attack as quickly as possible. Intelligence indicates that countermeasures developed by the United States and Russia are readily available.

President George W. Bush took a major step forward in addressing the programmatic issues when he announced on December 17, 2002, his decision to deploy a limited missile defense system for the protection of the American people, U.S. forward-deployed forces, and U.S. allies beginning in 2004. The initial system is to include ground-based and sea-based interceptors and sensor systems deployed on land, at sea, and in space—elements that will be tied together by a common command and control system. The initial defense system will be augmented on a continuing basis to provide a more robust defense over time. As the Administration moves to execute the President's plan, military planners and national security policymakers will need a fuller appreciation of the key technical and strategic issues involved.

In this study, Dr. Gregory H. Canavan, Senior Fellow and Science Advisor at Los Alamos National Laboratory, has surveyed the available technological options for ballistic missile defense based on the rapid development of the threat and on prior technological and organizational efforts to field missile defenses. His report makes specific recommendations on how policymakers and programmers should make the best use of existing programs to deploy the most robust missile defense possible. As one of the nation's premier scientists working in the area of weapons technology and national security, Dr. Canavan brings to this evaluation not only his broad technical expertise but also a knowledge of past efforts to develop and field ballistic missile defense systems.

The other studies in this series address such topics as contributions that missile defense systems will make to strategic stability in a multilateral environment.

— *Larry M. Wortzel, Ph.D., Vice President and Director of the Kathryn and Shelby Cullom Davis Institute for International Studies at The Heritage Foundation.*

List of Figures

FIGURES

List of Acronyms and Abbreviations

Acronyms[1]

AA	attack assessment
ABL	airborne laser
ABM	anti-ballistic missile
ABRES	Advanced Ballistic Re-Entry Systems
AFB	Air Force base
AHIT	Advanced Hover Interceptor Technology
AIT	Atmospheric Interceptor Technology
ALARM	Alert, Locate, and Report Missiles.
ALPS	Accidental Launch Protection System
AOA	Airborne Optical Adjunct
AOR	area of responsibility
APN	augmented proportional navigation
ASAT	anti-satellite
AWS	Advanced Warning System
BAMBI	Ballistic Missile Boost Intercept
BE	Brilliant Eyes
BM	battle management
BMC2	battle management, command, and control
BMC3	battle management, command, control, and communications
BMC4	battle management, command, control, communications, and computers
BMC4ISR	battle management, command, control, communications, computers, intelligence, surveillance, and reconnaissance
BMDO	Ballistic Missile Defense Organization
BMDS	Ballistic Missile Defense System
BMEWS	Ballistic Missile Early Warning Radar Systems
BP	Brilliant Pebbles
BPI	boost-phase intercept
BSTS	Boost-Phase Surveillance and Tracking System
C1	initial capability
C2	1. command and control; 2. intermediate capability
CBO	Congressional Budget Office
CC	Combat Commander
CEC	Cooperative Engagement Capability
CEP	circular error probable
CIA	Central Intelligence Agency
CINC	Commander in Chief (i.e., Combat Commander of a major command)

1. For a more extensive list of acronyms related to missile defense, see Missile Defense Agency, *MDA Glossary*, at *www.acq.osd.mil/bmdo/bmdolink/pdf/glossary.pdf* (March 26, 2003).

CMOC	Cheyenne Mountain Operations Center
CNO	Chief of Naval Operations
CONUS	continental United States
COTS	commercial off-the-shelf
CSO	closely spaced objects
DAB	Defense Acquisition Board
DACS	Divert and Attitude Control Systems
DARPA	Defense Advanced Research Projects Agency
DEW	directed energy weapon
DOD	Department of Defense
DRR	Defense Readiness Review
DSB	Defense Science Board
DSP	Defense Support Program
E2I	Endo-Exoatmospheric Interceptor
EKV	Exoatmospheric Kill Vehicle
EMD	Engineering and Manufacturing Development
EMP	electromagnetic pulse
ERINT	Extended Range Interceptor
ERIS	Exoatmospheric Reentry Vehicle Interceptor System
EW	early warning
EWR	early warning radar
FEWS	Follow-on Early Warning System
FLAGE	Flexible Lightweight Agile Guided Experiment
FOV	field of view
GBI	ground-based interceptor
GBR	ground-based radar
GBR-P	ground-based radar—prototype
GBX	Ground-Based Interceptor—Experimental
GCN	Ground Control Network
GEM	Guidance Enhancement Missile (PATRIOT)
GEO	geosynchronous earth orbit
GFC	Ground Flight Control
GMD	Ground-Based Midcourse Defense
GPALS	Global Protection Against Limited Strikes
GPS	Global Positioning System
GSTS	Ground-Based Surveillance and Tracking System
HEDI	High Endoatmospheric Defense Interceptor
HOE	homing overlay experiment
HTK	hit to kill
HWIL	hardware in loop
ICBM	intercontinental ballistic missile
IDO	Initial Defensive Operations
IFICS	in-flight interceptor control system
IFOV	instantaneous field of view
IFT	Intercept Flight Test
IFTU	in-flight trajectory update
IMU	inertial measurement units
IR	infrared
IRT	Independent Review Team
ISR	intelligence, surveillance, and reconnaissance
ITW/AA	Integrated Tactical Warning and Attack Assessment

JCS	Joint Chiefs of Staff
JNIC	Joint National Integration Center
JTAGS	Joint Tactical Ground System
JTAMDO	Joint Theater Air and Missile Defense Organization
KEW	kinetic energy weapon
KV	kill vehicle
LDS	Limited Defense System
LEAP	Lightweight Exoatmospheric Projectile
LEO	low earth orbit
LOADS	Low Altitude Defense System
LOS	line of sight
LSI	Lead System Integrator
LWIR	long wavelength infrared
MAD	mutual assured destruction
MDA	Missile Defense Agency
MDAP	Major Defense Acquisition Program
MEADS	Medium Extended Air Defense System
MIDAS	Missile Defense Alarm System
MIRV	multiple independently targeted reentry vehicles
MRBM	medium-range ballistic missile
MSP	Mosaic Sensor Program
MSR	Missile Site Radar
MWIR	mid-wavelength infrared
NAD	Navy Area Defends
NASA	National Aeronautics and Space Administration
NATO	North Atlantic Treaty Organization
NEAR	Near Earth Asteroid Rendezvous
NIE	National Intelligence Estimate
NMD	national missile defense
NNK	nonnuclear kill
NORAD	North American Aerospace Defense Command
NPB	neutral particle beam
NRL	Naval Research Laboratory
NTW	Navy Theater Wide
NWIR	near wavelength infrared
OPP	other physical principles
ORD	Operational Requirements Document
OSD	Office of the Secretary of Defense
PAC-2	PATRIOT Advanced Capability—2
PAC-3	PATRIOT Advanced Capability—3
PAR	Perimeter Acquisition Radar
PATRIOT	Phased Array Tracking Radar Intercept on Target
PAVE PAWS	Position and Velocity Extraction Phased Array Warning System
PLV	payload launch vehicle
PN	proportional navigation
R&D	research and development
RAMOS	Russian-American Observation Satellite
RCS	radar cross section
RMD	regional missile defense
RTS	Reagan Test Site
RV	reentry vehicle

SALT	Strategic Arms Limitation Talks
SAM	surface-to-air missile
SBI	space-based interceptor
SBIRS	Space-Based Infrared System
SBIRS-High	SBIRS high-altitude component
SBIRS-Low	SBIRS low-altitude component
SBL	space-based laser
SBX	sea-based S-band radar
SCR	Signal Corps radio [radar]
SDI	Strategic Defense Initiative
SDIO	Strategic Defense Initiative Organization
SDS	Strategic Defense System
SGEMP	systems-generated electromagnetic pulse
SLBM	submarine-launched ballistic missile
SMTS	Space and Missile Tracking System
SPAD	Space Patrol Active Defense
SSBN	ballistic missile submarine (nuclear)
SSTS	Space-Based Surveillance and Tracking System
STAR	System Threat Assessment Report
SWIR	short wavelength infrared
THAAD	Theater High Altitude Area Wide
TMD	theater missile defense
TOM	target object map
UAV	unmanned aerial vehicle
UCP	Unified Command Plan
UEWR	Upgraded Early Warning Radars
UHF	ultrahigh frequency
USSPACECOM	United States Space Command
USSTRATCOM	United States Strategic Command
UV	ultraviolet
VAFB	Vandenberg Air Force Base
WDL	weapon data load
WFOV	wide field of view
WMD	weapon of mass destruction
XBR	x-band radar

Abbreviations

C	Celsius
cc	cubic centimeter
cm	centimeter
db	decibel
ev	electron volt
g	gravity
GHz	gigahertz
J	Joule
kHz	kilohertz
km	kilometer
KT	kiloton
kV	kilovolt
MHz	megahertz

mJ	megajoule
mrad	milliradian
MT	megaton
MW	megawatt
psi	pounds per square inch
rad	radian
s	second
sr	steradian
W	watt

Introduction

Providing a system to defend the nation and its allies and friends from the terror of a ballistic missile attack is a national priority. Tens of countries now have ballistic missiles in the 1,000-kilometer range; several have missiles with ranges of several thousand kilometers; and North Korea has developed an indigenous intercontinental ballistic missile (ICBM) and has shown a willingness to sell its missiles and technology to states hostile to the United States. It has recently made clear that it seeks nuclear capabilities as well.

President Ronald Reagan established the Strategic Defense Initiative (SDI) to develop ways to blunt a possible Soviet missile strike. President George H. W. Bush relied on SDI research and development to develop the Global Protection Against Limited Strikes (GPALS) system, which would have provided high-confidence protection for the United States and its allies against accidental or unauthorized launches from Russia or China or rogue launches from elsewhere in the globe. GPALS was terminated for non-technical reasons by the Clinton Administration, which chose instead to concentrate funding and development on theater missile defense. The Clinton Administration was forced to return to a milestoned development of a midcourse national missile defense system after the unexpected North Korean launch of an ICBM in 1998. The results of that effort were mixed, with improvements in performance undercut by fundamental concerns about its robustness to countermeasures.

Since the current Administration formally withdrew from the Anti-Ballistic Missile (ABM) Treaty last year,[2] the United States has been free to select the best missile defenses available. The most effective way to destroy ballistic missiles once they are launched is to intercept them during their boost phase, before they can dispense multiple weapons or decoys. However, the development of such systems was delayed earlier by Treaty restraints and since by deliberate decisions. Today, the technology to defend against rogue, accidental, or unauthorized launches is largely in hand. A few hundred space-based interceptors (SBI) with a modest ground-based interceptor (GBI) underlay could produce high-confidence protection for America. SBIs were extensively developed and partially tested under GPALS. Further development should lead to interceptors that could address regional and theater missiles as well. Developing them would produce a fundamental shift from the offensive to defensive use of missiles.

During World War II, the German government developed rocket and guidance technologies that turned missiles into weapons and established the tactics that would make them effective. Subsequent U.S. and Soviet development extended their range and made them capable of intercepting offensive missiles. Since then, offensive and defensive developments in missile technology have been in competition, with offensive forces generally in the ascendancy.[3] This study examines the issues, decisions, and technologies that have governed the competition between offensive and defensive missile forces and the new developments that could shift the balance in favor of missile defense. It reviews the developments in offensive missile forces, proliferation, and threat that have increased the need for missile defenses. It then reviews the technological development of missile defenses over the last few

2. The United States formally withdrew from the ABM Treaty in June 2002, six months after giving formal notification of its intent to withdraw, per treaty requirements.
3. For more on the history and use of missile forces for strategic offense and strategic defense, see John L. Piotrowski, *Strategic Synchronization: The Relationship Between Strategic Offense and Defense* (Washington, D.C.: The Heritage Foundation, 2002).

decades and the basis of operation of current missile defense concepts. It describes elements of the current missile defense program and offers an assessment of its strengths and weaknesses. Technical appendices are provided to clarify important terms and estimates and to explain the specific technologies discussed in the text.

The author wishes to thank Dr. Baker Spring for his assistance and contributions to this study. He is also grateful to the late Dr. Edward Teller, his mentor and collaborator on missile defense from discussions in the early 1960s of whether it was possible to "hit a bullet" with a nuclear interceptor to those of the 1990s on how to hit a missile with a nonnuclear Brilliant Pebble and to those of recent years on how to progressively improve defenses in each layer. He would also like to thank Dr. Lowell Wood; John Darrah, Chief Scientist Emeritus of the Air Force Space Command; and Ambassador Henry Cooper, former Director of Strategic Defense Initiative Organization, for fruitful discussion of these matters throughout that period and for careful review and thoughtful suggestions on this text.

Technical Developments in Missile Defense

Missile defense is not mathematics, where a proof is reached directly from stated postulates. It is an empirical discipline that has developed through a series of contributions by many individuals and a series of technological advances that built on one another. Thus, a chronological discussion of those developments is a natural way to understand those developments, which leads to a fundamental appreciation of the strengths and weakness of the systems based on them. Such an understanding of earlier systems is essential, as many of their components are key elements of the systems under deployment today.

The need for missile defenses quickly followed the development of offensive missile forces and accelerated with the Cold War and the rapid proliferation of missile sizes and ranges. The first phase of their development used nuclear-tipped interceptors guided by radars. Those interceptors worked satisfactorily, but their radars degraded in the environment their intercepts produced, and their political issues ultimately led to the Anti-Ballistic Missile Treaty of 1972. It sought to limit the deployment of defenses and freeze the development of nuclear systems through a policy of mutual assured destruction (MAD).

The second phase developed lasers and nonnuclear kill (NNK) interceptors, which are essential new ingredients of current systems. The technology for lasers and SBI advanced significantly during the 1980s under the Reagan Administration's SDI. Under President George H. W. Bush, the survivable version of the SBI called the Brilliant Pebble (BP), the midcourse ground-based interceptor (GBI), and its advanced discrimination radar became the basis for Global Protection Against Limited Strikes (GPALS), which arguably could have survived Soviet and Chinese countermeasures and still provided high levels of protection for the U.S., its deployed troops, and its allies. GPALS had the highest performance of any defensive system to date. It was about halfway through engineering development when the Clinton Administration terminated development of BP, decimated that of national missile defense (NMD), and shifted emphasis to theater missile defense (TMD). The Clinton Administration was later forced by growing rogue threats to the United States and its allies to resurrect the ground-based elements of GPALS as the main component of its midcourse NMD program; however, it was strongly constrained by the ABM Treaty, which produced a strong sensitivity to unknown aspects of the threat that could not be predicted with confidence.

President George W. Bush's Administration established a process to formulate its program to achieve its goal, which is the development and deployment of effective defenses in each possible defensive layer for the United States, its deployed forces, and its allies against missiles launches, anywhere on the globe, as soon as possible using existing technology and systems. It is an appropriate but difficult goal. The midcourse defenses developed in the previous administrations are the most developed layer, so the new Missile Defense Agency (MDA) is developing an Initial Defensive Operations (IDO) Capability by integrating the GBI developed over the last two decades with radars and satellites of comparable vintage, which is to be completed to provide protection from Northeast Asia in 2004 and the Middle East in 2005. These defenses are to be upgraded in a spiral program with block modifications to produce maximum effectiveness and efficiency at each point in a two-year upgrade cycle.

Development programs for boost-phase systems were delayed by the Clinton Administration. A program for the development and testing of surface-based interceptors for deployment on land or ships has been formulated and initiated, but is not intended to produce a deployable system for six to eight years, so it does not impact the logic of near-term spiral developments. Such systems appear well suited to the few threats that afford them safe access to boost. A development program for SBI has been formulated, but not initiated, which would produce a system on an even longer time scale, although SBIs are preferred on cost and coverage grounds for large, multiple, or global threats.

Theater defensive systems entering production should provide adequate theater and regional defenses, although achieving their full effectiveness will require the removal of current command, control, and communication stovepipes. They might be integrated to provide protection for large cities from missiles launched from ships close to shore, although they could not adequately address ICBMs attacking cities. The nonnuclear interceptors that could have performed such engagements were terminated by the Clinton Administration and have not been restarted. There is no capability to defend cities from ICBMs in the terminal phase and no program to produce one.

These technologies could make effective defenses possible in each layer on the time scales desired, so there is a potential match between the goals of the current program and the technology available. However, only a portion of those technologies are under active development. The current MDA program is effectively still a midcourse system and is likely to remain so until well into the next decade. Barring fundamental improvement in the ability to discriminate midcourse threats, that system will be effective against a few missiles with a few simple decoys. It is aptly described as "better than nothing," but it offers protection to be used *in extremis*, which could fail catastrophically with sophisticated decoys or countermeasures. Thus, it does not represent a reliable military capability.

The MDA program is comparable in scope to that of previous national programs, but lacks the definite goals, short time horizons, strong leadership, multiple options, and deep research and development (R&D) that characterized successful previous programs. Its goal of producing an operating system in a few years is an improvement over that of the last decade, which had no such commitment, but that alone is not enough. It will deploy an initial system that lacks the robustness to deal with plausible countermeasures, which undermined its predecessors. If they can be remedied by spiral development, the problems in the IDO need not be debilitating, but they cannot be overcome by deploying more of the same or similar interceptors and radars. The broad R&D support included in earlier successful and missile defense programs was lost as resources were spread too thin, committed to current problems, and removed for political reasons. The R&D budget for the current program and the number of options it maintains are so small that previous successful national programs imply that it will run out of options before reaching a significant product.

One can argue how well the MDA program will perform against known countermeasures to which SBIs were shown to be insensitive. They could be developed and deployed on roughly the same time scale as the MDA program. Doing so would provide a capable and affordable boost layer that would reduce the threats reaching midcourse to levels that GBI could address. There was a window in which GPALS arguably could have used these advanced systems to eliminate the utility of offensive missiles. It was lost during the Clinton Administration's emphasis on the domestic economy, TMD, NMD, and the ABM Treaty. The current program could serve as the first step in a continuing spiral that could respond to the progressive improvements in offensive missiles. However, the appropriate long-term goal is to put missile defenses so far ahead of offenses that they will dissuade rogues and others from engaging in missile competitions altogether. That is not beyond the capability of missile defenses, even those attainable in the near term. The tools are now at hand, but not all are being fully developed. A balanced program must develop and use all available tools, including space-based sensors and interceptors.

Early Missile Defense and Offensive Proliferation

During World War II, the Radiation Laboratory at Massachusetts Institute of Technology and the U.S. Army Signal Corps developed the SCR 584 radar and coupled it with a 90 mm gun to produce the first radar-guided gun. Under the Lend-Lease Program, 200 were sent to the United Kingdom, where they were used against V-1 buzz bombs, with an effectiveness of about 95 percent against V-1s within the gun's lethal radius. The Royal Air Force devoted Spitfires and bombers to find and destroy V-1 launchers. When V-2 rockets were launched at the U.K., the SCR 584 was coupled to a Western Electric Bell Laboratories plotting board to infer V-2 trajectories in flight, which gave a few minutes of warning before the bombs hit their estimated impact areas. The radars could also backtrack their predictions to determine the launch areas in Holland and Belgium. That allowed bombers to attack the V-2 launch sites and fuel depots, which forced an interruption in launches until Germany developed a mobile launch system. The first attempt at ballistic missile defense was by a Spitfire, which happened on a V-2 rising out of trees and attacked it with its machine guns, to no effect. General Dwight D. Eisenhower reportedly remarked that if V-2s had attacked the Allied forces building up in U.K. ports, D-Day would have failed. Because of the above interruptions, the first V-2s were fired at those troops after the landings in Europe had succeeded.[1]

Immediately after World War II, the United States and the USSR both used captured German rockets and scientists to pursue the development of offensive ballistic missiles. Their developments were initially predicated on their carrying atomic weapons like those developed during the war, although they were not well matched to that application. Atomic weapons of the 1950s weighed 10s of tons, had yields of a few 10s to 100s of kilotons (KT, i.e., the energy of 1,000 tons of high explosives), and produced destructive overpressures to distances of a few kilometers (see Appendix A). The V-2's inertial guidance had an accuracy of about 15 km at its maximum range of about 370 km, or an angular accuracy of about 15 km/370 km = 40 milliradian (1 mrad = 0.001 radian = 0.06°). Such accuracies were adequate to attack a city the size of London from across the channel, but would produce a circular error probable (CEP) area of about 400 km at intercontinental ranges of 10,000 km, which is much larger than a city. Thus, its expected damage would be reduced by roughly the square of the ratio of its damage radius to its CEP to levels that were not strategically interesting.

A 10 megaton (MT, equal to 1,000 KT) explosion produces the pressures needed to destroy ordinary structures out to about 10 km. Early missiles could not carry such weapons, which weighed 10s of tons, to intercontinental distances. Thus, strategic bombers were developed that could to address the large cities and few hard military targets in each country. ICBMs were already a high priority of the Eisenhower Administration before the launch of Soviet Sputnik in 1957, but that stimulated the deployment of intermediate-range Jupiter and Thor missiles and accelerated the development of the intercontinental Atlas, Titan, Minuteman, and Polaris missiles. The hydrogen bomb ultimately provided a thousandfold increase in yield per unit mass of payload to about 1 MT/ton of payload. Parallel improvements in accuracy to about 0.1 mrad gave CEPs of about 1 km at 10,000 km distances, which were well matched to the MT weapons these larger missiles could carry. This combination of developments made large, accurate ICBMs with hydrogen bombs the dominant strategic system by the 1960s.

These improvements in range, accuracy, and yield contributed to a race in offensive arms, particularly after the Cuban missile crisis had demonstrated the benefits of strategic superiority—or more precisely, the weakness of a position of inferiority. During the crisis, the Soviet Union had 10 operational ICBMs, while the United States had over 100. The Soviet Union thereafter moved quickly to assure that it would not face such an imbalance in subsequent crises. Accuracy improved rapidly in the 1960s and 1970s to levels that could support attacks on military targets, which produced options for counter-

1. John Darrah, Chief Scientist emeritus, U.S. Space Command, private communication, March 6, 2002.

force attacks. The resulting concerns about their potential impact on stability were the backdrop for much of the discussion and analysis of defensive systems in later decades.

During this period most of the strategic community was focused on offensive uses of missiles, but important defensive developments were also initiated. Both the Army and Air Force started large programs, which were duplicative and competitive. The Defense Advanced Research Projects Agency (DARPA) was formed to consolidate control of these programs through its project Defender.[2] In 1958, the radar Ballistic Missile Early Warning System (BMEWS) system was approved, DARPA initiated the space-based Missile Defense Alarm System (MIDAS), and Lockheed Corporation proposed a missile early warning infrared (IR) satellite that became Subsystem G of WS 117L and evolved into the current Defense Support Program (DSP). General Bernard Schriever, who was already in charge of developing ICBMs, gave defensive elements a high priority in the Air Force's Air Research and Development Command,[3] which resulted in the deployment of both a functioning BMEWS and DSP satellite by 1963—despite numerous reviews by the Office of the Secretary of Defense (OSD).[4]

It was recognized that the same combination of launchers, nuclear warheads, and accuracy that threatened cities could be used to intercept ballistic missiles as well. The goal was to merge ICBM, hydrogen bomb, and guidance technologies with the radars developed during World War II for air defense to reduce the vulnerability of cities and military targets that their offensive application had produced. Those defensive goals stimulated research and development in the 1950s and 1960s. It accelerated during the 1970s and 1980s in response to threats from China and the USSR and focused on deployment in the 1990s in response to rogue missile threats. However, defensive applications had to await the development of sensors, rockets, warheads, and guidance to the levels required to detect and intercept an incoming reentry vehicle (RV), i.e., to "hit a bullet with a bullet." Those developments are discussed below, after a short summary of the evolving threats that stimulated them.

Offensive Proliferation

Although the number of nuclear weapons in the USSR and U.S. inventories grew into the 10,000s during the Cold War, their block discipline put some constraints on the proliferation of nuclear weapons, missiles, and accuracy during the period of bipolar confrontation. The number of nuclear powers grew slowly from one to seven and stabilized, and few countries could afford missile technology more advanced than the 300 km range SCUD missiles the Soviet Union developed from V-2 technology. That changed in the late 1980s as the growth in the number of theater-range missiles put deployed forces at risk, which potentially made the United States the "deteree rather than the deterer." It accelerated further after the fall of the Soviet Union. The number of nations with missiles grew from six in 1972 to 16 in 2001. SCUD missiles spread to the Middle East, Asia, and other parts of the world. In 1998 and 1999, Pakistan and India tested their 1,300 km Ghauri and Shahab-3 missiles, India tested its 2,000 km Agni II medium-range ballistic missile (MRBM), and China prepared to deploy its 8,000 km range mobile DF-31, which has the potential to carry multiple warheads. The political picture changed sharply in November 1998, when North Korea tested its Taepo Dong 1, which is capable of delivering a few hundred kilogram payload to intercontinental ranges and developed the Taepo Dong 2, which can carry larger payloads to intercontinental ranges.

A recent National Intelligence Estimate (NIE) on future threats notes that it is "unlikely that Russia or China would sell whole missiles to other countries," but a recent Central Intelligence Agency (CIA)

2. H. Beveridge, "Defender Introduction," Ballistic Missile Defense Program of the Advanced Research Projects Agency, A Review of Project Defender for DDRE, Vol. I, July 1960.
3. J. Richelson, *DSP Satellites and National Security* (Lawrence, Kan.: University Press of Kansas, 1999), p. 11. All comments on DSP and related early warning satellite systems are based on documents released in response to Freedom of Information Act requests cited in the references.
4. *Ibid.*, pp. 12–36.

report on the help China gave to Pakistan, Iran, Libya, and North Korea—and that North Korea gave to countries in the Middle East, South Asia, and North Africa—makes that statement less reassuring.[5] Missile states may be reluctant to sell whole systems, but they are apparently willing to do significant business in the components needed to assemble them. Missile and reentry vehicle technologies are now widely available.

The CIA assessment also states that countermeasures to defensive sensors that are now "readily available technology" include "separating RVs, spin-stabilized RVs, RV reorientation, radar absorbing material, booster fragmentation, jammers, chaff, and simple balloon decoys." Separating the RV permits the attacker to distance it from its booster and orient it to reduce its signature. Spin stabilization reduces signatures, eases reentry, and increases accuracy. Absorbing material reduces radar returns and interceptor range. Fragmentation, chaff sized to the radar frequency, and balloon decoys increase the number of "traffic" decoys a radar must examine, which consumes radar power, time, and defensive battle space. Jammers deny the radar information about RV range. The CIA statement that these technologies are "readily available" means the intelligence community agrees that techniques and systems developed over decades at considerable expense by the United States and USSR are now available at a fraction of that cost to rogues. That led to an assessment by the Deputy Director of the CIA that the likelihood of an attack against the United States or its allies using weapons of mass destruction (WMD, i.e., nuclear, chemical, or biological weapons) "is higher today than it was during most of the Cold War."[6]

It is increasingly difficult to project timelines for development and deployment of new threats. It is harder still to project missile performance, configuration, and countermeasures. The 1995 NIE predicted that the United States was unlikely to face a missile threat in less than 15 years. Because that estimate was criticized as politically driven, an independent commission to examine the threat was formed in 1998, which was led by Donald Rumsfeld, now Secretary of Defense.[7] Noting the intelligence community's lack of success in predicting the rate of development of the North Korean and other global ICBM programs, the rapid rate of proliferation, and the extensive cooperation between proliferators, the Rumsfeld Commission concluded that it would be better to base estimates on opponents' capabilities rather than intentions and to pace the development of defenses accordingly. It estimated that North Korea, Iran, or Iraq could have ICBMs within five years of a decision to acquire them. That estimate was criticized on its release as a worst-case estimate, but was given added credibility by India and Pakistan's tests in 1998 and North Korea's launch of a Taepo Dong 1 over Japan a few months later.

Early Technologies and Systems

Successful competition by defensive missile systems with offensive systems for funding in the 1950s and 1960s made possible the development of the technologies needed to make effective defenses possible. Given the developments discussed above, it was natural for them to be based on ICBM technology, air defense radars, and nuclear weapons. Progress in each was rapid, and useful combinations were found for intercepts in the atmosphere and in space. That led to a series of proposed developments, only one of which was approved for deployment. Advances in computation and guidance have made it possible to replace nuclear weapons with small nonnuclear hit-to-kill (HTK) interceptors, but

5. R. Walpole, National Intelligence Officer for Strategic and Nuclear Programs, "The Ballistic Missile Threat to the United States," statement, Subcommittee on International Security, Proliferation, and Federal Services, Committee on Governmental Affairs, U.S. Senate, February 9, 2000.
6. J. McLaughlin, "Remarks at the 4th Annual Conference by Deputy Director, CIA," Space and Missile Defense Center, Huntsville, Alab., August 21, 2001.
7. D. Rumsfeld, *Report of the Commission to Assess the Ballistic Missile Threat to the United States* (Washington, D.C.: U.S. Department of Defense, 1998).

the early interceptor and radar technologies are still retained. Each stage of development had its successes, whose legacies are imbedded in current systems.

NIKE Systems

In March 1955, U.S. Army Ordnance requested Bell Laboratories and Douglas Aircraft to perform an 18-month study of an advanced anti-aircraft system with some capability against ICBMs. The interceptor they derived from the surface-to-air NIKE AJAX missile was designated the NIKE II. This $1.65 million contract developed most of the concepts and issues that guided anti-ballistic missile defense through its first two decades.[8] It defined the basic interceptors and nuclear warheads, introduced atmospheric drag to filter threats, estimated target decelerations, and derived analytical predictions of the interceptor performance needed to engage them. It studied decoys and countermeasures, developed first-order counters to each, and sized and developed radars that are still in use. It ultimately performed 79 developmental tests and 68 systems tests, of which 57 percent were successful.

Defensive research and development in the 1950s and 1960s largely concentrated on whether a bullet can hit a bullet to within the few kilometer lethal radii of the nuclear explosives those interceptors could carry. In 1959, the NIKE-ZEUS system was proposed to defend the U.S. population. The accuracy required was modest; a miss distance of 1 km at an intercept range of 1,000 km is about 1 mrad, which was comparable to the 10 km/10,000 km = 1 mrad accuracy of the strategic offensive systems of the time. That accuracy was demonstrated in July 1962 when an Atlas launched from Vandenberg Air Force Base on the California coast was "intercepted" by a NIKE interceptor from Kwajalein Island in the Pacific Ocean test range. The NIKE interceptor flew within the roughly 2 km kill radius of the nuclear warhead that it could carry. A test in December 1962 came within a few hundred meters. Over the next two years, 10 out of 14 flyby "intercepts" of mock RVs were successful,[9] which proved that a practical interceptor could reliably hit a bullet—provided that it was armed with a nuclear bullet.

NIKE-ZEUS interceptors were large. A few KT nuclear explosive weighed a few hundred kilograms, so at the roughly 3 percent payload mass fractions of chemical boosters, the interceptors weighed over 10 tons, stood 17 m high, and cost approximately $1 million. While its interceptors were rugged, the mechanically steered radars NIKE-ZEUS inherited from earlier air defense systems could only be hardened to a few pounds-per-square-inch overpressures, so they could not survive direct attacks or even successful nearby intercepts. That produced an apparent Achilles' heel in the proposed system that was pronounced unacceptable by President Eisenhower, who rejected its deployment.[10]

The vulnerability of mechanically steered radars was overcome by the development of phased-array radars, which steer their beams by shifting the phases of many discrete radiating elements rather than by rotating their main face. That made it possible to move the beam very quickly so the radar could maintain track on objects already detected while continuing to scan for others. Phased-array radars were developed for the NIKE-X system, which was proposed in 1966 for the defense of the 52 largest U.S. cities against massive attacks. NIKE-X increased the range of its radars so they could detect and track targets outside the atmosphere, an approach that was carried over into the radar designs for subsequent systems. NIKE-X was also rejected, but it did demonstrate the integration of survivable radars and interceptors, which became important when it was not possible at the 1967 Glassboro summit to dissuade the USSR from its planned strategic offensive and defensive force build-up. President Lyndon Johnson then directed Secretary Robert McNamara to field a missile defense system to

8. ATT Bell Laboratory, "ABM Research and Development at Bell Laboratories, Project History," October 1975.
9. "ABM Chronology," in A. Chayes and J. Wiesner, eds., *ABM: An Evaluation of the Decision to Deploy an Antiballistic Missile System* (New York: Signet, 1969), p. 228.
10. "An Overview," in Chayes and Wiesner, *ABM: An Evaluation of the Decision to Deploy*, pp. 3–62.

defend the U.S. population against a light Soviet attack, which initiated a series of efforts to develop effective population defenses that have continued, with interruptions, to the present.

Sentinel

In 1967, Defense Secretary McNamara announced that the NIKE-X technology was to be used as the basis for a system called Sentinel, intended to provide a thin defense of U.S. population against a possible 1970s Chinese threat. Sentinel was a direct application of NIKE technologies to a light area defense of the whole nation, for which it was well suited.

Layered Defense. Sentinel had two distinct layers. The first was a long-range layer consisting of the Spartan nuclear interceptor supported by the Perimeter Acquisition Radar (PAR), which was a scaled-down version of the more capable radars envisioned by NIKE. The second layer was the short-range Sprint missile supported by the Missile Site Radar (MSR), which started as a modest adjunct to PAR but gradually increased in performance and autonomy. Spartan had an effective range of 600–800 km, which allowed it to intercept targets exoatmospherically (i.e., at altitudes above about 100 km) during the last few minutes of their roughly 30-minute midcourse phase during which they follow simple Keplerian trajectories. Sprint operated in the endoatmospheric terminal phase. That allowed it to take advantage of the atmosphere to filter out decoys, but compressed its battle space into the last few seconds before impact, complicating the kinematics of intercept. Sentinel had no boost-phase layer as no defense was then thought technically feasible in the first few minutes of powered missile flight.

Sentinel used layered defenses because they are necessary for high levels of attrition of large attacks. In such defenses, early layers progressively reduce the threat, which prevents the saturation of later ones. The effectiveness of a single-layer military system is typically 80–90 percent; thus, a defense composed of two independent layers, each of 90 percent effectiveness, would let about $(1 - 0.9)^2 = 1$ percent of the weapons leak through. That would be about one weapon for an attack of 100, which would be serious. However, if one layer failed, leakage would increase to 10 percent, which would be catastrophic. Robust multiple layers are needed for effective population defense. Sentinel had only two layers; so even at 90 percent effectiveness, it would have let about 1 percent of the attacking weapons leak through unless multiple interceptors were committed to a target in each layer.

Each Sentinel layer used radars to detect and track targets and command guide their interceptors. In command guidance the radar tracks both the incoming missile and outgoing interceptor, measures the range and angles to each, computes their separation, calculates the optimal interceptor divert, and transmits it to the interceptor, which then executes the maneuver. This process is repeated until the interceptor is at its closest approach, at which point it is instructed to detonate. Spartan and Sprint radars had accuracies of a few degrees, which produced miss distances commensurate with their kill radii. Their explosions only had to be timed to within the few tenths of a second of closest passage by the RV, which was not stressing. In command guidance the sensors and control remained on the ground, which was conceptually simple, but it introduced errors due to radar resolution, introduced delays, and exposed communication channels that were susceptible to interference and the effects of nuclear explosions.

Radar Search. The three principal radar functions—search, track, and discrimination—scale differently on the key radar parameters, which are power P, aperture area A, wavelength λ, and target radar cross section (RCS) σ. Search range R_{search} scales as $(PA\sigma/B)^{1/4}$ (See Appendix B), so achieving large search ranges requires large power-aperture product PA, even for large targets and narrow bandwidths, i.e., long search times. The power-aperture product required for search is independent of frequency, so search radars generally operate at ultrahigh frequencies (UHF = 0.3–3 GHz) where power is less expensive. However, there is a limit to the scaling advantages of low frequencies. If λ is much larger than the target, its σ is in the Raleigh region, where σ scales as $1/\lambda^4$, so that R_{search} falls as $1/\lambda$. Thus, wavelengths longer than the few meters of RVs and missile tanks are not useful. The PAR

frequency of about 0.7 m was a trade off between the cost of power, RCS, and the reduced sensitivity of higher frequencies to nuclear effects.

PARs were UHF radars with PA products of about 3 MW-m^2 and ranges of approximately 1,200 km against 1 m^2 targets.[11] PARs were two-faced phased arrays deployed primarily along the northern perimeter of the United States to meet minimum time and energy trajectories to the Continental United States (CONUS), for which they could give about 10 minutes warning. Their companion MSRs were not initially designed for search, as each could be collocated with a PAR and take tracks from it. MSR's main issue was survivability; therefore, it was hardened to overcome the NIKE radar's vulnerabilities. However, even strongly reinforced concrete faces could only be hardened to a few 10s of atmospheres, as it was necessary to transmit the radar beams through them, limiting how high an over-pressure they could withstand. MSR's design goal was to keep intercepts a few kilometers away, so that their hardness level would not be exceeded in successful intercepts.

BMEWS radars in Alaska, Greenland, and the United Kingdom radiate about 250 kW from each 25 m diameter face, which produces ranges of about 5,000 km against tanks with 10 m^2 RCS. That is roughly the longest range that is useful for objects on minimum energy trajectories, as radar beams propagate in straight lines while the Earth's surface curves away at long ranges. BMEWS are forward based, so they give about an additional 10 minutes of warning over CONUS–based radars. The newer Position and Velocity Extraction Phased Array Warning System (PAVE PAWS) radars in California and Massachusetts achieve similar ranges by radiating about 150 kW from each 22 m face.[12]

Radar Track. For track it is only necessary for the signal to be larger than the noise in the range-angle cell currently being irradiated. The track range R_{track} varies as $(PA^2\sigma/B\lambda^2)^{1/4}$, which scales more strongly on aperture than power (See Appendix B). Other parameters being equal, R_{track} scales as $R_{search}G^{1/4}$, where $G = (D/\lambda)^2$ is the track radar's antenna gain. To make the search and track radar ranges comparable, the power aperture product of the search radar can be made larger than that of the track radar by a factor of G over the gain of the search radar. That R_{track} scales as $1/\sqrt{\lambda}$ suggests a benefit for operating at shorter wavelengths, but it is limited by the frequency dependence of σ. RVs resemble cone-spheres of small nose radius, which have σ of about $\lambda^2/10$, which makes R_{track} independent of λ; thus, there is little benefit in wavelengths smaller than the RV's nose radius.

A radar of aperture A and effective diameter D of roughly $\sqrt{A}$ produces an angular divergence of about λ/D, which produces a beam diameter of $R\lambda/D$ at range R. The $\lambda = 0.7$ m PAR had D of about 40 m, so it produced an angular beam width of roughly 0.7 m/40 m = 18 mrad = 1°, which produced a physical beam width of about 10 km at 600 km.[13] That is larger than the interceptor's lethal radius, but the roughly tenfold beam division—possible with the 2^4 = 16-fold higher signal-to-noise ratios at half its 1,200 km detection range plus the additional observations during the weapons flight to there—produced roughly 1 km resolution, allowing PAR to support Spartan's roughly 1 km lethal radius.[14] The four-face MSR phased array radar could direct nearby intercepts at any azimuth. Its S-band (0.1 m) radar and 4 m aperture produced a 0.1 m/4 m = 25 mrad beam width, which with sixfold beam division gave a roughly 100 m diameter beam at 20 km commensurate with the lethal radius of its neutron warhead.[15]

BMEWS and PAVE PAWS have some capability to track successive hits in adjacent range and angle bins in their radar fences, but it is compromised by their limited resolution and computers. Their 25 and 20 m array diameters give beam widths of about 2°; thus, at the 5,000 km maximum range, their

11. J. Toomay, *Radar Principles for the Non-Specialist* (New York: Van Nostrand Reinhold, 1989), Chapter 1.
12. A. Sessler *et al.*, *Countermeasures* (Boston: Union of Concerned Scientists, 2000), p. 140.
13. "ABM Research and Development at Bell Laboratories," p. 8-7.
14. Toomay, *Radar Principles for the Non-Specialist*, p. 17.
15. "ABM Research and Development at Bell Laboratories," p. 7-5.

beams have cross-range resolution of $R\lambda/\sqrt{A}$ or about 150 km. That would determine an object's azimuth to within about 1.6°, which was not thought to be adequate to do more than cue the PAR to an azimuthal error basket of about 150 km/1,200 km = 7°. Radars scale differently for search and track, so in the first few decades after World War II, the two functions were usually performed with separate radars. Phased array radars can track objects while continuing to scan for others, but must do so at the same frequency. Thus, for them combining search and track requires a tradeoff between the two missions, which leads to a compromise on the wavelength that maximizes joint performance in search and track for a given cost or minimizes cost for a given level of performance.

Discrimination. PAR developed improved power tubes, supported research and development on solid-state technology, and tested improved discrimination waveforms and algorithms that are still in use. Search and track ranges are maximized by narrow bandwidths, but target range resolution scales as $c/2B$, which favors wideband operation. BMEWS radars have bandwidths of about 600 kHz in search and 10 MHz in track. The resulting range resolutions of about 250 and 15 m only provide information on the largest objects. PAVE PAWS radars have about 100 kHz in search and 1 MHz in track, so their range resolutions are 1,500 and 150 m, which provide neither useful metrics nor discrimination. However, both radars are being retrofitted into Upgraded Early Warning Radars (UEWRs) with modern digital exciters and computers. That should give them roughly 30 MHz bandwidths with resolutions of about 5 m in both search and track, which should separate large objects such as tanks from RVs and decoys. Their phase and temporal stability should also be improved by the upgrade, which could provide some discrimination capability as well.

The more recently developed x-band radars have bandwidths greater than 1 GHz and hence resolutions of about 10 cm, which are adequate to inspect RVs and decoys in enough detail for discrimination. Such resolution is particularly effective when combined with Doppler measurements (which image rotating or tumbling objects) and phase-derived range (which can measure features at the limit of range resolution). This combination of phase stability, amplitude, and computational power can exploit virtually every feature on which modern signal processing and discrimination are based.

Interceptors. Physics drove the design of the interceptors for both layers of Safeguard. Below altitudes of about 50 km, x-rays are absorbed strongly by the atmosphere, so most energy is deposited locally in a compact fireball, whose diameter scales as the cube root of the ratio of yield to air density. Thus, a 1 MT explosion produces a roughly 1 km fireball at sea level; a 100 KT explosion produces a fireball of that size at 15 km; and a 10 KT explosion produces one that size at 30 km. Fireballs rise, entrain air, and cool, but remain highly absorptive for 10s to 100s of seconds, obscuring significant portions of the solid angles radars must search. Neutrons can penetrate freely out to distances about 100 m in the dense atmosphere before scattering strongly, so KT-range weapons with enhanced neutron output can fill such volumes with lethal fluences while producing fireballs an order of magnitude smaller than those from x-ray weapons at the same altitudes. For that reason there is an advantage to using Sprints with KT-range neutron weapons in the 10–30 km intercept regime to avoid cluttering up the PAR battle space and using MT Spartan x-ray weapons at higher altitudes, which was the approach used by Sentinel, its successor Safeguard, and other nuclear concepts.[16]

Above about 100 km altitude most of the energy from a nuclear explosive is released as x-rays. The cross sections for absorption by air molecules are small, so the x-rays can escape and propagate to 100s of kilometers. The x-ray fluence at range R from a nuclear explosion is $F = Y/4\pi R^2$. For a yield Y of 5 MT, the fluence at 4 km would be about 10 KJ/cm^2, which would produce enough impulse to break structures such as RV heat shields. Such x-ray deposition is difficult to countermeasure and relatively insensitive to details of RV construction. Spartan used combinations of yield and accuracies in that range for its exoatmospheric engagements. It engaged at ranges up to 600 km after being

16. H. Bethe, "Countermeasures to ABM Systems," in Chayes and Wiesner, *ABM: An Evaluation of the Decision to Deploy*, pp. 130–143.

launched on PAR warning when the RV was at about 1,200 km range. Objects on minimum energy trajectories from Soviet or Chinese launch areas reenter at angles of 22.4°, so most Spartan engagements took place at altitudes of about 600 km x sin(22.4) = 230 km. Spartan reached them at near intercontinental speed using 5–6 g accelerations from three-stage missiles about 16 m in length, 1 m in diameter, and costing about $1.5 million.[17]

Sprint engaged at lower altitudes on MSR tracks or those handed over from PAR to MSR. It depended on atmospheric drag to discriminate decoys, which meant it had to wait until the objects were below about 100 km to launch. Thus, the top of its intercept altitude regime was about 30 km, although most of its intercepts were at 5–15 km. It could intercept at those altitudes and still protect targets below because they are shielded by the Earth's atmosphere. Large tests at high altitudes have shown that such explosions primarily eject mass into space rather than increasing damage on the surface below (See Appendix A). Sprint had to reach its intercept altitudes during the RV's roughly 40 s flight through the atmosphere, which it did using rockets with average accelerations of about 50 g. It was a two-staged cone 9 m long and 1 m in diameter shaped to survive hypersonic flight through the dense atmosphere, which cost about $1.1 million each. About 100 tests demonstrated its ability to intercept in its low-altitude design regime.

Siting and Cost. Sentinel was to be deployed in 15 sites by major cities in the continental United States, Hawaii, and Alaska. Each was to have ABM radars, although some Minutemen wings were only defended by Sprint. A Sentinel site consisted of a PAR and MSR, 40 Spartans, and 10 to 75 Sprints. A two-face PAR had an estimated cost of $160 million, and the MSR had a cost of $165 million. Sentinel had an expected total cost of about $5.5 billion, which established the pattern observed in later systems that the costs of interceptors to sensors to system usually fall in the rough ratios of 1:100:10,000, where the units are millions in current dollars. Sentinel's cost was on the order of the $10 billion canonical cost of offensive strategic systems of that time. Opponents estimated that its actual costs could be roughly double these government estimates, although uncertainties in cost were smaller and less controversial than those in performance.

Criticisms and Observations. Sentinel underwent significant analysis and criticism by technical and public interest groups governments. They were well documented and still provide a useful analysis of the weaknesses of systems based on its technologies, some of which are still in use.[18] In addition to the technical issues in reliability, computers, software, countermeasures, decoys, and blackout discussed in the section below on Safeguard, the criticisms introduced several more general issues such as cost effectiveness, adversary reaction, and escalation. The four generic missions Sentinel might be tasked to attempt were:

1. Population defenses against heavy attacks,
2. Protection of population against light attack,
3. Defending the deterrent forces, and
4. Protection against accidental launch.

The principal criticisms for Sentinel in these missions were, respectively:

1. Population defense against heavy attacks was too difficult because decoys and blackout could negate Spartan. Then Sprint could be avoided or exhausted. However, the attempt to develop such defenses would provoke escalation by the Soviet Union or China.

17. ATT Bell Laboratory, "ABM Research and Development at Bell Laboratories," p. 2-9.
18. "Summary," in Chayes and Wiesner, *ABM: An Evaluation of the Decision to Deploy.*

2. Protection of population against light attack was the mission for which Sentinel was designed, but was said to be impractical because countermeasures and blackout could negate Spartan and Sprint could be avoided by attacking unprotected cities.
3. Defending the deterrent was possible, but Sentinel was the wrong system. Spartan could be negated; thus, more numerous, cheaper Sprints and MSRs would be more effective—although more expensive than additional offenses or launch on warning, which was the critics' preferred option.
4. Protection against accidental launch was dismissed because such attacks could have enough missiles, decoys, and countermeasures to defeat both layers.

The common thread through the analyses was the assumption that Spartan would not work, which would leave Sprint vulnerable to saturation or avoidance. Under that assumption, the above criticisms can be reorganized sequentially as:

1′. PAR cannot discriminate in nuclear environments;

2′. Therefore, Spartan will collapse;

3′. Therefore, Sentinel will lose cost effectiveness, and

4′. Trying to build it would provoke an arms competition.

Critics were most concerned about Sentinel's arms control implications in the last step, which depended on a sequence of assumptions about the preceding arguments. As significant additional progress has been made on each in subsequent decades it is appropriate to review them.

These criticisms of Sentinel skip over a central point. While PAR and Spartan might have failed under the large attacks implied by missions 1 and 3, it is not clear that they would under the limited attacks of mission 2, for which it was designed. Current higher frequency radars would be even less likely to fail, as would IR satellites and the IR sensors on current midcourse interceptors. PAR might have been able to maintain adequate surveillance and track under a limited number of precursor or intercept explosions per site, so 1′ does not follow for the mission for which the PAR was designed. In that case, Spartan should have adequate support, so 2′ would not obtain. With PAR and Spartan effective, the upper layer should provide enough attrition to keep MSR and Sprint from being saturated or bypassed, so Sentinel should retain technical effectiveness, so assumption 3′ should not obtain. Then, for the costs and numbers cited above, Sentinel should be cost effective in defending cities, so it should not stimulate escalation, and statement 4′ should not obtain either. Thus, for the mission for which it was designed, Sentinel should have met all of the critics' objections.

These points were not disputed in the debates. Instead, critics shifted discussion from mission 2, protection against light attacks, to the other missions for which Sentinel was not intended and to threats for which it was not designed and to which its technology could not readily adapt. Such mission and threat escalation was also used in subsequent debates to undercut proposed defenses. Sentinel tests demonstrated the feasibility of coordinated two-layer systems and convinced the strategic community that such systems could address modest threats to population, but it failed to prevail in scientific and political debates, which highlighted the sensitivity of radar-based systems to nuclear effects in large attacks and the possible damage to those living near proposed interceptor bases. The former led to criticism in technical societies; the latter to strong public resistance, which forced a fundamental reassessment of the Sentinel system.

Safeguard

In 1969, President Richard Nixon and Secretary Melvin Laird renamed the Sentinel system as Safeguard, shifted its sites away from cities, and changed its mission to protection of the Minuteman

deterrent, the President and Secretary of Defense, and bomber escape routes. Defense against an attack from China was a second priority.[19] In terms of the possible missions discussed above, this amounted to shifting emphasis from the protection of population against light attacks (mission 2) to defending the deterrent (mission 3) while using the same hardware and two-layer system that had already been shown to be inappropriate for that mission against large attacks. Safeguard actually provided roughly the same coverage for cities that Sentinel had; it just moved radars and interceptors away from the cities toward military facilities. Spartan's long-range minimized the impact of that shift and retained the option to restore the defense of population later. Safeguard gave all sites Spartan and Sprint interceptors and radars, which increased costs to $6.6 billion for 12 sites or $7.2 billion to add Alaska and Hawaii.

Since Safeguard retained the key elements of Sentinel, its key technical issues remained as above. Whether Safeguard's exoatmospheric elements would provide the needed level of attrition against the large, structured attacks needed to negate the deterrent became a more important issue, which increased the importance of a careful assessment of technical feasibility—particularly in light of additional insights into the difficulties of operating in the nuclear environments from large nuclear explosions. Progress in the analysis of penetration aids and blackout was largely offset by additional uncertainties discovered in coupling, heave, and refraction, which complicate the environments with which the radars and interceptors would have to contend in a system with numerous high-altitude nuclear explosions.

Penetration Aids. The Safeguard PAR radars were constrained by countermeasures and blackout. Moreover, PAR had resolution of 10s of meters in range and several degrees in angle, so it could not see single objects or distinguish between several objects in its roughly 100 km^3 resolution cell. Since PAR could not reliably discriminate whether objects were weapons or decoys, Spartan's limited inventory could be exhausted by modest numbers of credible decoys. That reinvigorated discussions of decoys that had begun during the NIKE and Sentinel programs.

Traffic Decoys. Traffic decoys are countermeasures that greatly increase the number of plausible objects in order to overload the radars' ability and exhaust the time available to discriminate them. Safeguard had to deal with an increasingly sophisticated set of such penetration aids, some old, some new. The oldest was chaff. During World War II, the Allies discovered that long strips of metal cut to about half the radar's wavelength produced a dipole that gave large return signals that could saturate low frequency radars. Each piece of chaff could produce a return as large as an RV, and millions of pieces could be carried in exchange for a few 10s of kilograms of payload. In the atmosphere, dispensing uniform, dense chaff clouds was difficult, as air drag rapidly separated the chaff from the plane. In space, drag was absent although it was still necessary to dispense the chaff efficiently. During Safeguard, critics assumed that it could be dispersed efficiently, although that was only demonstrated a few decades later.[20] There was a similar delay between the assumption and demonstration of the feasibility of fragmenting boosters, i.e., wrapping explosives around their upper stages and detonating them after burnout to produce large clouds of fragments that could look like RVs to UHF radars.

Balloons. Balloons had been suggested during NIKE, but took on a new aspect when it was suggested that light balloons be covered with metal foil to make them resemble RVs more closely—or to conceal RVs within them.[21] While reflecting balloons were only a modest improvement in decoys, they represented an important change in strategic thought. Previous RV modification—from heat-sunk hemispheres to small radius ablators and pointed cones—followed a path of reducing the cross section presented to the radar to which they were oriented. It seemed natural to offensive defensive systems

19. "An Overview," in Chayes and Wiesner, *ABM: An Evaluation of the Decision to Deploy,* pp. 3–62.
20. Sessler *et al.*, *Countermeasures*.
21. R. Garwin and H. Bethe, "Anti-Ballistic Missile Systems," *Scientific American*, March 1968.

designers to make cross sections as small as possible in order to maintain the RV's stealth as long as possible to give the defense as little time as possible to react. Introducing balloons with large cross sections ran counter to that philosophy. The U.S. instituted an Advanced Ballistic Re-Entry Systems (ABRES) program to test such ideas, but balloons were not needed to negate the UHF radars then deployed, so few were actually used.

Discrimination and Elimination. The postulation of more innovative decoys also stimulated thought on how to discriminate and eliminate them. The first PAR built in South Dakota was equipped with the flexibility, bandwidth, waveforms, and computational ability needed to test most discrimination algorithms developed up to that time. However, with the limited temporal and phase coherence available from it, only the simplest metric discriminants could be implemented, and their effectiveness was limited by PAR's UHF measurements, which were essentially of the object's volume.

Traffic decoys like balloons are more effective against low frequency radars at long ranges. To discriminate them on the basis of atmospheric drag, the change in the object's velocity ΔV must be large enough to be observable, which occurs at higher altitudes for lighter objects, i.e., those of lower mass per unit area β. A radar with wavelength λ producing a T second burst of pulses has velocity resolution of $\lambda/2T$, which for PAR's $\lambda = 0.7$ m gives a ΔV of about 3.5 m/s (See Appendix C). Figure C.1 shows the slowing experienced by objects with β of 1, 10, and 100 kg/m^2, which indicates that PAR should discriminate $\beta = 1$ kg/m^2 balloons at high altitudes, 10 kg/m^2 light decoys at about 180 km, and 100 kg/m^2 heavy decoys at about 130 km. However, 10 to 100 decoys remained credible down too far for discrimination at 130–180 km to support Spartan's 150–250 km engagement range. The S-band MSR had a ΔV of 1 m/s, which with adequate sensitivity would increase its heavy decoy discrimination altitude to about 170 km. MSR was intended to support endoatmospheric intercepts, which would have made traffic decoys and balloons ineffective, but it was not designed to search at that altitude. X-band radars with ΔV of about 0.2 m/s shift discrimination altitudes even higher.

Active Discrimination. Given the difficulty of discriminating decoys, attention turned to means of eliminating them. In the latter stages of Safeguard, consideration was given to using Spartan's x-rays as an active discriminator to clear decoys as well as kill RVs. The basic idea derived from Spartan's x-ray kill mechanism. When x-rays are absorbed in a material—particularly an ablator such as a heat shield—material is heated, vaporized, and blown off at a velocity determined by the energy deposition per unit mass. That produces an impulse on the order of the absorbed energy fluence divided by its heat of vaporization. The recoil produced by the blow-off accelerates the object away from the explosion. If the deflection is large enough to be measured, the RVs could be identified by measurement of its differential displacement and attacked by a second interceptor.

At that time, ground-based radars did not have the requisite sensitivity, and fly-along sensors did not appear possible, so the concept was not pursued. It has resurfaced each time more efficient, nonnuclear means of irradiation (e.g., lasers or particle beams) or improved sensors have become available. Unfortunately, with each revisit, the offense and defense have both made progress. Given detailed knowledge of the proposed acceleration mechanism, it has always been possible for a surrogate attacker to develop some variant of the decoys that can reduce the separation of their response from that of the RVs below the threshold of measurement or detection. Active discrimination was generally susceptible to first order countermeasures possible at levels of technology comparable to that of proposed discriminants.

Trajectory Determination. A passive means for the defense of military targets can be provided by the precise determination of the trajectories of all threat objects. To threaten a military target, an object must be on a plausible trajectory toward it. If it is possible to measure an object's trajectory with sufficient precision to determine where it is headed, an object not headed toward a small, well-localized military target can be ignored, unless it is headed by chance toward a population center. As sensors improve, it should be possible to make the needed measurements with the requisite accuracy

from ground or space, so that the power of trajectory determination should increase. However, this approach is restricted to military targets. If the goal is to protect all population, i.e., territory, the knowledge that an object is not headed toward a military target does not relieve the defense of the need to engage it.

Electronic Countermeasures. Another countermeasure related to but distinct from decoys is fly-along electric countermeasures, or jammers, which take the violation of RCS stealth to yet another level. Rather than minimizing their return to maintain concealment, they actively broadcast their position in much or all of the bandwidth used by the radar. A jammer returns more power than the target for ranges greater than a crossover range that scales on the ratio of radar to jammer power and gain. For long ranges, jammers overwhelm the radar, and the target's position is concealed. UHF search radars have gain of about 200; therefore, for equal radar and jammer bandwidths and a radar processing gain of about 100, the jammer would dominate at ranges greater than a few 10s of kilometers. Thus, main-lobe jamming can be very effective against narrow band UHF search radars, which have broad main lobes susceptible to modest jammers. An x-band radar might have a gain of about 10^5, so main lobe jammers would dominate for ranges greater than about 300 km. Their side lobes are suppressed by factors of about 100, which could increase the burn through range to about 3,000 km, which would require precise jammer orientation. Such jammers were not observed during Safeguard, but are now widely available according to recent NIEs.

Radar Blackout was a serious problem for Safeguard, particularly for its PAR radars, whose UHF frequencies were strongly absorbed by ambient ionization, fireballs from intercepts or salvage fusing, and remote regions. Ionization due to auroral effects is strong enough to cause radar degradation in certain seasons. Fireballs are generally ionization regions centered on the explosion. Remote regions involve beta rays (electrons) and fission fragments from explosions at higher altitudes that deposit at 50–60 km, where they produce enough ionization and absorption to affect radar and communication systems. Although they can deposit at lower altitudes than the explosions that produce them, they have similar system impacts, so Appendix D treats them together.

Low-altitude nuclear bursts in the Sprint engagement altitude and yield region produce fireballs a few kilometers in diameter that quickly achieve pressure balance, radiate to temperatures of a fraction of an electron volt. Their initial absorption at radar wavelengths is very strong and is maintained for several minutes. They are essentially black to UHF and lower frequency radars throughout structured attacks lasting a few minutes. However, such fireballs need not completely block radar operation. The fireball from a MT burst at sea level is about 1 km across, as is that from a Sprint-sized KT range explosion at 45 km, which would exclude a solid angle of about $(1\text{ km}/45\text{ km})^2 = 0.001$ sr. Unless there were dozens of bursts in the radar's field of regard, performance should not be severely degraded. However, a MT explosion at that altitude would produce a fireball initially about 10 km across, which would block about $(10/45)^2 = 0.05$ sr. A dozen large explosions could block the radars for endoatmospheric intercepts and reduce the flexibility of those for exoatmospheric intercepts.

The uncertain coupling of energy into the low density ambient air at high altitudes by exoatmospheric explosions produces 10- to 100-fold uncertainties in predictions of the size of the regions affected by blackout and refraction from high-altitude explosions. Reducing these uncertainties would be difficult because of the lack of data. The U.S. detonated seven devices in the 10 to 250 km altitude region to be used for Safeguard defenses, but only two exoatmospheric nuclear tests relevant to Spartan. Neither tested the key coupling issues in those altitudes or the multi-burst phenomenology that would cause the greatest degradation and uncertainty in expected scenarios. Measurements were made of radar and communication degradations at various frequencies and ranges from the burst, but not of fireball inte-

rior ionization and absorption.[22] The tests were recorded photographically with films only sensitive in the visible; thus, there is little basis for IR background predictions.

Megaton explosions at altitudes of 150–250 km create hot, ionized fireballs 100s of kilometers across. Most ambient air molecules are stripped of some or most of their electrons, producing initial electron densities n_o of about 10^9 to 10^{12}/cc. At early times, the fireball would form a reflective region of a solid angle of about $(300 \text{ km}/600 \text{ km})^2 = 0.3$ sr. Placed in front of a PAR, an excluded angle that large would mask the trajectories of subsequent RVs the PARs would need to detect and track in 10s of seconds. Such obscurations would be unacceptable against attackers spaced at short intervals. As affordable basing allowed little overlap in coverage between adjacent PARs, these obscurations could not be overcome by internetting PAR measurements. After a few 10s of seconds, the fireballs' temperature should cool by radiation to temperatures of a few thousand degrees. As the fireball cools, the electron density falls. After that, the principal mechanism for the removal of ionization is radiative recombination, which is quadratic in electron density with rate coefficient $C_R = 10^{-12}$cc/s.[23] Recombination causes the electron density n_e to fall as $1/C_R t$. After a time of about 300 s, the electron density drops below the critical density n_c of about 3 x 10^9/cc that would cause complete reflection at PAR's UHF frequency.

Even later, when the fireball is no longer reflecting, absorption losses could still be unacceptable. When electron-ion interactions are the dominant source of collisions, the absorption coefficient α (db/km) is approximately $0.1(n_e/f)^2$.[24] Figure D.1 shows absorption as a function of time after a high-altitude explosion for frequencies of 0.5, 2, and 10 GHz. At the PAR frequency of 0.5 GHZ, absorption is over 1,000 db at short times. By 200 s, it drops to about 10 db/km, which would produce losses in propagating through a 100 km thick fireball of about 100 km x 10 db/km or 1,000 db, which is quite opaque. The losses drop to about 0.4 db/km by 1,000 s, but even that would give a one-way loss of 40 db, or 10^4, which is unacceptable. Thus, PAR would not recover during an attack executed over 10 minutes.

The situation was more favorable at the roughly threefold higher frequency of MSR, which would stop reflecting in 30 s and drop to 1 db/km after about 100 s. However, MSR was sized to take tracks from PAR rather than search for itself, so it lacked the sensitivity and range to take advantage of its reduced absorption. X-band radars developed subsequently have critical frequencies 20-fold higher than UHF. Their critical electron densities of 1.2 x 10^{12}/cc would only be reached only near explosions at 150 km, so x-band radars probably would not be reflected, and their absorption losses would drop below 1 db/km after about 20 s, 0.1 db after 100 s, and 0.01 after 400 s. However, x-band radars were not available during Sentinel and Safeguard, and even those available today are better suited to tracking than to searching large volumes. Potential nuclear environments were complicated by the range of options open to the attacker, who could use precursor bursts to straddle and reduce the PARs' effective viewing angle, and thereby reduce the value of its tracks to downstream radars and interceptors.[25]

22. Bethe, "Countermeasures to ABM Systems," pp. 130–143.
23. C. Blank, *A Pocket Manual of the Physical and Chemical Characteristics of the Earth's Atmosphere* (Washington D.C.: Defense Nuclear Agency, 1974), p. 147.
24. *Ibid.*, p. 247.
25. Bethe, "Countermeasures to ABM Systems."

Uncertainties

The two other major uncertainties in estimates of nuclear phenomenology are coupling and atmospheric heave. Coupling was known but not understood during NIKE. Heave was discovered a decade after NIKE. It was known during Safeguard, but its full impact was not fully assessed during the development of that system.

Coupling

Coupling is a term used to describe the uncertainty in dimensions of high-altitude fireballs due to the lack of detailed understanding of the mechanisms for sharing of a nuclear explosion's energy with the ambient atmosphere. At low altitudes most of the explosion's energy is contained locally and contributes to a fireball. Ragged but recognizable fireballs persist to about 100 km. At higher altitudes, x-ray mean free paths are so long that they can escape the fireball, which weakens its shock and alters the scaling of size on yield from that seen at lower altitudes, but their sizes and losses can still be estimated, since the weapon outputs and cross sections are known. At higher altitudes, the mean free path for collisions is larger than the scale height so conventional hydrodynamics is invalid. It is replaced by collisionless phenomena, whose understanding was not fundamental then and is not now.

If the expanding fireball picked up only the x-ray ionized air near the point of explosion, that would slow it only slightly, and the debris would continue to expand until it reached pressure balance with the Earth's magnetic field at dimensions on the order of 1,000 km. If instead the shock picked up all of the air it overran, a fireball at 150 km would be contained by a sphere of equal mass within a radius of less than 100 km. This order of magnitude uncertainty in fireball size produced 100-fold uncertainties in the excluded regions in the few hundred kilometers of altitude in the Spartan intercept region, with significant implication on its systems.

Expected radii depend on instabilities that might be generated by the interactions between expanding debris ions and the air ions that they swept over. It was predicted that instabilities could produce non-collisional interactions that would cause the expanding debris to pick up most of the air, which would lead to smaller, though brighter, fireballs. Plasma theory indicated that the interaction depended sensitively on debris ion velocity. Unfortunately, the test executed in this altitude region used a high mass weapon whose debris velocity was about a factor of three less than that expected from Spartan, which made extrapolation from the test data to operational conditions uncertain. As that uncertainty was not resolved by theory and there is no new data, it remains a fundamental uncertainty with which any defensive system that uses radars will have to cope.[26] It is mitigated in current systems, which are to intercept earlier in midcourse at higher altitudes, where uncertainties in coupling should have less impact. At intercontinental missile apogee altitudes, the air density is so low that fireball expansion is restrained only by the Earth's geomagnetic field, independent of the details of coupling. The resulting disturbed regions should be somewhat more predictable—although probably much larger.

Atmospheric Heave

The second uncertainty was discovered theoretically during the Safeguard debates when the first serious hydromagnetic calculations of multi-burst phenomena were performed. It was expected that the effects of the first few or few 10s of bursts in attacks would average out and simplify the background. Instead, they interacted in a nonlinear manner through a phenomenon called heave. X-rays that escape high-altitude explosions deposit at altitudes of about 100 km. This deposition increases the energy density and temperature there by roughly 10-fold. The air adjusts to the resulting 10-fold higher scale height by rising, or "heaving," upwards at speeds corresponding to its increased energy due to this energy deposition on a rough time scale of the ratio of the old scale height to the new sound speed,

26. Gregory H. Canavan, "Status of DNA High Altitude Phenomenology," invited presentation, *Proceedings of the 1971 Defense Nuclear Agency High Altitude Nuclear Effects Symposium* (Washington, D.C.: Defense Nuclear Agency, 1971).

which is roughly 10 km/1 km/s = 10 seconds. Thus, attacks occurring over longer time scales deposit their energy in a significantly altered atmosphere.

Heave lofts air from about 100 km to much higher altitudes, creating a new, denser atmosphere with an exponential scale height of about 100 km. Subsequent bursts take place in this higher density air, which is heaved to still higher altitudes. Earlier bursts prepare the environment for later ones in a non-linear cascade that depends on the detailed sequence of bursts, timing, yields, and spectra, which produces a complex space-time distribution of density and ionization in which radars and other sensors would have to operate. This sensitivity to details of scenarios made it difficult to predict the evolution of the environments in real time. It also complicated attempts to perform adequate analyses of expected systems performance.

Refraction

At electron densities below those required for reflection or strong absorption, disturbed environments can cause degradation through refraction, which is the bending of radar beams by variations in the index of refraction n, caused by ionization irregularities (See Appendix D). If a radar beam passes through an ionization region of varying thickness so that parts of the beam experience a different path length L, that produces a phase difference $(n - 1)L/\lambda$ that deflects the beam through an angle of Ldn_e/w, which falls as $1/wt$ rather than the $1/(wt)^2$ of absorption. Thus, even after 10s of minutes, UHF beams could be deflected through 10s of radians, which would be debilitating. At x-band, the bending would still be about 6° after 10 minutes. These deflections are effectively random and could persist to long times.

Figure D.1 shows the amount of refraction expected at frequencies of 0.5, 2, and 10 GHz as functions of time after a high-altitude burst. For 0.5 GHz the initial refraction is about 2×10^4 rad/km. It drops to about 200 rad/km after 1,000 s. Late-time refractions fall to about 50 and 10 rad/km for s- and x-band, but are still significant. PAR and UEWR beams are about 10 km across at 200–300 km altitude, so an electron density contour 1 km thicker for every 100 km horizontally would deflect the beam through an angle of about 20 rad, which would negate its measurements. At S-band, the late time deflection would be about 90°. At x-band, the late time beam deflection is about 10 rad/km x 0.01 km = 0.1 rad = 6°, which would still be a serious problem, though it would decrease with time.

A few-hundred-kilometer bubble like a nuclear fireball could act as a diverging lens of roughly that focal length. Intersecting shocks could act as converging lenses. The overall distortions from a random distribution of converging and diverging areas would act as a random phase screen, whose net effect would be to reduce possible resolution and introduce random directions in the apparent headings of targets viewed through it. Analysis of the loss of phase coherence due to such screens is well developed, but the lack of data on large scale ionization irregularities that cause gross distortions of the beam, let alone those on smaller scales that defocus it, makes its application qualitative. Uncertainty is complicated by heave, which lofts air from about 100 km to much higher altitudes, carrying pre-existing distortions with it and introducing additional ones. That produces sources of random distortion at all altitudes, including those where no explosions have taken place. Extending distortions to a thick region requires wave propagation rather than geometric ray tracing to predict cumulative effects. Such calculations were not feasible in non-real time analyses with the computers available during Safeguard and would stress today's computers.

Using multiple-frequency radars might make it possible to use a rough form of tomography to measure the three-dimensional distribution of distortions, which could provide the reciprocal phase distortions needed to support the radar's compensation of ionospheric and nuclear distortions. That would be analogous to using multiple L-band frequencies to correct for real time ionospheric distortions to improve Global Positioning System (GPS) accuracy. However, GPS only has to correct for electron density variations that are primarily horizontal. Correcting for nuclear effects would require the infer-

ence of three-dimensional ionization distributions, which could exceed the information that could be measured. It did not appear possible with the widely spaced frequencies of the L- and S-band PAR and MSR radars, but might be possible with the higher frequency, less distorted, x-band radars available today.

Weapon Detonation and Salvage Fusing

An important technical issue in estimating the strength of disturbed nuclear environments was whether an attacker would "salvage fuse" warheads that were intercepted to try to recover some value from them by complicating the defender's environment. That decision was complicated by the realization that weapons and heat shields could be harder than initially estimated, so that even nearby detonations might not require salvaging. Estimates of Soviet and Chinese yield-to-weight ratios indicated they might have more margin than U.S. designs; thus, it might be difficult to disrupt their warheads. That made salvage fusing less compelling for the attacker, as incorporating x-ray, neutron, electromagnetic, pressure, and other sensors on all weapons could introduce backdoor vulnerabilities, which might be exploited by the defense to defeat whole attack, if they were disclosed. Overall, the gains for salvage fusing did not clearly justify these risks. Over time salvage fusing became a favorite topic for defense analysts and scientists, but its impact on actual military design was less clear, as there appeared to be more leverage in a limited number of carefully placed, intentional precursor bursts.

Infrared

At the time of the high-altitude tests, films and cameras were not available with enough spectral range to cover much more than the visible portion of the spectrum. Thus, current IR sensors are designed to operate in the short to long wavelength IR (SWIR to LWIR), where there is little experimental basis to calibrate their expected performance in the nuclear backgrounds they will face. That should restrain advocates of sophisticated IR systems, but has not been the case, as nuclear effects have become less of a concern to analysts who were not involved in their measurement or interpretation.

Electromagnetic Pulse

A similar situation obtains in estimates of the strength and effects of the electromagnetic pulse (EMP) produced by nuclear explosions. Fermi suggested its existence during World War II, but its magnitude was not estimated until a decade later.[27] The calculation of EMP field strengths from high-altitude bursts is theoretically straightforward.[28] A nuclear explosion at an altitude above 50 km produces a pulse of gamma radiation a few nanoseconds in duration that carries about 1 percent of the weapon's energy. The gammas deposit at about 30 km, producing a forward-directed current of relativistic Compton electrons. They are turned by the Earth's geomagnetic field, producing a transverse current that amplifies an electric field that propagates along with them. The gamma source keeps the Compton current in phase with the electric field for a distance of about an atmospheric scale height, which amplifies the field to a strength of about 30 kV/m. The gamma and electron sources last 10s of nanoseconds. That gives the EMP pulse significant energy content up to frequencies above 100 MHz, which couple strongly into strategic radar and communication systems. Those field strengths could permanently damage the hard tube components of the time, let alone the softer integrated circuitry of today.

EMP was only observed in one high-altitude test, and even then only by an instrument whose signal saturated. There is anecdotal evidence about lights burning out at the time of the test, but laboratory experiments were unconvincing, and field tests on coupling to strategic missiles, aircraft, switches,

27. R. Garwin, "Determination of Alpha by Electro-Magnetic Means," Los Alamos Scientific Laboratory, Report LAMS-1871, September 1954, unpublished.
28. C. Longmire, "On the Electromagnetic Pulse Produced by Nuclear Explosions," invited paper, in *IEEE Transactions on Antennas and Propagation*, Vol. AP-26, No. 1 (January 1978).

and satellites were few, expensive, sparingly analyzed, and classified,[29] which led some to question the existence, size, and significance of EMP. However, detailed evaluations of each challenge generally led to the discovery of new theoretical effects that would make predicted fields even larger in amplitude and richer in frequency content. Current residual uncertainties would be difficult to resolve short of high-altitude atmospheric nuclear tests, which could have significant untoward effects. Tests by other countries do not appear to have produced better understanding.

Operating in predicted EMP environments could require expensive hardening or possibly redesign of strategic systems, as hardening existing components and boxes has generally been ineffective in achieving and maintaining isolation against electromagnetic interference. It is unclear how to achieve isolation in strategic systems such as radars, which must operate with their antennas open to the environment for effectiveness, or communication systems, which must be connected to unhardened commercial networks.

Systems Generated EMP (SGEMP) was discovered later, through theoretical calculations that predicted that the deposition of x-rays on satellites would eject photoelectrons that would turn in the magnetic fields they produced, which could generate EMP signals in the immediate vicinity of vulnerable components of lightly-shielded satellites.[30] Although there has been no direct field observation of SGEMP, it could be investigated through underground nuclear tests. EMP in its various forms is a phenomenon that seems certain to exist theoretically and has effects predicted to be too large to ignore. However, it is incompletely observed and characterized and would be expensive to shield against, so the defense community has a strong incentive to discount it, which it largely has. However, if a defensive system is intended to operate robustly in the nuclear environment produced by either it or the attacker, it will probably be necessary to explore EMP more thoroughly.

Artificial Radiation Belts

Electron injection into, or "pumping," the Earth's magnetic fields was examined in the Argus high-altitude test, after which the effect is named. It was observed again in the high-altitude Starfish test. Injection models roughly reproduce the results of those two tests, but have enough arbitrary assumptions to have limited credibility for other altitudes and yields. Those models are the basis for estimates that low-altitude satellites could fail within weeks of nuclear explosions in space. The uncertainties in those estimates are too large to be ignored. However, it is difficult to perform relevant laboratory experiments, and field experiments are strongly opposed because of their possible effects on U.S. and other satellites.

Summary of Developments in Nuclear Phases

Missile defenses flowed from the anti-aircraft and missile developments of World War II, were stimulated by the rapid buildups in offensive missiles thereafter, and further accelerated by the rapid proliferation of missiles and WMD after the Cold War. NIKE systems developed radar and interceptor technology to levels adequate for the defense of population against light attacks. To eliminate the vulnerabilities of mechanically steered radars, NIKE also developed the technology for phased array radars, which remain key components of current systems. Sentinel integrated those technologies and conducted the large number of tests needed to demonstrate that they were adequate for light attacks from China. However, technical criticisms of its deficiencies against attacks much larger than those for which it was intended and public concern over the collateral damage that might be produced by such attacks blocked its deployment and forced a redefinition of its mission.

29. J. Darrah, "Foreword," in *IEEE Transactions on Antennas and Propagation*, pp. 1–2.
30. W. Graham and T. Mo, "Probability Distribution of CW Induced Currents on Randomly Oriented Subresonant Loops and Wires," invited paper, in *IEEE Transactions on Antennas and Propagation,* pp. 107–117.

Safeguard was directed to address large attacks on the U.S. deterrent with the Sentinel hardware that had already been shown to be inappropriate for them. Technical and public debates over Safeguard exposed its sensitivity to countermeasures, unknown nuclear backgrounds, and the effects of high-altitude explosions, for which there is little data. Spartan's contribution to a layered defense was undercut by the strong blackout and refraction expected at the frequencies at which its PAR radars operated. Uncertainties about the effectiveness of its exoatmospheric layer led to arguments for shifting to an endoatmospheric terminal defense, but Safeguard's endoatmospheric MSR-Sprint layer could be saturated or bypassed. The number of MSR radars and Sprint interceptors was not adequate for a stand-alone terminal defense and could not be increased affordably. It was agreed that a viable, endoatmospheric defense could be developed based on larger numbers of smaller, cheaper versions of MSR and low-altitude versions of Sprint, but Safeguard was not a good approximation to such a system and could not be redesigned as one. These debates did not prove that no system based on radars and nuclear interceptors could meet this mission, only that Safeguard was not well suited to it.

Safeguard's boosters and warheads were successful, but its long-range radars would degrade unacceptably in the environments it was expected to produce, so it was completed, activated briefly, and then shut down due to perceived cost and lack of effectiveness. Those weaknesses are probably shared by the nuclear defensive system deployed around Moscow. Both illustrate the difficulties of operating in the unknown environments from defensive and offensive nuclear explosions. These nuclear effects were not calibrated, could not be predicted, and thus could not be compensated for in real time with the tools then available, so there was no way to predict or adapt to blackout and refraction irregularities. Those deficiencies identified battle management as a key issue. That was a strength of Bell Laboratories, the system integrator, but Bell withdrew from the program, which effectively ended that phase of development.

PAR and Spartan's expected contributions to attrition of the threat were too small and uncertain to be critical; however, they could have made a significant contribution to a defense against the limited Chinese attacks for which the Spartan hardware was originally designed. Those attacks contained roughly the numbers, yields, and countermeasures expected from Chinese or rogue threats today. However, the uncertainties in Spartan and PAR's performance increased sharply at about the number of weapons China could deliver in a first strike. As the predicted performance of radar-based systems is sensitive to the specific frequencies used, it would be useful to reexamine those uncertainties in light of intervening technological developments.

For the defense of deterrent forces, Safeguard's deficiencies were qualitative; they could not have been corrected with current levels of technology. For defense of population against light attacks, its deficiencies were quantitative; they might have been corrected through redesign, if the system had been more responsive. Performance could have been acceptable against attacks producing a few explosions per PAR. The higher frequency radars available today probably could have performed better. Unfortunately, those questions cannot be answered with confidence. The premature conclusion that it was not possible to make effective radar-based systems caused the United States to turn away from missile defense, advocate an ABM treaty that would make the development of such defenses impossible, and abandon the research that could have provided answers to these questions.

NIKE, Sentinel, and Safeguard's use of nuclear weapons was opposed on political and popular grounds, but in defending against nuclear attacks, the collateral damage due to a nuclear defense would not be significantly greater than that from a nonnuclear defense. If Sentinel had been deployed or Safeguard's operation continued, either would have provided adequate protection against the threats experienced up to the present, short of those from the Soviet Union. They could have provided protection from rogue threats comparable to that expected from subsequently developed nonnuclear hit-to-kill systems. The nuclear systems' fundamental deficiency was the limitations placed on their performance by the nuclear environments they produced, which bounded the threats they could

address. These nuclear backgrounds would also be encountered in attacks that contain precursor or salvage detonations, so these limits also apply to nonnuclear defensive systems that contain radars or IR systems as primary acquisition and track sensors.

Developments During the 1970s

The ABM Treaty

The 1972 ABM Treaty that ended Safeguard also limited subsequent research, development, and tests on concepts and systems with missile defense potential. Its immediate impact was to terminate research on the uncertainties discussed in the previous section. Funding fell by an order of magnitude over the next few years, effectively ending any organized research program on missile defense phenomenology and dispersing its investigators into unrelated fields. The Treaty was a U.S. creation. Having concluded that Safeguard would not work, the United States convinced the USSR that it was unlikely that any such system could work against large attacks, so that both countries would be better served by forswearing defenses and leaving their populations vulnerable to the other's retaliation, a policy aptly described as mutually assured destruction or MAD. Whether the USSR actually accepted that policy or used it as a tool to compensate for its strategic weakness is not known.

The United States withdrew from the Treaty in June 2002, but it is useful to review its major provisions as they strongly shaped the evolution of current defensive systems, controlled their development and deployment until recently, and are still reflected in some current strategic thought.[31] Article I prohibits national defense, i.e., defense of population, which was essential because allowing the defense of population would undercut the logical basis of MAD, which depends on leaving populations at risk. Article II narrowly defined ABM systems in terms of the interceptors, launchers, and radars then known. Article III restricted defenses to one site with 100 interceptors, for which the United States chose a missile wing and the USSR chose Moscow. Article V prohibited development, testing, or deployment of sea-, air-, space-, or mobile land-based systems as well as systems based on other physical principles (OPP), e.g., using lasers as interceptors or IR satellites instead of radars. Article VI placed restrictions on "testing in the ABM mode," which constrained subsequent research and development since system tests could be interpreted after the fact as having been in the ABM mode, which would impose Treaty limits on their subsequent development or deployment.[32]

Article IX prohibited transfer of ABM systems and components to other countries. The Treaty thereby permitted but did not define TMD, other than through a provision inserted by the U.S. in an unsuccessful attempt to prevent a Soviet surface-to-air missile (SAM) upgrade breakout. Compliance with these provisions significantly reduced the capability of the U.S. SAM-D (also known as Phased Array Tracking Radar Intercept on Target or PATRIOT) and later Navy air defense systems. Unratified attempts during the Clinton Administration at definition or "demarcation" of TMD would have further limited TMD interceptor speed and testing, space-based interceptors, and the components and technologies that could substitute for interceptors.

Article XII directs the use of national technical means for Treaty verification, but does not define them. They are now known to include U.S. and Russian early warning systems and related assets.

31. U.S. Department of Defense, Ballistic Missile Defense Office (hereafter cited as BMDO), *Harnessing the Power of Technology: The Road to Ballistic Missile Defense from 1983–2007*, September 2000, pp. 10–11.
32. H. Cooper, "Reviving Effective Programs to Protect America from Ballistic Missile Attack," *The Journal of Social, Political and Economic Studies*, Vol. 28, No. 1, Spring 2003.

Avoiding the mention of specific systems produced ambiguities in interpretation that precluded the effective integration of U.S. warning systems with other warning and defensive systems for several decades, making it necessary to define them as "adjuncts" to Treaty-defined radars, even as those satellites largely eliminated the need for such radars.[33] Agreed Statement B restricted the power-aperture product of ABM radars to no more than PA = 3 MW-m^2 at the defended site to prevent integration of the longer-range BMEWS's 160 MW-m^2 radars. The Treaty attempted to prevent future missile defenses because of the assessment that no defense based on any technology could work against large threats. The evolution of technology gradually altered the basis for that assumption and provided alternative approaches to defense, which reversed the negative assessment on which the Treaty was based. However, attempting to develop these technologies under Treaty constraints distorted research, development, testing, and policy for three decades. Treaty-imposed impediments to development and testing have now been removed, but their effects linger on.

U.S. Army Programs and Hit-to-Kill Technology

Research and development were not dormant in the decade after the Treaty and the deactivation of Safeguard. Two distinct, parallel development paths led to most of the technologies available today: Army development of nonnuclear hit-to-kill interceptors and DARPA development of space-based lasers.

In the mid-1970s, the Army studied the Site Defense and Low Altitude Defense Systems (LOADS) proposed as more effective and affordable defenses and successors to Safeguard to improve the survivability of silo-based ICBMs. They were essentially the endoatmospheric defenses with more and harder radars and Sprints suggested by critics during the Sentinel and Safeguard debates. The 1980 Defense Science Board (DSB) Summer Study found the preferential defense of Minuteman missiles in silos with LOADS to be the most cost effective way to address their survivability. LOADS was assessed to be cost effective in performing this critical mission for which it was designed, but it was rejected because of concerns over conflicts with the ABM Treaty. After studying and proposing several downsized nuclear systems, which were rejected for similar reasons, the Army addressed the fundamental problem and started the development of the sensors and guidance needed to make nonnuclear HTK possible.

A key enabler for HTK is the high energy density of hypersonic collisions. The energy per unit mass from the impact of bodies closing at 10 km/s is about a factor of 10 greater than that from high explosives; thus, at high closing velocities it is not necessary to provide an explosive kill package. There is enough energy in the collision to destroy any target, particularly a soft missile booster in powered flight, when any disruption is lethal. However, taking advantage of that energy requires the kill vehicle to actually hit the target, which required significant advances in guidance and control, particularly in proportional guidance with passive sensors (See Appendix E).

Command guidance for missile defense built on earlier unsuccessful German attempts to develop surface-to-air and air-to-air missiles during World War II. The United States successfully capitalized on it in postwar research that made command guidance practical.[34] The essence of those developments is proportional navigation (PN), which prescribes the acceleration, $y'' = KV_C\lambda'$ that an interceptor needs to hit a missile as a function of their closing velocity V_C, where the rate of change of the line of sight (LOS) between them is λ', primes denote differentiation with respect to time, and K is a constant "navigation ratio" that characterizes the responsiveness of the interceptor. The basic PN concept is illustrated by two cars approaching an intersection. If each observes that the other's headlights remain at a constant angle, they are on a collision course, i.e., constant LOS and decreasing range indicate an

33. A. Carter and D. Schwartz, *Ballistic Missile Defense* (Washington, D.C.: Brookings Institution, 1984).
34. P. Zarchan, *Tactical and Strategic Missile Guidance*, A. Seebass, ed., Progress in Astronautics and Aeronautics Series, Vol. 124 (Washington, D.C.: American Institute of Aeronautics and Astronautics, 1990), p. 17.

intercept trajectory. If λ is changing, the interceptor should accelerate toward the missile. PN is valid for non-maneuvering, accelerating, and evasive missiles and is the optimal guidance law under linear filtering and control in the presence of noise.[35] It is particularly appropriate for exoatmospheric midcourse engagements, where all objects follow ballistic trajectories that can be predicted and compensated for accurately.

For non-accelerating missiles, PN requires the interceptor to increment its velocity by the missile's transverse velocity V_M, after which their transverse separation decays exponentially on a time scale T/K, where T is the "time to go" until intercept. It is desirable to transfer to the intercept trajectory quickly, so interceptors typically use navigation ratios K = 3–5. Non-maneuvering targets require maximum acceleration at the outset. For typical midcourse values of $V_M = 0.1$ km/s, $T = 40$ s, $K = 4$, and the initial acceleration is $2V_M K/T = 2$ g, which is not stressing.

For targets with transverse acceleration A, the PN acceleration becomes $y'' = KV_C\lambda' + A$, which for zero initial velocity error integrates to $y' = KV_C\lambda + At$. For no initial LOS error, the initial acceleration is zero, but it increases to a maximum of $y'' = AK/(K-2)$ at the time of impact $t = T$. Figure E.1 shows the PN relative accelerations y''/A needed as functions of t/T for typical values of K. The PN curves are those that increase with t/T. For $K = 3$ typical of current designs, the relative acceleration at impact is $y''/A = 3$, which is the basis for the rule of thumb that an interceptor needs an acceleration about 3 times that of the missile for successful intercept.

Against accelerating targets, PN can be augmented to incorporate their projected acceleration. Figure E.1 shows the accelerations required to intercept targets with augmented proportional navigation (APN), which are the curves that fall with t/T. APN has maximum acceleration of $AK/2$ at $t = 0$. For typical values of K, PN and APN's maximum relative accelerations are similar. It is straightforward to show that their accelerations are linearly related. Optimal guidance produces smaller miss distances but requires comparable accelerations. Proportional navigation is also possible for boost phase, although missile accelerations are larger than midcourse and increase as their fuel is expended.[36] APN can also be used in the boost and terminal phases, although missiles have large and varying accelerations toward the end of powered flight and RVs have large and varying decelerations as they reenter.

Missile acceleration is a key issue in boost-phase intercept. Constant thrust missiles with 3–4 g average accelerations can reach 3–4-fold higher peak accelerations as their fuel is exhausted, which implies PN interceptor accelerations of 9–16 g near impact, which would place significant strains on interceptor design. However, if the missile's acceleration profile can be measured or predicted accurately, these varying accelerations can be compensated for with an interceptor trajectory that requires an acceleration equal to the average acceleration of the missile. PN then only has to cope with residual errors, which are much smaller. In predicting missile acceleration it is useful to have a set of acceleration and brightness profiles of known missiles to use as templates for identification. That could speed and improve predictions of accelerations, velocity profiles, and stage times and would reduce the interceptor's computational burden, errors, and velocity requirements. Thus, there is utility in gaining additional information on missiles for which templates are noisy or non-existent. The SBI sensors themselves could contribute to the accumulation of such a database.

HTK sensor and guidance packages were developed and tested in the 1984 homing overlay experiment (HOE), which was the first test of HTK in an exoatmospheric intercept. HOE used intercontinental ranges for both the test missile and interceptor, which is reported to have achieved a direct hit, although it deployed an umbrella-like kill enhancement device that expanded to several meters. HOE had a 1,200 kg kill vehicle (KV). Subsequent developments led to much smaller KVs. Those used in

35. A. Bryson and Y. Ho, *Applied Optimal Control* (Waltham, Mass.: Blaisdell, 1969).
36. Zarchan, *Tactical and Strategic Missile Guidance*, pp. 239–253.

current flight tests are an order of magnitude smaller than HOE. These reductions in KV weight and size directly reduced the weight and cost of whole interceptor, making defenses with large numbers of interceptors technically and economically feasible.[37] The Air Force anti-satellite (ASAT) was also developed in that period for a different purpose: deployment on F-15 interceptors to respond to time-urgent space threats. It was initiated on the final day of the Ford Administration and tested successfully in 1984. The ASAT was substantially smaller than HOE because it used more advanced technology. It underwent significant development and testing before being cancelled due to Congressional concerns about weaponizing space.

With active sensors such as radars, measurements of interceptor and missile positions and velocities can be filtered and used to estimate the ranges, LOS rates, and intercept times needed in the equations for PN. Radar measurements have limited cross-range resolution, but do observe range, and hence all of the elements of the missile state vector needed to solve for optimal acceleration. NIKE, Sentinel, and Safeguard used radar command guided systems that produced miss distances proportional to the intercept range. At ranges of 1,000s of kilometers, their mrad accuracies translated into kilometer miss distances, which necessitated nuclear warheads. To support the 1 m miss distances needed for HTK at 1,000 km, remote command guidance would need microradian accuracies. A hybrid approach used in early ground-to-air interceptors was to illuminate the missile with a ground-based radar, but let the interceptor use an on-board sensor and proportional guidance system to control its own intercept. That improved estimates of range and time to go, but required the radar to illuminate the missile for the whole intercept, which limited the number of missiles it could handle. In later interceptors such as PATRIOT Advanced Capability–3 (PAC-3), it was possible to package high frequency radars on board the interceptors with radomes calibrated sufficiently accurately to permit them to make radar-guided intercepts themselves with on-board computers, which freed their ground-based radars to continue search and track of other targets.

Passive sensors such as IR seekers only measure angular displacements, so they do not provide information on range or time to intercept, which must be provided from external sensors or estimated from multiple angular measurements. The key enablers for radar and IR homing interceptors were the advances in electronics, computers, and sensors that made it possible to package affordable homing sensors in the interceptors. With on-board sensors, HTK takes advantage of range. As it flies its pursuit trajectory, the interceptor's range to target decreases, so an on-board sensor of given angular resolution produces ever-improving spatial resolution as it approaches the target. From a separation of 10 km, the roughly 100 microradian angular resolution of a 10 cm IR optic produces a spatial resolutions of about 1 m, which can resolve targets and perform selection of vulnerable aim points.

Nonnuclear kill reduces the disturbed backgrounds from the levels produced by Spartan and Sprint nuclear engagements in proportion to the number of intercepts; however, an attacker could still use intentional precursors or salvage fusing to complicate the environment. Thus, nuclear issues can be reduced by NNK, but the amount of reduction is not known with confidence, and their residual level is uncertain because there were no IR measurements of detonations in the appropriate altitude regimes. For that reason Army research and development programs continued the development of higher frequency, power, and bandwidth radar components and algorithms better suited to the discrimination of complex threats in unknown backgrounds.

37. BMDO, *Harnessing the Power of Technology*, p. 4.

Space-Based Laser for Boost-Phase Intercept

DARPA also sought ways around the uncertainties and limited coverage of midcourse and terminal intercepts. It sought a way to address SS-9s in boost, when they were described as "titanium balloons full of gasoline." If the SS-9s could be reached during powered flight, their thin skin and volatile contents should make them vulnerable. DARPA had studied this problem two decades earlier under the Ballistic Missile Boost Intercept (BAMBI), which got as far as the design of a Space Patrol Active Defense (SPAD) system weighing about 30 tons with a nonnuclear KV weighing about 150 kg with a range of about 250 km. SPAD was not, however, able to resolve its issues in IR seeker, computers, and satellite complexity, so it lost out to NIKE-ZEUS.[38] DARPA reviewed the earlier BAMBI studies and rejected them as leading to interceptors as large as Minuteman ICBMs in orbit. A kinetic energy weapon (KEW) did not appear practical with current technology, and the improvements needed did not appear likely in the near term, so DARPA selected directed energy weapon (DEW) approaches for intensive development.

Having selected DEW, DARPA had to address the serious technical challenges in developing a space-based laser (SBL) with the large power P, optics diameter D, and beam quality necessary at the wavelength λ of about 3 micron wavelengths at which megawatt (MW) power level operation had been demonstrated. An ideal laser's beam divergence is about λ/D, so it can deliver a flux $F = P/(r\lambda/D)^2 = B/r^2$ on a target at range r, where $B = P(D/\lambda)^2$ is called the laser's "brightness." Irradiating a target for time T deposits a fluence, or energy per unit area, of $J = FT$, which must be sufficient to heat the target to failure. The brightness required for lethality at range r is thus $B = Jr^2/T$, where the laser determines B, the threat determines J and T, and the range at which the two times are equal is $r = \sqrt{(BT/J)}$ (See Appendix F). A laser with an average range to target r can cover an area of roughly πr^2. Covering the whole Earth's surface area $4\pi R_e^2$ would thus require $N = 4R_e^2/\pi r^2 = (2R_e/r)^2$ lasers, which would have an average spacing of $r = 2R_e/\sqrt{N}$. For example, $N = 18$ platforms would produce an average range $r = 2 \times 6{,}400$ km/$\sqrt{18} = 3{,}000$ km. It is necessary to place N lasers in orbit to assure that at least one is within range of the target at launch. This multiplying factor is called the "absentee ratio" of the constellation, which in this example is 18.

For rogue missiles, which are essentially launched from a point, a more accurate calculation is possible that accounts for the time it takes the laser to switch between targets T_s. The time to kill a missile at range r and switch to the next is $Jr^2/B + T_s$, which must be averaged over the range r to all satellites in range. Doing so gives the average kill rate $M' \approx (N/R_e^2)B/J$. Equating that to the launch rate M/T gives $M/T = (N/R_e^2)B/J$. An 18 satellite constellation of the above brightness and T_s of 1 s gives M/T equal to 0.2 kill/s, so it could address 100 missiles launched over 500 s. The roughly 1,000 Soviet missiles launched in its 600 s minimum interval could be addressed by roughly 150 lasers of that brightness. Alternatively, they could be addressed by 18 lasers of 8-fold higher brightness, i.e., 2.5×10^{20} W/sr, which could be produced by 12 MW lasers with 15 m mirrors. Since only the product of the number and brightness of the lasers is determined by the launch rate, the number and brightness of the lasers could be varied inversely to minimize the total cost of the constellation.

Such brightness levels seemed possible within a decade on the basis of direct scaling from power levels of a few MW, mirror diameters of a few meters, and brightness of about 3×10^{18} W/sr. It appeared that scaling power and diameter factors of 2 to 3 would provide adequate margin against expected missile threat. Thus, DARPA initiated parallel technology programs in lasers, optics, and microradian pointing and tracking. It was estimated that 10^{20} W/sr satellites could be built for about $1 billion each. If so, they would provide a roughly $20 billion counter to a threat costing about $1 trillion (1,000 missiles x $1 billion per missile), which would give SBL boost-phase defenses a cost leverage of about 30:1. However, before those programs came to fruition, the SS-9 was replaced by the SS-18, which was about 10-fold harder to laser radiation. Moreover, it was recognized that near-simulta-

38. D. Baucom, "The Rise and Fall of Brilliant Pebbles," Missile Defense Agency, U.S. Department of Defense, 2003.

neous launch could leave SBL only about 300 s for intercepting SS-18s in boost. That reduced the available time per missile to about 300 s/1,000 missile = 0.3 s/missile. That increased the brightness required by a factor of 30 to about 3 x 10^{21} W/sr, which was beyond the goal of DARPA's technology program. Since then, DEW technology has been in a race with the threat, with DEW usually about an order of magnitude behind in brightness.

Summary of Developments During the Interim Program

The Interim program developed both the NNK and DEW technologies that provided most of the new elements for current missile defense systems. The nuclear studies brought LOADS to the level of definition needed to assess its effectiveness in improving the survivability of silo-based missiles, if needed. The NNK programs developed and tested through HOE the levels of interceptor, sensor, computer, and guidance performance needed for effective midcourse and terminal kinetic kill systems. DARPA's DEW programs defined the levels of performance needed for effective boost-phase defense against massive Soviet attacks and started the technology programs needed to achieve them.

During this development program, it became clear that DEW platforms were vulnerable to suppression by the attacker before the main attack was launched. The SBL could defend itself by using some of its fuel to destroy the suppressors, but then its defensive capability would be degraded before the main strike was launched. This exhaustion could be avoided by basing the laser on the ground and using relay mirrors in space to redirect the beams toward the missiles, but the space-based mirrors would remain vulnerable to ASATs, the atmospheric links would be susceptible to weather, and the ground-based lasers would be vulnerable to attack or sabotage, so none of the proposed alternatives were completely satisfactory from an operational perspective.

These concerns led to dissent within the defense technical community, the DSB, and finally professional and public interest groups, which slowed the pace of DEW programs to the extent that their goals are unlikely to be met for decades. However, those debates also produced Congressional advocacy, so it is now difficult either to develop or to eliminate SBL. The divisions between the engineers in the program and the scientists criticizing it has persisted and isolated DEW from the expertise that DEW needs to realize its ultimate potential.[39] The major lesson from DEW was that it is not enough to have defenses that propagate fast; their platforms must survive long enough to perform their mission and they must be robust against countermeasures at the same level of technology. That is an extension of the lessons learned from NIKE, that radars must survive in order to be useful, and from Safeguard, that sensors must work in their expected environments. A corollary is that a defense that concentrates only on space-based boost-phase defenses gives the attacker a strong incentive to develop fast, hardened missiles and launch them from small areas to increase the number of space platforms needed for coverage, which the USSR attempted to do and which subsequent rogues have largely succeeded in doing.

39. G. Canavan, N. Bloembergen, and C. Patel, "Debate on APS Directed-Energy Weapons Study," *Physics Today*, Vol. 40, (November 1987), pp. 48–53.

Developments During the 1980s

Strategic Defense Initiative

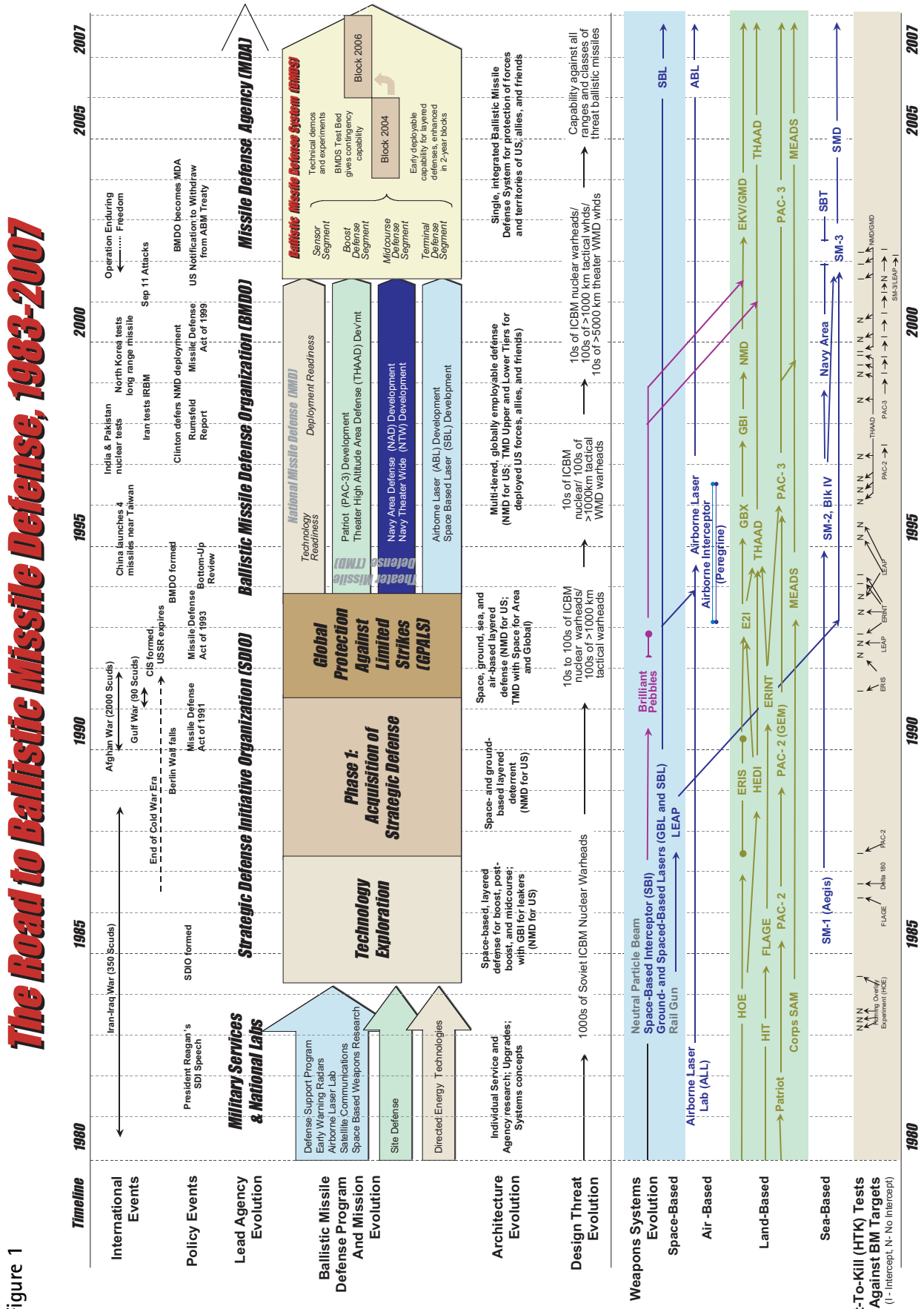
The SDI was formulated against this backdrop of opposition to nuclear intercepts, limited experience with HTK technology, and controversy over DEW. It was stimulated by two factors. The first was the Soviet Union's unconstrained buildup in missiles and weapons, which was characterized as "We build; they build; we don't build; they build." U.S. offensive measures had little apparent impact on a buildup that increased Soviet warheads from 2,400 at the signing of the ABM Treaty to 6,000 in 1980. The second factor was Soviet accuracy, which improved threefold between 1975 and 1980, and the "Team B" inference that this improvement could support a first strike capability.[40] This increase in numbers and accuracy cast doubt on whether the United States understood Soviet military and political strategy.[41] This combination of warhead numbers and accuracy appeared to represent a threat to U.S. silo-based missiles.

The first year of the Reagan Administration was spent in an extensive review of options for improvements in offensive systems and nuclear weapons. The second year involved a review of defensive options by the White House Science Council, which recommended research and development options on selected innovative concepts. In 1983, the Joint Chiefs of Staff recommended that defenses could introduce an important dimension into the interaction with the USSR. White House staff argued that defenses could outstrip the USSR economically. Defenses also had strong support from external groups such as the High Frontier. There are still arguments about which recommendation motivated President Ronald Reagan to make his speech on March 23, 1983, but it was probably his own decision. He came into office dissatisfied with deterrence through the threat of retaliation and the knowledge that the United States then had no other response. His speech called on the scientists who had created nuclear weapons to study ways to eliminate them, but there were few takers. SDI was primarily supported by those involved in earlier nuclear, Army HTK, and DARPA DEW defensive efforts. SDI led to an effort with several distinct phases that are used to organize the discussion below:

- 1984–1987 was an intensive exploration of technologies relevant to missile defense.
- 1987–1991, Phase I, initiated a phased National Missile Defense (NMD).
- 1991–1993 developed Global Protection Against Limited Strikes (GPALS).
- 1993–1996 primarily developed Theater Missile Defenses (TMD).
- 1996–2000 reinstated the milestoned development of NMD while acquiring TMD.
- 2000–present developed a spiral development program with definite deployment timelines.

40. Intelligence Community Experiment in Competitive Analysis, "Report of Team 'B': Soviet Strategic Objectives, an Alternative View," December 1976 (National Archives, declassified as NND 933009).
41. Intelligence Community Experiment in Competitive Analysis, "Report of Team 'B,'" part III.

Figure 1


The Road to Ballistic Missile Defense, 1983-2007
Timeline
1980
1985
1990
1995
2000
2005
2007
International Events
Iran-Iraq War (350 Scuds)
Afghan War (2000 Scuds)
Gulf War (90 Scuds)
End of Cold War Era
Berlin Wall falls
CIS formed, USSR expires
China launches 4 missiles near Taiwan
India & Pakistan nuclear tests
North Korea tests long range missile
Iran tests IRBM
Sep 11 Attacks
Operation Enduring Freedom
Policy Events
President Reagan's SDI Speech
SDIO formed
Missile Defense Act of 1991
Missile Defense Act of 1993
BMDO formed
Bottom-Up Review
Clinton defers NMD deployment
Rumsfeld Report
Missile Defense Act of 1999
US Notification to Withdraw from ABM Treaty
BMDO becomes MDA
Lead Agency Evolution
Military Services & National Labs
Strategic Defense Initiative Organization (SDIO)
Ballistic Missile Defense Organization (BMDO)
Missile Defense Agency (MDA)
Ballistic Missile Defense System (BMDS)
Ballistic Missile Defense Program And Mission Evolution
Defense Support Program
Early Warning Radars
Airborne Laser Lab
Satellite Communications
Space Based Weapons Research
Site Defense
Directed Energy Technologies
Technology Exploration
Phase 1: Acquisition of Strategic Defense
Global Protection Against Limited Strikes (GPALS)
National Missile Defense (NMD)
Technology Readiness
Deployment Readiness
Theater Missile Defense (TMD)
Patriot (PAC-3) Development
Theater High Altitude Area Defense (THAAD) Dev'mt
Navy Area Defense (NAD) Development
Navy Theater Wide (NTW) Development
Airborne Laser (ABL) Development
Space Based Laser (SBL) Development
Sensor Segment
Boost Defense Segment
Midcourse Defense Segment
Terminal Defense Segment
Technical demos and experiments
BMDS Test Bed gives contingency capability
Block 2004
Block 2006
Early deployable capability for layered defenses, enhanced in 2-year blocks
Architecture Evolution
Individual Service and Agency research; Upgrades; Systems concepts
Space-based, layered defense for boost, post-boost, and midcourse; with GBI for leakers (NMD for US)
Space- and ground-based layered deterrent (NMD for US)
Space, ground, sea, and air-based layered defense (NMD for US; TMD with Space for Area and Global)
Multi-tiered, globally employable defense (NMD for US; TMD Upper and Lower Tiers for deployed US forces, allies, and friends)
Single, integrated Ballistic Missile Defense System for protection of forces and territories of US, allies, and friends
Design Threat Evolution
1000s of Soviet ICBM Nuclear Warheads
10s to 100s of ICBM nuclear warheads/ 100s of >1000 km tactical warheads
10s of ICBM nuclear/ 100s of >1000km tactical WMD warheads
10s of ICBM nuclear warheads/ 100s of >1000 km tactical whds/ 10s of >5000 km theater WMD whds
Capability against all ranges and classes of threat ballistic missiles
Weapons Systems Evolution
Space-Based
Neutral Particle Beam
Space-Based Interceptor (SBI)
Ground- and Spaced-Based Lasers (GBL and SBL)
Rail Gun
LEAP
Brilliant Pebbles
SBL
Air -Based
Airborne Laser Lab (ALL)
Airborne Laser
Airborne Interceptor (Peregrine)
ABL
Land-Based
HOE
HIT
Patriot
Corps SAM
FLAGE
PAC- 2
ERIS
HEDI
PAC- 2 (GEM)
ERINT
E2I
THAAD
MEADS
GBX
PAC- 3
GBI
NMD
EKV/GMD
THAAD
MEADS
PAC- 3
Sea-Based
SM-1 (Aegis)
SM-2, Blk IV
Navy Area
SM-3
SBT
SMD
Hit-To-Kill (HTK) Tests Against BM Targets
(I - Intercept, N- No Intercept)
Nitrling Overlay Experiment (NOE)
FLAGE
Delta 180
PAC-2
ERIS
LEAP
ERINT
LEAP
PAC-2
THAAD
PAC-3
NMD/GMD
SM-3/LEAP

Each phase had distinct challenges and produced technologies, systems, and legacies that have provided the options on which the current program is building.[42] These phases and their products are discussed in order below (see Figure 1).

Exploration Phase

The bulk of the "Fletcher Study" of missile defense technologies—named for its director, former NASA Administrator Jim Fletcher, was executed in 1983–1984, but its emphasis on innovation continued through 1984–1987 and subsequent phases. Its recommendations were broad and fundamental. Critics characterized its proposed programs as "Star Wars," but its first phase actually emphasized conventional ground-based interceptors, radars, and space-based sensors. Its second phase added DEW lasers and neutral particle beams, but largely for active discrimination. Its third phase added advanced DEW for bus watching and other advanced applications. Those involved in the Exploration Phase recognized the need to use all three defensive layers to put as much pressure as possible on the threat, but judged the technologies for all three layers to be immature; thus, their recommendations primarily stressed research and development. The Army capitalized on the interest in HTK to develop the needed midcourse and terminal elements and convert earlier technology demonstrations into prototype HTK systems. FLAGE and PATRIOT Advanced Capability–2 (PAC–2) demonstrated significant improvements in low-altitude intercepts. Otherwise, the Exploratory Phase concentrated more on idea generation and laboratory demonstrations than technology and systems development.

Exploration Phase experiments revealed weaknesses in the understanding of DEW, KEW, and nuclear lethality and recommended broad experimental programs, portions of which were carried out in later phases. Studies of the effect of reflective surfaces and rotating boosters on laser lethality led to further increases in estimates of the brightness required and raised concerns that attainable systems would not be able to survive determined suppression attempts. The x-ray laser offered improvements in performance and survivability, but encountered technical and political barriers. However, external criticisms largely focused on estimates of DEW constellation sizes and countermeasures, to which SDI could successfully respond.[43] While the SDI program survived, these debates sensitized the Department of Defense (DOD) and public to the inherent vulnerabilities of large DEW platforms.

The Exploratory Phase led to a series of proposals for satellites to improve the DSP satellites that had provided early missile warning and track for several decades. Some proposals predated SDI. In 1977, President Jimmy Carter, concerned that trends in Soviet nuclear forces could require the ability to fight protracted wars, signed Presidential Review Memorandum 10, which called for a comprehensive net assessment of military force postures, and Presidential Directive 18 for its implementation. That led to acceleration of the Air Force's Mosaic Sensor Program (MSP), which was renamed the Advanced Warning System (AWS) as the first of a series of programs to improve the revisit time for missile observations by replacing spinning DSPs with satellites with mosaic arrays of detectors that could continuously stare at the whole Earth.[44] However, producing such satellites was slowed by technical and cost problems, so DSP continued to serve as the primary missile warning sensor.

When President Reagan signed National Security Directive 6-83 on eliminating the threat from ballistic missiles, the requirements for missile warning became more demanding. AWS's goal had been to count attacking boosters and determine where they were headed. To protect population it would also be necessary to detect and track their buses, which can deploy weapons and countermeasures on trajectories different than those of their boosters. Bus track became the mission of AWS, which was renamed the Boost-Phase Surveillance and Track System (BSTS), which was to have improved staring sensors and on-board computational ability to determine the trajectories of missiles and buses for

42. BMDO, *Harnessing the Power of Technology*, pp. 22–24.
43. Canavan *et al.*, "Debate on APS Directed-Energy Study."
44. Richelson, *DSP Satellites and National Security*, pp. 85–88.

attacks from any area on the globe. However, it was soon found to be too difficult to perform those tasks from geosynchronous altitudes, so its mission was reduced to roughly that of the DSP, and its track and assessment functions shifted to low-altitude satellites.[45]

These early SDI efforts also defined the semi-synchronous Space-Based Surveillance and Tracking System (SSTS) and its companion Ground-Based Surveillance and Tracking System (GSTS), which was to be popped up on warning. SSTS and GSTS were to have enough sensitivity to track missiles and buses through the powered and ballistic phases of their trajectories, identify "threatening clusters" of objects, discriminate weapons, and support improved interceptor allocation. The Airborne Optical Sensor, which was developed to test their IR optics, later evolved into the Airborne Optical Adjunct (AOA) used to gather IR test data.[46]

SSTS and GSTS were intended to identify threatening clusters containing weapons and to discriminate RVs from decoys. Most modern ICBMs carry independent buses that separate from the booster, thrust sequentially onto a number of separate trajectories, and release clusters of objects on each. Not all trajectories or clusters will contain weapons, but if the defense cannot determine which clusters contain weapons, it is forced to waste interceptors on empty ones. Identifying which clusters contain weapons headed toward critical targets improves interceptor allocation and supports efficient "shoot-look-shoot" firing doctrines that avoid wasting interceptors on targets that have already been killed. "Birth to death" observations and discrimination are essential for effective battle management and are consistent with a philosophy of maintaining pressure on all phases of the missile's trajectory in order to complicate attacks and minimize the attacker's options for overloading any given individual layer.

The initial phase of SDI screened a large number of technically plausible, but large, expensive, and redundant possibilities. It advanced most HTK and DEW elements then available, but did not develop them into systems. It greatly overestimated the potential of DEW and underestimated that of KEW ground and space-based HTK systems. Such errors were understandable in the quick, early study by DARPA, but not in the later and more deliberate Fletcher Study.

Strategic Defense System Phase I

The 1987–1991 Strategic Defense System (SDS) Phase I was initiated in response to the continuing Soviet buildup in heavy strategic missiles. Its goal was to blunt the leading edge of an attack, which was a step back from the protection of population to an earlier goal of assuring survival of the strategic forces to improve deterrence. As this was a military goal, Phase I could tolerate significant leakage, but it had to be able to negate the bulk of the highly accurate systems that threatened U.S. missiles in their silos. It was intended to counter the Soviet buildup at as soon as possible, so it had to use technologies developed in previous phases, and its choices were constrained by the levels of maturity they had achieved. It sought to develop them quickly to the levels needed for deployment decisions, which required significant experimentation and demonstration. As support for SDI was still uncertain, Phase I also had to use those experiments to gain political and international recognition, which required large-scale technology demonstrations.

The Delta Experiments were a good example. Their goal was to examine the plumes of thrusting rocket to determine if it was possible to see through them to locate the missile hard body well enough to support HTK intercepts.[47] The experiments were complicated, as plumes vary strongly with the missile and altitude. They can envelop the missile hard body and dwarf its infrared signatures at the altitudes of intended intercepts.

45. *Ibid*., pp. 180–182.
46. BMDO, *Harnessing the Power of Technology*, pp. 22–24.
47. Baucom, "The Rise and Fall of Brilliant Pebbles," p. 19.

Delta 180 was designed to do IR boost-phase detection in high clutter with KV-class optics against thrusting objects. It used passive sensors on the target to gather liquid plume data, a lidar on the target to track the interceptor and gather plume-to-hardbody handover data, and a Phoenix (AIM-54) x-band radar to guide the APN intercept of a 6 g accelerating target at the 3 km/s permitted by the ABM Treaty.

Delta 181's studied the phenomenology and discrimination of RVs, light RVs, decoys, balloons, spheres, and small motor plumes. It used a lidar to examine the plume-to-hardbody handover phenomenology of a solid engine and for the measurement of midcourse features, which were correlated with x-band measurements. It also examined countermeasures including chemical releases and performed ultraviolet (UV), visible, and IR feature detection. It was successful in all of them.

Delta 183's objectives were the acquisition of a missile's ignition plume through clouds, detection of its hot pad subsequent to launch, passive track of targets of opportunity (the liquid fueled boosters through staging), and demonstration of laser communication and attack warning—all of which were successful. It also deployed a free flying satellite inspection package for space qualification, although it was not released on orbit for policy reasons.

The Delta Experiments successfully demonstrated the ability to see through plumes at a range of wavelengths and altitudes and executed the only successful intercept of an accelerating missile; thus, they provide the essential design basis for all subsequent boost-phase intercept concepts. As each was executed in a matter of months for a cost of under $200 million, they also demonstrated the ability of the Strategic Defense Initiative Organization (SDIO) to carry out large, technically challenging experiments quickly and affordably, which was a key factor in gaining credibility with the DOD, Congress, and international partners.

Hit-to-Kill Technology

Phase I of SDI developed many of the components of current systems. HOE provided the first successful test of exoatmospheric HTK intercept. SDS Phase I had the task of converting it into engineering practice. Its efforts to do so took two directions. The first was the Exoatmospheric Reentry Vehicle Interceptor System (ERIS), which was the culmination of the earlier Army technology programs that produced a ground-based HTK interceptor. ERIS was intended to perform midcourse intercepts as far back along the RVs trajectory as possible with its simple discrimination sensor. Its successful tests provided information for the design of the later higher altitude Theater High Altitude Area Wide (THAAD) interceptor and served as the basis for PAC-3, the HTK successor to the decades-old PATRIOT of the Gulf War that had already been modified to PAC-2 and Guidance Enhancement Missile (GEM) in the Exploratory Phase. ERIS technology ultimately evolved into the current GBI. In the process, its KV weight fell over an order of magnitude from that used in HOE.[48]

HOE's second descendant was the High Endoatmospheric Defense Interceptor (HEDI), which adapted HTK technology into the 30–100 km altitude regime previously accessible only to nuclear interceptors and radar command guidance. Early HTK interceptor IR sensors were blinded by the strong aerodynamic heating at the hypersonic speeds needed to produce large defended footprints. High-speed air stagnating against seeker windows produced high temperatures, whose radiation could "redout" the KV's infrared detectors. HEDI's flow-cooled sapphire windows demonstrated that hypersonic intercepts were possible at altitudes down to 10s of kilometers, providing a nonnuclear alternative to Sprint. Other major enablers were Phase I's order of magnitude improvements in rocket thrust-to-weight ratios and improvements in the weight, cost, and efficiency of computers, communications, and inertial measurement units (IMU).

48. BMDO, *Harnessing the Power of Technology,* p. 18.

A third descendant was the Lightweight Exoatmospheric Projectile (LEAP), which was initially intended for railgun launch, but evolved into a small, light projectile for ground- and space-based interceptors. It provided the technology for the series of Advanced Hover Interceptor Technology (AHIT) proof-of-principle tests for hit to kill. Having gone through several stages of development, LEAP is now used in current Navy interceptor systems.[49]

Space-Based Interceptors

Phase I also introduced the Space-Based Interceptor, a constellation of small HTK interceptors pre-deployed in space to intercept ICBMs in boost. By maneuvering in front of the accelerating ICBM and letting the missile run over it, a few kilogram SBI could release the equivalent of about 100 kg of explosives, more than enough to destroy a 100 ton missile and cause its weapons and fragments to fall short of their targets. Intercepting missiles in boost maximizes their signatures and vulnerability. SBIs could destroy all of the multiple independently targeted reentry vehicles (MIRV) on a missile in one stroke without being distracted by decoys, as there are no simple surrogates for heavy, bright ICBM boosters. Boost-phase intercept is also consistent with the philosophy of denying the attacker a free ride in any phase of the attack. Reducing the probability of surviving boost-phase would make an attacker less likely to allocate payload to midcourse countermeasures or to incur delays to deploy them, as the probability of its being intercepted increases rapidly with the amount of time the bus takes for deployment. Intercept in the boost phase gives SBI the high ground, as missiles must climb a potential well to approach them, which leaves little opportunity for surprise. Space-basing and its associated omnipresence also provide adequate time for data gathering and the practice required to achieve readiness in a deployment that would be harmless to anyone but a deliberate attacker.

SBIs were designed to reach maximum divert velocities using efficient, high acceleration engines. Modeling their performance with the ideal rocket equation makes it possible to analytically execute the tradeoffs that determine the optimal number and speed for any given threat (See Appendix G). For high accelerations, velocities, efficiencies, and short delay times, the constellations needed for Soviet threats could be estimated with geometric arguments. Soviet SS-9 and SS-18s were particularly vulnerable and valuable during their boost phase, which lasted a time T of about 300 s. SBI with small release delay, high acceleration, and maximum speed V of 6 km/s could reach SS-9 and SS-18 launch areas from distances about VT = 6 km/s x 300 s = 1,800 km.

Soviet launch areas were distributed over an area covering much of European Russia, the Ukraine, and the trans-Siberian railway. The overall launch area was irregular, but had an effective radius R of about 1,600 km. Thus, SBIs could fly in from a ring of width VT around the launch area to supplement those within it, so the total radius of the area from which SBIs could contribute to boost-phase intercepts was about 1,600 + 1,800 km = 3,400 km. For a SBI constellation uniformly distributed over the Earth, that radius would give an absentee ratio of about $[2R_e/(R + VT)]^2 = (2 \times 6{,}400/3{,}400)^2 = 14.2$, so about 7 percent of the SBIs would be within range of the missiles during boost. Engaging each of the roughly 600 accurate Soviet heavy missiles in boost would thus require about 14 x 600 = 8,400 SBIs, although Phase I's deterrence goal would have been advanced by intercepting any significant fraction of them.

While the SS-18's boost phase ends at 300 s, the release of RVs from its final stage, which was called a "bus" because it deployed MIRVs, continued until a time T_{bus} of about 600 s. During the interval from T to T_{bus}, the bus remained an attractive target, although of decreasing value. During busing, SBIs could fly in from a distance of $R + VT_{bus}$ = 1,600 + 6 km/s x 600 s = 5,200 km, which would give an absentee ratio of about 17 percent through the end of busing.

49. R. Matlock, in Strategic Defense Initiative Organization, "Report to Congress," April 1987.

The bus-phase intercepts increase the number of SBIs participating by about a factor of 17/7 = 2.5, but the number of RVs killed does not increase proportionally, because most intercepts occur after some RVs have been deployed by the bus. The fraction of the constellation from which SBIs can reach the buses increases as the square of radius and hence time. Thus, many SBIs arrive toward the end of the engagement, when buses have released most of their RVs. If the rate of busing is constant, the number of RVs killed per intercept falls linearly with time. The number of SBIs added and RVs killed can be estimated as the ratio of the number of RVs killed to the number of intercepts during busing. For Soviet launches in which boost and bus durations were each about 300 s, bus kills were roughly half as effective as boost kills. Adding kills during busing increases SBI kills by about a factor of 1.6, which reduces the constellation size required to engage each missile or bus through the boost and bus phases to about 8,400/1.6 = 5,250 SBIs. That number is still large, but would have been feasible, if needed. The Soviet Union collapsed before it was necessary to attempt such blunting attacks.

SBIs were to be alerted by the DSP early warning satellites, subject to existing command and control (C2) and satellite communications systems; however, they used on-board suite of IR sensors to detect and home on ICBM boosters' large, bright plumes. While their on-board sensors were difficult to negate, part of their C2 support information was to come through the "garage" satellites, which were to house many SBIs to share overhead costs for housekeeping functions.[50] However, these garages introduced an unacceptable vulnerability into the SBI system, as they were vulnerable, attractive targets that made SBI C2 depended on BSTS and SSTS warning, information, and external C2 through vulnerable external nodes.[51] The SBI was not entirely new; it was anticipated by the early 1960s BAMBI concept and the 1970s DARPA boost-phase study. However, with early computer technology, BAMBI led to large KVs and rockets, so DARPA concluded that interceptors of that size that were dependent on external warning and C2 would be vulnerable to suppression by Soviet ASATs. The F-15 ASAT demonstrated new designs and technologies that could have been exploited in SBI, but were not. While the SBI ultimately failed of deployment, it did refocus attention on boost-phase intercepts and the leverage they could provide in reducing the number, types, and times for countermeasures available in attacks on survivable, predeployed boost-phase defenses.

Sensors

Phase I developed the high frequency and efficiency components for the ground-based radar (GBR), which is the basis for the highly competent x-band (10 GHz, 3 cm) wideband radars subsequently used for both tactical and strategic systems. The significance of x-band is threefold. First, its higher frequency permits it to form 10-fold narrower beams for correspondingly improved cross-range track resolution. Second, its higher carrier frequency permits 10-fold higher bandwidths than those possible with UHF radars, which produces the roughly 10 cm range resolution needed to discriminate precision decoys. Third, its higher frequency reduces absorption and refraction in nuclear environments about 1,000-fold, which increases their robustness in engagements involving nuclear weapons.

The Exploratory Phase led to the definition of the BSTS satellites for improved detection, track, and attack assessment. Phase 1 exposed the difficulty of realizing them with existing hardware and defined their successor systems—Follow-on Early Warning System (FEWS) and Alert, Locate, and Report Missiles (ALARM)—which were designed for limited roles at less cost. The new configurations were to have improved sensors and enough on-board computational ability to assess the nature and intent of attacks. However, practical designs could not be shown to reliably detect the tracks of buses and other objects with similar signatures or infer the intent of attacks from measurements they could make, which made their mission unclear and delayed development.

---

50. "SBI/GBI Concept," Report KEWT001, Aerospace Corporation Concept Design Center (Space and Missile Systems Center, Air Force Material Command, March 22, 2000); Aerospace *Crosslink*, Winter 2000/2001, pp. 5–11.
51. D. Baucom, "Ballistic Missile Defense: A Brief History," (Washington, D.C.: Missile Defense Agency, U.S. Department of Defense, May 2000).

Phase 1 attempted to develop affordable versions of LWIR focal plane detectors at about 10 micron wavelengths needed to detect the peak of radiation from 300°C black bodies. Such detectors are expensive, as are the large optical telescopes required to capture enough photons for detection at long ranges.[52] Compromises on cost and control led to a primary mirror diameter D of about 30 cm, which produced angular resolution λ/D of about 30 microradians and hence spatial resolutions of about 100 m at the average satellite spacings of about 3,000 km dictated by economics (See Appendix I). At those ranges, these sensors were unable to resolve, let alone discriminate, closely spaced objects, so their mission was reduced to that of a midcourse tracker of undiscriminated clusters of objects. That limitation was compounded when the satellites were given additional missions. That produced further compromises, delayed the schedule, increased cost, and impacted funding. Sensor activities during Phase 1 led to a number of useful technology demonstrations and systems insights, but not to designs that could adequately discriminate projected threats.

Directed Energy Weapons

Directed energy weapons initially appeared to be an attractive alternative to kinetic energy systems because their speed-of-light intercepts made defenses in the boost phase possible. They also held significant promise as active midcourse discriminants. However, the key parameter in determining a laser's effectiveness was not the time to arrive, but the time it took to penetrate the missiles, which depends on their hardness and configuration. Hardening and spinning upper stages could increase lethal fluences by factors of 3 to 10, which made DEW's already difficult technical tasks that much harder. Fast-burn missiles increased laser absentee ratios, although their impact on lasers was less than that on KEW (See Appendix F). In Phase I, it became clear that that to be effective against Soviet launches, which could dedicate 100s of MT to defense suppression prior to launch, defenses would have to be survivable to be useful. Requisite levels of survivability are difficult to achieve with large space-based lasers and optics, whose coatings are susceptible to nuclear radiation and whose light structures are ill suited to withstanding large impulsive x-ray loads. Techniques for mitigating these modes of attack were immature, which meant DEW would develop on longer time scales than Soviet threats.[53]

Neutral particle beam (NPB) DEWs were also developed extensively during Phase 1. Their main advantage was their ability to penetrate thick shielding, which was useful for midcourse discrimination. The neutrons ejected from irradiated materials provide an effective means to measure mass, which is the most difficult quantity for an attacker to simulate. Particle beams' key issues were their source brightness and beam handling technologies, which were less mature than those for lasers, and their large, heavy, and complex structures, which could not readily be hardened, maneuvered, or decoyed for survivability. DEWs suffered by comparison to SBI because their technology was less mature, and they were packaged poorly and more vulnerable to attack. That gradually shifted DEW into a secondary position to SBI in the near term, although their favorable scaling against fast-burn missiles, compact launches, and rogue threats justified continuing research and development as long-term options.

Battle Management, Command and Control

Battle management and command and control (BMC2) efforts for Phase I started in 1985. By 1988, they had demonstrated efficient, robust C2 to manage threats larger than the ones that overwhelmed Safeguard two decades earlier. Software, coordination, and integration problems that had appeared insuperable earlier were successfully simulated at the National Test Facility. The ground-based BMC2 to support them made adequate progress, but the integration of the midcourse sensors needed

52. Sessler, *et al.*, *Countermeasures*, Appendix B. Comments in this paper on the characteristics and performance of satellite detection and track sensors are based on the unclassified estimates in that document.
53. Canavan et al., "Debate on APS Directed-Energy Study."

for effectiveness against countermeasured threats developed slowly, and the satellite technologies needed to make boost phase effective were immature and of questionable survivability. Phase I indicated that the elements of a layered system were feasible, but also showed that key components were large, complex, and expensive, which raised questions about whether they would justify the investment needed to develop and deploy them. That was particularly so with the potentially significant but unproven SBI boost-phase layer. It appeared to be the key to an effective multi-layer system, but its large, expensive multi-SBI "garages" appeared vulnerable to ASAT attacks, a critical weakness.

These concerns were focused by preliminary DOD estimates that the full deployment of all key concepts in each of the three intercept phases could cost about $119 billion, which was felt unacceptable by the SDI, DOD, and Congress, who felt that a SDI deployment should satisfy the Nitze criteria, i.e., that they should be cheaper than the offenses at the margin, so that the USSR would not have an incentive to overwhelm them simply with numbers. With an average cost per interceptor of $119 billion/5,000 SBIs or about $24 million/SBI, it seemed unlikely that Phase I would meet that criteria. The radical rethinking and redesign needed was undertaken, which led to a different approach to space-based interceptors called Brilliant Pebbles, which promised to be adequately effective and affordable.

Brilliant Pebbles

The renewed emphasis on survivability in Phase 1 led to a series of studies that produced BP, a new version of the space-based interceptor, which appeared to be capable of surviving Soviet suppression and reducing Soviet leakage to levels modest midcourse underlays could address. BP used sensor, computer, and control technologies similar to those developed for the SBI, but did so in a manner that produced boost-phase interceptors that were survivable by design and affordable through mass production. It based them individually rather than in garages, giving each BP enough autonomy to be survivable and relatively independent of external BMC2.

Elements

BP had four essential elements. The first was a stand-alone interceptor, which could operate for long periods without external support. The second was a "lifejacket" that permitted cost effective, long-duration operation by providing the power, connectivity, and awareness needed during its dormant phase prior to attack, but was shed prior to intercept to minimize the weight accelerated to high speed. The lifejacket also provided the shielding and countermeasures needed to survive Soviet ASATs and nuclear attacks.[54] The third was advanced star trackers and computers that allowed each BP to quickly orient itself in space, select its optimal target, determine the optimal trajectory to it, and maneuver to it autonomously. The fourth was a set of integrated sensors with the range of spatial and spectral resolutions needed to pursue targets from detection through impact without external assistance. The first feature made them affordable; the second made them survivable; the third eliminated the Achilles heels from external early warning and C2; and the fourth made them capable of intercepts in the boost phase.

With this combination of attributes, the BP boost-phase layer could survive long enough to provide the level of attrition needed to reduce leakage to levels that a modest underlay could address, particularly if the weapons that leaked did not have time to deploy sophisticated penetration aids. BP was not a totally new concept, but used new technologies and designs that made earlier concepts feasible. BAMBI had postulated boost-phase intercept, but technology of its time made its interceptors large and vulnerable. SBI had modern technology, but depended on external sensors and C2 and was not survivable. BP represented design, architecture, and technology catching up with earlier concepts to produce a practical, effective, and survivable interceptor.

54. G. Canavan and E. Teller, "Strategic Defense for the 1990s," *Nature*, Vol. 344, (April 1990), pp. 699–704.

Design. The key to BP effectiveness was reducing the KV mass by an order of magnitude. That meant reducing booster, lifejacket, and deployment masses and costs by like amounts, which opened a number of options for improving responsiveness and speed. The KV mass reduction was largely achieved by miniaturizing the size of the visible, IR, and laser radar sensors needed for successful boost-phase intercepts. That made it possible to integrate them on a common optical telescope with a relatively large 15 cm aperture and integrate the outputs from their focal planes with on-board computers with computation rates comparable to those of the large mainframe computers of the time.

BP was thought to involve radically new technologies, but its real distinction was in its design philosophy. The payload mass fraction of ground-based interceptors is typically under 3 percent. That is small in absolute terms, but large enough to allow BP to use fairly heavy sensors, structures, and thrusters, which still fit on boosters of conventional size, weight, and cost. SBIs are subject to not only this 30-fold penalty for insertion into space, but add another 30-fold penalty in generating the roughly 6 km/s divert velocities needed for coverage with modest constellations. While ground-based interceptors could afford to maximize performance and let booster mass float, BPs had to minimize mass and maintain performance through advanced technology. As a result, BP designers counted grams as GBI designers counted kilograms, which led to the roughly 30-fold difference in mass between the 90 kg GBI and the 3 kg BP KVs. This mass discipline had to be applied to all BP elements, including sensors, computers, structures, and engines.

Both design philosophies are feasible. Each has had successes; each has had problems. GBI KV mass grew to a level that stressed its booster, which caused problems in its readiness tests. BP achieved the required reductions in sensor, computer, and structural masses to meet its interim weight goals, but it could not achieve the velocities, accelerations, and payload mass fractions needed from the small engines then available. That was only achieved at the end of the next GPALS phase.

Figure G.1 shows the range of frequencies and sensor fields of view (FOV) that BP used to progress from detection to impact. The abscissa is the time before impact, which progresses toward lower values to the left. The ordinate is the diffraction-limited resolution possible at each time in the BP UV/visible, SWIR, mid-wavelength infrared (MWIR), LWIR, and laser radar (lidar) sensors. BP used its wide-field-of-view 3–5 micron MWIR camera to detect missiles in boost. It then shifted to its visible and UV cameras to guide it toward the rocket's bright, compact vacuum plume. It then shifted to its 1° FOV, 10 micron LWIR imager in conjunction with its 1° cofocal lidar to separate the missile hard body from its plume. Finally, it used its lidar with a 0.1° FOV to provide range information for intercept and imaging for aim point selection.

That sequence makes it possible to reliably shift from detection, to plume, to the booster while maintaining track through the missile's strong and varying accelerations during its final seconds of powered flight. BP sensors spanned the spectrum from visible to LWIR. Their visible, SWIR, MWIR, and LWIR focal planes were updated frequently with the best available commercial detector arrays in a quick turnaround approach since adopted by other space missions, notably NASA's successful asteroid and Mars missions. These optical elements, including the BP telescopes, cameras, and lidar were used in the DOD Clementine experiment that remapped the Moon at high resolution in wavelengths from the visible through the LWIR. Clementine also performed an experiment conceived en route to use scattered communication radio signals to detect water on the Moon's south pole.[55]

The unique challenge of the boost phase is intercepting strongly accelerating missiles in large, bright plumes. Delta 180 showed that it is possible to separate the hard body from the plume, and Delta 181 and 183 showed that it is possible to intercept from a range of geometries. Doing so requires PN to be corrected for the missile's increasing acceleration at end of boost. In boost-phase intercepts, the target acceleration A generally has a significant component transverse to the LOS. The corrective accelera-

55. S. Nozette, "Clementine Mission," *Science*, Vol. 226, (December 1994), pp. 1835–1862.

tion needed depends on the guidance law used (See Appendix E). For PN, the interceptor's relative acceleration is initially small but increases to $K/(K-2)$ near impact. For $K = 3$, the interceptor could need an acceleration about 3 times that of the missile. APN has its maximum acceleration initially, where it is $AK/2$, which would give a relative acceleration of about 1.5.

The BP flies in from a distance r or about 1,000 km at a velocity v of about 7 km/s, which takes a time $t = r/v = 140$ s. It would have to supply a velocity increment a fraction f of the missile's velocity $V = 7$ km/s, so its transverse acceleration would be on the order of $fV/t = fV/(r/v) = fV^2/r = 1$ g for $f = 0.2$ and $v = V$. *A priori* or measured knowledge about the missile would reduce the acceleration required to about the average acceleration of the missile. Optimal guidance has smaller miss distances but similar accelerations. The velocity increment needed to intercept an accelerating target with PN is $\Delta V = KAT/(K-1)$. For $K = 3$, $A = 10$ g, and $T = 50$ s, $\Delta V = 1.5 \times 0.1$ km/s^2 x 10 s = 1.5 km/s, which would be stressing. That for augmented proportional navigation is precisely half that, or 0.75 km/s.

If the missile's acceleration is known or can be estimated accurately, APN can be supplemented with that knowledge to support intercepts with maximum accelerations roughly equal to the missile's average acceleration. Such an approach is robust, if supported by reliable data. *A priori* information can provide stage times and intervals and prevent discontinuities in estimates. Delta 180, 181, and 183 demonstrated that such information could be gathered and used for the specific boosters tested. It remains to be shown that the approach is generally applicable and sufficiently accurate for rogue missiles, for which less design and propulsion information is available. The high frequencies involved in intercepting accelerating missiles put a premium on accurate measurements or estimates of range and time to go, which determine the guidance loop response frequencies needed. To ease those tolerances, BP used a lidar to reduce errors in range, closing velocity, and time, which was tested on Clementine active measurements of the Moon's surface. BP computers had enough capacity and flexibility to use *a priori* and real time measurements of missile type, expected acceleration profiles, and inter stage times.

The key to reducing optics and sensor sizes and masses was the short BP range to target due to the small separations implied by the large constellation scaling discussed above. For $N = 2{,}000$ BPs, their separation was about $2R_e/\sqrt{N} = 300$ km, which placed BP sensors 100 times closer than those of early warning satellites in geosynchronous earth orbit (GEO). Thus, their apertures could be reduced 100-fold without impacting their resolution or radiometrics, which reduced the weight of their sensors 100^2-fold to about 1 kg. Reductions in computer size and weight were the result of Moore's law that processing speed and memory density double every 1.5 years, which had not been systematically exploited in SBI. The result was a few kilogram KV on a roughly 100 kg booster that could maneuver in front of an accelerating ICBM before burnout. That was a 100-fold improvement over the masses achieved by BAMBI and a 10-fold improvement over those achieved by SBI. The BP KV used in the tests performed before the cancellation of GPALS had a mass of about 4 kg, which was about 50 percent over its ultimate objective of 2.5 kg and a factor of 15 less the 60 kg of a GBI KV. The reduction resulted from improved technology and design philosophy.

Performance. As the performance expected of BP differs from that of other concepts, including SBI, it is discussed in detail in Appendix G. The key parameters affecting performance and effectiveness are flight time T_{flight}, delay time T_{delay}, maximum velocity V, acceleration A, and cost C. BP initially intercepted missiles while they were in boost or their buses while they were deploying multiple weapons. Their engagements were later extended to the midcourse and terminal phases. For the two original missions, the BP's maximum flight time was determined by the difference between the missile's boost or bus time and the BP's release time. BP's principal target, the liquid-fueled SS-18, had a boost phase T of 300 s and a roughly equal bus phase, so they were accessible in boost and bus phases for a maximum of $T_{bus} = 600$ s. The solid fueled SS-25 deployed later and in lesser numbers had a boost

phase of 150 s, roughly the minimum then possible. It was a single-warhead missile with no bus. Other Soviet missiles fell between these extremes.

BP's release times had three major components: the time for sensors to detect the missile, the time for them to establish and report a valid track, and the time for the C2 system to release the BP and communicate that decision. DSP satellites have revisit times on the order of 10 s and operate in SWIR water bands to reduce ground clutter, so they detect missiles as they emerge from the water vapor at about 10 km or break through clouds slightly higher and later.[56] Its C2 system acts on that information to classify strategic missiles and release a missile alert in about a minute according to experience in the 1991 Gulf War, where prompt missile warning messages were released for the SCUD missiles fired at both theaters.[57] The sensors on its Space-Based Infrared System (SBIRS) successors should operate faster and could see deeper into the atmosphere, so release times of about 30 s should be possible. Because of its on-board sensors, BP could detect and track faster, and because it was at a lower altitude, its sensors could rapidly gather enough signal for detection using bands that see deeper into the atmosphere. Because each BP had a small field of regard, it could revisit detections more rapidly, which allowed it to verify detection and establish track in seconds. The only delay would be for release, which could be based on BP detections rather than DSP or SBIRS.

BP's flight time to a SS-18 booster was $T_{flight} = T - T_{delay}$, which with $T_{delay} = 60$ s would be about $T_{flight} = 240$ s. Its flight time to the bus was $T_{flight} = 540$ s. The maximum velocity V from the engines initially available was about 6 km/s. Absent other delays, that would give ranges to the booster and bus of 1,440 and 3,240 km, respectively. The relatively long flight times for heavy Soviet ICBMs produced BP ranges comparable to those of SBIs and midcourse GBIs. However, the engines available were also limited in acceleration, which imposed additional delays. The time to reach maximum speed $V = 6$ km/s at average acceleration $A = 6$ g is $T_{accel} = V/A = 100$ s. During that time the BP's average velocity is $V/2 = 3$ km/s, which reduces its range by $V^2/2A = 300$ km. Accounting for finite acceleration, the BP's maximum range to a booster was $V(T - T_{delay} - T_{accel}/2) = 6$ km/s(300 – 60 – 50 s) = 1,140 km, and the range to its bus was about 6 km/s(600 s – 50 s) = 2,940 km.

The penalties for finite acceleration are more serious intercepting the booster than the bus, but are only 10–20 percent for heavy missiles like the SS-18, so early BP could function effectively against them even with modest accelerations and velocities. Figure G.2 shows BP range to booster versus BP velocity and acceleration for 30 s delays. For a velocity V of 4 km/s, the maximum range from a SS-18 booster was about 800–1,000 km, weakly dependent on acceleration. For today's $V = 6$ km/s, ranges increase to 1,200 to 1,400 km for 4 to 8 g accelerations. For the $V = 8$ km/s possible with advanced technology, ranges increase to 1,300 to 1,750 km, although additional velocity would not improve the range of the 4 g BP, which would spend little time at maximum speed.

Constellation Scaling. Boost-phase intercept requires enough interceptors within range to reach the missiles before boosters burn out. Precise calculations of the constellation sizes and inclinations needed to satisfy this condition requires computation of optimal trajectories for all BP, but it is possible to provide reasonably accurate estimates of constellation sizes with geometric considerations. At the lower accelerations and velocities of early BPs, delays for release and acceleration were significant, so it is necessary to include them at least approximately in the analysis.

A BP with divert velocity V could reach any point within a roughly circular area of radius $r = V(T_{flight} - T_{accel}/2)$ around its initial trajectory. The number of such circles required to cover the Earth is roughly $(2R_e/r)^2$. Soviet missiles were distributed over an area of effective radius $R = 1{,}600$ km, so interceptors up to about $R + r = 1{,}600 + 1{,}200 = 2{,}800$ km away could make boost-phase intercepts, which gave an absentee ratio of $(2 \times 6{,}400 \text{ km}/2{,}800)^2 = 20$, which would require 12,000 BPs for sin-

56. Richelson, *DSP Satellites and National Security*, p. 10.
57. *Ibid.*, pp. 161.

gle coverage of 600 missiles launched from this distributed area. Figure G.3 shows constellation size as a function of maximum velocity. The distributed launch area makes the impact of V and A smaller than that on constellation size. At V = 4 km/s, the constellations require about 25 BPs in orbit per simultaneously launched missile. At 6 km/s, they require about 18–22. At 8 km/s they require about 14 for 6 to 8 g and 18 for 4 g. Thus, extending BP range 13 percent with 6 g interceptors would reduce the absentee ratio about 30 percent, and increasing acceleration to 8 km/s would reduce it 50 percent. For 600 missiles, an absentee ratio of 15 would require 9,000 BPs, which is 10 percent larger than the number of ideal SBIs discussed earlier due to this calculation's inclusion of the BP's finite acceleration.

BP could also intercept in the bus phase, which would increase the range from which BP could intercept to about 1,600 + 3,000 = 4,600 km, which would give an absentee ratio of $(2 \times 6{,}400 \text{ km}/4{,}600)^2 = 8$. However, some buses would have already released RVs. Adding bus-phase intercepts would increase the total number of RV kills by a factor of about 1.6 over those from boost alone, which reduces the number of BPs needed to about 9,000/1.6 = 5,600. Soviet deployments sought to stress space-based defenses by reducing missile boost time, missile launch area, and warheads per missile, especially with the short burn SS-25 in concentrated basing, but its attempts did not significantly impact these estimates before dissolution of the Soviet Union, as its force was still dominated by heavy, distributed missiles.

Cost. Cost effectiveness at the margin was a requirement for U.S. defenses against Soviet threats because significant offensive escalation was thought to be an economically viable option for the USSR. Thus, interceptor cost was a key design variable. The KV mass was to be reduced by miniaturizing sensors and computers to about 2 kg. Early test programs achieved a mass of about m = 4 kg, which is used for estimates below. BP engines had a specific impulse I_{sp} of about 300 s, i.e., an exhaust velocity $c = gI_{sp}$ = 3 km/s, for which the ideal rocket equation gives an estimate $M = me^{V/c}$ = 30 kg for the initial rocket mass M needed to accelerate a m = 4 kg KV to V = 6 km/s. Real engine performance, staging, and structures could roughly double that estimate. The KV cost is estimated to be about C_K = \$500,000, and the cost to launch such payloads to low Earth orbit is C_V = \$20,000/kg, so the total cost of a 6 km/s BP on orbit is roughly $C = C_K + MC_V$. = \$500,000 + 30 kg x \$20,000/kg = \$1.1 million, which is roughly the cost the DOD carried through the Defense Acquisition Board (DAB) discussed below. Costs vary with BP speed. A low speed BP would cost about \$500,000, about half that of the nominal BP, but would require much larger constellations. A 9 km/s BP with appropriate technology would cost about twice as much. With current technology, they could cost much more.

Figure G.4 shows the on-orbit constellation costs per missile launched for boost-phase engagements. They are proportional to the product of constellation size N and BP cost C. For large distributed launch areas, costs are insensitive to A and V less than about 8 km/s. For BP velocities of 4 to 6 km/s and accelerations of 4 to 8 g, the cost is about \$20 million per simultaneously launched missile engaged in boost, so engaging 600 Soviet missiles in boost would cost about \$12 billion. Engaging them in the boost and bus phases would cost about \$12 billion/1.6 = \$7.5 billion. That would give the defense roughly 8:1 leverage over an offense costing 600 missiles x \$100 million/missile = \$60 billion, which would justify GPALS's decision to address a large fraction of the threat with BP. These estimates scale with the cost of the BP KV. Those used above are based on analysis, laboratory tests, and limited field tests that went into the DAB. Cost estimates from programs at this stage are not always reliable, so the costs for a production BP will be uncertain until development and tests are performed. The cost exchanges above are uncertain to a similar extent.

Brilliant Pebbles in Midcourse and Terminal

BPs that are unable to reach the boost phase can still engage in midcourse. The BPs used would have lower absentee ratios and effectively zero cost in midcourse, since they are the BP that were in the wrong place to reach the boost-phase defense. However, they could face countermeasure and discrimination issues similar to other midcourse system. For intercontinental launches the time available for midcourse engagements is about 1,800 s, so about half of the BP constellation could contribute (See Appendix G). Figure G.5 shows the number of kills in each phase out of 3,000 RVs launched on 300 heavy missiles with 10 RVs each with the 300 s boost and bus times of SS-18s launched from distributed launch areas. The BPs have maximum speeds of 6 km/s, accelerations of 10 g, and delays of 30 s. The number of boost-phase kills K_{boost} increases linearly to about 1,000 RVs (i.e., 100 missiles) at N = 2,000 BPs, which gives an effective boost-phase absentee ratio of about 2,000/100 = 20, in accord with the above estimates. The number of kills in the bus phase is about 1,750, which is about 70 percent larger. At N = 2,000 BPs, the total number of boost and bus kills is about 2,750.

The downward slanting curve is the number of RVs that leak through the boost and bus phases, which is about 2,700 at N = 250 and approaches zero at N = 2,000. The bottom curve is the number of RV kills in midcourse, assuming that BPs have the same 0.9 kill probability there as in other phases. It lies on top of the K_{boost} curve for N < 1,600 and on top of the leakage curve for larger N. At N = 1,600, the boost, bus, and midcourse layers contribute comparable numbers of kills. About 800 RVs leak through boost and bus, but enough BPs engage in midcourse to reduce leakage through all three layers to 70 RVs. By N = 2,000 the leakage is too small to compute accurately.

Figure G.6 shows the number of RV kills for 20 missiles with 10 RVs apiece, which is roughly the number of RVs possible from a single nuclear ballistic missile submarine (SSBN). The threat is about the same whether it is on station, in bastion, or in port, because in any case it is effectively a point launch, which is stressing to BP constellations. For that reason K_{boost} is only about 20 at N = 180 because the leverage of boost-phase intercepts falls as the launch area decreases. However, number of bus-phase kills increase to about 160. As the number of boost-phase kills is reduced, there are more opportunities for kills in the bus phase. The two phases together could kill all 200 RVs by N = 200, which would give an absentee ratio of 200/20 = 10. At N = 120, where K_{mid} peaks, the bus kills are about 7.5 times those in boost.

The RV kills in the boost, bus, and midcourse layers are 14, 104, and 60, which add to about 90 percent of those launched. About 80 RVs would leak through the boost and bus layers but only 20 RVs through all three. By N = 180, none would leak through all layers. Because about 2,000 BPs would be required to blunt the 300 ICBMs included in the Phase I threat, the SBI constellation would be oversized to handle the number of RVs from a single SSBN, which was a lesser included threat. Because SSBNs on station or in bastion would be widely separated, they would be addressed by independent parts of the BP constellation, so the fact that a number of them were at sea simultaneously would not impact these results. However, if SSBNs in port could launch, the fact that a number of them would be in close proximity in a port would multiply the SBI required for a given level of performance in proportion to the number in one port.

Figure G.7 shows the number of boost and midcourse kills for the launch of 100 single-weapon, 150 s burn time missiles like Soviet SS-25s. Because their launch area is not widely distributed, boost-phase kills are reduced. Because the SS-25 is a single-RV missile with no bus, the only other defensive layers are midcourse and terminal. At N = 450 BPs, the boost phase contributes about 12 kills for an absentee ratio of about 450/12 = 38, which is about twice that for distributed heavy missiles. The 100 SS-25 missiles could be addressed in boost, but would require constellations of about 100 x 38 = 3,800 BPs with 10 g accelerations and short delay times. The 2,000 BPs constellation for the heavy, distributed missiles would engage about 100 x 2,000/3,800 = 53 of them. The SS-25 only had a single weapon and no bus, so it could not release decoys and could be intercepted in midcourse. Figure G.7

shows that about 400 BPs would intercept all of them in midcourse. However, its replacement the SS-27 has a bus with 3 RVs that can deploy credible decoys, so addressing it is more problematical.

The estimates above lump midcourse and terminal BP kills together. Although they involve different phenomena, their total depends on the number of BPs that can reach the RV in midcourse or terminal, which is a largely geometrical calculation that combines their total contribution, which can be divided between midcourse and terminal depending on the problems expected in each. Midcourse BP intercepts resemble those in boost, with detection and track from boost providing the error basket needed to begin the intercept. Preliminary analyses indicate that BP's sensors would be relevant for midcourse, although the LWIR camera and lidar should be adjusted for the smaller signatures expected from RVs. The main complication is the presence of decoys in midcourse, for which it is not clear that BP's UV/visible, SWIR, and MWIR sensors would suffice. Its LWIR sensor was simpler than the GBI's which might be partially compensated for by its lidar.

For terminal intercepts, the BPs would fly to the RV's predicted atmospheric pierce point, aerobrake there, and use their transverse acceleration to maneuver in front of the RV. Depending on the altitude of engagement, the BPs might be able to take advantage of atmospheric drag for discrimination. Its velocities should be too small to degrade its sensors, so it might represent a version of IR sensor that could intercept over large footprints without redout.

Sensors, Command and Control

BP presented unique command and control problems and options. In boost-phase intercepts, time is of the essence, so it is necessary to detect launches, report them, and receive release authorization in delays of 10s of seconds under all conditions in the presence of deliberate interference. That appeared difficult to accomplish with heritage sensors and C2 assets, but the BP provided intrinsic capabilities that could reduce dependence on outside detection and C2. Its visible-to-IR detection and track sensors could detect launch earlier and reliably, support warning and release, and form missile tracks quickly. Each BP could evaluate its optimal allocation strategy, independent of information from or calculations by vulnerable external sensors or computers. And its dense constellations afforded the opportunity for a built-in set of communication nodes well suited to highly redundant communication networks. All BP needed for C2 was a short release message from the human in the loop, which could be delivered by survivable, redundant, low bit rate, distributed communication systems carried by the BPs themselves.

Depending on the threat, BP constellations were optimally deployed 400–600 km above the surface of the Earth in order to reach missiles during their boost or bus phases. Since BP sensors were about 100 times closer than geosynchronous satellites, 100-fold smaller and cheaper sensors could perform the same detections with comparable resolution and radiometrics. Their adequacy was assured by direct scaling on altitude from earlier sensors. These capabilities eliminated the need for the few large SSTS, GSTS, and BSTS warning and track satellite sensors. Their vulnerabilities were recognized by the shift to BP, so those satellites from earlier phases were replaced by "brilliant eyes" (BE), which were essentially BP sensors deployed on distributed constellations of small satellites.

Survivability

Survivability was a key issue in addressing the Soviet threat, which could devote several hundred MT warheads to suppressing defenses before launching its main strike. Survivability is a key feature of many space systems, so the approach used to assure BP survivability is discussed in detail in Appendix H. The principal mechanisms available to BP were hardening, maneuver, decoys, and self-defense. The last was best employed after the others had been optimally combined.[58] Autonomous situational awareness and response were required due to the short attack times of direct ascent nuclear

58. Canavan and Teller, "Strategic Defense for the 1990s."

ASATs.[59] This discussion concentrates on the results of survivability tradeoffs, but the process that led to them is also of interest.[60]

A nuclear explosion of yield Y at a distance K from a BP produces fluence $J = Y/4\pi K^2$ that the satellite must survive to be effective. If it is hardened by applying a thickness Δ of material over its surface, the hardening material mass is $M_{hard} = 4\pi\rho D^2 = CM^{2/3}/K^2$, where C depends on Y and the specific hardening material. The cost to harden against nuclear ASATs increases as $1/K^2$, where the miss distance K is largely under the control of the attacker. These penalties can be large, particularly for heavy satellites with large areas.

A BP can maneuver to make K larger, at the cost of additional fuel. To miss the ASAT by a distance K through a divert executed at range R, the BP must deflect from its initial trajectory by an angle $dV/V = K/R$, which takes fuel mass $\Delta M = 2MK/R$. The BP could reduce ΔM by maneuvering at large R, but that would also give the ASAT more time to correct for the divert.

BP can use light decoys, although it has to maneuver enough to hide itself in them. If N decoys are separated distances of K to force the ASAT to devote a nuclear weapon to each, the overall diameter of the cloud is $N^{1/3}K$, so the mass to maneuver the BP and decoys is $2N^{1/3}K/R(M + mN)$. The masses for the BP, hardening, and maneuver all involve BP mass, so the expected loss is reduced by minimizing its mass.

The attacker must commit a warhead with a mass M_{wpn} of about 100 kg to each decoy, and even light decoys are credible on the short time scales available to the sensors on small ASATs. Thus, the attacker must expend an effective mass of $(1 + N)M_{wpn}/E$ to negate the BP, where division by the absentee ratio E recognizes that negating an ASAT overhead also negates E BPs elsewhere in orbit, which is important in extended engagements and deployment. BP's goal is to make the mass to attack much greater than that to defend. Figure H.1 shows these masses as functions of N for $K = 1$ km, $R = 300$ km, and $E = 10$. The ratio of attack to defense masses is less than unity for small N, but reaches 4 by $N = 30$ and 10 by $N = 100$. For larger N, the ratio saturates. For light decoys the ratio of attack to BP mass M_{att}/M_{def} is about NM_{wpn}/EM. Thus, the defender should use many light decoys, maneuver at long ranges, and maintain low absentee ratios, while the attacker should discriminate, minimize range, and use small yields.

Even for large N, the BP mass is a significant part of the defender's expected loss, because by expending $N + 1$ ASATS, the attacker can be assured of killing the BP. The BP defending itself with hardening, maneuver, and decoys alone surrenders advantages in ground-based sensors, C2, and range to the ASAT, so the outcome of such engagements is not clear. However, if the BP uses hardening, maneuver, and decoys as indicated above to complicate the ASAT's attack, self-defense can then be used effectively to reduce the impact of ground-based sensors and C2. The BP can use a small self-defense interceptor to negate the simple ASAT aimed at it. Then the attacker's cost is unchanged, but the BP survives, which removes the BP's mass from its expected loss. That reduces its expected loss to the sum of the masses for decoys and hardening, which is $mN + M_{hard} \approx M_{hard}$. In the examples above, the BP mass was the dominant loss. Preventing its loss increases exchange ratios by about an order of magnitude for a range of numbers of decoys, which gives it significant margin over the attacker.

This approach works for satellites of modest mass. Figure H.2 shows how the components of the mass for defense varies with satellite mass M for $N = 30$ light decoys and an attack mass of 300 kg. The defense mass increases from about 20 kg at $M = 1$ kg to 1,000 kg at $M = 1{,}000$ kg. For small M, the attack/defense mass ratio favors the defense by about an order of magnitude. It drops to unity at a mass M of about 300 kg, and to about 0.2 by $M = 1{,}000$ kg. Achieving favorable mass ratios at large

59. Baucom, "Ballistic Missile Defense," pp. 25–26.
60. R. Bennett, "Brilliant Pebbles: Amazing New Missile Killer," *Reader's Digest*, September 1989, pp. 128–133.

M would require the deceptive deployment of large numbers of large decoys, so there appears to be a limit of a few hundred kilograms for which hardening, maneuver, decoys, and self defense provide spacecraft survivability.

Summary of Developments During SDI

The Exploratory Phase of SDI supported a wide-ranging investigation of the technologies that could be useful for missile defense. SDI Phase I responded to the continuing Soviet offensive buildup by attempting to develop a system that could blunt the leading edge of a Soviet counterforce attack. While that defense could allow significant leakage and still have a positive effect on stability, it could not be vulnerable to surprise or suppression or it would have the opposite effect. Phase I started but did not develop midcourse HTK interceptors. It did make fundamental advances in the design and development of endoatmospheric HTK interceptors that provided alternatives to earlier nuclear systems. It also made advances in the high frequency radars and satellite sensors needed to give those interceptors adequate trajectory information and in the BMC2 technology required to command them.

Phase I also made significant advances in DEW for boost and midcourse intercepts, but its time scale was not commensurate with that for the development of likely threats. Moreover, large DEW platforms were found to be vulnerable. They could not survive the large attrition attacks that the USSR could execute. Pop-up versions might reduce some of those problems, but would be heavier and take longer to develop. Although NPB and SBL could be capable midcourse active discriminators, neither could defend itself or maneuver out of the way of ASATs directed toward their predictable orbits.

SBI was developed from earlier concepts by introducing the computer and propulsion systems lacking earlier. Being small and inexpensive, SBI had the right characteristics for boost-phase intercept by dense constellations of affordable interceptors. It evolved rapidly from SDI to service to contractor development, but that progression followed the usual military satellite management process, which resulted in large satellites that relied on external warning sensors, communication, and C2. That introduced unacceptable vulnerabilities, because it made the SBI buses large and expensive enough to be valid targets and made their release and C2 dependent on large, vulnerable sensor and communications satellites that were less survivable than the SBIs themselves.

The BP was created as a small, separate program to address the vulnerabilities that the SBI program recognized but lacked the flexibility to correct. When those vulnerabilities became debilitating, BP was substituted for SBI. BP was the one concept developed in Phase I that appeared capable of surviving long enough to engage Russian missiles in boost. Its autonomy, achieved through new levels of sensor, computation, and system level integration, was essential to its survivability and effectiveness. BP carried its own detection sensors and C2, which freed it from dependence on current satellite systems. Adequate survivability was assured through hardening, maneuver, decoys, and self-defense, in which BP's small mass was the essential element. Large, non-maneuvering, non-deceptive satellites have little chance of survival or effectiveness in the boost phase under strong suppression. BP is the one space system that was survivable against determined suppression attacks, primarily because that was its principal design objective.

BP could detect and pursue missiles with a capable suite of passive and active sensors and robust algorithms for identifying and tracking missiles hidden by plumes, which were tested in the laboratory and field. BP was predicted to perform well against distributed Soviet heavy missiles of the time and to degrade gracefully against missiles that burned faster, released MIRVs faster, or were launched from more concentrated areas. It was predicted to be cost effective by an order of magnitude for the mission of blunting counterforce attacks for which it was designed. It did not require foreign bases, over-flight negotiations, status of force agreements, or base use negotiations. Its ground facilities and C2 were minimal. It did not need large crews, commands, ships, or planes. It did not depend on the

large warning and battle management command, control, and communications (BMC3) systems and satellites that midcourse and terminal systems require, which have increased their cost.

BP significantly advanced defense and space technology. It developed a common set of technologies for lightweight KVs that could have been used to reduce the size and weight of interceptors for the other layers. It incorporated all of the passive and active sensors needed for reliable intercepts in all layers, which could have been used to extend the range and capability of land- and sea-based boost and midcourse interceptors as well. Using the BP as a common KV could have reduced development expense and time. Since BP KVs masses were a few kilograms, they could even effectively engage precision decoys using antisimulation that remained credible to midcourse discrimination sensors. Thus, BP was well suited to act as the centerpiece of the high quality protection system needed for residual threats after the Soviet Union collapsed, which are discussed next.

Global Protection Against Limited Strikes

During the Cold War, the main threat was the massive launch of the 1,000s of missiles and weapons in the Soviet arsenal. With the end of the Cold War, that threat receded, and the focus of missile defense shifted to accidental or unauthorized attacks by Russia or China and limited attacks by rogue nations. Thus, in 1991–1993, SDI was redirected to Global Protection Against Limited Strikes.[61] Adequate defenses against such limited strikes could be constructed from the elements of the SDI program, but significant redirection was needed. Two words in GPALS's title were particularly important: "Global" meant missiles were to be negated no matter from where on the globe they came, as it was to cover not only threats to the United States, but also threats to its deployed forces, allies, and friends. The objective of Phase I was deterrence, for which significant leakage was acceptable. Phase I could have enhanced deterrence of the launch of 1,000s of weapons even if half of them could have leaked through. "Protection" meant GPALS had to achieve low leakage. Thus, every attacking missile and weapon had to be engaged, some many times. In doing so it had to provide "high confidence of very low leakage," which has generally remained the goal of successive phases.

Scope

GPALS emerged from a March 1990 review of SDI for Secretary of Defense Richard Cheney by Ambassador Henry F. Cooper, who had lead the Defense in Space negotiations in Geneva and subsequently became the Director of SDI and had the responsibility for implementing the recommendations of the review. President Bush endorsed GPALS in January 1991. The Gulf War validated its recommendation to make TMD part of a global architecture against the coercive threats expected to become an increasing problem. GPALS included accidental and unauthorized launches in part to incorporate Senator Sam Nunn and Representative Les Aspin's support for an Accidental Launch Protection System (ALPS). That merger overcame several political hurdles, but it did not lessen their opposition to space-based systems, which kept BP from being one of the elements identified for immediate deployment. Once GPALS became global, the Navy accepted a serious missile defense role for the first time, which it maintained throughout GPALS and the Clinton Administration, with the help of Congress.[62]

While the size of needed defenses could be reduced along with the threats, significant deployments were still needed in each defensive component because even the residual threats were still of significant size. While the launch of a single rogue missile without decoys might not require more than a few midcourse interceptors, an advanced rogue threat or a single heavy Soviet missile could generate 10s of RVs and 100s of decoys, and the launch of a dozen such ICBMs was possible given the centralized Soviet launch control system. A single Russian SSBN could launch 20 missiles with 10 RVs each and dozens of decoys per RV. Figure G.6 shows the number of boost- and bus-phase RV kills for a SSBN launch as a function of BP constellation size. RV kills would largely come from the bus and midcourse phases, as SSBN launches are essentially point launches, which minimize the contribution from boost. The launch of a dozen heavy ICBMs would be a lesser-included threat because the

61. *Report to Congress: Conceptual and Burden Sharing Issues Related to Space-Based Ballistic Missile Defense Interceptors* (Washington, D.C.: U.S. Department of Defense, March 1992).
62. H. Cooper, private communication, April 1, 2002.

ICBMs are spread over a larger territory, which allows more kills in boost. Figure G.5 shows that 2,000 BPs could kill 3,000 RVs, which scales to about 70 BPs for 100 RVs. The launch of SS-25s would require a constellation comparable to that for a SSBN, although its lack of a bus would permit more RV kills in midcourse. GPALS's stressing threat was the number of RVs a Russian submarine could launch from the waters off Bermuda, about 1,200 km from Washington and New York.

Concerns with possible qualitative and quantitative improvements in missile threats from the People's Republic of China, proliferation of missile and countermeasure technologies to rogues, reduced timelines for threats from new threat areas, and the possibility of multiple ICBM and SLBM launches required defenses to retain layered systems with significant capability in each layer. The BP boost layer retained its significance because of its ability to reduce accidental or limited launches to levels that modest midcourse underlays could negate. GPALS represented a shift in emphasis to layered defenses against quantitatively reduced but qualitatively undiminished threats. It was based on the integration of previous SDI technologies. Its primary goal was to complete development and deployment of their key elements as soon as possible, with an initial deployment by 1996.[63]

Elements

GPALS's main elements were a Ground-Based Tier for Homeland Defense (Limited Defense System or NMD), a TMD in each region, and a Global Defense Layer[64] to be supplied by BP, supported by a fully integrated joint global command and control system. These elements have been recognized by the subsequent program, although not all programs have continued the development of all elements.

Limited Defense System

The principal elements of the Limited Defense System (LDS) were the Ground-Based Interceptor and Ground-Based Radar. While the GBI was not completed during GPALS, related HTK interceptors underwent significant development. ERIS and HEDI were combined into an Endo-Exoatmospheric Interceptor (E2I), intended to use HEDI's advances in aerodynamic window cooling to provide an interceptor that could operate in both the late midcourse and early terminal phases of a fully layered system. The Exoatmospheric Kill Vehicle (EKV), which was put under contract in 1991 and flown in 1997, is the basis of current KVs.

GPALS initiated development of the GBR, which provided the primary track and discrimination for the later TMD and NMD programs. GBR was the first missile defense radar to shift to the higher frequencies that should be less susceptible to nuclear blackout and refraction issues that had degraded earlier systems. It was not completed during GPALS, but became the centerpiece of later systems.

Satellite systems underwent significant restructuring, largely to take advantage of the greater survivability and performance of small, low-altitude platforms and sensors. The FEWS early warning satellite, derived from the earlier BSTS, was modified into a still simpler ALARM system to support TMD. ALARM was in turn later redesigned into the current SBIRS-High. This history reflects several competing objectives. At the outset of SDI, BSTS was intended to detect and track buses and RVs from geosynchronous orbit. When it was realized that was not possible with affordable optics, that requirement was dropped and its design simplified. FEWS and ALARM sought to provide better theater warning through improved technology.

Watching theater missiles all the way to burnout would require a departure from DSP's spinning configuration, but the technology for sensors that could simultaneously stare at the whole Earth was not available. SBIRS-High was designed to watch missiles to burnout and see dim missiles under poor

63. U.S. Congress, Missile Defense Act of 1991 (Washington, D.C.: U.S. Government Printing Office, 1991).
64. *Report to Congress: Plan for Deployment of Theater and National and Ballistic Missile Defenses* (Washington, D.C.: U.S. Department of Defense, March 1992).

meteorological conditions by using focal planes of the sizes available and repointing its telescope to successively view all theaters of interest with a revisit time of a few seconds. Tracking dim targets to burnout requires staring sensors, but seeing missiles in bad weather proved to be less stressing than previously thought. These revisions were driven partly by a better understanding of target and background signatures and partly by a desire to reduce costs.[65]

To capture the improved performance, cost, and survivability possible with many small satellites at low altitudes (See Appendix I), large SSTS tracking satellites were reconfigured into the non-interceptor version of BP called BE. Subsequent to GPALS, BE evolved as an Air Force program back into larger Space and Missile Tracking System (SMTS) satellites intended to support TMD, which then evolved into the current SBIRS-Low element of the MDA program. GSTS was terminated when GPALS was initiated, because its pop-up capability was not needed to complement the equally or more survivable BE.[66]

An application of early warning satellites important to the Gulf War was DSP's ability to detect and track theater missiles. It had been known for decades that the DSP was sufficiently sensitive to detect dim Soviet N-6 SLBMs, whose amine fuels radiated poorly in its detection band. DSP also saw tactical missile launches in the Yom Kippur war and had a high detection probability for SCUDs in the eight-year "war of the cities" between Iran and Iraq.[67] Thus, it was not a surprise that it had a high probability of detecting SCUDs and predicting their rough aim points in the Gulf War. That capability improved markedly when the observations from two or more satellites were fused to increase the number of observations, improve look angles, and reduce predicted error baskets and ambiguities.[68] That combination of improvements provided adequate warning to PATRIOT and troops in theater. While PATRIOT performance as an interceptor was marginal, DSP's performance demonstrated that wide area warning and missile defense should be an essential part of the support to any future expeditionary force.

DSP's technical performance was adequate in the Gulf War, but there were complaints about delays in the distribution of its data to theaters. The Chief of Naval Operations (CNO) wrote to the U.S. Space Command about its failure to provide the crypto keys needed to decipher data. The Army started developments of mobile ground stations that could receive downlinked DSP data directly. The Army and Navy later combined their effort to produce the Joint Tactical Ground System (JTAGS). As Space Command maintains a central fusion center called TALON SHIELD, each service now has its own theater reporting system. Each gets the other's theater reports as well as those from Space Command's central processing system, which was created to collect, fuse, and process the observations from all DSP, and later SBIRS, warning satellites. Thus, each launch is now reported by several theater stations, not all of which receive the data from all satellites. The benefits of fast and assured access are said to outweigh the confusion that can arise from multiple, conflicting inputs.[69]

Theater Missile Defense

The main elements of TMD were Corps SAM, Extended Range Interceptor (ERINT), PATRIOT, THAAD, TMD-GBR, and ship defense elements, which formed a progression in intercept range and altitude capability that permitted layered defenses in theaters for the first time. The low-altitude endoatmospheric interceptors underwent significant modifications, and ERINT provided the technology for both THAAD and PAC-3, which were completed in the subsequent TMD phase discussed below.

---

65. J. Richelson, *DSP Satellites and National Security*, p. 198.
66. *Ibid*., pp. 117–119.
67. *Ibid*., pp. 67, 72, and 159.
68. *Ibid*., pp. 161 and 172.
69. *Ibid*., pp. 189–190.

Global Defense Layer

The main elements of the global defense were BP, BE, and their supporting BMC3. Although BP was not approved as part of the initial deployment, it had the highest leverage of Phase I systems so it became the highest priority element of GPALS. A BP KV weighing about 50 percent more than the ultimate design was developed and underwent significant testing. BP's integrated sensor package weight was reduced to roughly its design goal. The remaining needed reductions were in structural and propulsive elements. Lightweight structures with integral tanks were developed with the requisite rigidity for maintenance of LOS stability during repeated thrusting. Early BP performance was limited by the engines available, which had poor payload mass fractions at small size. Thus, an effort was started on the development of more efficient engines. Preliminary results were achieved during GPALS. Subsequently, such engines have been scaled to sizes large enough to power rockets about the size of BPs to several kilometers per second while maintaining high payload fractions.

Development and Testing

GPALS's primary goal was to deploy defenses against accidental, unauthorized, and rogue threats. Components inherited from SDI needed significant development and testing to bring them to the level of confidence needed to support a decision to deploy. BP, as the newest element, required the most development, which is discussed below. Development and testing of the ground-based systems is discussed in the context of the later programs within which they became mature.

Space-based interceptor technologies had achieved significant levels of development in the laboratory and field during the BP phase of SDI Phase I. In 1988, BP was studied in detail by the DSB, JASONS, and other DOD review groups, on the basis of which it passed its DAB and became a Major Defense Acquisition Program (MDAP).[70] These reviewers commended the BP technology and design philosophy and recommended exporting BP technology to other systems. That was never accomplished, possibly because of institutional resistance. BP developed a set of technologies for common, lightweight KVs that could be used to reduce the size and weight of other interceptors, with the explicit intent of feeding the more advanced technology from BP into the GBI as it matured.

These reviews and their interpretation by the press were an object lesson in how not to conduct technical assessments of advanced systems partly in public, which has not yet been fully internalized by the DOD. The JASON's summary statement that the BP had "no fundamental flaws" was interpreted by the DOD as a very positive assessment, but was interpreted by the press that it must have many less than fundamental flaws.[71] That was further confused by ambiguous separate statements by study members.[72] They culminated in ad hominem attacks on the group's leaders, which led to their disaffection from the process.[73] The DSB review was similarly positive, but its recommendation that the "SBI should be retained in the SDI architecture for up to two years" was interpreted by some as a lack of confidence in BP. The DSB also observed that with BP, the SSTS and BSTS were no longer needed and that BP should be capable of supporting a distributed surveillance systems. These recommendations were later implemented, as discussed below, which caused some additional strains in the program, as those systems had been accreting support over several decades.[74]

Engineering and Manufacturing Development Test

BP completed about half of its Engineering and Manufacturing Development (EMD) Test by 1993. It underwent significant laboratory simulation and integrated field tests, but it did not perform intercept

70. Baucom, "Ballistic Missile Defense," pp. 36–41.
71. *Ibid*., pp. 47–50.
72. *Ibid*., p. 69.
73. *Ibid*., pp. 71–80.
74. *Ibid*., pp. 51–2.

tests in the completeness or numbers required for statistical confidence. Laboratory tests successfully demonstrated the effectiveness of its mechanical and electrical design, sensors, computers, and guidance algorithms. Air table, rail, and drop tests showed that they could operate at the high speeds involved in actual intercepts. The space experiments were more complicated. The ABM Treaty prevented its testing in realistic conditions or geometries, so it was necessary to convert the BP into a conceptually equivalent GBI. That was accomplished by having its booster loft the BP to a higher altitude, launch a target missile, and then let the BP intercept the target while were both on a downward trajectory.[75] That was a complicated process compared to that used later in GBI tests, so it naturally encountered some problems.

There were three BP EMD tests, all from Wallops Island, Virginia. In the first test on August 25, 1990, BP sensors were launched to an altitude of 200 km to acquire and track a target missile. However, one of the explosive bolts that held the shroud in place misfired, and it caused the loss of telemetry. As that was not due to BP elements, it did not impact the BP program. However, the BP test program was then interrupted by the Gulf War.[76]

In the second test on April 17, 1991, the sounding rocket had a BP package on board. Its star tracker properly located the BP in inertial space, stabilized its attitude, calculated and performed a series of intercept maneuvers, and gathered Earth background data. All but one of its planned measurements were successful, so the contracts for EMD were released on the basis of this test. The two successful contractors ultimately had development programs approaching $500 million.

The third test on October 22, 1992, was a non-intercept test in which a BP was to detect and track the target and close to within 10 m of it. The test used a refurbished booster that had to be destroyed early for range safety. That loss was not due to the BP, but came at a time when BP was under strong criticism for political reasons and needed positive results. BP's EMD program was to have taken about four years. It was roughly halfway through when it was cancelled, which suggests that it could complete development in roughly two to three more years.

BP system and constellation costs scale with those of the KV. The DAB cost estimate of roughly $1 million per BP on orbit was derived from analyses, laboratory tests, and these limited field tests; however, some organizations involved in cost estimation and technical support of the MDAP argued that KV costs could be several times higher. Cost estimates from programs at early levels are not generally reliable, and the program was terminated before the required production cost data was accumulated, so the costs for BP production are uncertain and will remain so until detailed engineering development and tests are performed. The DAB value of about $1M per BP on orbit in volume is used for estimates below, as it is the only authoritative value available. The appendices explicitly show the sensitivity of BP and constellation costs to KV costs. They can be used to estimate the impact of component cost variations. In general, KV cost increases of an order of magnitude would not change the qualitative conclusions below.

In 1992, the BP was reduced to a "robust technology program." In 1993, that program was also abolished and BP's technological elements were eliminated.[77] The DOD Inspector General stated that the reduction was for non-technical reasons, that its fully approved MDAP had been managed "efficiently and cost-effectively within the funding constraints imposed by Congress," and that the Clinton Administration's termination of key contracts "was not a reflection on the quality of program management."[78] It has been suggested that the decision represented the result of a technical evaluation, but the DOD IG's assessment is consistent with the former director's position that the BP program

75. Cooper, "Reviving Effective Programs to Protect America from Ballistic Missile Attack."
76. Baucom, "Ballistic Missile Defense," pp. 116–117, 130, and 152–153.
77. *Ibid.*, pp. 149–157.
78. U.S. Department of Defense Inspector General, letter, April 1994.

was eliminated for political reasons,[79] which is consistent with the Clinton Administration's position that the ABM Treaty was essential for stability and its interpretation that BP represented a mobile launcher, which was prohibited by the Treaty.

Sensors

Despite the cancellation of the interceptor program, the BP sensors were tested thoroughly in the Clementine program, a joint program between the SDIO, Naval Research Laboratory (NRL), and Lawrence Livermore National Laboratory that mapped the entire surface of the Moon at high resolution. Clementine was the DOD's first and only deep space mission. A NRL rocket containing the BP sensors orbited the moon for several months and re-mapped it, including the hard-to-view polar regions, where Clementine used reflected communication signals to detect water on the moon for the first time.

Clementine's sensors included visible, near wavelength infrared (NWIR), MWIR, and LWIR cameras originally designed for missile detection, track, characterization, and hard body handover, mounted on a telescope with a large beryllium BP primary. Clementine used a 100 mJ/pulse lidar for range-gated imaging in Clementine, which tested its intended use in BP for range measurement, imaging, and aim point selection.[80]

Clementine used the BP sensors as they existed in 1993 with few updates. Subsequent programs and experiments ranging from asteroid intercept to tactical missiles and autonomous space navigation have given opportunities for updates of the sensors and other BP elements. While opportunities for intercept tests have been infrequent, sensor and satellite components have been continually updated to take advantage of advances in detector materials, electronics, IMUs, software, and computers.[81]

Engines

It was also possible to perform further demonstrations on efficient, small engines, which are essential to BP because accelerating a 4 kg KV to 6 km/s only requires an initial interceptor mass of about 30 kg with an efficient engine. Current engines have poor payload fractions in the few kilogram BP KV payload region because they are pressure-fed and driven by high-pressure bottles, whose thickness is determined by the pressure, not the amount of fuel they hold. The bottle-to-fuel mass ratio scales as the reciprocal of the cube root of the fuel mass, so for small rockets and fuel masses, most of the mass is in the bottle rather than fuel. For a given payload, if the engine or fuel mass increases, the total mass that must be orbited and the cost for orbiting it increases proportionally.

One way around that is to use pump-fed engines in which the fuel tanks remain at low pressure and the high pressure for the combustor is generated by small turbo pumps. That eliminates the thickness penalty for small high pressure bottles, but requires the development of light and efficient small pumps. The Astrid rocket flown in 1994 was roughly a 2 m by 15 cm diameter sounding rocket that developed 50 kg thrust for 30 s from 13 kg of N_2H_2. It produced a vacuum equivalent ΔV of about 2 km/s while maintaining a ratio of propellant mass to propulsion system mass (engine plus fuel) of about 0.85.[82] In that mass region, cold gas thrusters typically produce ratios of 0.2–0.3; the smallest (20–30 kg) liquid spacecraft engines produce ratios of 0.5–0.6; and small (200–300 kg) spacecraft such as NASA's Lunar Prospector, Clementine, and NEAR produce ratios of about 0.8. While current small pressure-fed engines have mass fractions factors of 3–4 below BP's needs, pump-fed engines are approaching them. Most effort of late has been on pumping non-toxic monopropellants such as

79. H. Cooper, "Testimony for the Record," Vermont Committee on Housing and Military Affairs, Vermont House of Representatives, March 18, 2002.
80. Nozette, "Clementine Mission," *Science*, Vol. 226, December 16, 1994, pp. 1835–1862
81. Baucom, "Ballistic Missile Defense," pp. 160–161.
82. *Ibid.*, p. 165.

H_2O_2 for Navy interceptor applications, but their viscosities, densities, and technologies are similar for the higher specific impulse 280–320 s nontoxic and toxic bipropellant fuels preferred in space.

Navy Contributions

GPALS saw the first significant Navy contributions to missile defense. In earlier periods it had been preoccupied with improved defenses for carrier battle groups, but when a layered system became the goal, it recognized that the existing Navy Area Defends (NAD) interceptor and the Navy Theater Wide (NTW) with the Standard Missile 3 and a modified SDI LEAP KV might provide significant boost- and ascent-phase contributions to theater defenses and could be extended to contribute to strategic defenses as well. However, such development would immediately run into the mobile platform provision of the ABM Treaty, so those capabilities were not advertised openly, although they stimulated technical developments in radars, guidance, and interceptor technology important to later phases.[83]

International Discussions

While the essential technical components of GPALS were available from those the U.S. developed in SDI, it was recognized that other technologies such as the ground-based midcourse and terminal systems developed by the Soviet Union and improved by Russia could complement them. It was also recognized that protection for allied forces was an essential element of alliance integrity and that international cooperation could provide additional support for deployment. There were a number of attempts to secure such international cooperation, starting with a series of meetings on strategic issues held at Erice, Sicily, in the 1980s and 1990s, which sought a common understanding of strategic stability and the impact of various types of defenses on it. The key common recognition was that as bilateral antagonisms receded, it should be possible to reduce offensive forces significantly and that under those conditions the introduction of the modest defenses needed for accidental, unauthorized, and rogue threats need not significantly adversely impact stability.[84]

While that overarching principle was generally recognized, there was less agreement on the details of proposed national or joint defenses—space versus ground basing being a particularly difficult issue. U.S. participants were convinced that space-based interceptors had the most promise for global coverage, performance, and cost. Russian participants generally preferred ground-based systems that had shorter ranges and were nuclear. All participants stressed the advances that could be made if the antagonisms of the Cold War could be forgotten and all nations interested in missile defenses could regard each other as friends and develop them together. However, they could not readily explain how that friendship could be rationalized with the maintenance of large strategic offensive forces by both sides, nor could they readily describe how the two previously antagonistic sides could demonstrate this new, non-antagonistic relationship. This inability to agree on the details of joint defenses made U.S. efforts to develop and deploy GPALS appear unilateral.

There were a number of exchanges of information on Russian ground-based and U.S. HTK midcourse systems and generic space-based interceptors, but no fundamental agreement on the impact of space systems on stability. Russian game theoretic analyses indicated that the impact of space-based systems was quantitatively, but not qualitatively, different from that of ground-based systems. U.S. analyses indicated that BP should incentivize a shift from MIRVs to more survivable single-weapon missiles, but there was a residual concern that space-based systems might somehow reduce stability.[85] The Russian General Staff's emphasis on the maintenance of the ABM Treaty prohibition on space,

83. Cooper, "Reviving Effective Programs to Protect America from Ballistic Missile Attack," pp. 12–13.
84. A. Zichichi, ed., International Seminar on Nuclear War, 11th Session: Planetary Emergencies (London: 1991, World Scientific).
85. *Ibid*.

mobile, and OPP elements effectively stopped progress during the Clinton Administration. A less autonomous version of BP that would decouple defenses against rogue threats from strategic systems and the Treaty, used transferable technology, and could be produced and controlled jointly could not overcome these concerns, which persisted in Russia and to a lesser extent in the United States.[86]

Confusion among international participants and disagreement within the U.S. strategic community over whether to simplify the BP were significant barriers to progress. Some thought the United States should hold to the full BP's advantages against countermeasures and suppression that could accompany accidental and unauthorized attacks. Others argued that it would be useful to start with a lower level of technology that would be less threatening to others and reserve the full BP technology for the United States development in case it was needed later for resurgent strategic threats, proliferation of countermeasures, or rogues.[87] There was a meeting of minds at the level of technical scientists and analysts, less so at the level of arms control advisors, and still less at the level of military and political advisors.

Nevertheless, Russian arms control advisors took the proposal for a joint missile defense against, accidental, unauthorized, and rogue attacks to President Boris Yeltsin, who accepted the ground-based elements as a starting point for joint development and control and considered the principle of letting satellite "sensors go free," i.e., excluding them from the ABM Treaty. He did not choose sides in the argument over space- versus ground-based interceptors. In a UN General Assembly speech on January 31, 1992, he proposed that SDI take advantage of Russian technology and that the United States and Russia work together to build a global defense for the world community. Unfortunately, there was a six-month delay before the United States responded at the June Summit, which set up the Ross-Memedov talks to explore these issues. The talks were positive, but were interrupted by the change of administrations.

The Clinton Administration discontinued the Ross-Memedov talks, reversed course, reaffirmed the primacy of the ABM treaty, and decimated the GPALS program. When President Yeltsin offered to continue the high level talks at his first meeting with President Bill Clinton in Vancouver in April 1993, apparently no one in the U.S. delegation was familiar with the concept. The Russian factions who had supported the initiative for cooperation were undercut; they lost ground with their colleagues who had consistently opposed defenses—or at least U.S. defenses. The opportunity for joint defenses was lost in first two years of the Clinton Administration while it concentrated on the domestic economy. The international contacts, agreements, and openings developed during GPALS still exist, but strategic thought has apparently regressed to earlier, limited concepts of cooperation.

Summary of Developments During GPALS

GPALS was perhaps the high water mark of missile defense. It started the development of deployable systems in all layers and a global BMC2 system to integrate them. Those elements could have provided protection of the U.S., its deployed forces, and allies. It had support from the Administration, military services, allies, and friends. It brought a set of key components close to the level required for a decision to deploy. The main elements of GPALS could arguably survive suppression attacks on their sensors and interceptors, operate in the resulting backgrounds, and still provide multi-layer protection. The boost-phase BP layer could reduce the number of penetrating weapons to about 10 percent of those launched, which a modest midcourse layer, to which BP itself could contribute, could reduce leakage to acceptable levels. Terminal defenses could reduce that to levels acceptable for protection of population.

86. J. Woolsey, "The Way to Missile Defense," *National Review*, June 19, 2000, pp. 36–41.
87. P. Smith, "A Geopolitically Acceptable Rogue Missile Defense System? Space-Based Boost-Phase Interception with Restricted Latitude Coverage," white paper, Aerospace Corp., August 24, 2000.

GPALS's key element was the strong, robust, inexpensive boost-phase attrition that BP appeared capable of providing. That would reduce both the size and sophistication of the expected downstream threats, as strong pressure on the boost phase gives the attacker a strong incentive to deploy weapons quickly and decoys simply. BP were projected to cost roughly $1 million apiece on orbit, so deploying 1,000s of them in a boost-phase defense would not greatly increase the cost of the overall system. Its midcourse layer should be affordable, particularly if BPs that were out of range for boost were used to augment GBIs to address RVs that leaked through boost and bus layers. The DOD estimated that about 2,000 BPs could support a 99 percent effective two-layer defense against the most stressing GPALS threat and that it could be deployed and operated for a decade for about $11 billion (1990 dollars = $15 billion in 2002), which is an order of magnitude less than the cost estimates for Phase I. That allowed BP to meet the Nitze criteria with considerable margin, which reduced concerns about Soviet escalation. GPALS was a complete, credible response to the requirement for high quality protection for the U.S. and its allies. However, the developments of the following decade represented a step back from its goal of global defense, as discussed in the next section. The current program is an attempt to reintegrate the ground-based midcourse elements of GPALS.

Theater Missile Defense

Theater missile defense has been a priority element of missile defense for several decades. The necessity to defend deployed troops was recognized in the European theater and reflected in defensive developments such as the Corps SAM. The 1990 review conducted by Ambassador Cooper before he took over SDI program emphasized the importance of defenses for U.S. and allied forces deployed to various theaters of operation, which was underscored by deficiencies noted in the 1991 Gulf War. Prior to hostilities, missile defense was not a high priority, as maneuver alone was thought adequate to protect military forces from inaccurate theater missiles.

The Gulf War reintroduced an aspect of missile war that had been seen earlier in World War II and the Iran-Iraq War of the Cites: the use of missiles as terror weapons against civilian populations as well as military forces. The first operational engagements of Iraqi SCUDs by PAC-2 indicated a need to modify their software to engage the leading element of the fragments produced when extended-range SCUDs broke up during reentry. Those modifications were executed while the PATRIOT batteries were under attack by SCUDs. Because of these "accidental" but effective penetration aids and the deficiencies in fusing due their detuning for ABM Treaty compliance, PATRIOT had limited operational effectiveness during the war, although it maintained Alliance cohesion, kept Israel out of war, and showed some promise for more robust solutions.

The first year of the Clinton Administration changed SDIO's name to the Ballistic Missile Defense Organization (BMDO) and initiated the Bottom-Up Review, which ostensibly shifted emphasis from national to theater missile defense. In practice, it cut the missile defense budget in half, reduced NMD 80 percent to a technology program, cut TMD 25 percent, and eliminated global defenses altogether.[88] The Clinton Administration returned the GBI development proposals unopened. Technical Base Research, which included the E2I, Red Teams, Raptor Talon, and a large basic science and technology program, was cut by 90 percent. Little science or technology survived the Clinton Administration reductions. These budget cuts produced a set of stovepiped TMD systems, each of which had to sort out its own BMC3 or impose it retroactively, neither of which has worked well historically. Dropping GPALS's integrated global BMC3 was arguably the most unfortunate legacy of the Clinton Administration.[89]

Key Elements

The key initial elements of the TMD program were THAAD, PAC-3, and the NAD. The first two were to provide the exo- and endoatmospheric layers of a tiered defense for ground forces in theaters like those previous systems would have provided against intercontinental missiles. NAD was to provide coverage for carrier battle groups so that they could operate close to shore and could augment the coverage given by PAC-3.[90]

PAC-3

PAC-3 was the culmination of the SAM-D to PAC-2 to GEM guidance progression. It incorporated the advanced ERINT technology developed in SDI Phase I, which had produced two successful inter-

88. Baucom, "Ballistic Missile Defense," pp. 158–159.
89. Cooper, "Reviving Effective Programs to Protect America from Ballistic Missile Attack," pp. 48–49.
90. "Terminal Defense Segment," BMDO Link: The Ballistic Missile Defense System, January 2002.

cepts. That technology enabled PAC-3 to guide itself to HTK intercepts rather than being command guided, as earlier low-altitude interceptors had been. The limited accuracy of command guidance required precise fuse timing, which adversely impacted PATRIOT's performance in the Gulf War. PAC-3 was command guided to within range of its on-board radar by a competent C-band radar, but then used its on-board radar to reacquire the target, track it, and home in on it by itself, which produced miss distances so small that it did not need explosives or fuses. PAC-3's HTK was made possible by a compact, powerful on-board Ka-band (30 GHZ) radar with a carefully calibrated ceramic radar dome that reduced guidance errors from refraction to the milliradian levels required for the roughly 10 cm miss distances required for high kill probabilities.[91]

PAC-3 is intended to serve as a low-atmosphere component of a system to intercept both air-breathing threats and SCUD-derived missiles with ranges of 300–600 km using intercept ranges on the order of 15 km. On maximum range trajectories, such missiles have apogees of 75–150 km, so they cannot make effective use of light decoys, which would be stripped out by atmospheric drag. However, such missiles remain in the atmosphere, so they can maneuver. Thus, PAC-3 has advanced aerodynamics for more robust intercept of maneuvering targets. Logistics problems in the Gulf War indicated the need for greater transportability, reliability, and accuracy, so an integral part of PAC-3 development was the reduction of its size and weight for efficient transport to forward locations by tactical lift or truck.

Theater High Altitude Air Defense

The Theater High Altitude Air Defense interceptor is the high-altitude component of a layered theater defense. It is supported by an x-band radar with adequate power-aperture product for track-while-scan at ranges over 400 km; thus, it can detect many theater missiles before apogee and intercept out to ranges of about 180 km. Its performance is improved in passing engagements where it can observe weapons at off-boresight angles, which increases their cross sections. The GHz bandwidth of the THAAD radar would make it difficult for missiles to make effective use of even moderately heavy decoys for ranges under 600 km, as light decoys could be identified by their differential slowing due to atmospheric drag over most of their trajectory.

THAAD has a strong heritage in the successful endoatmospheric Exploratory Phase HIT and FLAGE interceptors as well as advanced sensor technology from the high endoatmospheric Phase I HEDI. It has lightweight interceptor technology from the exoatmospheric Phase I ERIS, which was matured in the endo-exoatmospheric GPALS E2I, whose technology was available to it. However, THAAD did not fully capitalize on that technology, so it had problems with weight, volume, transportability, and engineering during development. It failed six flight tests before improved design, quality control, and testing produced two successful flight tests, which were determined to be an adequate basis for advancing it into acquisition. Deployment is scheduled for 2007, although costs could still cause delays.

Navy Area Defends

The Navy Area Defends was derived from Aegis and the existing interceptors originally designed for defense of the carrier battle group from air threats and to address the higher closing velocities of short-range theater missiles. The NAD interceptor is derived from existing ship-defense missiles with a more capable KV. It serves as the lower tier of a layered system, which is described as essentially "PATRIOT on a boat." Since ships are hardened to conventional explosives, it resembles a nonnuclear version of Sprint and MSR's defense of strategic deterrents. However, for short-range threats, attacks can be closer to the surface, so propagation through moist atmosphere is a problem, and on-board power and aperture are constraints. Thus, Aegis radars currently operate in S-band, which is a

91. "PATRIOT Advanced Capability-3," BMDO Link: The Ballistic Missile Defense System, January 2002.

longer wavelength than that of the x-band GBR and THAAD; so they generally have detection ranges well under 1,000 km with modest bandwidths that are inadequate for discrimination. NAD has had recurring issues in integration, testing, and cost and is currently in research and development.

Additional Elements

A fourth element was to be added from a competition between the Corps-SAM, Navy Upper Tier, and a boost-phase system such as the airborne laser (ABL). Instead, all of those systems found a role. The Corps-SAM was integrated into the preliminary interceptor for the NATO Medium Extended Air Defense System (MEADS) low air defense, for which the United States later advanced the PAC-3 as an interim interceptor. Navy Upper Tier, later named Navy Theater Wide, was part of GPALS, but it had a problem with the ABM Treaty. If NTW developed as expected, it would quickly violate Article V of the ABM Treaty, which would block further development. Thus, Navy leadership emphasized NAD, although it understood that NTW had more potential in the long run.

NTW evolved into an overlay for the battle group with residual capability for projecting defenses over adjacent theater land areas. It used the Navy Standard Missile 3 with the LEAP KV, which had enough acceleration and velocity to intercept theater missiles in ascent from ships near the launch area. NTW sensors and interceptors could have some capability against RVs in midcourse and terminal, but it was prohibited by the ABM Treaty, so it was not openly pursued. NTW was fully funded in the last GPALS Future Years Defense Program. The Clinton Administration tried to terminate it, but Congress restored it. NTW's ABM Treaty issues were partially settled for the rest of the Clinton Administration when the DOD declared that a reduced capability version was Treaty compliant—although those reductions put unfortunate limits on its capabilities for later phases. With those developments, NTW gained enough service and Congressional support to advance to a MDAP, which brought the total number of TMD systems to five.

Battle Management, Command and Control

These TMD systems were largely developed as independent stovepipes, each with its own sensors, interceptors, and command and control elements. Integration was primarily within elements rather than across systems. The results resembled simple fire control systems for each system more than actual BMC2 systems for the whole TMD system. There were efforts to improve fire coordination to improve interceptor allocation by passing tracks and firing solutions over common theater data links as well as to provide early satellite warning data from systems through direct downlink and in-theater processing of data to the fire control systems of the individual elements.

This situation represented a step back from GPALS, in which the open BMC3 architecture that BP naturally imposed on the system was an essential element of C2. GPALS followed the 1986 Eastport Study Group's recommendation to avoid the "appliqué" approach to BMC3, but the services opposed that approach as contrary to their custom of buying BMC3 with each system. The Air Force preferred a centralized computer that could gather all information, infer optimum battle plans, and direct the battle. For space elements, that approach implied one operator per satellite and the opposite of pre-delegating authority to the lowest possible level, which is essential for timeliness and efficiency. While GPALS made progress in open BMC3, when BP was killed, the thrust toward decentralized execution was lost. A preference for centralized C2 might have been justified during Phase I, but by TMD, the commercial Iridium and Teledesic satellite programs had demonstrated that many satellites can be controlled effectively by a few operators, so decentralized systems were not only feasible, but preferred for cost and efficiency.

Advanced Interceptor Technology

Many of the technologies from the Exploratory, First, and GPALS Phases that were not incorporated into TMD system—particularly interceptor technologies—were placed in the Atmospheric Interceptor Technology (AIT) program, which was to develop them for block upgrades to the technology in THAAD, MEADS, Boost-Phase Intercept (BPI), and other endo-atmospheric systems. The key issues were dynamics, aerodynamics, seekers, thermal control, and shroud deployment, but there were also issues in LOS measurement, guidance and control, and aim point selection on targets with specific vulnerabilities or high value components such as submunitions.[92] The number of issues included was large because the elimination of BP, which had been a major technology driver during SDI and GPALS, left large gaps in development programs, particularly those for lightweight sensors and interceptors. The integration of improved space sensors with C2 was particularly difficult because of the large budget cuts in each and the complexity of making proposed approaches compatible with the Treaty. All areas except milestoned satellites and their sensors were supported at levels far below their promise and difficulty of execution.

International Efforts

The main international thrust of TMD was MEADS, which became the principal vehicle for international cooperation with the adoption of the PAC-3 as the putative interceptor. Japan entered into a formal agreement with the Navy for Aegis-based defenses. Cooperation continued on the Russian–American Observation Satellite (RAMOS). Its primary function was technology development and transfer, but it also helped Russian scientists find commercial projects that would keep them in their design bureaus, but working on non-offensive systems, rather than seeking military support abroad. Satellite cooperation also had some potential to improve Russian early warning satellite and radar systems, which were degrading rapidly with the loss of territories, facilities, and support.

The U.S.–Israel Arrow cooperative project supported the development of an interceptor that has a defended radius about half that of THAAD, although it achieves kills with explosive lethality enhancers. Arrow completed its development tests and is operational, so Israel now has a missile defense that is arguably adequate to the threats it faces for an investment of a few $10 billion, which was shared by the two countries. While the United States would classify Arrow as a TMD system, it is clearly strategic from Israel's perspective. Arrow was started much later than comparable U.S. systems that are still a decade away from deployment. It is the only operational missile defense system outside of Russia today, so it is arguably the principal delivered product of U.S. missile defense efforts to date.

Summary of Developments During TMD

TMD focused on improvements on deficiencies in missile defenses for U.S. and allied troops and threatened populations that were made clear from PATRIOT's performance in the Gulf War. The technology emphasized during TMD is now moving toward deployment, and it later supported NMD applications. However, the TMD program reduced NMD to a technology effort and eliminated BP altogether, which amounted to a unilateral retreat from the strategic capability the United States had sought to develop during GPALS for accidental or unauthorized launches from Russia or China and for rogue threats. Subsequent events have shown that those capabilities are needed at least as soon as they could have been produced by GPALS's; thus, from that perspective, the TMD program was a five-year slip in the key elements of missile defense that the United States could ill afford.

92. BMDO, "Atmospheric Interceptor Technology," *Fact Sheet* 113-00-11, November 2000.

National Missile Defense

National Missile Defense, which had been relegated to a technology program during the previous five years of the TMD program, was reactivated in 1996–1999 by four events. The first was Congress responding to the threats seen in the Gulf War by writing into the Missile Defense Acts of 1991 and 1992 that NMD was required to defend all 50 states at the earliest technically feasible time.[93] The second was the 1994 election and the Contract with America, which gave missile defense top priority. When the Clinton Administration offered a muted response to this mandate, Congress probed into the state of missile defenses. The CIA responded with the 1996 NIE, which stated that there was no threat to the lower 48 states for 15 years. Congress's dissatisfaction with that estimate led to the third event: the formation of the Rumsfeld Commission and its documentation of the increasingly unpredictable scope and timing of rogue threats.

The Clinton Administration criticized the Commission's July 1998 report as "worst case threats," but those threats were given support by the fourth event: On August 31, 1998, North Korea launched a three-stage Taepo Dong 1 missile over Japan, almost reaching Hawaii. It was probably a failed space launch, but the fact that North Korea could launch a hitherto unsuspected, and initially undetected, third-stage with the characteristics of a separating bus that could put a few hundred kilogram payload anywhere on Earth largely discredited intelligence community estimates of missile capabilities, which forced a more deliberate NMD program.

Scope

The Clinton Administration complied with the Missile Defense Acts (MDA), but restricted the mandated NMD program to a single site to protect the ABM Treaty as the "cornerstone of strategic stability." It could do that because the MDA of 1991 was a compromise. On the one hand, it called for an effective multi-site defense using space sensors; on the other, it stated that the deployment had to be ABM Treaty compliant. Thus, the Administration could choose the wording it preferred; it chose the latter.

Prior to the Rumsfeld Commission, the NMD program did not have a specific deployment date; it was described as a "3 + 3 deployment readiness" program that was intended to bring technology to a level within three years that could support a decision then to produce a deployment in another three years, if needed. Critics argued that it could only meet a 3 + 5 objective.[94] After the Rumsfeld Commission and Congress's response to it, NMD changed from a deployment readiness program to a milestoned development program with a 2005 to 2007 deployment date. In the spring of 1998, the NMD program selected Boeing as the Lead System Integrator (LSI) to direct all elements of the program. In December 2000, the NMD contract was restructured to support a 2006 deployment after it was determined there was excessive risk in the developing and deploying an essential x-band radar (XBR) in Shemya for a 2005 deployment.

The NMD program was principally oriented toward an initial capability (C1), which was to address a few missiles with primitive penetration aids. It was to do so by upgrading the existing BMEWS

93. U.S. Congress, House of Representatives, *National Defense Authorization Act for FY 1992 and 1993*, 102d Cong., 1st Sess., Conference Report to Accompany H.R. 2100, Report 102-311, November 13, 1991, p. 34.
94. BMDO, "National Missile Defense Architecture," *Fact Sheet* 106-00-11, November 2000.

radars, integrating the observations from the DSP satellites into GBI fire control, deploying 20 GBI interceptors at Ft. Greeley in Alaska, building an x-band radar in Shemya (which is roughly midway between North Korea and Alaska), and adding the BMC2 needed to tie them together. This C1 phase was expected to cost a few tens of billions of dollars, be approved by the end of the Clinton Administration, and be completed by 2006. The new elements were the x-band radars, which were later seen to be unnecessary for C1, and the GBIs, whose development problems delayed NMD deployment.

Coverage

NMD faced the difficult task of providing coverage for the whole U.S., including CONUS, Alaska, and Hawaii, from the single site allowed by the Treaty, which it was required to respect. However, NMD's Treaty compliance was suspect. The Treaty required that the interceptors be within 150 km of the missiles they protected, while in C1 they were 1,500 km away, but it was assumed such details could be negotiated. A more fundamental issue was Article I of the Treaty, which prohibits each party from "a defense of the territory [i.e., population] of its country," whereas protection of population was the explicit objective of Congress's 1991 and 1992 Acts. That fundamental conflict was never resolved.

Another concern was the status of early warning satellites and radars, which the U.S. interpreted as "adjuncts" to radars, which could thus be integrated into BMC3. Russia did not concur. The issue was surfaced, but not resolved, in the first summit between Presidents Yeltsin and Clinton. Thus, the proposed deployment apparently would have clashed with the ABM Treaty in several ways, some of which were fundamental. However, because of the restricted nature of the proposed defenses being developed and the delays in their development, these issues did not have to be resolved during the Clinton Administration. However, they did resurface in later summit discussions, including the one in which President Bush announced the U.S.'s intent to withdraw from the Treaty.

The applicability of the Treaty to these defenses was unclear. NMD was intended primarily for rogue threats. It made use of radars and satellites that Russia had demonstrated the ability to suppress several decades earlier; thus, it was clear that they would have no impact on the Russian retaliatory strikes contemplated by the Treaty. However, the NMD system would use radar and satellite components that could also be used as elements of a U.S. defense against a Russian strike, and NMD was intend to protect the entire population of the United States, so Russia claimed that its deployment would violate the Treaty. Russia's implicit extension of the Treaty to rogue threats was not resolved. That left an ambiguity that impeded the integration of common elements that was needed to overcome the problems caused by restricting NMD to a single site.

The requirement that NMD protect Alaska and Hawaii dictated that the system be based in Alaska, as CONUS-based radars could not detect attacks on Alaska from Asia in time for interceptors based elsewhere to engage before impact. However, placing the site in Alaska meant the radars there would not be useful in defending against attacks on the United States from the East, Mideast, or other locations. For the East Coast of the United States, warning would have to come from satellites; confirmation and track from the unprotected BMEWS radars in Thule, Greenland, and Fylingdales, U.K. (assuming that Denmark and the United Kingdom approve their use for NMD); and discrimination from the GBI interceptors themselves.

Integration

Integration of essential information from satellite systems was essential to support the short timelines and restricted radar coverage of rogue threats. As noted above, the U.S. finessed that issue by declaring those satellites to be adjuncts to the radars allowed in the Treaty. That was legally satisfactory as an interim measure, but made the integration process contingent on an interpretation that Russia never accepted. Maintaining it prevented open discussion of the integration that was actually needed.

NMD also had several difficult non-Treaty requirements. One was a "high probability of zero leakage," which is mathematically ambiguous, but captures the public's expectation that everyone, anywhere in the United States should be protected equally from attacks from any quarter. It is a stressing requirement, but one that citizens have chosen in polls for decades—and apparently think the DOD has already provided. A second requirement was "human in control," which is desirable, but in conflict with the short timelines for effective operation of the BMC2, especially for the boost-phase and regional missile defenses discussed below, which require near real time decisions and release. A third, which seemed less critical at the time but proved awkward later, was that defenses be able to defeat all of the "design to threats" included in, derived from, or implied by the System Threat Assessment Report (STAR) that conveys the intelligence community's estimates of the missile threat to the United States to the research and development community. The STAR contains the intelligence community's best estimates of rogue threats, Russian and Chinese accidental and unauthorized threats, and plausible excursions to them.

There were also several important features that NMD did not attempt to integrate. GPALS had aspired to the development of a global capability to protect allies and other like-minded countries, not just the United States. TMD dropped that goal and concentrated on defense of deployed U.S. and allied combat forces with a residual capability for civilian populations in theaters. Defense of allied populations was left to their governments. In reestablishing NMD and the reintegration of TMD elements, the United States ignored intermediate range, or regional, missiles and with them the need for regional missile defense (RMD). NIEs made it clear that such a capability was needed, but pursuing them could have blurred the distinction between NMD and TMD, which was not defined by the ABM Treaty. Thus, RMD was omitted in an attempt to create a rough firebreak between NMD and TMD. Attempts to better define the distinction between NMD and TMD through un-ratified but informally enforced Demarcation Agreements did little to define needed RMD, but further limited TMD sensors and interceptors.

Elements

The key elements of the NMD program were the radars, satellites, HTK interceptors, and BMC2 left from GPALS, as modified, delayed, or terminated by the TMD program. The radars and satellites had undergone significant development during previous phases, but the GBI program had to be restarted from what little was left from the research started by GPALS. Another challenge was developing BMC2, which was simplified by separating TMD and NMD and ignoring regional defenses together. However, that produced separate BMC2 stovepipes for each theater system and an independent one for NMD, which proved inappropriate for later systems.

Ground-Based Interceptor

The essential element of NMD was the Ground-Based Interceptor (GBI) and its Exoatmospheric Kill Vehicle (EKV), which were the products of the progression from the Exploratory Phase's HOE to the SDI Phase's I ERIS and HEDI and to GPALS's E2I.[95] At the end of GPALS, the Clinton Administration ceased development of the E2I and returned the GBI proposals unopened. However, BMDO was able to keep the Ground-Based Interceptor–Experimental (GBX) active, so the Army was able to continue its development during TMD. Progress made in updating and testing key components reduced delays when it was necessary to restart GBI in 1998. This interim development produced a number of scientific and engineering developments in structures, communications, guidance and control, computers, sensors, cryogenic focal plane technology, and discrimination algorithms. GBI took advantage of some of the advanced technologies demonstrated by BP during SDI Phase I and GPALs. That produced an EKV whose primary sensor was its large cooled LWIR focal illuminated by a large, wide

95. BMDO, "National Missile Defense Weapon Element," *Fact Sheet* 125-00-11, November 2000.

FOV optic that was sufficiently sensitive to detect cold RVs at distances of about 500 km and detect small temperature differences between them.

Altogether, these developments led to a 20-fold reduction in mass from the HOE KV to the 60 kg GBI EKV. However, the transfer was not complete. The slowing of GBI development from the end of GPALS in 1993 to the resumption of NMD, plus the need for the GBI to enter tests as soon as possible thereafter, forced a design freeze that left the GBI with roughly 1993 technology and a mass almost two orders of magnitude heavier than the BP had demonstrated earlier in vehicles of greater functionality.

One element of this development that seemed innocuous at the time but proved important later was the decision to use existing boosters for the developmental GBI and test targets.[96] NMD was restructured at the time when the commercial off-the-shelf (COTS) philosophy was at its peak in the DOD, so it was natural to use adaptations of commercial boosters for the GBI and target boosters. Each decision led to problems. The payload launch vehicle (PLV) for the GBI range test booster used the first two stages of decommissioned Minuteman II strategic boosters, which was permitted under the Treaty given significant care, because they could be refurbished for only $20 million. However, the resulting PLV did not have the performance expected of the GBI booster, so its use placed restrictions on the GBI test envelope and led to criticisms that the NMD program was not ready for a deployment decision.

Those criticisms increased when the COTS booster failed in developmental tests. The reasons for the failures were typical of that period, during which U.S. missile programs from TMD to Titan 4 all experienced problems due to the loss of key personnel, technical competence, and quality control that were not solved until the end of the decade. The choice of a COTS booster was reasonable. Its acceleration, velocity, and accuracy requirements were not stressing, and the reasons for its failure were familiar. Since GBIs must be stored in canisters in a passive state for up to a decade and then must act almost instantaneously, it is useful for them to have minimum maintenance components. But to intercept, they need large divert velocities. Thus, they use high performance solid engine divert attitude control systems (DACS) designed close to the limits of their performance to meet the baseline EKV weight. When the GBI's weight increased to provide improved performance, the additional divert thrust required forced the DACS into unexplored regions where they failed. Such problems were also encountered in booster developments of previous decades, but when the largest and best U.S. defense contractors appeared to be unable to repeat their successes of previous decades, it made all advertised NMD timelines suspect, which led to the restructuring of the GBI booster program. The impact on schedule probably could have been accommodated, but the loss of confidence came at an unfortunate time for the NMD deployment decision, as discussed below.

Exoatmospheric Kill Vehicle

The exoatmospheric kill vehicle (EKV) also had significant developmental problems, which were revealed by the testing program discussed below.[97] The EKV was plagued by a series of errors on key tests, which reduced confidence in quality control and led to demands for extended testing programs. The failures mostly reflected the eroding industrial base, which had been unable to renew the competence gained in previous decades. Expanded and effective use of ground tests of key components and software with massive hardware in loop (HWIL) simulation restored confidence as it had in THAAD. However, it led to cost estimates of about $35 million per EKV, which with a booster estimate of $20 million gave a cost of roughly $50 million per GBI. That was an order of magnitude more than the cost of earlier nuclear interceptors, although the price should fall in production volume. However,

96. BMDO, "National Missile Defense Multi-Service Launch System," *Fact Sheet* 102-00-11, November 2000.
97. BMDO, "National Missile Defense Weapon Element."

they made it appear that the substitution of HTK for nuclear kill vehicles could come at a significant price.

However, this high price was not just for HTK. The EKV is essentially a one-element NMD system. Operating from satellite warning and crude radar data, the EKV could reacquire targets from any azimuth or launch point; correlate its observations with those from satellites and radars, sort out the elements of the threat cloud; discriminate RVs using advanced passive sensors, computers, and algorithms; and intercept them with its on-board computers. Thus, its cost could be justified on the basis that it reduced NMD's dependence on satellites and ground-based radars for search, track, and discrimination, which also eliminated any sensitivity to their degradation by unknown nuclear effects.

Moreover, cost effectiveness at the margin was not a constraint on NMD defenses against rogue, accidental, and unauthorized launches. For them the primary metric was the defense's ability to provide high confidence protection against small threats. NMD's economics were more likely to be determined by the cost of BMC2 and operations than that of the GBIs, no matter how expensive. However, as rogue threats grow, or as more attention is paid to accidental or unauthorized threats that could produce 10s to 100s of credible warheads and 10s to 100s of plausible decoys, the number of GBI could grow to 100s. At that level of threat, an adequate number of $100 million/GBIs could approach several billion dollars, which would no longer be negligible.

Sensors

In NMD the key acquisition, track, and discrimination sensors were those on the EKV itself. However, for full effectiveness and global reach those sensors needed prompt, accurate, and reliable launch warning from satellites and reliable search, detection, and preliminary tracks from ground-based early warning radars. The supporting sensors available could actually provide much more than that.

Satellites. DSP satellites could provide warning messages on a time scale of about a minute, which is only about 3 percent of an ICBM's roughly 30 minute flight time. DSP is to be replaced in mid-decade by SBIRS-High, but that should not cause any problems as the transition is to be accomplished by blocks that are meant to be seamless. SBIRS can draw on a heritage of decades of DSP operations, as well as the extensive redesign for more complex missions that had been accomplished in earlier decades for BSTS, FEWS, and ALARM, which can be used to reduce the uncertainties encountered in the transition between the two.

Even in optimized constellations, the accuracy of DSP measurements of missile azimuth is limited by its roughly 1 km pixel size.[98] For a typical ICBM burnout about 500 km downrange, the error in its azimuth prediction would be about 1 km/500 km = 2 mr, which would give an EKV a midcourse error basket of about 5,000 km x 0.002 rad = 10 km. That is small enough for the EKV to search and reacquire the RV from an initial range of 500 km with a sensor with a FOV of about 10 km/500 km = 1°. Such a sensor with a 100 x 100 detector array would have a 0.2 mrad instantaneous field of view (IFOV), which would give a resolution at 500 km of about 500 km x 0.2 mrad = 100 m, which is adequate to start the target detection process. Detection at that range would leave a time of about 500 km/10 km/s = 50 s for divert. Clouds of objects on the order of 10 km across would require divert angles on the order of 10 km/500 km = 0.02 rad and divert velocities of about 0.02 rad x 10 km/s = 0.2 km/s, which GBIs could produce.

Thus, DSP is a reasonable match for the acquisition characteristics of the GBI. SBIRS-High should be a somewhat better match because of its higher frame rate and resulting better resolution of missile burnout time. The handovers provided when the satellite data is augmented by UEWR or x-band radar observations is even better. If the EKV continued to observe the approaching cloud of objects to a dis-

98. Richelson, *DSP Satellites and National Security*, p. 118.

tance of 50 km, it would have a resolution of roughly 10 m, which would be much less sensitive to closely spaced objects (CSO) and much better suited to discrimination, although the deflection required to cover the whole cloud would be about 2 km/s, which is beyond the capability of most interceptors.

The NMD phase had some impact on the program for SBIRS-High to replace DSP as the primary early warning satellites. NMD added a requirement for the direct downlink of missile warning data for processing in theater as well as at the central ALERT's global processing center. The downlink's goal was to improve the assuredness, timeliness, and quality of aim and launch point predictions to theaters. Its byproduct was the addition of cost and complexity to SBIRS's development, which contributed to twofold increases in cost and schedule that put its completion at risk. NMD also changed the BE satellites from GPALS. They had already been redesigned during TMD into SMTS, but mission creep made SMTS unaffordable, so SMTS was reconfigured into SBIRS-Low, which was intended to support missile defense, intelligence, and Space Command's missions in missile intelligence and battlefield preparation. Their addition resulted in delays and cost growths, so the Air Force returned BE to the BMDO, with a primary justification of NMD discrimination. When SBIRS-Low was found to have problems with CSO in that role, it dropped back into development. Thus, the progression in low earth orbit (LEO) satellites dating back to SDI does not contribute to the spiral development discussed below, which does not contain a satellite discrimination sensor. Restoring a discrimination capability to a LEO constellation with the technologies and designs under development would be difficult and expensive.

Radars. When the BE satellites slipped, GPALS started a gap-filler program to modify the BMEWS radars into what are called UEWRs. Those upgrade programs were terminated at the end of GPALS but restored by NMD to provide search and initial track for XBR, which did not have enough power to search large threats by itself. It is planned to upgrade the UEWRs' computers, databases, and communications enough to enable UEWRs to perform track-while-scan with enough capacity to search for NMD targets without impacting existing North American Aerospace Defense Command (NORAD) missile and aircraft search missions. The upgrades would increase the UEWRs effective bandwidths to about 30 MHz, which should give them the ability to measure ranges to an accuracy of about 100 m. That, together with improved temporal coherence and phase stability, should provide some level of discrimination in undisturbed environments.[99] However, UEWRs remain UHF radars, so they have the same sensitivity to natural and disturbed ionospheres and jammers, which compromised earlier systems based on them. UEWR search and XBR track is a natural combination for undisturbed environments and moderate threats, but if the UEWRs are degraded somehow, the XBRs have inadequate power to search for themselves, unless purchased in large numbers and deployed globally.

With these improved metrics, UEWRs could support defenses against ICBM launches from Northeast Asia toward the United States. Because of the forward position of the UEWR at Clear, Alaska, it could provide detection well before weapon apogee. That would leave about 20 min for the GBI to fly out, which would support ranges of about 7 km/s x 1,000 s = 7,000 km, which is adequate for launches against Alaska or CONUS. However, launches from Northeast Asia to Hawaii could not be intercepted on the basis of UEWR warning, as even the apogee of such a trajectory would be below the roughly 3,700 km effective range-to-targets of the Clear and the Beal UEWRs.

Even if the Clear UEWR could detect the weapon at apogee, it could not support an intercept. The weapon is at maximum speed and only has to fly another 3,800 km to Hawaii, while the GBI starts from rest after some delay and would have to fly about 4,500 km. Thus, there would not be enough time left after detection for the GBI to fly from Alaska to Hawaii before impact.

99. BMDO, "National Missile Early Warning," *Fact Sheet* 113-00-11, November 2000.

The existing Cobra Dane or the proposed XBR radar at Shemya would see such launches almost immediately after launch and could release GBIs to intercept the weapons about midway to Hawaii. GBI could readily make such intercepts, given accurate time and azimuth measurements. Such information could come from the UEWR, Cobra Dane, XBR radar, a theater radar such as THAAD, an accurate tracker such as the ABL sensors, or an Aegis ship off Japan—provided that they were on station at launch and their information was immediately provided to the GBI BMC2 system.

For the special case of Hawaii, even an IR warning satellite could provide the information, because the missile's azimuth and range are pre-determined to adequate precision by the compactness of the North Korean launch areas and the Hawaiian islands and their known geometry. All the GBI needs to know is the missile's time of arrival, which could be estimated with sufficient accuracy from satellite observations. It is customary to rely on more than one sensor phenomenology, such as IR and radar, to warn of attack, but in this case the IR signature and trajectory together should be so unambiguous as to allow response on IR alone, if the other sensors were occupied elsewhere.

For similar geometric reasons, the BMEWS in Thule, Greenland, would not see a launch from Libya to the Atlantic seaboard of the U.S. The trajectories of such missiles would remain at latitudes below 45°, which is the lowest latitude at which Thule can see objects at 1,000 km apogees because of the Earth's curvature. Even if the EWR at Thule could see the RV, a GBI could not intercept it from Alaska in the time remaining. The weapon would have about 5,000 km remaining and the GBI about 5,600 km from rest after delays. For successful intercept, the GBIs would have to be launched on the heading from the U.K.'s Fylingdales and would have to operate with the roughly 2 mr x 10,000 km = 20 km error basket produced by the radar's accuracy propagated to the GBI's maximum range.

At long range, there is not enough time for a shoot-look-shoot engagement strategy, so it would be necessary to commit several GBI simultaneously. To achieve a high probability of no leakage, e.g., a leakage of 0.01 percent, using independent GBIs with 90 percent kill probabilities would require about 4 interceptors per missile to achieve that compound penetration probability. Thus, the launch, or reported launch, of 4 to 5 missiles could trigger the release of 16 to 20 GBI, which could exhaust the inventory of an initial defense. To avoid launching all of them on a false alarm, it is necessary to know each potential adversary's missile inventory and types, which will require the coupling of intelligence assets into BMC2. Doing so also makes possible the integration of other assets that could be brought to bear to destroy missiles before launch on the basis of unequivocal warning. This "pre-boost" phase is an essential aspect of any effective missile defense system, but little more needs to be said here other than to note the need for the integration of the BMC2 for such operations with those for missile defense in a composite C2 system that provides the requisite information and options to those at the level to use it.

The BMEWS and UEWR deficiencies discussed above largely result from their line-of-sight range due to the Earth's curvature, not from their detailed operating frequencies or bandwidths. However, the need to provide prompt detection and metrics on launches from countries such as North Korea toward Alaska and Hawaii produces a need for a gap filler to provide coverage for them. Deploying a new XBR on Shemya Island at the extreme western end of the Aleutian Islands was the approach chosen by NMD to maximize such coverage and give some experience with and data from these new high bandwidth x-band radar sources. A XBR with a phased array antenna mounted on a pedestal that can scan mechanically to cover a wide range of azimuths was added to the radar suite in NMD.[100] While it was a good choice from the perspective of gaining track, knowledge, and experience, placing the XBR on Shemya, where the construction season is only a few months and the weather is hostile even then, put its construction on the critical path to NMD deployment. In reality, the ground-based radar–prototype (GBR-P) at Kwajalein could have provided the needed experience during the scheduled

100. BMDO, "National Missile X-Band Radar," *Fact Sheet* 119-00-11, November 2000.

GBI's development tests, while the Cobra Dane radar on Shemya could provide both the gap filler and threat RV measurements needed at a slightly longer wavelength.[101] Delays in the XBR construction project at Shemya also directed attention away from the radar's development, which later led to questions about its technological maturity.

Battle Management, Command and Control

Battle management is the system used to execute the intent of the commander. Command and control is the process the commander uses to assess the information his intelligence, surveillance, and reconnaissance (ISR) assets have gathered about the threat and formulate his actions. The two are often combined into BMC2, BMC3 if communications are included, or BMC4 if computer networks are also included for technical discussions. The acronym BMC4ISR is used to indicate the flow of information all the way from intelligence to execution, but is so broad that it has limited operational utility. NMD BMC2 contained the elements requiring integration to make the previously developed satellites, radars, and GBI into an effective weapon system.[102] The main tasks involved were upgrading the UEWRs, completing the GBR and GBI, and providing a BMC2 system for them that suited the commander. However, it was also necessary to integrate the result with the Integrated, Tactical Warning, and Attack Assessment (ITW/AA) that U.S. Space Command (USSPACECOM) provided to U.S. Strategic Command (USSTRATCOM, previously Strategic Air Command), the Secretary of Defense, and the President to support their decisions on the appropriate responses to evolving threats. Executing the former tasks was difficult; executing the latter was more difficult than anticipated.

Integration

Under NMD, TMD assets were not integrated with each other or with CONUS defenses, which was a step back from GPALS, in which they were to have been integrated into a global BMC3 supported by BP. That integration was lost in the Clinton Administration transition to TMD. If GPALS's integration had been retained, that would have removed the President from decisions about the allocation of defensives between strategic and theater threats, which could not be handled at lower levels by NMD's BMC2. Eliminating the need for detailed allocations at high levels would have made the concept of pre-delegation of authority—which is essential for effective responses to theater and boost-phase threats—familiar, so that it would not now be necessary to connect existing stovepiped systems with BMC3 appliqués.

There were technical problems in transmitting information from the primary warning, search, and track sensors to the GBI that was to use it to intercept approaching weapons. DSP and its SBIRS-High replacement were the primary sensors for search, early warning, and coarse track, particularly for launches beyond the horizon of radars. IR satellites had global coverage, but limited ability to type or track missiles. UEWRs had wide coverage, but limited discrimination capability, which might degrade during the engagement. X-band radars had limited search capability, but their coherent precision measurements exploited most known physical phenomena to improve discrimination.

These sensors' measurements of detection, track, typing, position, dimensions, absorptivity, polarization, spin rates, etc. were all thought to be valuable for discrimination, so they were to be transmitted to the GBI as *a priori* information about the threat objects it would see when it opened its sensors and began to search the threat cloud from a distance of about 500 km from. This information was to be included in the weapon data load (WDL) provided to the GBI before launch and updated with the data and guidance commands provided to the GBI periodically as part of the in-flight trajectory update (IFTU) sent to it by the in-flight interceptor control system (IFICS). The transmitted information was to include a target object map (TOM) generated by the fusion of the data from radar and other mea-

101. BMDO, "National Missile X-Band Radar."
102. BMDO, "Battle Management, Command, Control, and Communications," *Fact Sheet* 122-00-11, November 2000.

surements up to that time. The TOM was to be the primary input to the GBI for its initial attempt to recognize the objects in its field of view, on which the GBI would add its measurements to improve its confidence in discrimination in the roughly 50 seconds remaining.

In practice, the process has not worked quite that way. Radars have large bias errors as well as poor cross-range resolution, so the transfer of their information to GBIs is complicated by the fact that satellite and EKV sensors have different biases as well as viewing geometries. Thus, it is difficult to associate objects in the radar TOM with what the EKV sees unambiguously. The radar TOM contains three-dimensional estimates of the positions and sizes of the objects as seen at radar frequencies, while the EKV looks at a two-dimensional projection of them by its visible and IR focal plane arrays. Moreover, the EKV observes them from a different orientation, which is not known with precision. Only if the radars can measure, translate, and communicate their observations accurately, with little bias and with trusted covariances will they resemble the projections the EKV sees when it opens.

Radar TOMs have thus far been only moderately reliable in communicating their track and discrimination information to non-radar sensors. Since the EKV's sensors have good spatial and spectral resolution and its computers can quickly perform their own identification and discrimination, currently the optimal approach is for the EKV to use the TOM's rough feature measurements, reject its position estimates, and perform its own discrimination. That works effectively against primitive objects. However, it seems plausible that using all available data, including that from radars, should produce better handover and discrimination, so the fusion of all information sources continues to be explored for the long term. Collection of satellite and radar early warning information to the GBIs is conceptually straightforward, as the messages from them could be copied at the satellite ground stations and radar communication nodes and sent to the GBI BMC3 center for deconfliction and incorporation. In practice that could interfere with planned hardware and software upgrades, so it necessitated the development of an interim fusion facility, although it should not impact the timelines for initial defenses.

Command and Control

Command and control is a conceptually straightforward issue that was complicated by the lack of a decision on the NMD control center's ultimate location. Early in NMD, the commander in chief (CINC) of USSPACECOM was given the mission to develop the requirements for NMD, but the Army was given the responsibility to develop its main components, the midcourse GBI and XBR from GPALS. Which of the two would ultimately be the operator for NMD was to be decided after the decision to deploy, which was not made during the NMD program. As it is difficult to develop a generic command center that does not have a location, activity shifted to a combination of development in USSPACECOM's Joint National Integration Center (JNIC) and the Army fire control center for the GBI, which became the de facto BMC3 for NMD.

Attack Assessment

Strategic warning feeds were clear and relatively insensitive to the specific point of fusion, but there was a related issue that introduced additional complexity that is not yet fully resolved. For several decades, USSPACECOM had the mission of providing early warning (EW) and ITW/AA to Strategic Air Command, and then to its successor USSTRATCOM. ITW/AA was made up of two separate functions. ITW was used to differentiate the post-launch information provided by DOD sensors from the pre-launch information provided by intelligence community assets through other channels. EW and ITW were used to flush the alert bombers to improve survivability against large Soviet launches and to support decisions by the President and Secretary of Defense. It was important that the Attack Assessment (AA) provided to STRATCOM, the President, and the Secretary of Defense be timely and consistent, so they would have a common picture to support their decisions on responses, which were to be made within the half-hour flight times of ICBMs. As defenses were developed, it was thought important to integrate them into this process in a way that took advantage of the synergisms

between offensive and defensive responses without upsetting decades-long patterns of strategic decision-making.

At a minimum, CINC USSPACECOM and CINC USSTRATCOM needed to see consistent pictures of the missile threat. That could be accomplished by capturing all early warning feeds, deconflicting those that arose through multiple paths or reports, feeding them into a common display, and providing it to CINC USSPACECOM, CINC USSTRATCOM, the Secretary of Defense, and the President. That is essentially a database management problem, although a complicated one. A more complex issue arose from CINC USSPACECOM's mission of providing attack assessment, his estimate of whether an attack is on bombers, missiles, bases, command structure, the President, the Secretary of Defense, cities, or other targets. ITW/AA is difficult in that it must use imprecise, incomplete, and noisy sensor and human data to make a real-time estimate of the attacker's intent, which in reality might not be completely clear at the outset of the attack, at any point in it, or ever. Thus, attack assessment (AA) involves integrating information from many sources, sensor measurements, and human judgments, which must be performed while the missiles are in flight to have value. To achieve the needed speed, over a period of decades, AA was imbedded in partially documented computer codes written with varying objectives. AA is a logically separate function from missile warning and has developed largely independent of it. AA has significant human intervention to assure that it does not produce information that conflicts with that from intelligence and warning channels. As providing AA was a continuing mission of Space Command, while the ultimate control of NMD had not been decided, their relative priorities were not necessarily the same in USSPACECOMCOM, USSTRATCOM, and BMDO, the developer.

AA is intended to infer the intent of large attacks on significant parts of the U.S. military and/or civilian infrastructure. Thus, they must use approximations that are intended for large attacks, which may lose accuracy when applied to smaller attacks of concern today. AA models have overestimated the size of small attacks. For attacks of interest there could be significant discrepancies between the predictions of such models and direct sensor readouts of their actual size and direction that are used to direct missile defense. That makes the integration of AA displays with the direct pictures used to defend against small attacks potentially confusing. Constructing BMC2 that presents a consistent picture of the missile threat is straightforward if it is not required that it be consistent with that from AA. Constructing BMC2 that is required to be consistent with AA is complicated, involves importing legacy software, introduces strict configuration control, and requires stringent testing because of its offensive nuclear implications. Imposing AA on NMD BMC2 would be a major burden, which was not resolved. Uncertainty about which option to pursue caused delays in deciding how to present a common picture and how to operate on it. After the Cold War, the absence of alert forces, the reduced dependence on options that need AA, and the insensitivity of missile defense operations to AA all suggest that it should not be a constraint on NMD BMC2, but in practice, those constraints continue to exist and have led to delays, costs, and disconnects.

The maintenance of a common picture is dependent on the choice of the operating CINC. If USSPACECOM was given the role of commanding NMD, it would be natural for the CINC to use the picture resident in the Cheyenne Mountain Operations Center (CMOC), where he already exercises his NORAD Air Defense mission. There, the missile threat CINC USSPACECOM sees in executing his missile defense mission should logically be consistent with that he sees in his NORAD mission, which would imply that AA should be integrated into the BMD2 for NMD. If the responsibility for assessing global threats was given to another CINC, he might feel less pressure for integrating AA into his picture, although he would provide a picture of the missile threat, which he could accomplished by sending a copy of his display to USSPACECOM. The issue was not resolved operationally or technically during NMD. The current division of responsibility for the defense of CONUS discussed below has reduced the confusion that might have resulted from other assignments.

Direct Downlink

The provision of direct downlink information to theater CINCs gained their support for SBIRS-High to replace DSP, but introduced multiple reporting for each event. Direct downlink allowed theater JTAGS processing centers to receive downlinks from all satellites to which they had a clear LOS. Thus, multiple sites and Space Command's central facility could produce multiple reports for each event. That generates additional "ghost" tracks for each report, which requires duplicative communication and processing capabilities in each theater. A corollary problem that is potentially more serious is that the reports from theater JTAGS could feed into the central system that collects information from all satellites. That can produce discrepancies because JTAGSs only see a portion of the satellites, while the central ALERT accesses all of them.

ALERT should ultimately produce a more accurate picture, but theater CINCs face missiles with short ranges and flight times, so they need fast responses. ALERT's roughly 60 s strategic warning time could use a significant fraction of a theater missile's few minutes flight time—and almost all of the boost phase during which it is most vulnerable. JTAGS stations seek to minimize delays in warning by operating at low detection thresholds, accepting many false alarms, and screening them through intensive operator involvement. ALERT concentrates on the longer flight times of the brighter missiles that are threats to CONUS, so it can operate at higher thresholds where the false alarm rates are lower.

Those differing thresholds can cause the two systems to report detections and false alarms at different rates. Theater systems reporting launches that ALERT does not report could degrade ALERT's credibility. That could force it to operate at theater detection thresholds, which could lead to false alarm rates that are inappropriate for a strategic system. While it should be possible to find a compromise that satisfies the requirements of both systems, it does not appear to have been done. Direct downlink was accepted due to assurances that its cost and complexity could be contained and its introduction would not impact the SBIRS-High development schedule. There have been problems with both.

Testing

Testing missile defense components and systems has invariably been controversial. Decades of research, years of development, months of preparation, and weeks of execution are generally distilled down to the simple question "Did the interceptor hit?" That loses most of the detailed information on subcomponent performance, integration, and BMC2 that range experiments are intended to test. Thus, a test that seems successful to the developer can seem a failure to a critic, who may feel that a broken cable, failed cooler, stray object, or possible unintended cue suggests a fundamental weakness in the test or the system. So it was with NMD tests, which had three major phases: sensors and discrimination, intercept, and confidence testing.

Sensors and Discriminants

Range test series generally start with a set of flybys like those executed by NIKE, Sentinel, and Safeguard to test system integration and range procedures. NMD's test series was similar, though shorter. NMD performed three flyby tests to check control, communication, reporting, search sensors, and procedures under expected conditions. In the process it gathered useful data on sensor performance against a range of representative light decoys. Such tests are not usually intended to produce primary data, but in this case they provided useful results. Intercept Flight Test 1A (a repeat of aborted flight IFT-1) and flight test IFT-2 both demonstrated a significant ability to distinguish the balloon decoys expected from rogues from each another and from simulated RVs.[103]

103. E. Gerry, "National Missile Defense Update," presentation to Technology Working Group, Washington, D.C., January 26, 2001.

The tests were primarily designed to demonstrate that practical infrared sensors could detect, acquire, track, and discriminate a set of objects like those expected in the rogue threats for which they were designed. Their ability to discriminate light decoys was consistent with earlier experience that "simple" decoys are not effective against capable sensors. The objects deployed roughly spanned the space sizes, shapes, and absorptivities accessible to light decoys, so their results indicated that such decoys would not be credible under the conditions tested. Objects were discriminated on the basis of size, shape, temperature, and their temporal variation. The discriminants used were computed quickly in real time on the basis of distant observations, so they would not have been saturated by larger numbers of objects.

IFTs tested decoy separations characteristic of missiles without separating buses. Expanding or contracting the size of that cluster of objects to further separate or concentrate the spacing of objects would not introduce fundamentally new problems. It would advance or retard the time when the GBI sensors found a single object in their IFOV. That would not necessarily shift those times outside of those needed for divert, but it would aggravate CSO problems, which could force later, larger diverts. The tests used the same times and illumination geometries. The impacts of adjusting decoy absorption and reflection to match the RV's mean temperature and temporal variation for other conditions were to be examined in subsequent tests.[104]

The IFTs contained certain features for range safety and to simulate NMD elements that were not ready for testing. The additions were analyzed and publicized in advance.[105] Other illumination and trajectory conditions were to be tested in a series of 20–30 flight tests. Typical conditions involve transit through the terminator between darkness and light, which exposes light decoys by the rapid temperature change caused by their lack of thermal inertia. The flyby tests indicated that the EKV sensors, computers, and algorithms were appropriate for the design threats, which was an important result in itself. However, as in the previous systems, public discussions tended to bypass such rogue threats to discuss decoys of the levels of sophistication possibly developed during the Cold War. In those discussions, it was argued that Russia would proliferate such countermeasures to rogues, although international trends suggest that if it did so, those countermeasures might be used on Russia itself.

IFT-3 on October 2, 1999, was a successful intercept of a mock RV in a test with a modest number of associated objects. It was logically equivalent to the 1962 "intercept" of an Atlas by a NIKE-ZEUS, although IFT-3 was more complete in that it actually did execute a HTK intercept. All of the NMD elements were internetted and met their objectives, except the IFICS, which was not yet completed. Elements (and functions) tested included DSP (detect, acquire, track), GBR-P (detect, acquire, track, discriminate, and hit assessment), BMC3 (communicate, planning, weapons release, human in control), and GBI/EKV (deploy, navigation, acquisition, discrimination, divert, home, intercept).[106]

As in the 150-experiment range tests of nuclear systems in the 1950s, 1960s, and 1970s, NMD's two subsequent attempts to repeat this success failed, each for minor reasons.[107] IFT-4 on January 18, 2000, failed because its IR focal plane did not cool due to a blockage in the small throttling throat in its cryocooler. That was apparently due to the premature transfer of key engineers to other projects and the lack of quality control by their replacements. Thus, the EKV sensor warmed and lost sensitivity as it approached the target, which resulted in a 100 m miss, which was in itself unexpectedly close for those conditions. That reinforced the need for improved ground testing, but it also showed the

104. Sessler *et al.*, *Countermeasures*, Appendix A, p. 121.
105. L. Gronlund, D. Wright, and S. Young, "An Assessment of the Intercept Test Program of the Ground-Based Midcourse National Missile Defense System," Union of Concerned Scientists Working Paper, November 30, 2001.
106. Gerry, "National Missile Defense Update."
107. Gronlund, *et al.*, "An Assessment of the Intercept Test Program of the Ground-Based Midcourse National Missile Defense System."

problems testing could cause in a streamlined program as the cryocooler problem might have been caused by its excessive thermal cycling during ground tests. That was necessitated by inadequate test items, which were minimized in an attempt to cut cost and schedule. IFT-5 on July 8, 2000, failed because the KV did not separate from its booster. It gave useful system information, but underscored the need for better quality control. Subsequent range tests were delayed until all elements could be thoroughly tested on the ground with hardware-in-loop simulations. The five tests above represented the basis for the deployment decision on NMD.

Decision and Confirmation

The results of these analyses, simulations, HWIL simulations, and range tests were analyzed by several government and independent groups in the fall of 1999 and early 2000. The results of the reviews were mixed. NMD did not pass its Deployment Readiness Review (DRR) and was not forwarded to EMD. The reasons given were numerous and of uneven weight. Those highlighted by the Independent Review Team (IRT) that had reviewed NMD several times in previous years were typical. It stressed the failure of the COTS GBI booster, technical concerns about GBR development, and incomplete analysis of STAR design threats. The COTS booster failed because the evolving GBI design overloaded its DACS. It probably could have recovered without overly impacting critical path. Radar concerns were resolved and later shown not to impact the critical path. The analysis of the remaining threats and countermeasures was completed successfully; however, it left residual concerns that further antisimulation could undermine discrimination by the EKV sensors,[108] so the DRR decision was negative.

After the delays for reviews, the negative deployment decision, and the insertion of stronger simulation and quality control, IFT technical performance improved markedly. IFT-6 on July 14, 2001, was successful. A buffer on the GBR-P computer used to guide the GBI to acquisition overflowed after the completion of its portion of the mission, so to check the posited solutions to that and other concerns, the same configuration was tested in IFT-7 in December 2001, which discriminated a single large balloon from mock decoy and hit the target successfully. These successes gave greater confidence in the repeatability of such intercepts, much as the 14 NIKE-ZEUS tests under similar conditions had improved confidence in the reliability of nuclear intercepts.

Testing repeatedly under similar range conditions rather than moving rapidly to tests of multiple decoys, weapons, or conditions was the unanimous recommendation of external reviewers, who argued that a rush to test more complex decoys and geometries had diluted the careful effort needed to improve the GBI, test the EKV sensors, and unequivocally demonstrate the reliability of HTK intercepts. This emphasis on small, careful steps paid off in terms of understanding and performance, but did not satisfy external groups who wanted faster and more aggressive tests—whether positive or negative. It caused criticism of the program both by those who felt the first two IRTs had demonstrated all that was needed about rogue decoys and by those who felt tests should have shifted immediately to more stressing decoys.

Concerns about the use of surrogate components such as GPS signals to simulate UEWR search and a C-band beacon to make up for GBR-P's limited range will be resolved as prototypes are added to the test range, range safety issues permit other geometries and illuminations to be tested, and multiple launchers are completed.[109] Such gradually improving tests are compatible with a spiral development. Overall, the delay of NMD by reviews and the DRR decision may have been a blessing in disguise. The proposed NMD program was on an unrealistic schedule characterized as a "rush to failure," which could have frozen technical deficiencies into an inflexible program. NMD needed time to regroup; the negative DRR decision gave missile defense time to do that.

108. Sessler, *et al.*, *Countermeasures.*
109. K. Gildea, "Independent Panel of Scientists Endorses Missile Defense Test Plans" *Defense Daily*, January 10, 2002.

Countermeasures and Discrimination

Countermeasures and discrimination were important inputs to the Deployment Readiness decision, so it is appropriate to comment on them. It is not possible to give a full discussion of them, but it is possible to indicate which ones were the most stimulating. The challenges raised by the IFT's test conditions, which were largely governed by the ABM Treaty, were discussed in the previous section. Correcting them to allow for other geometries and illuminations will primarily take time, expense, and the testing freedom now permitted by withdrawal from the Treaty. This section concentrates on the technical issues raised by unclassified reports, which fell into the three major areas. There were a number of issues that were not all independent, which can be categorized as balloons, shrouds, and submunitions.[110]

As noted above, the balloon decoys tested in the first few IRTs were not effective, which was consistent with previous experience that such decoys are generally not effective against capable sensors.[111] The possibility of a balloon that could envelop the RV to minimize its contribution to external observables was not tested in the IFTs, because it is mechanically difficult to do remotely. If it was possible to mask the RV's IR and radar signatures with such enclosures, that would significantly degrade discrimination. However, deploying such structures in space is difficult and could, through contact between the RV and balloon, produce unintended observables.

Shrouds to cool an RV's surface to temperatures low enough to significantly reduce the sensitivity of conventional IR detectors would be effective on trajectories that remained in the dark.[112] If the trajectory crossed the terminator, the transition from night to daylight would improve detection by existing IR and visible sensors. Such shrouds might be possible at the level of space technology arguably achieved by Russia, but do not appear to be compatible with the space technology of lesser powers and rogues. If shrouds were employed in advanced threats, they would shift of the peak in the IR spectra to longer wavelengths. That would necessitate sensors with peak sensitivity at such wavelengths, which would probably require more cooling and larger apertures. Such modifications are quantitative rather than qualitative, so it is not clear whether shrouds would cost the defense or the offense more.

Submunitions represented a serious challenge to unitary GBIs, as they were big enough to be detected readily, but too small to represent attractive targets with unitary HTK KVs. GBI demonstrated that it is possible to intercept small targets with the current 80 kg EKV. To achieve a rough mass and cost balance with a 20 kg submunition, it would be necessary to reduce the KV mass by about a factor of four. That is possible with multiple KVs; indeed, it is possible to go farther. The tested BP KV weighed about 4 kg; carried an adequate suite of sensors to acquire, track, and discriminate small targets; and had enough divert capability to engage submunitions dispersed in a several kilometer cloud. A booster the size of the GBI's could carry about 20 such KVs, which could engage a like number of 20 kg rogue submunitions at a favorable mass exchange ratio of 20 kg/4 kg = 5:1. Such multiple KV interceptors, which were called "genius sand" during GPALS, were prohibited by the ABM Treaty and suppressed during the Clinton Administration. Now that the Treaty is no longer binding, such multiple mini-BPs are permitted and could be developed rapidly due to prior design and testing.

The questionable effectiveness of these countermeasures to NMD system is consistent with previous experience that terms like "simple," "easy," and "cheap" rarely apply to space systems. In all likelihood, any nation that wanted to use such countermeasures would find it necessary to test them to have any confidence in their performance, which would give the U.S. opportunities to observe them in action. Public discussions about countermeasures did reinvigorate the discussion of techniques for antisimulation. They had been discussed in Sentinel and Safeguard, but antisimulation was not needed

110. Sessler, *et al.*, *Countermeasures*.
111. *Ibid.*, pp. 35–39.
112. *Ibid.*, p. 84.

to defeat their UHF radars, so it fell into disuse. NMD brought it back to the fore and stimulated useful discussions of its strengths, although there could be little public discussion of its weaknesses.

Sophisticated antisimulation, which could make RVs look like decoys, would degrade midcourse discrimination and performance. For decades, U.S. and Russian RVs have been spin stabilized to improve accuracy and stealth. Doing so produces distinctive temporal signatures that are difficult to reproduce with light decoys. Tumbling the RVs would remove those temporal signatures, making it necessary to discriminate on the basis of other features.[113] However, some critics have argued against that approach, saying that in future tests "dummy warheads will intentionally be deployed so as to tumble end over end. This simulates the most primitive ICBM technology, where the ICBM is not spin-stabilized [which] causes its signal brightness to scintillate wildly [which] would never be used by any adversary, but would make it possible to distinguish warheads from decoys in flight tests."[114] Thus, there are opposing opinions on this countermeasure. Light decoys might still be distinguished by methods discussed above; heavy precision decoys probably could not. However, precision generally means mass, so it should be possible to intercept heavier decoys at favorable exchange ratios with the small, multiple KVs described above. The ultimate effectiveness of antisimulation is a matter of quantitative assessments of technical issues that cannot be done in the open literature because revealing the expected performance of discriminants would also reveal how to defeat them.

Summary of Developments in NMD

NMD reintegrated elements of defense against ICBMs that had been started in GPALS but deleted, de-emphasized, or reduced to technology during TMD. NMD also initiated programs to complete, test, and integrate its interceptors, sensors, and BMC2. Using partially operational interceptors, the testing program produced convincing evidence of the capability of HTK against rogue threats. The tests had a significant number of failures, which caused a delay in the DRR decision. They showed the need to improve quality control, test components on the ground, and simplify IFTs, which NMD did. That made subsequent range tests more useful for the assessment of system reliability. It also stimulated interest in more aggressive testing against uncertain and unexpected threats. Some assumed that NMD had demonstrated the ability to address the threats for which it was designed and should move to countermeasures that equaled or surpassed those from Russia. These discussions tended to overestimate the effectiveness of simple countermeasures, but drew attention back to the antisimulation issues discussed in previous decades, which enriched discussions in defense circles.

At the change of administrations, the future path of NMD was unclear, partly because of concern about the technical issues discussed above and partly because of cost. The widely discussed option called C1, a defense against a few missiles with rudimentary decoys, was estimated by the Congressional Budget Office (CBO) to cost about $45 billion and be completed by 2007.[115] C2, an intermediate step with more capability against countermeasures, was expected to cost $90 billion and be completed in 2010. C3, which had more capability against decoys and enough interceptors for 20 missiles was to cost $100 billion and be completed by about 2012. These costs were approaching the $119 billion of SDI Phase I, which were found excessive a decade earlier for more stressing threats. Moreover, the CBO estimates did not include the costs of radars, satellites, and command and control, without which midcourse defenses could not function.

Since the deployment decision was delayed on technical grounds before cost was considered, it was not necessary for the DOD to come to grips with the overall cost of a NMD deployment. NMD prima-

113. P. Coyle, "NMD Testing," seminar at Center for International Security and Cooperation, Stanford University, December 4, 2001.
114. T. Postol, "Why Missile Defense Won't Work," *Technology Review*, April 2002, pp. 42–51.
115. *Estimated Costs for National Missile Defense Deployments* (Washington, D.C.: Congressional Budget Office, U.S. Government Printing Office, 2002).

rily demonstrated, through a combination of tests and analyses, that nonnuclear kill could effectively address the modest threats for which it was designed in an arguably affordable way. Some argued that it might be able to address larger and more sophisticated threats, but the uncertainties about unknown threats and countermeasures undercut confidence in purely midcourse defenses.

Spiral Development Missile Defense

The Bush Administration did not immediately describe its missile defense program in specific terms. Only the broad outlines were given, which appeared to be consistent with most of the lessons learned from previous programs. Its overall goal was to develop defenses in all three layers as soon as possible, while retaining the flexibility to find early and effective combinations of existing defensive elements. That approach could lead to early deployments at levels approaching those advertised for earlier concepts, which were in danger of being lost through constant restudy and redefinition. For successful execution, it would need to select elements that could work together effectively in an initial deployment of useful effectiveness that could be followed by a succession of increments of progressively increasing effectiveness. The Administration embarked on such a process of deliberate selections, although the net effect of doing so was that for the first few years it essentially continued the main elements of the Clinton Administration's program.

These goals and the freedoms and responsibilities they implied were set out in a memo from Secretary of Defense Rumsfeld dated January 2, 2002, which directed that henceforth the Missile Defense program would be exempt from the Operational Requirements Documents (ORDs) and other requirements viewed as impediments to rapid decisions and progress.[116] The program was to be planned on the basis of "incremental realization of meaningful levels of capability," i.e., spiral development. That initially led to uncertainty over how the objectives of various spirals were to be defined, which was largely settled in the formulation of the initial capability discussed in the next section. Removing MDA from the ORD process left the CINCs with only an indirect voice in MDA's program formulation, so the Joint Staff commissioned the Joint Theater Air and Missile Defense Organization (JTAMDO) as the lead agency for communicating CINC recommendations to the MDA and vice versa. After this process had worked for a year and had apparently succeeded in accelerating the development of missile defenses, the DOD suspended the ORD process and went to a capability-based approach on all major developments.

Summary and Discussion of Earlier Approaches

Implementing a requirement-free development process requires flexibility and selectivity. For it to work, the MDA must have freedom to choose and the technical taste to make good use of it. That combination has been difficult to achieve in the past. Some insight into the difficulties comes from an examination of the earlier programs discussed above, which contain examples of both successful and unsuccessful approaches. NIKE successfully used a DOD requirement-driven development system to produce innovations in radars and boosters that are still in use by current systems, although it did not lead to a fielded system. Sentinel used a similar system to successfully integrate those advances in radars, boosters, and warheads into a system for limited threats, although it had vulnerabilities that prevented its deployment. Safeguard also used that system, but applied it to a mission to which those technologies were not suited. Thus, its system was technically successful, but failed in its overall objective. LOADS used a similar system to define the LOADS nuclear defense of Minuteman pre-

116. D. Rumsfeld, memo to the Missile Defense Agency, January 2, 2002.

ferred by critics of Safeguard. The development was successful, but LOADS was not deployed because of Treaty concerns, to which that management structure could not respond. It should be noted that the requirements-driven system under which the nuclear systems were produced three to four decades ago was a much more flexible one than the ORD process that had to be suspended for MDA.

The Interim programs were primarily R&D, so they used systems appropriate for that level of maturity. The Army and DARPA both used systems that stressed competition and innovation to develop the nonnuclear HTK and DEW programs that are the basis for both near- and long-term defenses. The Army and DARPA programs were primarily driven by technology, which resembles a spiral approach, but at a lower level of development. The Interim program demonstrated that a combination of adequate research, technical competition, and selective management could be productive, although its research did not lead directly to fielded systems until they were swept up by SDI and implemented in GPALS.

SDI used a flexible system like that of the Interim program to accelerate the advance of technology. In its later phases SDI shifted to an emphasis on maturity, cost, and a structured approach to counterforce threats, but it did not produce deployable systems. BP, which could potentially meet the levels of performance needed to address SDI's goal, emerged toward the end of SDI from a small, very flexible, compartmented program that was largely isolated from SDI's management.

GPALS used a largely requirement-driven management structure to redirect SDI technologies to post–Cold War threats. It effectively down-selected existing options in ground- and space-based sensors and interceptors and formulated programs for them that were arguably adequate for the residual threats. GPALS also correctly anticipated the need for and started the interceptor and satellite developments to address emerging theater and rogue threats. It was an example of a requirement-driven, but flexible system that was successful in decisions, innovation, and deployment. It was the only missile defense management approach capable of performing all three. It was only undercut by an Administration that was fundamentally opposed to its main elements on policy grounds.

Through deliberate choices executed by a standard largely military requirement-driven TMD program, the Clinton Administration surrendered the U.S.'s lead in advanced strategic defenses. TMD consolidated and tested needed capabilities in theater interceptors and sensors with inappropriate BMC2, which demonstrated that system's tendency to efficiently pursue sub-optimal goals. NMD used a similar approach to revive and milestone GPALS's midcourse technologies. The resulting system was not approved for deployment for reasons that it could have addressed had it been more flexible, which again demonstrated that system's inflexibility and tendency to efficiently pursue sub-optimal goals.

In summary, the initial phases of NIKE and Sentinel were requirement-driven but innovative. Safeguard used their system and technologies to inflexibly pursue an inappropriate objective, which produced an unacceptable system. The Interim Phase used a flexible system to stimulate innovation, but produced no actual systems. SDI institutionalized that innovation. GPALS used a conventional but flexible system to produce defenses arguably adequate for projected threats. TMD reverted to a conventional system to manage the theater system. NMD used a similar system to revive GPALS' midcourse defense and did not integrate them with TMD or restore the global elements needed for effectiveness.

Thus, earlier programs fall into three groups. The first is the requirement-driven nuclear systems, which were initially innovative but became rigid and unproductive when directed to an inappropriate objective, that are joined by the TMD and NMD nonnuclear systems, which were rigid from the outset. The second group contains the Interim program and SDI, which were innovative, but did not lead to deployable systems. The third contains only GPALS, which was requirement-driven, but flexible. It programmed deployable systems and anticipated the need for new technologies. One success out of

six is not completely reassuring, but it is encouraging that the success was from a flexible management approach rather than one of the requirement-driven systems. However, the Interim and SDI programs were creative without producing deployable systems, while GPALS was both creative and productive under comparable funding, technology availability, scientific manpower, and policy constraints, which suggests that harnessing creativity to productivity may be as much a function of the leadership and judgment of program leaders as the technology available to them.

For the last five decades, missile defense has oscillated between innovative and conservative approaches on time scales of 3–5 years, which is so short that neither phase has produced useful systems. The history consists of a series of compressed programs, most of which apparently could have addressed the threats for which they were designed—except the one that was actually deployed. GPALS stands apart; it made major shifts to introduce the new technologies needed to address evolving global and theater threats. The recent TMD and NMD programs can be interpreted either positively as appropriate responses to emerging theater and rogue threats or negatively as hostages to the ABM Treaty reflecting a loss of confidence in the technologies that could support effective global defenses. A broader vision and more flexible approach will be needed for projected threats.

Spiral Development

Spiral development initially moved toward better integration of existing technologies into defenses against the threats defined by the previous administration. Its goal was to steer between the trap of technology without deployment and that of premature deployment for inappropriate threats to find a sequence of progressive steps appropriate to evolving threats. Doing so would require the reintegration of TMD with NMD and the restoration of the global layer needed for robustness against sophisticated threats. MDA initially attempted only the first, with a spiral development program intended to produce "sockets" into which MDA could plug new technologies as they became available. Its goal was to deploy a defense built on NMD as early as 2004, with block upgrades roughly every two years thereafter. It was argued that deployment on such a timescale would have to start with the NMD's midcourse defense based on GBIs in Alaska, with bases, interceptors, and radars to be added later to improve coverage and performance. It was recognized that it would ultimately be necessary to restore the global component if the program was to be more than a renaming of NMD, with all its known weaknesses. However, it was argued that its introduction would have to be delayed until later blocks.

MDA was given the freedom to assemble a useful, evolving program from the elements of previous programs. Doing so does not require perfect people, but it does require a competent group of capable and experienced professionals who can select the appropriate combinations and who are determined to see them through to deployment. The MDA has addressed that need by establishing a National Team composed of representatives from each of the major contractors to determine the appropriate technologies for inclusion in each block of spiral development. Their work is not completed; only the initial step discussed below is known in detail. However, it is possible to review the options MDA has at its disposal for that selection, which are conveniently done by starting with theater systems and working back toward the boost or global phase, which moves in the direction from more to less developed systems.

Theater Segment

When tested, affordable, and available in sufficient numbers, existing HTK systems should be appropriate theater components of a spiral development program. PAC-3 and THAAD could be integrated into a two-layer system into which Arrow and Allied systems could be integrated, given appropriate and flexible BMC2. Sea-based Navy Area Wide and NTW have the potential to be similarly complementary with one another and with ground theater systems. The ABL should be an effective theater defensive system in later spirals and could have significant impact as a high precision boost-phase

sensor for ICBM track in the first or second spiral. The transfer of PAC-3, Navy Area Wide, and MEADS to the services would appear to conflict with plans for integrated defenses, but those systems are already at stages in development where service testing and procurement policies are appropriate. Their transfer need not impact effectiveness if MDA asserts its role in assuring that their BMC3 supports the rapid flow of global warning and track data to them, processing of the data from theater and regional sensors, and promptly transmitting the data to the commanders responsible for intercepting missiles that leave the theaters.

While HTK interceptors can be supported by their intrinsic radars, their performance could be improved by satellite information, warning, and track. Current theater systems are limited by their radars to coverage of point targets or small areas that scale with warning time. Perhaps the best way to improve their performance is by internetting their sensors and information at the data or track level and by efficiently allocating interceptors through integrated BMC3. The challenge is to avoid the approaches that drive theater systems into separate stovepipes, which would lose the synergisms that would result from their integration within theaters and with external systems.

Figure J.1 is for the intercept of a theater ballistic missile launched from a distance of 600 km with a maximum velocity of 2.5 km/s as a function of interceptor velocity v for release delays of T_{delay} = 15, 30, and 60 s (See Appendix J). The missile is assumed to be detected a distance D = 300 km from the interceptor launch point, which is greater than the range a PAC-3 radar could search with such a handoff from another sensor and slightly less than the range a THAAD could search without a handoff. For a 60 s delay, the intercept time drops from 145 s at v = 1 km/s to about 125 s at v = 3 km/s. For a 30 s delay, it drops from 135 s to about 105 s by an interceptor speed of v = 4 km/s. For 15 s delay, it drops to from 130 s to about 95 s by 4 km/s. This 50 s overall reduction could increase the defended range of a 3 km/s interceptor by up to (3 km/s/√2) x 50 s = 100 km, which is significant tactically.

The figure shows the forward extent of the defended footprint of a ground-based interceptor, which is the most difficult dimension of the footprint to achieve. For an interceptor velocity v of 1 km/s, a 60 s delay gives a forward defended radius of about 50 km, which is about half the kinematic capability of an interceptor of that speed. Its radius increases with v to about 85 km by a speed of 3 km/s, where it saturates due to delay time. The curve for a delay of 30 s gives a radius of about 70 km for 1 km/s, rising to about 120 km at 4 km/s before saturating. The top curve for a 15 s delay increases from about 75 km to 135 km at 5 km/s. Thus, in intercepting SCUD C and D class missiles with 300 km sensors, there is about a factor of two benefit for increasing speed by a factor of two at each delay and a factor of two for reducing delay times by a factor of four at high speeds. However, there is only a 30 percent benefit for decreasing delay times at low speeds. A 0.8 km/s PAC-3 could thus defend a forward radius of 45–70 km and the 2.7 km/s THAAD could defend out to 80–125 km, depending on their delay times. As the forward footprint scales roughly as the square of the radius, THAAD should defend an area about $(125/70)^2$ = 3.2 larger than PAC-3 for the minimum delay time for each, which is appropriate for a two-layer defense.

Figure J.2 shows the effect of increasing the detection range to D = 600 km against 600 km range missile for interceptors of speed v = 0.8, 2.7, and 5 km/s launched with T_{delay} = 60 s. Such an effective detection range would require launching the interceptors on warning and trajectories from external sensors. The radius for the PAC-3 increases to about 60 km by a range of 400 km, where it reaches the kinematic limit of its booster. The higher speed missiles' radii continue to increase with detection range out to about that of the launcher, where they reach values on the order of 250 km. Doubling interceptor speed at large detection distances only increases the defended radius about 50 km, or 20 percent. Those values only increase about 10 percent if the delay time is reduced by a factor of two. The area covered by THAAD or a longer-range interceptor is greater than that of PAC-3 by about a factor of $(250/60)^2$, which is also appropriate for a two-layer defense.

The extension of the threat from SCUD derivatives to missiles with 1,000 km ranges requires HTK interceptors with improved sensors, guidance, and control for intercepts at higher closing velocities, which requires modifications of known technologies more than new developments. Control bandwidths are proportional to closing velocity and inversely proportional to flight time, which leads to more sensitive and less stable guidance. At some speeds these quantitative improvements become so large they become qualitative. That defines what will be included in the initial block and what will be delayed for later upgrades. Terminal systems will probably start with the existing terminal systems, internet them for improved compound kill probabilities, improve their communication with external BMC2 for better cuing, and ultimately shift in later blocks to commit on external sensor detection and track to achieve roughly $(250/80)^2$ = 10-fold greater defended areas against long-range theater missiles.

Missiles with ranges of less than 1,000 km spend only a fraction of their flight time above the atmosphere. A 500 km missile spends less than 40 percent of its time above 125 km, so competent theater radars could observe it for most of its trajectory and discriminate light decoys on the basis of atmospheric drag (See Appendix C). Longer missile ranges provide opportunities for exoatwospheric or high endoatmospheric intercepts where missiles have little opportunity to maneuver, which could be exploited by sensor and interceptor technologies developed in SDI and GPALS. Terminal systems offer limited coverage against such missiles, as they are saturable, penetrable, and avoidable, but for the intercepts they can address, terminal systems could provide low cost, localized protection.

Terminal Segment

There is a growing concern about ballistic missiles launched from ships or submarines positioned a few hundred kilometers off the U.S. shore. Defense Secretary Rumsfeld has described the threat as "Countries have placed ballistic missiles in ships…all over the world. At any time there's any number off our coasts—coming, going. On transporter-erector-launchers, they simply erect it, fire off a ballistic missile, put it down, and cover it up. Their radar signature's not any different than 50 others in close proximity."[117] This threat is magnified by the fact that about half of the U.S. population and value lie within about 100 kilometers of its shoreline. Systems suited to theater terminal intercepts should be suited to defense of coastal areas of the U.S., as offshore missiles rely on technologies, trajectories, and countermeasures similar to those for which those theater defenses were designed. Thus, little additional development would be required. Providing such defenses would largely be a matter of producing additional interceptors and radars and deploying them around large coastal cities.

Theater Ballistic Missiles

The figures in the previous section show the expected capability of ground-based systems such as PAC-3 and THAAD acting separately or in concert, with or without external cueing, against short-range missiles launched from the ocean nearby. Acting separately with the 60 s delay times characteristic of current systems, they could provide forward footprints of 40–80 km radii, which would require basing close to each major city. With external cueing, reduced release times could probably increase those footprints to 60–120 km, which would still require basing close to each major city but could usefully increase the distance from interceptor bases to their centers. With external warning and track they should be able to increase footprints to 60–250 km, which could cover either coast from less than a half dozen sites.

PAC-3 and THAAD could provide coastal cities a natural two-level defense, which is appropriate for such high value targets. The main development required for such defenses would be to internet PAC-3 and THAAD radars with satellite and UEWR warning information. Their prompt determination of launch position could reduce their organic radars' search requirement, increase their track ranges, and

117. D. Rumsfeld, press conference, Pentagon, September 16, 2002.

support early interceptor release and efficient firing doctrines. The time scales for such intercepts are short, but the only C2 required from higher echelons is release. There are no critical issues in interceptor allocation that require inputs from higher levels. The missiles and interceptors have short ranges compared to those between major defended areas and continents, so each engagement is effectively local. The main issue is the efficient allocation of interceptors between U.S. cities and deployed forces, which could be adjusted in response to changing estimates of the threat. Siting could initially present a problem, as the range of PAC-3 acting on organic radar is short, so for effectiveness it would have to be deployed close to or within the defended cities themselves, which could have undesired impacts similar to that of Sprint. THAAD has a much larger defensive radius, so it could be deployed well outside of cities and still achieve full coverage, so it should have less impact, when available.

Sea-based systems offer a useful adjunct to the defense of cities because they are positioned some distance off shore, which allows them to take advantage of the large footprints that modest interceptors can generate in protecting targets behind them (See Appendix J). Figure J.3 shows the backward extent of the footprint sea-based interceptors of various velocities and release delays of 60 s could generate against 3 km/s missiles launched from 1,000 km offshore. The intercept times range from 100 to 170 s for detection at D = 200 km; 160 to 460 s for D = 300 km; and 210 to 760 s for D = 400 km, with most of the increases coming at v greater than 2 km/s. Such detection ranges are appropriate for detection and commit on organic radars or external warning only. At an interceptor v of 1 km/s, the backward footprints are 30, 80, and 100 km respectively. By 2.5 km/s they increase to 150, 625, and 625 km, respectively, as the 400 km detection range is undercut by the kinematic limits of the interceptor. In addition to providing useful coverage by themselves, these footprints would provide useful overlays to ground-based defenses ashore.

ICBMs

However useful such defenses could be against missiles close to shore, it is not clear that they would be effective against longer-range missiles such as rogue ICBMs, nor is it clear that they lie on the path to defenses that would be effective against such threats. ICBMs approach at higher velocities and altitudes and can use countermeasures that could remain effective until well into the atmosphere. If the weapon's decoys have been discriminated or if the missile did not deploy any in the first place, theater interceptors like THAAD might be restored to the velocities and sensors required to intercept them.

The performance of current ground-based defenses against rogue ICBMs can be treated with an extension of the models used to estimate their performance against theater missiles in the previous section (See Appendix J). Figure J.4 shows the forward defended radius as a function of interceptor velocity for delays of T_{delay} = 15, 30, and 60 s after the launch of ICBMs with minimum energy velocities of about 7.2 km/s and reentry angles of 22.4°. Their RVs are assumed to be detected at a distance of 600 km from the interceptor launch point, which assumes external cueing and targets with no or only discriminated decoys. For T_{delay} = 60 s, the interceptors have a radius of about 15 km for all velocities, which would require proximate basing (and deposit most debris) within the cities defended. For a delay of 30 s, the radius would increase to about 30 km for PAC-3 and 55 km for THAAD. For 15 s PAC-3 would increase to about 40 km and THAAD to 80 km. Thus, for ICBMs, THAAD's defended area is reduced by a factor of $(250/80)^2$=10 from that for offshore SCUDs from Figure J.2.

Figure J.5 shows the effect of cueing from external sensors against undecoyed or discriminated RVs, which is represented by increasing the detection distances to intercontinental values. The 1 km/s interceptor reaches its kinematically limited radius of 100 km at a detection distance of 1,500 km. The 2.7 km/s interceptor reaches its kinematic limit of 625 km at a distance of 3,500 km. The 5 km/s interceptor can reach 1,400 km using the full detection range of 4,500 km. Thus, the higher speeds of ICBMs greatly reduce the defended radii of ground-based interceptors designed for theater missions, but the use of external track information to cue or commit them could more than offset those reductions to

produce defended footprints of the size of the coastal regions of the U.S. In such configurations, the ratio of the THAAD and PAC-3 forward radii is about 6.3, which is large but still useful. It would be advantageous to use still higher velocities against intercontinental targets without decoys. There are programs to achieve such velocities either by using much larger missiles or much smaller KVs, but it is not clear either is a near-term program that would impact early spirals, as now pursued.

Sea-based systems would also have significant backward coverage against undecoyed ICBMs. For detection ranges of D = 600, 1,200, and 2,400 km, interceptors with delays of 30 s would intercept in about 100, 200–300, and 400–700 s, respectively, for v = 1 to 5 km/s. Figure J.6 shows that they would produce backward coverage of about 40–140, 100–840, and 100–2,230 km, respectively. NTW interceptors with a maximum speed of 3 km/s operating on organic assets might produce coverage of about 100 km; external cueing might produce 400 km; and external warning, track, and commit might produce about 850 km. The first would provide a useful local overlay; the second could cover the coasts from two to four sites; and the third could generate footprints covering much of the interior of CONUS with deployments off each coast. The principal needs for full exploitation of these footprints would be higher interceptor speed, integration of BMC2, commit on non-organic sensors, and adequate external sensors to discriminate threats that could use decoys.

Sea-based systems also have a useful capability to provide defenses against theater missiles or ICBMs in boost, which is discussed below along with other follow-on systems, as that capability is not likely to be implemented in early spirals on current programs.[118] For maximum forward footprints or backward coverage, ground- and sea-based interceptors need speeds in excess of 3 km/s, which would cause strong heating that could redout current IR seekers. To avoid it, current interceptors would have to intercept exoatmospherically. Intercepts at apogee would do so, at a cost of about a factor of two in interceptor range and footprint radius. While the resulting footprints would still be useful, they would largely offset the advantages of higher velocities.

If the threat decoys had already been discriminated, adding exoatmospheric intercepts by local theater interceptors would just add a few more intercepts at the end of midcourse, which the GBI already addresses competently for non-decoyed threats. Thus, a logical question is whether it would be preferable to add more GBIs and bypass the expensive development of theater interceptors with less competent sensors for late midcourse intercepts. If a modified version of THAAD was given better sensors for such intercepts, its sensor and booster would probably converge to something like the GBI.

If threats were not discriminated before reentry, they might be discriminated on the basis of atmospheric drag, as MSR discriminated for Sprint in the earlier Sentinel and Safeguard systems. For such intercepts, the factors discussed above that governed Sprint's design were primarily kinematic. An RV cannot reliably be distinguished from its decoys by drag at altitudes much above 100 km. At a speed of 7.2 km/s, that gives the interceptor about [100 km/sin(22.4)]/ 7.2 km/s = 36 s to react, which pushes its intercept altitude down to a few 10s of kilometers. For useful footprints, the interceptor must reach that altitude quickly, so it must achieve high acceleration and velocity in the dense atmosphere. That produces heating that would blind IR sensors in development, which forces the interceptor to radar command guidance. Since its accuracy is incompatible with that needed for hit to kill, its warhead must be nuclear, so the interceptor reduces to Sprint and the system approaches the LOADS system developed at the end of the nuclear phase. While that approach is technically feasible, it did not find public acceptance when it was advocated for Sentinel in an era when nuclear explosives were viewed more favorably.

HTK interceptors could arguably be developed that could operate on the basis of discrimination on high-altitude drag using today's capable x-band radars. The Army's early HEDI explored the technical issues in that altitude regime and achieved successful intercepts. GPALS's E2I was to have carried

118. Cooper, "Reviving Effective Programs to Protect America from Ballistic Missile Attack."

it further, but it was terminated by the Clinton Administration. Reviving them could take a decade, so such interceptors do not appear to impact early missile defense spirals. The terminal phase could be a useful layer, and much the same technology and C2 could suffice for cued and layered theater and coastal launches. However, it is not clear that exoatmospheric terminal intercepts would be cost effective relative to additional GBIs for discriminated threats. High endoatmospheric intercepts could be more effective, but nonnuclear versions are at an early stage of development. The terminal phase does not appear to have a nonnuclear near-term solution for countermeasured threats. Such systems appear to be about a decade away, but the leverage they offer appears to justify continued research and development.

REGIONAL SEGMENT

Regional missile defenses were not clearly defined by the ABM Treaty; thus, they were not developed aggressively. Indeed, they were opposed. In the mid-1990s, Paul Warnke, former Director of the U.S. Arms Control and Disarmament Agency and Strategic Arms Limitation Talks negotiator, testified against giving PATRIOT even the limited capability of a PAC-2 because it would "violate the ABM Treaty." The Treaty's ambiguity caused the U.S. engineering community's system development to stagnate, although it continued to conduct limited BMD R&D. In past decades regional defense was less of an issue, because there were few missiles with 1,000 to 5,000 km ranges other than SLBMs and SS-20s. The NIEs cited earlier indicate that there are now significant numbers of such missiles, and their number is likely to grow if no effective counter to them is developed. Recognizing the growth of these regional missile threats complicates the discussion of technology, but could simplify the acceptance of missile defenses. Regional missile ranges correspond closely to the areas of responsibility (AORs) with which regional CINCs now defend, and C2 issues become clearer when intra- and inter-regional defensive activities are discussed in the same terms as other force elements.

The sensors and interceptors developed for terminal and midcourse systems could be applied to RMD with appropriate modifications. Figure J.2 shows that theater PAC-3 and THAAD interceptors operating on organic radars would be restricted to forward footprints of 25 to 50 km. With external cueing they could achieve about 60–120 km. With external track and commit at detection distances of 600 km, they could achieve coverage on the order of 60–250 km against 600 km missiles. Longer detect ranges would not be useful as the kinematic limit of the THAAD is only slightly larger than that.

For maximum RMD interceptor range, the fundamental requirement is rapid launch detection and accurate trajectory determination. With adequate cues, their defended footprints are proportional to the product of interceptor velocity and flyout time, which can reach regional dimensions.[119] Figure J.7 shows that for missile detection immediately after launch, PAC-3 would be limited kinematically to 64 km; THAAD's range would increase from 150 to 730 km before limiting; and a 5 km/s interceptor would reach about 1,200 km, absent sensor redout or other limitations. These footprints would support useful layered regional defenses, for which their ratios are appropriate. Thus, by using internetted radars and external sensors with regional ranges, appropriate interceptors could close with higher velocity regional missiles, defending footprints twofold to fourfold larger than those attainable with their organic radars. Some of the needed improvement in interceptors could come from restoring the performance removed in their descoping in the TMD phase. Such improvements are compatible with block upgrades. Reaching their full capability would require a fully integrated BMC2, which should be introduced in accordance with a long-term plan, probably extending over several blocks.

Regional missiles differ from theater missiles in that the deployment of their warheads and decoys takes place at ranges that are not observable by radars. Moreover, their warheads spend enough time above the sensible atmosphere to make decoys effective. While observation in midcourse by precision

119. *Technologies for Ballistic Missile Defense: An American Institute of Aeronautics and Astronautics Assessment* (Washington, D.C.: American Institute of Aeronautics and Astronautics, 1993).

x-band radars could eliminate light decoys, heavier decoys could still be effective on missiles with separating buses, fast deployment, and antisimulation. At that level of technology, which might appear in regional as soon as in global threats, the measure-countermeasure competition could be as stressing as that in ICBM engagements. If so, constant observation of the weapons and decoys and integration of information from multiple sensors will be critical for discrimination. The key element is data integration, which will take significant development, which is why regional systems are likely to emerge in subsequent blocks rather than in early deployments.

Midcourse Segment

Midcourse offers the longest time for engagements as well as efficient areal coverage for a given number of interceptors; although for the reasons discussed above, those advantages have been undercut in recent decades by concerns over countermeasures and performance in disturbed environments. However, when early deployment is at a premium and the timelines for rogue threats are uncertain, midcourse defenses are an appropriate near-term response. Given the momentum of recent programs, it was natural that the spiral program started there with the reuse and improvement of existing sensors and interceptors.

Space-Based Sensors

DSP is essential for prompt, robust launch detection and booster track for ICBMs. SBIRS-High, its intended replacement, should extend track more accurately to burnout for both ICBMs and theater missiles. Satellite IR sensors phenomenology is more closely related to that of HTK interceptor sensors and is thought to be less sensitive to details of missile design, disturbed backgrounds, offensive countermeasures, Earth curvature, and non-sovereign basing than radars now in use. The DSP/SBIRS qualitative and quantitative measurements are the first indication of anomalies in the threat, so they are key ingredients in the formulation of the extended state vector for each threat element. While the ability to incorporate this information into radar and EKV TOM is currently limited, it should improve in later stages of deployment. Observations and tracks from SBIRS-High will also be available to SBIRS-Low when it becomes available in a form useful for track and discrimination.

Ground-Based Sensors

While satellite early warning is essential for prompt regional and global coverage, UEWR and XBR provide detection and coarse trajectories that are good enough to commit GBIs and attempt discrimination. UEWRs have known problems with backgrounds and discrimination, but have demonstrated an ability to cope with them to some extent at a modest cost. In time, their measurements could reach a level that justifies fusion into the multi-phenomenology TOMs provided to GBIs. While UEWR, GBR, theater radars, and IR sensors each have range, horizon, frequency, and bandwidth limitations, fusion of their observations could fill most elements of the extended state vector of the target cloud that are of value to the EKV, and each orthogonal mode of observation complicates the generation of credible countermeasures.

While BMEWS radars will approach their performance limits in the UEWR upgrades, GBRs still have much to gain by proliferation, additional metrics and discriminants, and experience. They are likely to be the main element of discrimination until LEO satellites emerge in the next decade. GBRs are much less affected by natural and disturbed backgrounds than UEWRs, although there are concerns about nuclear effects in attacks with significant numbers of detonations. As those uncertainties could not be removed short of testing, they argue for the retention of multiple sensor phenomenologies and platforms, particularly in early phases of deployment.

In the near term, the combination of DSP warning, UEWR coarse track, and EKV discrimination allows the GBR at Shemya to be taken off the critical path to deployment. Cobra Dane can act as the needed gap filler, and GBR-P can be used for experience with x-band. XBR could be deployed on the

East Coast of the United States, in Europe, or in the Mideast to provide early and accurate trajectories and simplify GBI timelines for the defense of allies. These improvements in track and discrimination can only be fully realized when the information from strategic, regional, and theater sensors is transmitted, fused, and disseminated promptly, which would probably have higher payoff in the near term than improvements in the sensors themselves.

Interceptors

The GBI and its EKV are appropriate technologies and systems for near-term threats, which are expected to have few weapons and only rudimentary decoys. While each could be improved with availably technology, their overall performance is likely to be improved more by improved basing and BMC2 than by technical details. Their effective use of long-range satellite warning and radar track for prompt commit should permit global coverage with efficient interceptor allocation. Significant improvements in performance and cost are possible through the use of lighter sensors and components developed for space-based and boost-phase interceptors. In time, more effective use of the information available through radar TOMs should improve object correlation and discrimination beyond that possible with the EKV's on-board sensor. Such improvements are well suited to subsequent block modifications.

Discrimination

Discrimination was an acknowledged weakness of Sentry and Safeguard, which used some of the same radars as the current systems. It was even more of a concern for NMD due to the proliferation of countermeasure technology over the intervening decades, particularly after the end of the Cold War. The classes of decoys discussed above with respect to NMD discrimination should not be stressing to the suite of on- and off-board sensors and discriminants discussed above. Balloons seem less effective in practice than in theory; shrouds are technically difficult and inconclusive; and submunitions imply their own counters now that Treaty prohibition on multiple interceptors has been removed. It is not excluded that more effective countermeasures will be found, as the limits of antisimulation have not been explored, at least not in public. However, if threats conform to expectations about decoys and countermeasures, it should be possible for developed elements to cope with them.

The combination or fusion of the radar and IR discriminants from XBR, DSP, SBIRS, and the EKV itself appears to provide significant margin against uncertainties and improvements in the threat. Until theater or space-based sensors and interceptors are developed that can observe all phases of deployment and collapse the deployment time and space available to the attacker, weaknesses in current discrimination sensors will continue to produce uncertainty about the performance and possible abrupt degradation of the overall midcourse defense. These issues will be a source of continuing controversy, although they should be reduced when it becomes possible to inspect all missiles, buses, and countermeasures in flight, which is an auxiliary capability that should be produced by the satellites and SBI deployed in later blocks.

Test Bed

Recognizing that BMC2 integration was a key element to any future missile defense deployment, in 2002, the MDA defined a "Test Bed" to integrate a subset of the elements that would be needed in subsequent deployments.[120] It included a number of components that were already in use or development for testing. Key elements included the GBR-P at the Reagan Test Site (RTS) at Kwajalein in the Marshall islands, two silos there for testing multiple simultaneous intercepts, an IDT communication site to provide commands and IFTUs to GBI in flight, and a RTS Ground Flight Control (GFC) sys-

120. R. Kadish, "The Missile Defense Program: Fiscal Year 2003 Budget," testimony submitted to Senate Armed Services Committee.

tem to command and control GBI intercepts. These components were needed to relax the inflexible geometry used in earlier tests and to start the substitution of operational elements for the test components used earlier. Additional test bed elements included the Beale Air Force Base (AFB) UEWR and the Cobra Dane radar at Eareckson Air Station (Shemya), whose use as a gap filler took the XBR off the critical path. Adding these radars to the GBR-P at RTS afforded an opportunity to internet one of each type of radar relevant to later configurations.

An interim DSP satellite feed was provided at Boulder, Colorado, which began the full integration of the inputs from warning satellites into missile defense system. That capability had been declared a valid adjunct role in GPALS and NMD, but had not been exercised in the Gulf War or implemented in the earlier NMD system. It also provided the opportunity to start the process of fusing the data from multiple radars with the IR data from warning satellites as the first step in upgrading the construction of multi-phenomenology TOMs.

As expected, the most numerous elements were those for C2 and communication, which involved an extended, largely fiber Ground Control Network (GCN) and its IDTs terminations at Vandenberg AFB (VAFB), Ft. Greeley, and Eareckson Air Station. GFC nodes were also added at Ft. Greely and the Joint National Interim Command at Schriever AFB in Colorado. Together they provided much of the global net needed to transmit data from all relevant sensors to the sites for theater or central fusion. They provided at least interim versions of each level of the C2 systems required to support later missile defense systems.

The Test Bed also includes two novel elements. The first is a dedicated Aegis ship, which is added to study the issues involved in coupling information from high bandwidth sea-based radars into a largely Link-16 GBI BMC3 system. It should also be useful in studying the problems in integrating data from a Navy CEC system that has a distinctly different architecture and data flow than other elements of the Test Bed. It should also produce a mobile capability for repositioning a gap filling radar for detection and track of missiles from regions that were not adequately covered by other radars, which could be useful for northerly trajectories from North Korea or for interim coverage should there be delays in gaining approval from the U.K. or Denmark for using the Fylingdales or Thule BMEWS as components of a missile defense system. In addition to these sensor ships, the Test Bed is to contain up to 20 sea-based interceptors for boost and midcourse intercepts, which is discussed below.

The second novel element is a floating sea-based x-band radar (SBX) and supporting IDT to be deployed in the Pacific to refine the trajectories from satellite warning sensors and to study missiles during their boost and ascent phases to gain more information for discrimination. Such platforms could host much more capable radars, which could avoid the commitment of valuable Aegis ships to picket duty and provide more accurate detection and track.

These Test Bed elements were given a higher priority than research and development programs that could only deliver results for later blocks, and the National Teams MDA recruited to assist in planning its long-range activities were given the role of mapping out the systems and BMC3 programs needed to provide these capabilities at some early, but unspecified date. The MDA then reoriented its priorities to reflect those plans. The Test Bed was later largely incorporated into the interim deployment that is discussed below.

Command Structure

While these programmatic elements were being defined, work went on in parallel on the command structure needed to efficiently employ such defenses. While there was significant progress, there were several impediments. One was the inefficiency caused by the lack of a decision on who would ultimately command the defenses. The other was the difficulty of defining how to command defenses that had not yet been fully defined. These issues were impacted by several parallel developments. One was

the decision to disestablish USSPACECOM and redefine USSTRATCOM's missions, which impacted the choice of a CINC. The second was the realignment of the AORs of the CINCs to eliminate overlap and define clearly which commander was responsible for each area of the globe. The third replaced the term CINC with Combatant Commander (CC) to better reflect those realigned missions and AORs. A key principle in the realignment was that each CC had the responsibility to defend the forces within his AOR, which impacted the subsequent decisions on missile defense discussed below.

During 2002, USSPACECOM was disestablished and its missions distributed between the newly established USNORTHCOM and the redefined and expanded USSTRATCOM. In accord with the above principle, USNORTHCOM had the primary missions of defending the North American Continent from missile or aircraft attacks. However, the Unified Command Plan (UCP) assigned several new missions to USSTRATCOM, one of which was providing planning and support for global missile defense. The apparent overlap of assignments between the two Commanders (previously CINCs) seems in conflict with the principle that each CC is responsible for defending his AOR, which gives the Commander USNORTHCOM the responsibility of defending North America. However, Commander USNORTHCOM has the freedom to delegate the execution of some parts of his mission to USSTRATCOM.

The CCs are apparently moving toward an understanding in which USSTRATCOM will control the planning for missile attacks, manage the EW sensors through AFSPC, and use their observations to provide warning to CCs with geographic AORs, including USNORTHCOM. Commander USNORTHCOM will then use that information to assess attacks on the U.S. for the President and Secretary of Defense and to decide whether to provide release authority for the GBIs. The precise nature of the forces that will execute that release will not be clear until the documents that specify which CCs get the "forces for" specific missile defense missions are released, but the details of who "pulls the trigger" for them is less important than a clear demarcation of who is in overall control for each part of the mission, which the above partition should clarify for the main issues.

While the responsibilities of the CCs for missiles that remain in one AOR are clear, they are less clear for missiles that transit several AORs—including launches from rogue countries toward CONUS, which offer opportunities for multiple engagements by theater boost and midcourse defenses. Those issues will be further complicated when theater and global systems are added. Games and simulations have indicated that the clarification of those command relationships could significantly improve the allocation of limited defensive assets in the near term and that they will be essential when fully multi-tiered, integrated global defenses are deployed in the next decade.

The Initial Defensive Operational Capability

Under the pressure of external events, the spiral development program and its emphasis on Test Bed activities evolved further at the end of 2002 into a more concrete Ground-Based Midcourse Defense (GMD) System with firmer deployment timelines. The basic concept is a fast-paced GMD imbedded in a broader, global Ballistic Missile Defense System (BMDS), which is to evolve from it on a longer time frame. GMD in turn contains an Initial Defensive Operational Capability (IDO), which is to be in place by the end of 2005. On December 16, 2002, the IDO's charter was laid out in National Security Presidential Directive 23, which directed that "The Defense Department…shall proceed with plans to deploy a set of initial missile defense capabilities beginning in 2004."[121] The following day the Secretary of Defense further defined the scope of the IDO with the statement that "It will be an evolutionary program…improve as you go along.... It would be a very preliminary, modest capability."[122] Thus, the Administration certainly cannot be accused of raising excessive expectations for the IDO.

The IDO's key elements are familiar from the Test Bed.[123] The main difference is the specific date for their activation and the requirement that they be integrated as an operational system rather than as a test. The Test Bed elements described above are to be integrated with a few others by the end of 2004 to provide the U.S. with an initial defense from attacks from Northeast Asia and supplemented the following year to provide an initial defense against attacks from Southwest Asia and the Middle East by the end of 2005. These steps should provide the whole U.S. with roughly the level of protection the NMD C1 system was to have produced somewhat later.[124] While their superficial descriptions are similar, there are significant differences between NMD C1 and the IDO. NMD was basically a few GBIs and its fire control system. It paid excessive attention to the XBR, which could not be built on schedule, and too little attention to BMC2 and the integration of the sensors needed to make the system effective. If successful, IDO should provide an entry-level midcourse defense with a BMC2 that could incorporate new sensors and interceptors as they become available and a backbone for extending defenses to other systems and theaters.[125]

These integrations are to be performed while maintaining a robust testing and development program, bringing technologies to the levels of maturity required for the deployment of GMD, and preparing to field those GMD elements as part of an overall BMDS. This combination of operational and developmental activities involving technologies at several different levels of development presents challenges that have rarely been successfully met in previous programs. It calls for continuing the use in tests of a number of key components while they are being simultaneously integrated into an operational system. Once the IDO is activated, it can and will be used for tests, but in the interim there will be competitions for those assets needed to both complete key confidence demonstrations and activate the IDO.

---

121. G. Bush, National Security Presidential Directive 23, December 16, 2002.
122. D. Rumsfeld, press conference, Pentagon, December 17, 2002.
123. DOD Missile Defense Agency, "Missile Defense Operations Announcement," December 17, 2002.
124. J. D. Crouch (Assistant Secretary of Defense), "Missile Defense Deployment Briefing," *Space Daily*, December 17, 2002.
125. Lt. Gen. R. Kadish (Program Director of the Missile Defense Agency), "Missile Defense Deployment Briefing," *Space Daily*, December 17, 2002.

Elements for 2005

The main element the IDO adds for 2005 integration is the UEWR at Fylingdales, whose integration is contingent on approval by the U.K. government and additional funding. Fylingdales is needed to detect and determine the trajectories of launches from the Middle East toward the Eastern seaboard of the U.S., for which the radar at Thule is too far north for detection and too far west for useful warning. Without Fylingdales or an alternative, the 2005 increment could not defend against launches from the Middle East. The Aegis is under development for 2004 for use in North Asia, largely to refine trajectories toward Hawaii. However, other means appear adequate for that, so the Aegis could be repositioned in the Mediterranean to provide adequate tracks, given reliable warning and cueing from satellite sensors. The Aegis provides useful reliability in defending Hawaii, although trajectories from North Korea can be determined with adequate accuracy for GBI commit with DSP.

Assuming that early warning and coarse track are available, the other elements to be added in 2005 are 15 additional GBI (4 at VAFB and 11 at Ft. Greely) and an IDT on the East coast to guide them. With adequate warning, GBIs from VAFB and Ft. Greely can intercept missiles headed to the East Coast, so the additional GBIs need not be deployed there, although doing so would provide more margin and allow more efficient allocation. The BMC3 network described above should be adequate to support intercepts on both coasts, although details remain to be worked out between the CCs, as do the details of the C2, computer, and communication systems they will use.

Integration

Integration challenges in the IDO are significant because they involve the synchronization of a large number of elements that must be completed and tested in parallel. Development, construction, environmental impact, and government contributions are well underway, but there are several remaining large hardware and software builds that extend essentially up to the date of IDO's activation. That allows little time for checkout and integration, which are typically pacing items. There is little slack in the software builds, which frequently delay large programs of this type. On the surface that would appear to be a serious weakness, but a competent fire control system was produced for the GBI tests, which could be used as an interim C2 system for the initial activation, provided that it is acceptable to operate it with some level of pre-delegation, minimal BMC3, and flexible rules of engagement. Such a process would not be acceptable with offensive strategic systems, but should be acceptable for the release of GBIs that would cause little or no damage on the ground.

Tests

There are four major tests on the path to activation of the IDO: booster, flight, discrimination, and readiness tests. The failure of the test of the GBI booster was one reason for the negative decision on the NMD IRT. The new prototype booster has recently been successfully tested, and another test is planned for 2003. If successful, those tests should be an adequate basis for proceeding with the production of the boosters for at least the initial GBI for IDO. The booster is the one element that could impact the critical path; thus, it is reassuring that its testing will be completed early and that the production booster will be further evaluated in subsequent flight tests.

The failure of IFT-4 and IFT-5 for minor technical and quality control problems were essential elements of the negative IRT recommendation on the DRR decision on NMD deployment. Subsequently, IFT-6 and IFT-7 flew successfully, demonstrating that the solutions to those problems were successful. IFT-8 and IFT-9 were flown to complete the demonstration that hit to kill is reliable in practice. At the end of 2002, the flight test program had five successes for seven trials for a 71 percent success rate, which was somewhat better than the nuclear intercept tests of previous decades. NIKE was 57 percent successful in development tests and 10 of 14 or 71 percent in mock intercepts, which was regarded as more than adequate.

IFT-10 was to have demonstrated the reliability of GBI discrimination against basic decoys, but the EKV did not separate from its booster, although for different reasons than on IFT-5. That leaves a gap in the demonstration of discrimination that will not be filled until IFT-14 is flown at the end of 2003. However, as noted above, IFT-1A and IFT-2 already demonstrated much of the required discrimination against expected rogue decoys in related environments, and IFT-6 through IFT-9 performed as expected against a succession of surrogate objects. At present, GBI flight tests of discrimination are thus five for eight or 63 percent. After IFT-14, they will be either 67 or 56 percent successful, depending on whether its outcome is positive or negative. There is little doubt that the IDO will be activated either way, as either is close to the levels found acceptable to activate previous systems. This level of performance is far short of the unrealistic "99 percent probability of 100 percent coverage" requirement demanded of previous systems, but it certainly meets the common sense requirement of being much "better than nothing," as with a 70 percent single shot kill probability, two GBI fired simultaneously would give useful 90 percent probability of kill.

Initial tests of Test Bed hardware and software readiness will be made with IFT-13A, IFT-13B, and IFT-14, supported by radar, communication, and C2 nodes. The demonstration of IDO readiness in 2004 will involve IFT-15 and IFT-16, a full set of sensor and C2 nodes, and five GBIs in Ft. Greely, which should provide an initial defense against attacks from Northwest Asia by the end of 2004. At that point, the IDO should provide the functional equivalent of the C1 defense against attacks from that area envisioned by the Clinton Administration NMD program for a somewhat later time. The modification and integration of the Fylingdales UEWR and the execution of IFT-17 and IFT-18 in 2005 should serve as the basis for the activation of the IDO against attacks from Southwest Asia and the Middle East by the end of 2005.

Capability

Those tests and the activation of IDO should provide roughly the functional equivalent of the C2 defense against attacks envisioned by the Clinton Administration NMD program for somewhat later than 2010. The IDO differs in that it will actually provide the sensors and BMC3 needed to make the GBIs effective and their allocation efficient. The IDO is intended to provide modest midcourse protection against attacks by a small number of missiles with rudimentary decoys from a limited number of sites. It will largely make use of the satellites and radars developed for previous systems. Thus, it will be able to draw on their extensive development, but it will also inherit their known uncertainties and vulnerabilities. One of IDO's primary contributions will be to flesh out the BMC3 elements that were recognized as critical elements of, but ignored by, programs subsequent to GPALS. As such, it should restore the midcourse core of GPALS, stripped of its terminal and space elements and globally integrated BMC3.

A question often asked about IDO is what countermeasures it can address. The Administration's answer is that it "will be capable against the countermeasures we expect in the time frame that we're talking about,"[126] which is reassuring, but logically equivalent to the ability asserted by the intelligence community in the previous decade regarding its definition of NMD threats. The IRT for the Clinton Administration's NMD showed that such systems could not be extended into robust missile defense simply by just adding more GBIs. A logical question is whether or not the IDO could serve as the nucleus for a system that could address more robust threats, which is addressed in the next section, along with the extensions required to provide defenses that cover the whole globe and U.S. allies.

126. *Ibid.*

Follow-On Phases

The IDO has the advantage of building on developed technologies and components and their successful demonstration, but also has the disadvantage of inheriting their known susceptibilities to countermeasures. Previous analyses have indicated that it is not possible to provide an adequate global defense by adding GBIs to such midcourse deployments. Some problems will be reduced by time. The EMD and deployment programs for the PAC-3 and THAAD should be able to support deployment as the theater elements of the 2006 or 2008 BMDS blocks as well as providing the number of terminal interceptors required to provide localized protection for high value targets. GBIs can provide additional midcourse intercepts, but they cannot afford to intercept all objects credible to their sensors. To reduce their numbers to affordable levels, it will be necessary to reduce the number of objects they must intercept. That will require either greatly improved discrimination, so GBIs can efficiently intercept the weapons revealed, or effective boost-phase defenses, so fewer missiles survive boost to deploy decoys and weapons that otherwise would have to be discriminated and intercepted by downstream layers.

Both approaches look promising. With the passage of time, maturing technologies should make it possible to observe missiles and weapons from birth to death, which makes it difficult for attackers to achieve surprise in the decoys they use or secrecy in their method of deployment. That should also make it possible to develop and deploy advanced sensors and communication systems that could support more and better observations, improved discrimination, and fusion of observations throughout boost and midcourse. It is not possible to discuss them further here, but it appears that time and technology are arguably on the side of the defense.

Boost-phase intercept concepts are also maturing. Two options that have undergone significant development are intercept from the surface and from space. Each involves significant advantages and challenges. Intercepting from the surface allows the interceptors to be based close to the threat, which minimizes the number of platforms required, but places significant technological challenges on interceptor performance. Intercepting from space distributes the interceptors over much of the Earth's surface, which dilutes their local concentration, but reduces their mechanical design challenges. The global distribution of interceptors is a penalty in defense of a single site, but is an advantage when global defenses are considered. Thus, these technical performance issues are given differing weights depending on the number and size of the areas, deployed forces, and allies to be protected.

Boost-phase defenses serve two essential functions. The first is that they are insensitive to the decoys and countermeasures that are stressing to midcourse systems, because there are few credible boost-phase decoys or countermeasures. Thus, they are likely to intercept the missiles rather than the decoys in boost. That reduces the number of decoys that the downstream layers must face both by killing missiles before they can deploy decoys and by forcing the missiles that do survive boost to deploy their decoys rapidly and poorly. Their second function is to form an independent layer, whose attrition compounds that from the midcourse and terminal layers to produce overall defenses of high effectiveness, which complicate an attacker's attempts to degrade any given layer. For these reasons boost-phase defenses contribute a uniquely effective layer; however, they must operate quickly, which requires fast release, responsiveness, and nearby basing.

Surface-based systems can produce the velocities, accelerations, and responsiveness needed for effectiveness against threats whose geometry provides interceptors prompt access to boost and their

launchers safe basing. Sea-based systems are advocated for small launch areas that are surrounded by water, where ships can approach closely without exposing themselves to unacceptable risk. It is argued that near-term systems could be based on the integration and deployment of current technology, although the quantitative arguments supporting that assessment have been questioned. While airborne sensors and interceptors have limited defensive radii and endurance, both could be expand through the use of lighter sensors and unmanned aerial vehicles (UAVs). The ABL competes favorably on the basis of its fast response, range, and precision metrics. Its precision tracker alone could be valuable in determining the trajectories of rogue ICBMs from countries of moderate size, which it could cover without need for overflight. It will be introduced in later spirals of development, depending on the laser and propagation milestones it must meet. The SBL was scheduled for a space demonstration in 2012, but is paced by technology developments that are difficult to accelerate, which are compounded by funding. While a seemingly ideal system for global coverage of limited missile threats, SBL has been returned to R&D because of the long development times involved. Thus, the discussion below concentrate on generic surface-based systems of the type that might emerge in later blocks.

SBI and BP underwent significant development in SDI and GPALS, so their earlier levels of technology could be recovered and updated rapidly. Their sensors have been kept current through Navy interceptor programs, but their engine technologies have undergone significant improvements, which have not been incorporated. BP should probably be updated to the simpler but faster operation appropriate for rogue missile threats. The BP's ability to detect and track missiles against their bright plumes, which is the most difficult step in boost-phase intercept, was demonstrated in the SDI Delta 180 tests. Thus, a useful BP system could probably be developed on about the same time scale as the second MDA spiral by leveraging off this technology and handover database. Development and testing could apparently be completed in about three years, with deployment of the constellation of about 100 BPs needed for rogue missiles taking two to three years more. However, given the current low level of funding of BP and the lack of focus on space-based versus surface-based boost-phase interceptors—and of boost-phase versus midcourse intercepts—it is unclear that either will be available in early blocks. Boost-phase intercept is clearly essential for defenses that produce high overall attrition and complicate the deployment of countermeasures to midcourse and terminal layers. Thus, if the current program does not develop them, history suggests that it is likely to be succeeded by one that does.

Surface-Based Boost-Phase Concepts

Surface-based boost-phase systems are possible for launches from areas such as North Korea, Iraq, and Libya, where trajectories toward the United States allow ready access to missiles in boost from either ships in international waters[127] or land bases controlled by allies.[128] An interim capability to defend against launches from such bases could be developed with existing technology by moving ships forward, although in general that would place them in harm's way. Improved sensors and interceptors with higher acceleration and velocity could support defenses against more capable missiles and deeper basing from safer locations, although such capabilities will only be available later. Several studies of boost-phase defenses of such areas have led to similar conclusions.[129]

The analysis of boost-phase intercepts by surface-based systems is somewhat different than that of ground- or sea-based terminal, midcourse, and SBI systems, so it is briefly summarized below and in Appendix L. The kinematics of intercepts of rogue ICBMs can be approximated simply. In the near

127. G. Canavan, "Boost Phase Defenses for Regional Missile Defenses," Heritage Foundation presentation to Japanese Ministry of Defense, July 22, 2002; Los Alamos Report LA-UR-02-4497.
128. G. Canavan, "Missile Defense Capabilities for the Defense of Europe," ed. A. Zichichi, International Seminar on Nuclear War and Planetary Emergencies; 27th Session (London: World Scientific, 2002; Los Alamos Report LA-UR-02-5303.
129. G. Canavan, "Space-Based Missile Defense and Stability," invited presentation, *Proceedings of the American Philosophical Society*, Vol. 145, No. 3, September 2001, at *www.aps-pub.com/star_wars/canavan.pdf*.

term, they are likely to use conventional chemical rockets, which require a time T of about 200 s to reach intercontinental speeds V of 7.2 km/s with average accelerations under 4 g. An interceptor with an acceleration A of 6 g reaches a top speed of V of 6 km/s in about 6 km/s/0.06 km/s^2 = 100 s. If it is launched to maximum range after a delay T_{delay} of about 50 s after the missile, the interceptor could reach the missile's burnout point from anywhere within a circle of ground radius $(V/\sqrt{2})(T - T_{delay} - V/2A)$ = 6 km/s(200 s – 50 s – 50 s) /$\sqrt{2}$ = 430 km. If the interceptor approaches the missile head-on, the missile's flyout distance of about 400 km would allow the interceptor to meet it from a total range of about 830 km along the missile's track. That is much greater than the 530 km from North Korean launch areas to the Russian port of Vladivostok, which has stimulated suggestions that ground-based interceptors be sited there in a cooperative program.[130]

Interceptors from naval platforms in favorable locations could reach missiles by burnout that are launched over the ocean. However, geometries involving tail chases, i.e., where the missiles are launched away from the interceptor platform, the missile flyout and interceptor ranges subtract, so the standoff distance would be reduced to about zero, which means the interceptor could not close during boost with such delays. More detailed calculations of missile and interceptor trajectories do not fundamentally alter these scalings. Rather than depending on sensor and interceptor performance alone, arguments about the utility of surface-based boost-phase intercepts depend on assumptions about higher accelerations and velocities by interceptor advocates and on longer operational delays by critics. These arguments have not been resolved, but have narrowed the range of parameters to which they are sensitive.

Airborne interceptors have similar scaling, with two important modifications. The first is that they can reduce the range to the intercept point by flying over the opponent's territory, but doing so prior to hostilities is an act of war that makes the aircraft and interceptor vulnerable to suppression attacks prior to launch. The second is that the interceptors an aircraft can carry are generally lighter than those on ships or ground vehicles. Their higher release altitude does not compensate for their reduced fuel load, which restricts their operating and defended radii. Manned aircraft cannot cover even modest territories without overflight. Thus, airborne interceptors, like surface-based interceptors, provide options for the near-term protection of a limited number of threat locations, but do not provide a clear path to protection against the more stressing threats expected in the future.

Achieving the fast response required for effectiveness requires either good on-board sensors and BMC3 or good communication to an external system that is connected to such sensors. The sensors on board current naval assets are marginal for detection and track at useful ranges, as is their connectivity to sensors that could provide it. On-board sensors and connectivity can be fixed, but doing so for an adequate number of ships could take much of a decade. For launch areas further inland, the benefits of surface-based systems are reduced. Interceptors for launches in central Iran or Iraq would face difficult tail chases, basing in insecure territory, or long overflights. For such threats surface-based systems require interceptors with faster response and higher accelerations, which are susceptible to false alarms and spoofing. Their numbers increase in proportion to the number of large or widely separated launch areas. They become intractable when applied to the nonnuclear submarines planned for use as missile launchers in the next decade. For large or many threat regions, space-based interceptors are preferred for both coverage and cost.

The quantitative requirements for surface-based boost-phase intercept from a launcher in the most favorable position under a missile's trajectory are estimated in Appendix L. Its main results for theater and regional missiles are summarized in Figure L.4, which shows the interceptor velocity v required as a function of launch delay time T_{delay} for interceptor acceleration a = 4, 5, and 6 g. For T_{delay} = 0 there is no restriction on v, but by a delay of 60 s, a 4 g interceptor would need a maximum velocity of

130. R. Garwin, "Space Weapons or Space Arms Control?" invited presentation, *Proceedings of the American Philosophical Society*, Vol. 145, No. 3, September 2001, at *www.aps-pub.com/star_wars/garwin.pdf*.

about 12 km/s, which is incompatible with mobile basing. A 5 g interceptor would need about 8 km/s, which would require long development, and a 6 g interceptor would need about 6 km/s. The gains for higher accelerations diminish rapidly; and would be difficult to realize in any case. The interceptor acceleration increases with missile acceleration A, as $A/(1 - A/\sqrt{2a})$, so an interceptor with acceleration under 2.9 g could not intercept a 4 g missile with any velocity.

For an interceptor speed v = 6 km/s, a 4 g interceptor could only intercept in boost with a maximum delay of about 30 s; a 5 g interceptor with 45 s; and a 6 g interceptor with 55 s. At these delays, the interceptor would have essentially zero cross range; it would have to be directly under the missile's burnout point and fire straight up. For nominal missile ranges and interceptor parameters, delays over 60 s are unacceptable. A rough summary is that an interceptor speed of 6 km/s, acceleration of 6 g, and less than 60 s release delay are needed for boost-phase intercept of theater and regional missiles from the surface. That combination does not appear consistent with the goals of systems in development.

The requirements are less stressing for rogue ICBMs because they burn out at a shallower flight path angle, which provides a longer time for interceptors to reach to them. Figure L.5 shows the sensitivity of interceptor cross ranges to delay time for 6 km/s interceptors with accelerations of 6, 10, and 14 g. For 6 g and a 30 s delay, the cross range is about 550 km, as above, but by 60 s it falls to 350 km, and by 90 s to 0. Increasing interceptor acceleration to 10 g increases the range at 30 s to 670 km, but it still falls to 480 at 60 s and 250 km at 90 s. Increasing the acceleration to 14 g increases cross ranges only marginally. The total range for a 6 g interceptor in line with the missile's trajectory falls from roughly 1,300 km with 0 delay to 950 km at 60 s and 600 km at 90 s, which is the approximate missile flyout distance. At 60 s, increasing the interceptor acceleration would increase the total range by about 30 percent. Increasing acceleration further would only increase it slightly. Conversely, for 6 km/s and 6 g, increasing the delay from 0 to 60 s decreases the interceptor cross range by 50 percent, which is almost 1 percent/s. For longer delays, the decrease is over twice as fast.

Thus, surface- and air-based systems can have distances between the missile and interceptor launchers on the order of 1,000 km, if the interceptor launcher lies along the track of the missile and has the appropriate combination of parameters to intercept. If the missile trajectory is to the side of the location of the interceptor launcher, cross ranges are typically 300–500 km. If the launcher is further away than that, it has no opportunity for engagement in boost, as is the case for launches that are directed away from the interceptor or over hostile territory. If the missile burns out before reaching the interceptor's range, it is still possible to engage it in the bus phase, but such engagements are of lesser leverage because they only kill a fraction of the missile's weapons and decoys and must contend with the same range of missile decoys and countermeasures that midcourse systems would face.

Space-Based Boost-Phase Concepts

Space-based interceptors can survivably overfly threat launch areas and engage missiles in the boost phase; thus, they are not subject to the azimuth and range limitations that restrict surface-based interceptors. The SBIs are already in space; the missiles have to climb a large potential well to reach their altitude, rather than the interceptor having to climb one to reach the missile, as surface-based systems do. Moreover, as the SBIs are in the vacuum of space rather than in the atmosphere, they can orient their thrust in the optimal direction for intercept rather than facing the dynamic pressure and erosion considerations that limit ground-based systems or taking the drag losses associated with accelerating through the dense atmosphere.

GPALS tests demonstrated that on-board sensors and processing could support the response times required for boost-phase intercepts from space, and the analyses above indicate that their economics should be favorable in that role. Their main disadvantage is that at any given time, most of the SBIs are somewhere else in their orbit. Thus, for launches from small areas such as rogues, SBIs are penal-

ized by their absenteeism, which is reflected in reduced effective performance and increased cost. However, if the launch area is large enough, or if there are enough geographically dispersed threats to require global coverage, absenteeism turns to the BPs' advantage, because they then provide global defenses at no additional cost. That suggests a progression from surface- to space-based defenses as the number and size of launch areas grows and the number and speed of missiles increases.

Performance

Early BP performance was limited by the small engines available to them, which had limited accelerations, velocities, and payload fractions. Early BPs had average accelerations of about 6 g; current liquid fuel engines could support about 12 g. Pump-fed engines have now been scaled to the few kilogram payloads needed for BPs while maintaining high efficiency and payload fractions. Laboratory and limited flight tests have demonstrated performance that would be useful against rogue ICBMs. Current high-acceleration solid engines developed for theater applications could increase acceleration to 25–50 g with comparable efficiency for shorter-range theater and regional missiles at a two- to fourfold weight penalty. However, such accelerations are not needed for liquid-fueled rogue ICBMs, which only have about 4 g accelerations and burn times of 200 s.

For a release delay T_{delay} = 30 s, the range of a 6 km/s BP against a 240 s rogue ICBM is 960 to 1,200 km, varying about 25 percent with interceptor acceleration. Range variations are similar at higher velocities. For 8 km/s BPs, the ranges increase to 1,200 to 1,500 km. Constellation sizes can be estimated by setting the launch area to zero in the above analyses. For launch of a single missile, the constellation size needed to cover the 40° latitude rogue or "SCUD belt" is the same as the absentee ratio of $2(R_e/r)^2$, which is a function of SBI acceleration, velocity, and delay through r. The $1/r^2$ scaling of point launches penalizes poor interceptor acceleration, velocity, or responsiveness. This scaling differs from that for distributed launches, in which the addition of the radius of the launch area to that of the SBI reduces the sensitivity of constellations to SBI speed and flight time. While constellations are large for small interceptor speeds, by a V of 6 km/s they fall to 90, 70, and 60 BPs for 6, 12, and 24 g accelerations. The 25 percent range variation noted above maps into a 50 percent variation in constellation size.

Because of the short duration of the boost phase, it is generally not possible to use efficient shoot-look-shoot approaches. If instead two SBIs are released simultaneously to reduce leakage through the boost phase to a few percent, that doubles the constellation size. If 5 missiles were launched in a time short compared to the replenishment time r/V = 1,000 km/7.5 km/s = 130 s for an additional SBI to rotate into place over the launch area, it would be necessary to increase the constellation density by a like amount. Taken together, these two factors would increase the constellations about 10-fold to 900, 700, and 600 BPs, respectively. However, if there is a single launch area or the rogue countries are at roughly the same latitude, inclining the SBI constellation over that latitude can decrease the size of the constellations needed by a factor of 4 to 10. Doing so can also increase the density of BP over that latitude enough so that the SBIs arrive more rapidly than missiles can be launched. If so, the constellations are reduced by factors of about 4 by concentration and 5 by replenishing the SBIs faster than missile launch for a total reduction of 4 x 5 = 20 to about 45, 35, and 30 SBIs.[131]

Even so, for rogue ICBMs, SBI absentee ratios are larger than the 5:1 to 10:1 predicted for SDI and GPALS because of the difference in size of the launch areas from which missiles were launched. Soviet deployments of newer missiles attempted to stress BP defenses by reducing the boost time, area, and warheads per missile, particularly by concentrating the basing of short burn single-weapon SS-25s. Those changes only reduced BP effectiveness a limited amount, because the bulk of the Soviet RVs remained in heavy SS-18 missiles, which were slower and widely distributed. However, rogue deployments achieve the characteristics needed to stress SBI constellations automatically, as

131. G. Canavan, "Concentration of Space Based Interceptor Constellations," Los Alamos Report LA-UR-02-5739, July 2002.

rogue launch areas in small countries are necessarily reduced effectively to a point, have access to missiles with modest burn times, and use single-weapon missiles, for which the bus time is zero. Together, these factors produce the 10-fold higher absentee ratios for rogue launches discussed above.

It would be desirable to cover both rogue threats and accidental or unauthorized launches from Russian and Chinese ICBMs and SLBMs remaining from the Cold War with the same constellation, and it is generally possible to do so. The requirements for the latter are estimated above in the section on GPALs. For ICBMs and SLBMs, the dominant concern is the launch of all 20 missiles on an SSBN, which according to Figure G.6 would require a constellation of about 2,000 BPs for negation in boost, 250 for negation in boost and bus, and 150 for intercept in the boost, bus, and midcourse phases. Thus, apart from issues in the inclination for SBI constellations to cover this range of threats, rogue, ICBM, and SLBM threats could be treated by a common constellation when intercepts in boost, bus, and midcourse are included.

Parametric Cost

Parametric cost estimates can be based on those for SBI components discussed earlier. For a cost of $500,000 for a 4 kg SBI KV, $20,000/kg for launch, and ideal engines with specific impulse I_{sp} of 300 s, the on-orbit KV cost would be about $1 million for a 6 km/s SBI and $2 million for an 8 km/s SBI. Figure M.1 shows constellation costs, as estimated by the product of constellation size and SBI costs, for 6, 12, and 24 g constellations. The top curve for 6 g has a minimum at about 6.5 km/s. That gives roughly $0.1 billion for single coverage of each missile launched simultaneously, which is the product of a 90 SBI constellation and a KV cost of about $1 million. The cost for coverage with a 12 g SBI has a minimum of roughly $0.07 million at about 7 km/s. That for a 24 g SBI is roughly $0.06 billion at a speed of 8 km/s. Higher accelerations decrease sensitivity to off-optimal velocities and reduce total costs by about a factor of 2, although the reduction saturates at about 12 g and costs could be higher for real high acceleration engines. As noted above, these space hardware procurement costs vary with KV costs; thus, they will remain uncertain to within factors of 2 to 4 until the appropriate detailed design and testing are completed.

Cost effectiveness at the margin is not a requirement for defenses against rogue threats as unlimited escalation is not an option for them, as it was thought to be for the Soviet Union. The main requirement is that SBI costs be competitive with those of other elements of the defense. If surface-based boost-phase systems are not compromised by azimuths or delays, they should be able to negate a $1 billion missile for the roughly $100 million cost of an interceptor, which would give it a favorable cost-exchange ratio of about 10:1. If the space-based system could negate the missile for 1 SBI at a cost of $1 million, that would give a cost exchange ratio of 1,000:1. The exchange ratio of missile and BP masses in space is about 1,000 kg/4 kg = 250:1, which is also favorable. However, BP absentee ratios are about 100, which would decrease their overall exchange ratios to about 10:1, which is the basis for the observation that if ground-based boost-phase system can engage a missile, it could be favored on a cost basis. If not, or if there are several launch areas, cost considerations favor the BP.

Using BP to Kill Decoys

While the discussion above concentrates on RV kills, rogue ICBMs would have little effectiveness if they did not employ decoys or countermeasures. Without them their warheads would be readily detected and intercepted by midcourse and terminal defenses. Since decoys are an essential element of rogue effectiveness, eliminating them is essential in reducing their overall effectiveness. The impact of SBI on decoys depends on whether or not the rogue uses separating buses. If not, the decoys are deployed in a single cluster with the RV somewhere inside. Even if the decoys are deployed quickly, SBI has some leverage in that it could detect the rocket thrusts needed to hide the RV in the cluster, take the first shot, and transmit the result to support efficient allocation of downstream defenses. Put-

ting pressure on the bus to force rapid deployment would also have a positive effect. SBI could also be used in midcourse, although effectiveness would depend on their discrimination ability.

If the missile uses a bus to deploy a weapon and clusters of objects along several separate trajectories, each trajectory will contain decoys, but only one will contain the RV. It typically takes buses 10s of seconds to minutes to transfer between 5–10 trajectories. Thus, they remain targets, although of decreasing value, for SBI during the bus phase. If the bus is designed to release clusters on 10 different trajectories over a period of 300 s and the RV was to be released on the fifth trajectory, intercepts at up to 150 s into busing would kill the RV. Thus, the SBI intercept would be fully effective for the 200 s of boost phase plus 150 s of bus phase, which would increase the time during which the weapon could be killed in ascent by 75 percent. The extent of the improvement depends on when the RV is to be deployed. If it was in the first trajectory, it would have to be killed within the first 30 s of busing. If it was in the second, within the first 60 s, and so on. Game theory determines the trajectory the attacker should put the RV in and when the defender should try to intercept. The result is related to the partial busload calculations described above, but is not essential here.

Even if the RV was deployed in an early trajectory, the BP intercepting the bus before it completed deployment would mean the BP would destroy the clusters remaining on the bus. If it did not, downstream defenses would have to deal with them—probably by committing GBIs to them. SBI's role in destroying decoys in the bus phase thus corresponds roughly to that in killing RVs in partially filled buses of MIRVed missiles. Figure M.2 shows the number of RV and cluster kills possible as a function of the total time the bus takes to deploy 10 clusters of decoys and one weapon. For a 15 BP constellation, the number of clusters killed in boost is about 3. The number of clusters killed at the end of boost at 300 s is about 3. The number killed in the bus phase increases quadratically with time during deployment to 4 at 600 s, so the total number of clusters killed in the boost and bus phases is about 7. If rogue ICBMs depend on decoys for effectiveness and use separating buses for their efficient deployment, that creates an important defensive role for post-boost intercepts that SBI can execute effectively.

Theater and Regional Missile Defense from Space

The U.S. must protect its deployed forces, allies, and friends from a growing array of theater and regional missiles. Such missiles present additional challenges due to their short flight and boost times. Figure M.4 shows the ranges of SBI with 6 km/s, accelerations of 6, 12, and 24 g and 30 s release delays as functions of missile range. For 4,500 km regional missile ranges, SBI ranges vary from 525 to 750 km. For 1,000 km theater missiles they fall to 70–220 km, and for 500 km SCUDs to 20–80 km.

SBI constellations scale as $1/r^2$. For 5,000 km range regional missiles, constellation sizes are 100–250, in accord with those estimated above for ICBMs. For 2,500 km intra-theater ranges, they increase to 350 to 1,100. For 1,000 km ranges, they increase to 1,000 to 10,000, which are larger than those for GPALS. For theater range missiles with ranges less than 1,000 km, even SBIs with velocities of 8 km/s and accelerations of 24 g would not be attractive for boost-phase intercepts. Constellation costs also vary with missile range. The 6 g constellation cost passes $1 billion for single coverage at missile ranges of 2,700 km and $26 billion by 1,000 km. The 24 g constellation crosses $1 billion at about 1,500 km and $2 billion at 1,000 km. Higher velocities do not significantly reduce these costs because for short ranges the missile burn times are so short that even high acceleration SBIs spend little time at maximum speed.

However, theater missiles spend a small fraction of their flight above the atmosphere. A 1,000 km missile spends about 70 percent of its time above 125 km, but a 700 km missile spends only 60 percent, and a 500 km missile never goes above 125 km. Short-range theater missiles cannot deploy or use decoys and countermeasures effectively, so midcourse and terminal theater systems should not

face large numbers of decoys. Regional missiles with ranges of 1,000 to 5,000 km can take advantage of the countermeasures ICBMs use, which makes boost-phase intercept attractive for them. However, a 4 g 3,000 km missile only accelerates for about 140 s, so BP would need fast response. Theater missile defense programs have developed and tested solid axial-thrust motors that provide accelerations of 25–50 g at specific impulses of 260–280 s, which would provide adequate interim regional missile capability, although current versions are about fourfold heavier than the ideal scaling assumed above. Single coverage of 3,000 km regional missiles with 6 g BPs is optimized by $V = 5$ km/s, which gives constellation costs of about $0.7 billion per missile simultaneously launched. At 12–24 g, SBI would optimize at 7–8 km/s and cost $0.34–0.24 billion. For such ranges, accelerations greater than 6 g would reduce costs by about a factor of two. Higher speeds would decrease constellation size but increase costs.

Since theater and regional missiles have difficulty making use of decoys and countermeasures, SBI should be able to engage effectively in midcourse. Figure M.5 shows the boost and midcourse kills out of 5 simultaneously launched missiles as functions of range for constellations of 500 and 1,000 SBIs with 6 km/s, 25 g, and 30 s delay. At 4,500 km range, the expected boost-phase kills are 1.6 and 3.1 missiles, respectively, and the number of midcourse weapon kills are their complements, 3.4 and 1.9. At 2,000 km the boost-phase kills drop to 0.5 and 1 missile. At 1,000 km they drop to 0.1 and 0.3. Even the midcourse engagements are not sufficient to kill all of them. At 500 km range, there is about 20 percent leakage through both layers. For ranges under 1,000 km, the constellations required for theater missiles are as large as those for rogue ICBMs, even with higher SBI accelerations and short delay times. DSP and SBIRS should provide adequate detection, downlink, and processing rates for regional missiles. For theater missiles, the numerous, low-altitude SBIs themselves could carry the detection and track sensors needed to detect and pursue their targets without external sensors or C2 beyond weapons release. They could carry advanced sensors that see to the ground, which would provide near-instantaneous detection.

SBIs are almost as effective against regional missiles as they are against ICBMs. For intra-theater missiles SBI ranges shorten, constellations grow, and costs increase accordingly. Increased acceleration and reduced delay times can partially offset those trends, but for theater missiles with ranges under 1,000 km, their costs and constellations become unattractive. Fortunately, such missiles spend little time above the Earth's atmosphere, so they can be engaged effectively by SBIs in midcourse or by ground-based systems in theaters where they are available. The effectiveness of SBI against missiles of progressively shorter ranges should improve with improvements in their technology for other missions. It does depend on improvements in engines, release times, detection, and BMC2 that go beyond those needed elsewhere.

Directed Energy

For rogue missiles, the essential scaling for laser constellations remains $M/T = (N/R_e^2)B/J$, although for point launches it is useful to rewrite it as $BN = R_e^2JM/T$, since only the product of the number of satellites and their brightness is important, not their individual values. Rogue missiles are given credit for burn times T of about 250 s, hardness levels J of 10^8 J/m^2, and the simultaneous launch of $M = 5$ missiles. That could require a total DEW brightness $BN = MJR_e^2/T = 5 \times 10^8$ J/m^2 x (6,400 km)^2/250 s = 8×10^{19} W/sr, which could be achieved by 10 lasers of 3 MW and with 5 m mirrors. This brightness would decrease the cost to defend against rogue missiles by about a factor of 6 from that for Soviet launches. However, the initial investment would remain significant because their full constellation must be placed in orbit to address a few missiles. DEW is attractive for theater and regional threats because of its speed-of-light intercept, but the key parameter is its time to penetrate, which depends on the hardness of the missile. Theater and regional missiles can increase theirs significantly by using the shielding and spinning techniques strategic missiles use with much the same technology,

which is apparently available to them. Hardening missiles further would increase *BN* proportionally. The impact of spin depends on its rate relative to kill times.

Reducing burn times increase *BN* as $1/T$. That increase is less than the $1/T^2$ of KEW, which is the reason for their large constellation costs for short-range missiles in the previous section. The 60 s burn time of current SCUDs would increase *BN* by a factor of 4, which could be accommodated by increasing laser power by a factor of 2 and mirror diameter by 40 percent to roughly 6 MW lasers and 7 m mirrors. It would increase the size and cost of SBI constellations by a factor of 16. Survivability is difficult to achieve with large space lasers and optics, but it might not be required for rogue ICBM and theater missiles.[132] For these and related reasons, DEW concepts have been retained in R&D although their immediate uses are unclear.

Stability

Ideally, systems should be chosen on basis of which defends best for least. It is unlikely that strategic choices will be made that rationally in the near term, given the residue of Cold War logic. Stability is an important military consideration that was quite distinct from the ABM Treaty, but the connection between the two became confused, perhaps intentionally. Some have viewed the Treaty's constraint of zero defenses as a convenient way to prevent any competition in the defensive dimension, rather than as a potential way to reduce the intensity of competition and risk of conflict in both. Stability need not be an impediment to safe deployment of global or space-based systems.

John von Neumann gave a rigorous foundation for games in which one side's gain is the other's loss.[133] John Nash extended it to the non–zero-sum games that are more relevant to strategic stability and discussed the equilibrium solutions that are each side's best response to the other's rational counters.[134] Thomas Schelling popularized game theory as a tool for the analysis of the stability of strategic conflicts.[135] Others have applied qualitative versions of it to historical crises.[136] Due to the difficulty of the mathematics and prevalence of the massive retaliation strategy at the time of its development, the originators of game theory primarily applied it to unconstrained wars in which populations were the ultimate targets of both sides. That was inconsistent with traditional military thought and morality; moreover, it led in extreme cases to predictions that no crisis could ever lead to war, which caused those who had recently fought several to conclude that the theory was flawed or inapplicable. Those impressions were amplified by approximations to the analyses that made even modest defenses appear destabilizing.[137]

Stability is a legitimate concern for defensive systems. Imperfect defenses can be useful, but not if they invite escalation or attack. Predicting when they do not is the essence of game theory applied to crisis and arms control stability. Appendix K reviews the essentials of game theory and Nash optimal solutions. Figure K.1 shows the elements of game theory for crisis stability. It defines the graph of play, gives the decision nodes, determines which side decides at each node, and specifies the payoffs for each path.[138] The nodes represent decisions whether to strike; thus, the two sides' first and second strike costs are the appropriate payoffs. The two sides are identified only as U and P as identification

132. Canavan *et al*., "Debate on APS Directed-Energy Study."
133. J. von Neumann and O. Morgenstern, *The Theory of Games and Economic Behavior* (Princeton, N.J.: Princeton University Press, 1944).
134. J. Nash, Jr., "The Bargaining Problem," *Econometrica*, Vol. 18 (1950), pp. 155–162.
135. T. Schelling, *The Strategy of Conflict* (Cambridge, Mass., Harvard University Press, 1960).
136. G. Allison and P. Zelnikow, *Essence of Decision: Explaining the Cuban Missile Crisis* (New York: Addison Wesley Longman, 1999).
137. G. Canavan, "Transition from Adversary to Cooperative Interaction," *Proc. International Seminars on Global Emergencies*, ed. A. Zichichi (London: World Scientific, 2001
138. R. Powell, *Nuclear Deterrence Theory* (Cambridge, U.K.: University Press, 1990).

with specific countries would presume more knowledge about their objectives than is generally available.

Figure K.2 shows the cost to side U in a bilateral interaction with side P at START I level offensive forces as a function of U defenses and the probability *u* that U could strike first in a crisis.[139] The figure indicates that modest numbers of interceptors would not change strike incentives. Margins for error erode, but the optimal decisions for both sides do not change below 600 interceptors. Larger numbers produce large U costs at small *u* where P has an incentive to preempt. At large *u*, U could strike first and use its defenses to negate the other's second strike, as summarized by President Mikhail Gorbachev's statement that "the United States cannot develop defenses that could negate our first strike, but could develop defenses that would mop up our ragged retaliation." At large defenses, U's costs fall below those of inaction for all *u*, which represents defensive dominance; however, for undiminished strike objectives, this reduction would entail reciprocal strikes causing large costs to P, and possibly U.

It is possible to exchange offenses for defenses while improving stability. Figure K.3 shows the impact of U reducing offensive forces while deploying roughly equal numbers of defenses.[140] The top and bottom curves are U and P's first strike costs if U unilaterally reduces its offensives without deploying defenses. By about 100 U offensive weapons *W*, the discrepancy would be a factor of 5, which could stimulate a strike by P. The two central curves are their first strike costs if U increases its defenses as it reduces its offenses, which are equal at $W = 100$ and within a few percent of each other at all stages of the reduction. Thus, it is possible to trade offenses for defenses without impacting stability. In this example, 1,600 defensive interceptors are traded for 1,900 offensive weapons at an exchange ratio of 1.2:1.

The transition from adversarial to cooperative interaction, as represented by the progressive reduction of damage preferences, improves stability monotonically, and the proper inclusion of high value targets reduces strike incentives in bilateral and trilateral interactions, with or without defenses. It also stabilizes trilateral configurations by permitting smaller sides to hold significant value at risk with modest forces. A trilateral analysis raises the concern that China could perceive a U.S. defense against Russia as applying to China, which could provoke fear of attack, pressure to preempt, or offensive action. However, it appears possible to trade the offensive missiles allocated to China for defenses with no impact on stability much as they could in the bilateral interaction with Russia. It is possible to perform such trades between countries with greatly unequal force levels. Those interactions are naturally imbedded into a multilateral defensive framework, although these problems could be avoided altogether if it was possible to return to a policy of developing and deploying global defenses jointly with like-minded nations.

It would be useful to develop and deploy defenses with allies and others as a confidence building measure and as a way of demonstrating commitment. Some defensive technologies could be developed with other nations. That could be useful, given the importance of communicating capabilities and objectives in increasing stability. Different defenses have varying impacts on offenses, proliferation, and international cooperation. Terminal systems have little impact on strategic systems because they can be saturated or penetrated; however, they have modest potential for cooperation on technology other than military sales. Midcourse systems have little impact on strategic systems because of the vulnerability of their sensors and launchers, so their impact on stability is more apparent than real. They have some potential for cooperation with regional allies, but are too geographically constrained for significant cooperation with others. Surface-based boost-phase systems have little survivability

139. G. Canavan, "Analysis of Decisions in Bi- and Tri-Lateral Engagements," U.S. State Department Stability Workshop (Institute for Defense Analysis, November 2000); Los Alamos National Laboratory Report LA-UR-00-5737, November 2000.
140. G. Canavan, "Freedom to Mix Defenses in Modest Forces," Los Alamos National Laboratory Report LA-UR-01-4563, June 2001.

and hence little impact on stability. They could suppress proliferation in areas they could reach and have potential for technical cooperation, although their requirements for BMC3 and responsiveness are too short for direct Allied involvement. Vulnerable SBIs that depend on external sensors and C2 could suppress rogue ICBM and theater threats, which would remove proliferation incentives. Because they can be suppressed, they should have little impact on strategic forces or stability, but could offer options for cooperation on unclassified, non-proprietary technologies.

Brilliant Pebbles could impact strategic forces and proliferation. As they are survivable by design, they could have significant impact on offensive launches, although the analysis above indicates that they need not impact stability. BP are not subject to the catastrophic failures possible in other layers that could lead opponents to inaccurate, unfortunate assessments of net capabilities. BP could be operated under joint control with allies and others to enforce no-fly zones against unauthorized missile launches, but their advanced technology for effectiveness and survivability could impede cooperation. BP is the only defense appropriate for sophisticated or deliberate attacks. A BP overlay with a modest midcourse layer and a surface-based boost deployment could provide the maximum impact on protection, proliferation, and cooperation with the minimum impact on stability.

It has been argued in the past on qualitative grounds that even modest defenses could significantly reduce stability, but it now appears possible to deploy defenses large enough for rogue, accidental, or unauthorized launches by large forces without impacting stability. Bilateral stability between the United States and Russia is not impacted by any proposed level of defenses, whether surface- or space-based, midcourse or boost. The stability of the current offensive configuration is high and insensitive to the reduction of offensive forces, the deployment of defenses, and the exchange of offenses for defenses. Much the same situation obtains with respect to China and would-be proliferators. If U.S. offensive and defensive changes are made in appropriate relationships, they need not lessen stability with respect to Russia, China, and others. Thus, it does not appear that stability should be a constraint in force alterations as long as a few scaling relationships are satisfied in force modifications.

Considerations on Boost Phase and Follow-on Phases

Effective layered defenses could both address emerging threats and provide insurance against accidents during the drawdown of current strategic forces. Defenses of high effectiveness in every layer would produce compound attrition of threats and complicate the problem of attacking any given layer. Boost-phase defenses are particularly effective in such layered defenses, because they are insensitive to decoys. However, they must operate quickly, which means fast response and nearby basing. Surface- or air-based defenses are appropriate for threats that are accessible from safe launch areas, but require their interceptors to be based nearby and to achieve very high accelerations and velocities in the atmosphere. The performance required for surface-based interceptors is arguably feasible but appears to be beyond that of any in development.

Inland, multiple, and sophisticated launches favor SBI. The key parameters determining SBI effectiveness are delay, velocity, acceleration, and cost. For ICBM launches, the response times provided by current DSP satellites are adequate. For faster threats, on-board sensors could support the detection and release times needed. SBI sensors have demonstrated adequate performance and weight margin in space in long-duration missions. Small, efficient engines have been developed that are appropriate for rogue ICBMs. Theater missile programs have developed high acceleration engines that should suffice for regional and theater threats after some development.

Preliminary cost exchange estimates based on the information available indicate that SBI should be competitive with or superior to other means of intercept in stressing engagements. They could support boost-phase defense that would be an appropriate first layer for rogue, accidental, and unauthorized launches. GPALS experience suggests that SBI could be developed in two to three years at a cost of a

few percent of the current MDA budget. Before resuming their testing, it would be appropriate to update sensor, computer, and guidance technology and test integrated interceptors carefully under realistic conditions. An important element in that is the ability to develop and test SBI rapidly in a realistic environment, which is possible now that the restrictions of the ABM Treaty are no longer a constraint. The integration of defensive capabilities with offense forces could maximize the overall protection of the country, minimize the likelihood and consequence of incidents in the transitional period, and maintain stability during the transition from offensive to defensive forces.

Impact of the ABM Treaty

The ABM Treaty was mentioned in several sections above because it has impacted every program since Safeguard. It was initially hoped that the Treaty would slow the growth of offensive missiles; instead, offensive missiles grew more rapidly after it was signed. Nor has it slowed proliferation of offensive missiles, which has made the need for defenses greater. It had no impact on NIKE, Sentinel, and Safeguard because they completed their development and deployment before it went into force, but it was based on an incorrect interpretation of their limitations. The Treaty killed LOADS, which could have been an effective defense of the deterrent, whose deployment might have eliminated the need for further development of the defenses that ultimately undercut the Treaty. Its prohibitions on nuclear systems indirectly stimulated R&D on new DEW concepts, although it was understood at the outset that their results could not be deployed while the Treaty was in force. They would violate its prohibition on mobile launchers, which also inhibited work on space-based sensors and interceptors.

SDI did not propose a specific system, but the existence of the Treaty limited its scope to concepts that might be viewed later as somehow Treaty compliant. Most of its studies were of radars and rockets in compliant modes. Satellites were considered, but defined as "adjuncts" to marginal radars, which inhibited thought on new technologies or their effective integration into battle management. The Navy stayed out of SDI altogether as long as its platforms could not be shown to be compliant for fear that they would be constrained by the Treaty.

GPALS had to assume that the Treaty would be amended or go away, because the core of its defenses was BP, which was mobile. GPALS succeeded in thinking "out of the Treaty box," but was terminated for doing so. The TMD program explicitly eliminated NMD and global elements on the argument that the "ABM Treaty was the cornerstone of stability." The NMD program was the minimal response allowed to the North Korean launch over Japan. It confined itself to one site and argued that its interceptor basing was compliant, although it violated Article I among others.

On June 13, 2001, the U.S. withdrew from the ABM Treaty with little notice or complaint from domestic and international critics who had argued that it would destroy the foundations of stability. Domestic groups were largely quiet; NATO reaction was minimal; Russian comments were circumspect. Even China, which tried to make itself an unofficial signator of the Treaty in its final months, could gain little support for its opposition. Leaving the Treaty gave researchers the freedom to explore new ideas and technologies essential for the effectiveness of current and proposed systems.

Interestingly, until recently the MDA program appeared to act as if the Treaty was still in force. Its main developmental elements are still the GBI and midcourse NMD from the Clinton Administration. Elimination of the limitation to one site has not yet been exploited. Satellites, BMC2, and theater sensors have not yet been integrated, although there are plans to do so in the IDO. Navy mobile elements have only recently been added. Surface-based boost-phase interceptors are under consideration, but only for long-term development, and SBI has not been restarted, which is surprising given that it was the key element of the GPALS, which was assessed to be much more capable that the current midcourse system under development.

The ABM Treaty left undefined the status of theater and regional systems, thereby inhibiting the development of the mobile sensors and interceptors that would be most effective in regional defenses, which delayed the Navy's entry into strategic defense for about a decade. Thus, its elements are still relatively immature. Satellites were adjuncts to Treaty-defined ABM radars that are marginal in an era when interceptors largely manage themselves. For effectiveness against short timeline theater threats, tight integration of sensors and interceptors is essential. At a minimum, they must be coordinated to support the efficient allocation needed to maximize the effectiveness of limited numbers of interceptors. The need for the integration of sensors and systems for maximum efficiency was recognized at the beginning of SDI, but was inhibited by Treaty considerations. Thus, the needed move to their necessary fusion was lost, and individual systems were allowed to develop their own stovepipes. The challenge for the current system is to reintegrate them without having to rebuild them from scratch.

Attributes of Successful Previous Programs

In assessing the probability that MDA's current program will produce a broader and more robust program, it is helpful to review the attributes of successful past programs of this scope. Doing so suggests that the options under consideration by the MDA are too narrow and are supported by inadequate R&D. Successful previous national programs—Manhattan, Polaris, Minuteman, Oxcart, U-2, Corona, DSP, Apollo, etc.—shared certain characteristics. Each had a definite goal, a time span of 3–5 years, and a small core team that remained with it throughout. Each had leadership from within and little external control, review, or security impediments. Their early stages were generally the most successful—producing most of the strategic, tactical, and intelligence systems of today. By contrast, missile defense has undergone a number of disjointed programs. Each had some success. Nuclear systems produced rockets, radars, and the technology for surveillance systems and an appreciation for reliability in large systems, nuclear phenomenology, and BMC2. The Interim programs stimulated research on new concepts that SDI developed further. GPALS selected the best of them and advanced global elements to within a few years of deployment. BMDO developed TMD components, and NMD restarted national defenses.

Successful early national programs were rich in options and research, which was fortunate. Many options failed, but they always had an alternative ready. Missile defense started that way, but gradually focused its research in support of its main concepts which generally had one main approach per program element. That led to programs with little breadth off the current critical path. That made their probability of success the compound probabilities of many steps, which has produced a large supply of failure modes for study. For NIKE it was soft radars; for Sentinel, it was inadequate MSRs; for Safeguard, it was PAR blackout; for SBL, SBI, and hence SDI, it was survivability; for TMD and NMD, it was boosters, quality control, and countermeasures.

The one approach that did not fail for technical reasons was GPALS, which benefited from adequate funding of technology and multiple options for each of its key technologies. It failed because the high level of capability and autonomy to which it aspired disturbed critics. It clashed with their concept of what was stabilizing, although that concept ultimately proved indefensible. GPALS can be said to have failed because of the flawed understanding of the positive impact it could have had on stability at the time.

Successful early programs were identified with key individuals: General Bernard Schriever for missiles and warning satellites, Dr. Albert Wheelon for Corona, Dr. Edward Teller for the hydrogen bomb, Admiral Hyman Rickover for Polaris, and so on. Each recruited a core cadre who understood both the need and the possible solutions and stayed for long enough to see them through to completion and deployment. Program management is now a profession. Few managers stay for more than a fraction of the 15–20 year duration of current programs, which means there is tendency for them to take

their promotions and leave problems to their successors. Management reforms appear to aggravate that tendency.

The final point is competition. There is now little competition in ideas and demonstrations past the inexpensive initial paper phases. Each element is quickly narrowed to a fixed approach. NMD became a ground-based midcourse system based on existing EWRs, DSP, and GBI, whose designs proved resistant to new technology when found inadequate to countermeasures. Missile defense is an important field of study that deserves enough research and development on alternate approaches to assure that such failures will not happen again. Restoring that level of research would be relatively inexpensive and would represent a return to the mode of operation used in previous successful national programs.

Such R&D is important for another reason. NIEs confirm the attractiveness of missiles and WMD to rogues and others and indicate that the tools to improve their effectiveness are already in commerce. Current defenses might be able to provide defenses against rogue threats' indigenous countermeasures, but in the coming years, sophisticated countermeasures developed by the former Soviet Union may become available to them. If so, it will be necessary to address a progression of ever more capable and sophisticated countermeasures that will continue until offensive missiles are convincingly devalued. Until that happens, a progression of defensive measures will be needed, which will have to draw on an equally robust set of R&D programs.

A cloud on the generally positive picture above is that missile defense does not have the support of the scientific community. It had the support in the immediate post–World War II period, when many scientists who worked during World War II remained active, but that support was lost in the divisive debates over Sentinel, Safeguard, and the ABM Treaty. Few scientists who engaged in those debates responded to President Reagan's call for research to eliminate offensive missiles. Some thought it was a resumption of the attempt to deploy Safeguard, which led to the reluctance of academic scientists to discuss the new approaches under study except in polemics. In a sense, this polarization appears to be a continuation of debates going back to the hydrogen bomb, as few scientists have changed sides on missile defense since the Sentinel, Safeguard, and hydrogen bomb debates. Scientists in National Laboratories, military managers, and industrial engineers should be able to meet the near-term technical needs of the IDO, but if missile defense is to meet its full potential, it will have to draw fully on the talents of the full set of scientists who could contribute. Convincing them to do so will require a full and honest discussion of the technical strengths and weakness of the program and a willingness to support the new concepts that arise from it.

Summary and Conclusions

Developing missile defenses has been a long and difficult process. The steps have built on one another, although not always in a clear and logical way. Defenses have historically been about one step behind offenses—perhaps by design over the last decade. However, offensive developments have now slowed. There are few obvious benefits from higher yields, accuracies, or numbers. Conversely, defensive technologies continue to multiply and mature. The early NIKE, Sentry, Safeguard developments significantly improved missiles and ground-based radars, including the phased-array radars essential to any proposed deployment. They started the development of the metric sensors required for highly effective defenses. The specific systems they developed were susceptible to newly discovered nuclear phenomena, but their basic technology was a significant advance. While they were criticized on the basis of countermeasures that had not been demonstrated, if they had been deployed as intended, they would have been as effective as current systems against the rogue threats under consideration.

The Interim LOADS developed effective nuclear defenses for improving Minuteman survivability that would be effective today if such improved protection was needed. That it is not is a measure of the progress in international cooperation over the last two decades, to which defenses have made an important contribution. Interim developments also produced HTK and DEW, which made it possible to access for the first time the high leverage possible in boost-phase intercepts. DEW has not yet found an appropriate application other than the Airborne Laser, but it is a resource for downstream threats that require faster responses and non-ballistic missile threats that require the prompt delivery of energy all the way to the ground anywhere on the globe. Army research and development in KEW produced options for effective and affordable defenses in all stages of missile flight.

The SDI took these new technologies through assessment and testing in a program whose strong research component provided many of the concepts being converted into systems today. It also fostered the treatment of missile defense as a system, rather than as a set of isolated technologies. GPALS selected the best and most mature of those technologies to formulate a layered system that could give high levels of protection against global threats to U.S. citizens, deployed forces, and allies. It was built around BP, which provided both defenses and an integrated, global BMC3 to all layers of defense. The GPALS architecture was sufficiently credible to win the support of scientific, military, and international communities. It was opposed by some who were concerned that its effectiveness might erode strategic stability, a concern that appears to linger today among both missile defense critics and advocates.

TMD was a deliberate step back from the goal of protection of population. It recognized the weaknesses in defenses for deployed forces identified in GPALS and demonstrated in the Gulf War. It concentrated on the improvement of PAC for defense of deployed forces and Arrow for defense of Israel, but weakened NMD and killed global defenses. The subsequent NMD program had to resurrect a subset of the GPALS building blocks for midcourse defense. NMD introduced little new technology, although in its later stages it did introduce more effective procedures for testing the components it inherited. Against simple threats, it should have achieved roughly the same effectiveness as earlier nuclear systems. It did not advance the level of countermeasures that it could address, which was their limiting phenomenology.

The current program is essentially the NMD system of the Clinton Administration with two important modifications. The first is that it is actually building the BMC2 needed and attempting to integrate it with the C2 for other defensive layers and regions, which NMD did not attempt. The second is that it is working toward a fixed date for activation of the IDO. It will be a modest system, but certainly worth having, given the pace of missile and weapon proliferation. The IDO and the GMD into which it fits are ground-based, midcourse, radar- and IR-driven systems with the problems those technologies entail, but it is argued that it should be adequate for the threats of this decade, given appropriate caveats on the difficulty of predicting those threats. The IDO faces significant challenges in construction, international cooperation, activation, and testing, but there appear to be adequate workarounds for the major hurdles in each area, so it is likely that it will meet its IOC. GMD will extend this defense quantitatively, but not qualitatively without some fundamental advance in discrimination, which has proved difficult, fragile, and resistant to progress over the four decades during which it has been investigated strongly.

The key issue for the follow-on steps is to find some combination of them that can rapidly and affordably overcome the known weaknesses of GMD. Boost- and ascent-phase intercepts are reasonably well-developed possibilities. The two main approaches have strengths and weaknesses that are largely complementary. Surface-based systems have limited geographic coverage and demanding technology, but can concentrate on single launch areas. Space-based systems have good coverage for large areas. They are intrinsically global, which counts against them for single small threats, but counts to their advantage when the goal is to protect America, allies, and friends from missiles launched anywhere. SBI technology was developed extensively in GPALS, but has not been supported for the last decade. Surface-based systems appear to be adequate for a few geographic areas, but space-based systems are for the rest.

While somewhat disjointed, these developments have produced most of the pieces needed to make useful missile defenses. Theater systems are progressing, apart from cost growths and resulting delays. Regional systems are lacking, but could evolve from existing theater systems. Midcourse systems are relatively mature. They should be effective against design threats, although they need greater robustness against countermeasures. Boost-phase defenses are in disarray programmatically, but were developed to a significant level in the recent past and have been kept updated by related programs, so it should be possible to restore them relatively quickly, now that their main political impediments have been removed. There is little foundation as yet for integrating global, national, regional, and theater systems. There are provisions for direct downlink from satellites into theater ground stations, but not for providing accurate satellite and radar information on launches to the CCs responsible for using them. Introducing it could complicate technology and integration programs in the near term, but would improve their acceptance by CCs in the long run.

Navy contributions to missile defense need a broader view. In both CONUS and the theaters, their applications overlap the protection from land-based systems, which is more developed and arguably cheaper and easier to integrate. In the near-term defense of CONUS, current Navy systems would only contribute more GBI-like intercepts. Search, track, and discrimination would have to be provided to them by external sensors. For simple threats, their additional interceptors would not be needed. For complex threats, the external sensors would be overloaded, so the search, track, and discrimination they need would not be available. Shipboard discrimination of threats will require new radars, which could take a decade. Use of existing Aegis in picket duty to refine the trajectories of missiles detected by other sensors and intercepted by other systems is an important but unpopular and expensive role.

Existing Navy Aegis ships might be used to detect launches from ships close to the U.S. and commit existing Navy missiles with current or improved KVs; however, the integration involved is on the order of that for GMD. Current systems do not have a clear growth path to theater boost or ascent phase, although one could be formulated. They should be a useful complement to SBI, which could be

more costly in areas where sea-based systems have access to boost. Now that Navy defenses have had several successes, a strategy is needed to achieve their long-term potential. Getting support seems more difficult than in the past. Navy programs survived the Clinton Administration with congressional help, but the current Administration's actions have undercut them. The Navy now faces the choice of building the modest systems left from the Clinton Administration or shifting to big missiles on dedicated picket ships, neither of which has been endorsed.

Boost-phase defense development programs are uneven. Surface-based concepts are early in development and years from testing, let alone deployment. Space-based interceptors were the most mature elements and were designated the "first to deploy" during SDI and GPALS, but were deliberately delayed for the last decade. Given their significant development in earlier programs and updates by KV programs, their testing could arguably be completed in a few years, and their deployment could be executed on roughly the time scale of the second spiral. Some strong organization would have to lead that effort. In SDI, the Air Force led development of SBI, although not without strains. In GPALS, SDIO led development through the National Laboratories with more success and fewer strains, although ultimately no product. Given the recent reorganization to centralize management of space programs in the Air Force Space Command, it might be capable of executing a fast-paced development program through its Space and Missile Center. To take advantage of that opportunity, the Air Force would have to reassert its claim to SBI, which complements its current focus on space-based IR and radar sensors.

Defenses that could be deployed in a few years should be adequate for design rogue threats. They could probably handle a few missiles with light decoys, even operating in the environments the intercepts would produce. However, the next step is more difficult. Increasing the number of midcourse GBI should suffice for basic threats, but there is little confidence that midcourse systems could handle an attack of the size that China could now mount. Such defenses would at best lead to a MAD relationship with China like that the United States is now ending with Russia. Accidental or unauthorized attacks of the size considered in GPALS would suppress or saturate such midcourse defenses. A dozen ICBMs or a boatload of SLBMs would be beyond their capability, even if midcourse defenses could be made survivable. There is a natural complementarity between surface-based midcourse defenses for light rogue threats and space-based boost-phase defenses for larger threats and accidental or unauthorized launches. Developing one and not the other would reflect a lack of balance. Developing both should disabuse Russia and China of any benefit of developing more offenses and others from seeing any value in the development of offensive missiles before they ever started.

Layered Defenses for America

MDA has as its goal the development and deployment of effective defenses in each possible defensive layer—boost, midcourse, and terminal—for the United States, its forces, and its allies against missiles launches anywhere on the globe as soon as possible using existing technology and systems. It is an appropriate but difficult goal that is not close to realization. Midcourse is the most developed layer. MDA is developing the IDO Capability by integrating the ground-based GBIs developed over the last two decades with radars and satellites of comparable vintage. That integration is to be completed to provide protection from missiles from Northeast Asia in 2004 and from the Middle East and Southwest Asia in 2005. Their BMC3 will be integrated across systems and with USSTRATCOM's BMC2 system for the planning of U.S. defenses and control of offensive missiles. These systems will be upgraded in a spiral modification to produce the maximum effectiveness and efficiency possible at each point in time. MDA's proposed increased use and pace of testing is essential to the timely realization of such defenses.

The development programs for boost-phase systems were delayed and confused by the policies of the previous administration. A program for the development and testing of surface-based interceptors that

could be used on land or ships has been formulated and initiated, but is not intended to produce a deployable system for 6–8 years, so it does not impact the logic of near-term spiral developments. Surface-based systems appear to be well suited to the few threats that afford them safe access to boost. A program for the development of SBI has been formulated, but not initiated. It is intended to produce a tested system on a time scale even longer than that for surface-based interceptors, which appears to ignore their significant development during GPALS. Space-based systems are preferred on cost and coverage grounds for large, multiple, or global threats.

PAC-3 and THAAD should provide adequate theater systems. They might also be able to provide protection for large cities from missiles launched from short ranges off shore, although there are a number of issues to be addressed in providing timely information from off-board detection sensors and C2. Those missiles could not adequately address ICBMs with decoys attacking those cities. They do not have the acceleration or velocity to produce useful footprints against discriminated threats. The accelerations and velocities required approach Sprint's, which performed in similar range-altitude combinations. However, Sprint was nuclear, so it could be command guided. The NNK interceptors that could have performed such engagements were terminated by the previous administration and have not been restarted. Thus, there is no terminal phase capability to defend cities from ICBMs and no program to produce one.

The technologies described in the previous sections could make effective defenses possible in each layer on the time scales desired; thus, there is a potential match between the goals of the current program and the technology available to support it. However, only a portion of those technologies are under active development. The current MDA program is effectively a single, midcourse system and is likely to remain so until well into the next decade. Barring fundamental improvement in the ability to discriminate midcourse threats, that system will be effective against a few missiles with a few simple decoys. It is as described "better than nothing," but primarily represents protection *in extremis*. As the protection it affords could fail catastrophically with the development of more sophisticated decoys or countermeasures, it would not represent a reliable military capability.

The missile defense program is comparable in scope to that of previous national programs—the Manhattan Project, the hydrogen bomb, ICBMs, Corona, etc.—but the current program differs from them in significant ways. Previous successful national programs had definite goals, short time scales, strong internal leadership, multiple options, and deep R&D to support each. R&D options were so deep that it seemed they could not fail, which was fortunate, because most of them had to call on most of their options to achieve success. Several approaches were often needed at each step in the Manhattan project, including five for the production of the material for the uranium weapon and a totally new approach to the design of plutonium bombs.

The current program, to the extent that it is defined, is an open-ended progression of steps designed by a committee. Its goal of producing an operating system in a few years is an improvement over the programs of the last two decades, which had no such commitment, but that alone is not enough. Safeguard deployed a system, but had to deactivate it after a day because it clearly lacked robustness against countermeasures, including its disturbed backgrounds. The current program will also deploy an initial system that lacks robustness against plausible countermeasures—indeed, many of the same countermeasures. If that can be remedied by spiral development, the problems in the IDO need not be debilitating. However, those problems cannot be overcome by deploying more of the same or similar interceptors and radars. The improvements currently under discussion are of that type.

The broad supporting R&D included in earlier successful missile defense programs was lost as total resources were spread too thin, committed to current problems, or removed for political reasons. The damage was serious. A decade ago, GPALS invested over $1 billion per year on its technology base. The Clinton Administration cut it to a few $10 million per year, so key industry and university teams left the field. Under the current Administration, MDA is spending less than $100 million per year on

key technologies and systems out of a budget that is twice GPALS's. The research budget for the current program and the number of options it maintains are so small that previous experience with successful national programs implies that it will run out of options before reaching a significant product.

The current program has both a lack of options and too narrow a focus on the few systems it includes. One can argue how well it will perform against the jammers, decoys, backgrounds, and countermeasures discussed in public and technical literatures, but it is difficult to argue that it will not fail at some level of these known threats. A decade ago during GPALS, a number of independent review committees concluded that BP was not susceptible to such countermeasures, an assessment that has not since been disputed. Adding rogue ICBMs, regional, and theater missiles to the accidental and unauthorized launches of GPALS introduced some new issues, but the above analysis indicates that SBI should be able to address them with modest development and could introduce important options into the spiral development process on relevant time scales. BP was about halfway through EMD before it was cancelled, which suggests that SBI could be developed and deployed on roughly the time scale as the second MDA spiral. Doing so would provide a capable and affordable boost layer that would reduce the threats reaching midcourse to levels that GBI could address. The current program does not develop SBI on that time scale, which suggests that MDA does not grasp the limitations of midcourse, does not understand the positive impact of SBI on stability, or implicitly respects the ABM Treaty.

Current technical developments are well suited to block modifications, but the integration of global, regional, and theater BMC3, which is as important to the performance of the overall system as that of the individual sensors and interceptors, is immature and untested. Connecting and integrating them with the attack operations and passive defenses essential for overall effectiveness is the greatest challenge at present. For integration to proceed, early resolution of command structures is required.

The millennial competition between the offense and defense is not likely to be settled by a single stroke, nor is any defensive solution likely to remain static. There was a window in which GPALS arguably could have eliminated the utility of offensive missiles. It was lost during the Clinton Administration's emphasis on the domestic economy, TMD, NMD, and the ABM Treaty. The current program could serve as the first step in a continuing spiral that could respond to the progressive improvements in offensive missiles, but the appropriate long-term goal is to put missile defenses so far ahead of offenses that they will dissuade rogues and others from engaging in missile competitions altogether. That goal is not beyond the capability of missile defenses, even those attainable in the near term. The tools are now at hand, but not all are being fully developed. A balanced program must develop and use all available tools, including the full potential of space-based sensors and interceptors.

Appendix A
Effects of Nuclear Weapons

Atomic weapons have yields Y of a few 10s to 100s of kilotons and produce lethal radii of a few kilometers. The pressure p at distance r from a ground explosion of yield Y is about $Y/2\pi r^3$, so the pressure p_D required for damage of a given type of structures is achieved at a radius of roughly $r_D = (Y/2\pi p_D)^{1/3}$. A 1 MT explosion produces a crater about 100 m deep, a fireball out to about 1 km, and the 1 psi (lb/in^2) needed for destruction of ordinary structures out to about 4 km.[141] Fireball diameters D scale on density relative to sea level as

Equation A.1

$$D(km) = \left(Y(MT) \frac{\rho_0}{\rho} \right)^{\frac{1}{3}}$$

Thus, a 1 MT explosion at sea level produces about a 1 km fireball. Air density falls by a factor of 10 every 15 km in altitude, so a 100 KT explosion a produces a 1 km fireball at 15 km, and a 10 KT explosion produces a 1 km fireball at 30 km. As fireballs rise, they entrain air and cool, but remain sufficiently ionized to absorb strongly around their periphery and reflect from their core, which suggests the presence of a tight, hot, unmixed torus. As their size increases, low-altitude fireballs obscure solid angles of about 1 sr at 1 km, $(1\text{km}/15\text{km})^2 = 0.04$ sr at 15 km, and 0.001 sr at 30 km. Using KT-range devices in 10–20 km intercepts avoids cluttering up the space through which the radar must search for or track the weapon.

Surface damage due to a high-altitude explosion falls because the shock from the fireball redirects energy upward, reducing damage below, as predicted theoretically and observed experimentally in the Soviet half-yield test of a 100 MT weapon at 10 km.[142] The shock propagating downward at a distance R below the burst is initially strong, so it has a velocity v of approximately $\sqrt{(Y/\rho R^3)}$. As R increases (i.e., as the shock propagates to lower altitude), air density increases as $e^{R/H}$, where the scale height H is roughly 7.5 km. Thus, v falls as $e^{-R/2H}$, which is a factor of 10 every 30 km, so the downward shock weakens by a factor of 100 every 30 km, producing much less damage on the ground than it would in a uniform atmosphere. The upward going shock increases in strength as $e^{R/2H}$, so it essentially blows the top off the atmosphere, leaving a hot, under-dense bubble behind. As large explosions at high altitude primarily eject mass upwards, they do less damage below, so it is possible to shield ground targets from large explosions, providing the intercepts can be kept at altitudes of a few scale heights.

This scaling breaks down above about 75 km, where an increasing fraction of the warhead's energy escapes, and is altered sharply above 100 km, where most of its x-rays escape to long distances, which reduces the residual energy and size of the fireball. The scaling alters again at about 150 km, where weapons with high yield-to-mass ratios are not thought to pick up the neutral fraction of the air. That increases the expected size of the fireballs in that altitude region. At altitudes of 300 km and higher,

141. S. Glasstone, *Effects of Nuclear Weapons* (Washington, D.C.: U.S. Government Printing Office, 1964).
142. Y. Zeldovich and Y. Raizer, *Physics of Shock Waves and High Temperature Hydrodynamic Phenomena* (New York: Academic Press, 1967), Vol II., 812–819.

the expansion of weapon debris ions is primarily contained by the Earth's geomagnetic field, which produces diameters for MT bursts approaching 1,000 km.

Appendix B
Radar Cross Sections and Technology

Radars perform three distinct functions—search, track, and discrimination—that scale differently on power P, aperture area A, and wavelength λ, and target radar cross section σ. Scaling for the three functions is discussed first. Then, the tradeoffs involved in using the same radar for multiple functions are discussed.

Radar Cross Sections

Radar cross sections (RCS) are difficult to compute theoretically; in practice, they are measured experimentally, but the cross sections of some simple shapes can be determined analytically, which gives insight into the scaling of more complex shapes.[143] Spheres of radius r much larger than λ have $\sigma = \pi r^2$, the familiar limit from geometric optics. An RV abruptly truncated at its base also has a cross section of that form with a radius of its base radius. An RV with a rounded nose have a similar cross sections with a radius equal to its nose radius. A sphere with a radius much smaller than λ has $\sigma =$ volume$^2/\lambda^4$, which is the familiar Raleigh limit for volume scattering.

A flat plate of diameter D seen broadside has gain for backward scattering of about $(D/\lambda)^2$; so it has a cross section $\sigma = (D/\lambda)^2 D^2 = D^4/\lambda^2$. As one transverse dimension d becomes smaller than λ, the plate's gain in that direction approaches 2π, which produces the broadside cross section $\sigma = 2\pi d(D/\lambda)D = 2\pi dD^2/\lambda$ of a cylinder. A long wire has area λD and gain D/λ, which produces a cross section $\sigma = \pi D^2$. For $\lambda = 2D$ this reduces to the $\sigma = \pi(\lambda/2)^2$ of a resonant dipole. When both dimensions are smaller than λ, they are replaced by λ and both gains by 2π, which produces the $\sigma = \lambda^2/10$ of a sharp cone within a few 10s of degrees of its nose, which is the orientation an RV tries to maintain with respect to high frequency tracking radars.

Search. A radar of power P radiating into a solid angle Ω for time t produces an energy per unit area $Pt/\Omega R^2$ at range R, so a target of RCS σ intercepts energy $\sigma Pt/\Omega R^2$. If the target scatters isotropically, the energy density scattered back at the radar is $\sigma Pt/\Omega R^2(1/4\pi R^2)$. A receiver of area A there collects an energy of about $\sigma tPA/4\pi\Omega R^4$. For target detection, this received energy must be about a factor of 10 larger than the background noise BkT, where for matched filtering the radar bandwidth $B = 1/t$, k is Boltzman's constant, and T is the background temperature. Thus, effective search requires $PA \propto \Omega R^4 BT/\sigma$. The key parameter is the power-aperture product PA, which increases with search solid angle Ω, range R, bandwidth B, and noise temperature T, and falls with target RCS σ. For given PA product, inverting this equation shows that the search range scales as

Equation B.1
$$R_{search} \propto \left(\frac{PA\sigma}{\Omega BT}\right)^{\frac{1}{4}}$$

143. Toomay, *Radar Principles for the Non-Specialist*, p. 74, Table 4.1.

Thus, for long ranges, large power-aperture products are required. Diffraction sets the beam width as approximately $\lambda/\sqrt{A}$, which produces a cross-range beam width of about $R\lambda/\sqrt{A}$. That is generally fairly coarse from UHF radars, but angular resolution can be improved at high signal-to-noise ratio (S/N) with averaging. The variation of radar return with beam angle can determine angular position to about $(\lambda/\sqrt{A})/(S/N)$. If each observation can be treated as independent, M observations over the few hundred seconds of the RV's flight would improve accuracy by a factor of about $\sqrt{M}$.[144] These improvements do not obtain at maximum range. Accuracy improves with S/N, the number of observations, $1/R^4$, and $1/T$ as the RV approaches, so the full benefit is not realized until the RV approaches the radar.

Track requires that the signal to be larger than the noise in the cell being irradiated. A radar of aperture A and wavelength λ deposits an energy fluence $\sigma Pt/(R\lambda/\sqrt{A})^2$ in time t on a target of RCS σ at range R. A fraction $A/4\pi R^2$ of the scattered energy is collected by the radar for a total of $\sigma PA^2t/\lambda^2 4\pi R^4$. That must be larger than noise kT, so $PA^2 \propto \lambda^2 R^4 BT/\sigma$, where the key parameter is PA^2, which places greater emphasis on aperture than power for high resolution. PA^2 increases with R, B, T, and decreases with σ. It scales as λ^2, so shorter wavelengths are preferred for track. Range scales as

Equation B.2

$$R_{track} \propto \left(\frac{PA^2\sigma}{BT\lambda^2}\right)^{\frac{1}{4}}$$

The $\lambda = 0.03$ m (10 GHz) x-band radars radiate about 170 kW through roughly 13 m faces to produce ranges on the order of 4,000 km, where their 0.13° beams produce cross resolutions of 9 m.

Discrimination. Search and track ranges scale as $1/B^{1/4}$, so both favor narrow bandwidths. Range resolution scales as $c/2B$, which favors wideband operation. Thus, BMEWS radars have bandwidths of about 0.6 GHz in search and 10 MHz in track, which give resolutions of 250 and 15 m that are not useful for discrimination. PAVE PAWS have 0.1 MHz in search and 1 MHz in track and resolutions of 1,500 and 150 m, which are not useful either. Both will be upgraded to 30 MHz bandwidths to give resolutions of about 5 m in both modes, which could separate large objects such as tanks from RVs and decoys. They can also revisit targets and measure their temporal fluctuations as rough temporal discriminants. X-band radars have bandwidths of about 1 GHz and resolutions of about 10 cm, which can inspect RVs and decoys in detail, particularly when combined with phase and Doppler imaging.

Countermeasures. For a jammer with bandwidth b, power p, and gain g in the direction of the radar, its energy density at the radar is pg/bR^2, which is generally much larger than thermal noise kT, which can be ignored. The power density from a target of RCS σ at range R is about $\sigma PG/BR^4$, where its gain is $G = A/\lambda^2$. The ratio of the energy density from the target to that from the jammer (S/J) is about $\sigma bPG/BpgR^2$, which decreases with R. Thus, beyond some range the jammer will deliver more power than the target, producing $S/J < 1$, which occurs for $R > \sqrt{(\sigma bPG/Bpg)}$. When the jammer overwhelms the radar, the radar only sees the jammer, and the actual target range and position are concealed. If the jammer is in a side lobe of the radar, it can still inject noise, but its efficiency in doing so is reduced by the radar's gain in that lobe. Successive lobes are generally suppressed by about 20 db, i.e., successive factors of 100, so the burn through ranges for successive lobes are increased by successive factors of 10.

144. Sessler *et al.*, "Countermeasures," p. 136.

Appendix C
Atmospheric Reentry

On reentering the atmosphere, objects are slowed by drag. In the hypersonic regime, their deceleration is described approximately by Newton's second law $mdV/dt = -C\rho AV^2$, where C is a drag coefficient on the order of unity, A is the object's projected area, V is its velocity, and m is its mass. These parameters can be collected into a ballistic coefficient $\beta = m/CA$, the RV's mass per unit area. Along its trajectory $d/dt = -V \sin\phi\, d/dz$, where ϕ is the angle between the objects path and the local horizontal, the deceleration equation becomes $\beta \sin\phi\, dV/dz = \rho V$, which has solution $V = V_o e^{-H\rho/\beta \sin\phi}$. High objects decelerate in the lower atmosphere, where the density has an approximately exponential distribution with scale height $H = 7.5$ km. Strong deceleration occurs at the altitude where the exponent is of order unity, i.e., at

Equation C.1
$$z_R = H \ln\left(\frac{H\rho_0}{\beta \sin\phi}\right)$$

which is about 27 km for a $\beta = 500$ heat sunk RV, 22 km for an RV with a nominal $\beta = 1{,}000$ kg/m^2, and 18 km for a $\beta = 2{,}000$ precision ablator RV. The reentry angle ϕ is about 22.4° for minimum energy intercontinental trajectories, so the distance the RV flies in the atmosphere after discrimination on drag at 100 km is about $r = 100$ km/sin(22.4°) = 260 km, and the time it takes to do so is about r/V_o = 260 km/7.2 km/s = 36 s. For discrimination on atmospheric drag, the change in the object's velocity must be large enough to be observable. That occurs at higher altitudes, where the exponent is small, which gives a $\Delta V = V_o H\rho/\beta \sin\phi$, which is shown in Figure C.1. A given ΔV occurs at altitude

Equation C.2
$$z_{decoy} = H \ln\left(\frac{V_0 H\rho_0}{\Delta V \beta \sin\phi}\right)$$

A radar operating at wavelength λ in the pulse burst mode (string of pulses) for T seconds has a velocity resolution $\lambda/2T$.[145]

145. Toomay, *Radar Principles for the Non-Specialist*, p. 96.

Figure C.1 Ballistic Missile Defense

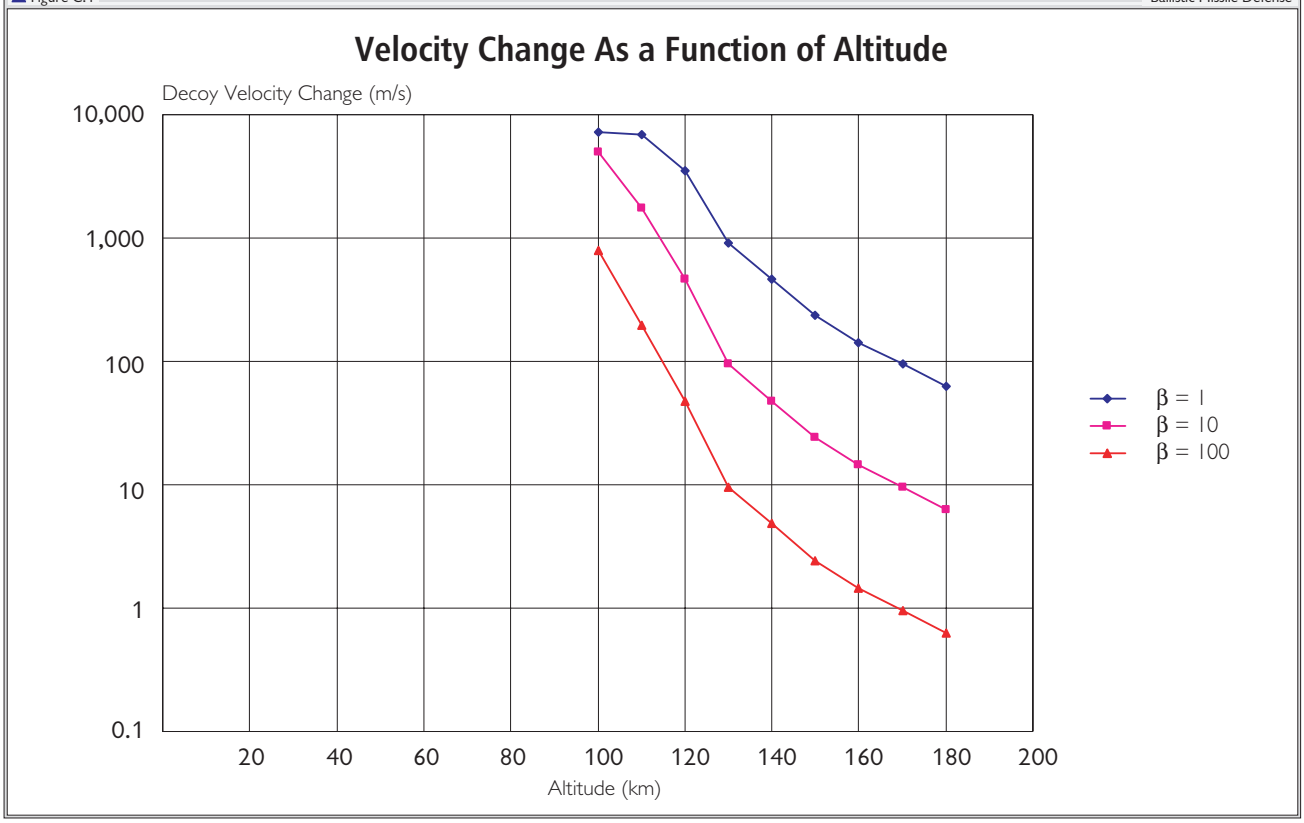
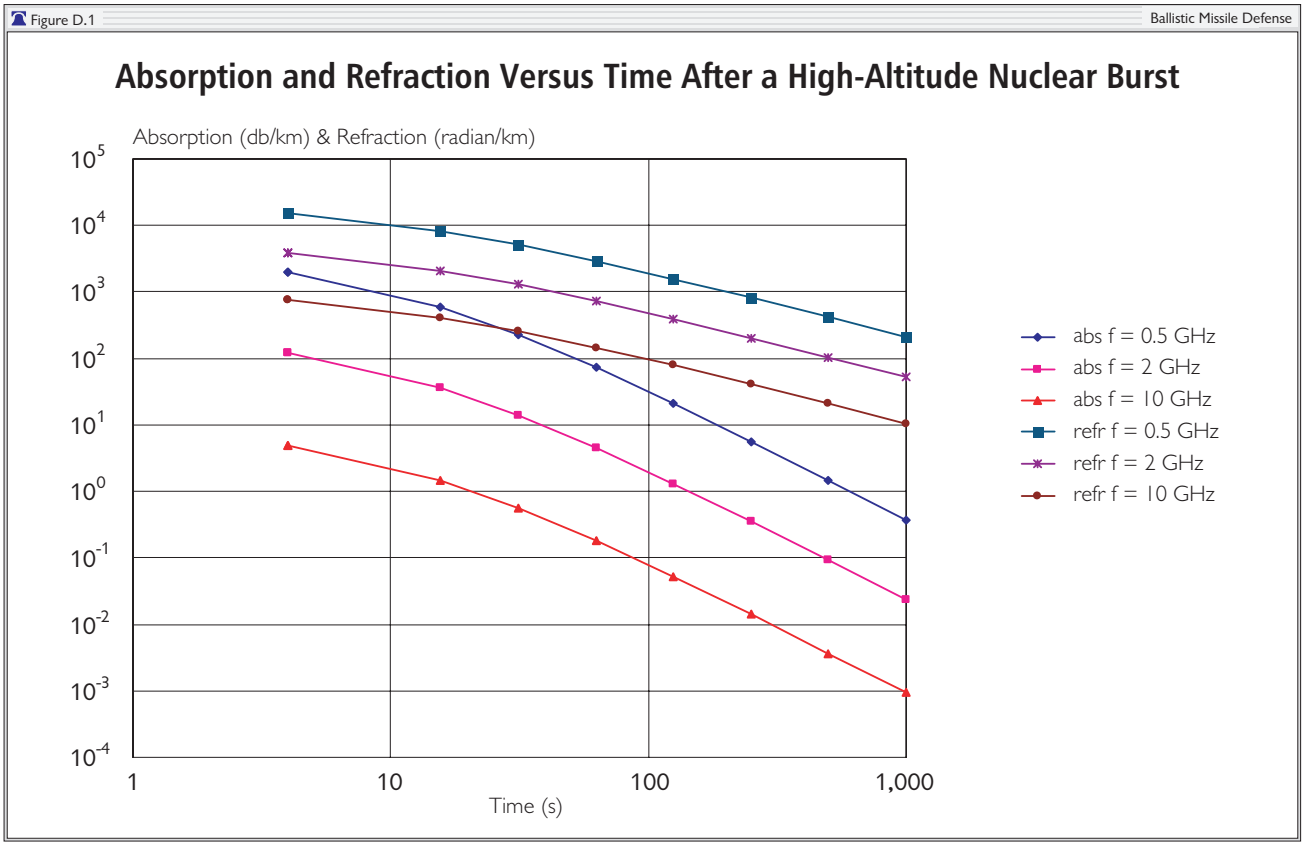


Figure D.1 Ballistic Missile Defense



Appendix D
Nuclear Effects

The dominant effects of nuclear explosions below 50–100 km are the reflection or strong attenuation of radio frequencies by the fireballs of modest size as they rise, entrain air, mix, cool, and stabilize. Despite several 10s of tests and multiple measurements on each, there are still uncertainties in each of these stages, which lead to uncertainties in the observable phenomena expected from bursts in this regime. The dominant effects of higher altitude explosions are blackout, refraction, and atmospheric heave, which are complicated by uncertainties in hydrodynamic coupling and energy transport. Blackout was recognized and quantified at the time of the Safeguard debates. Refraction was recognized but not quantified, particularly for multiple bursts. Heave was recognized later and was not quantified for multiple bursts during the time of the debates.

Blackout. At sea level, MT bursts produce fireballs about 1 km across. Their diameters scale as the cube root of the ratio of yield to density, so a 10 KT burst at 30 m would also produce a fireball of about that size. Fireballs reach pressure balance in a few 10s of seconds. While expanding they cool to temperatures around 1 ev, at which thermal radiation can escape. Over a few 10s of seconds the fireball rises to about a scale height H of 7.5 km before stabilizing.[146] It entrains several times its initial mass in the process, which cools much of the torus to 0.5–0.3 ev. Subsequent motion and cooling are slower. The fireball is a hot, underdense gas, in which high temperatures produce significant ionization, particularly at high altitudes. The absorption per unit length of such a hot plasma is

Equation D.1

$$\alpha = C\left(\frac{\omega_p^{\ 2}}{\omega^2 + \nu^2}\right)\nu \quad \text{(db/km)}$$

where $C = 1.5 \times 10^{-5}$, $\omega = 2\pi f$ is the radian frequency, $\omega_p^{\ 2} = 4\pi n_e e^2/m_e$ is the plasma frequency corresponding to electron density n_e, and ν is the electron-neutral collision frequency, which scales on neutral density as $\nu = \nu_o(\rho/\rho_o)$ with $\nu_o = 2 \times 10^{11}$/s at sea level. If $\omega < \omega_p$, the radar wave is reflected. Maximum absorption occurs at the altitude where $\omega = \nu = \nu_o(\rho/\rho_o) = \nu_o e^{-z/H}$, which is at $z = H \ln(\nu_o/\omega)$. The maximum absorption is at sea level for $\omega = 2 \times 10^{11}$/s ($f = 32$ GHz), where $\alpha = C\omega_p^{\ 2}/2\omega \propto n_e/\omega$. For $n_e = 10^9$/cc and $f = 500$ MHz, the maximum $\alpha = 10^4$ db/km at 30 km, falling by a factor of 10 each 15 km in either direction. The 1 km stabilized fireball from a 1 KT burst at 45 km gives a two-way attenuation of about 2 x 10 db/km x 1 km = 20 db, which would strongly degrade the radar's measurements. If n_e is proportional to the neutral density,

Equation D.2

$$\alpha = C\left(\frac{f_i\left(\frac{\rho}{\rho_0}\right)\omega_{po}^{\ 2}}{\omega^2 + \nu^2}\right)\nu$$

146. Zeldovich and Raizer, *Physics of Shock Waves and High Temperature Hydrodynamic Phenomena*, pp. 611–628.

where ω_{po} = 2.8 x 10^{14} rad/s is the plasma frequency corresponding to single ionization at sea level, corrected for fractional ionization f_i and neutral density at that altitude by ρ/ρ_o. That gives a maximum at sea level of $\alpha = Cf_i\omega_{po}^2/(\omega^2 + \nu_o^2)]\ \nu_o$ for all ω. For $f_i = 10^{-5}$, α is about 10^8, which would produce very large attenuations. For $\omega >> \nu$, α is about $Cf_i(\rho/\rho_o)\ \omega_{po}^2\nu/\omega^2 \propto (\rho/\omega)^2$, so attenuation falls by a factor of 100 every 15 km, giving α = 10 db/kb at 90 km. These large levels of attenuation are likely to persist for 100s of seconds, so fireballs from low-altitude bursts will remain excluded angles of about (3 km/30 km)^2 = 0.01 sr each for much of the attack.

MT explosions at altitudes of 150–250 km make hot, ionized fireballs a few hundred kilometers across, within which much of the ambient air molecules are stripped of some or all of their electrons, producing initial electron densities n_o of about 10^9 to 10^{12}/cc. At temperatures of several thousand degrees, the principal mechanism for removing the free electrons that cause absorption is radiative recombination, which is binary in electron number density and has a rate constant C_R measured to be about 10^{-12} cc/s. The electron density falls from its initial value as

Equation D.3
$$n_e = \frac{n_0}{1 + n_0 C_R t} \approx \frac{1}{C_R t} = \frac{10^{12}}{cc/t(s)}$$

for $t(s)$ greater than 10^{12}/cc/n_o. After about 300 s, the electron density drops to the critical density n_c = 3 x 10^9/cc needed to cause complete reflection at UHF.[147] At earlier times, the fireball would produce an obscured or reflective region of about (200 km/600 km)^2 = 0.1 sr between the PAR and the trajectories of subsequent RVs, which the PAR needed to detect and track at much shorter times. Such system obscurations were unacceptable, but there was little overlap between adjacent PARs, so they could not remove it by covering for each other.

At later times, when the fireball is no longer completely reflecting, absorption is still a concern. When electron-ion collisions are dominant, the absorption coefficient is $\alpha = C[\omega_p^2/(\omega^2 + \nu^2)]\nu$ (db/km), which scales as $\nu = \nu' n_e$, where ν' is about 10^{-4} cc/s at a temperature of 1 ev. At high altitudes, ν is much larger than ω, so α is about $C(\omega_p^2/\omega^2)\nu$. Since $\omega_p^2 = 4\pi n_e e^2/m_e$,

Equation D.4
$$\alpha = C\left(\frac{4\pi e^2}{m_e}\right)\nu'\left(\frac{n_e^2}{\omega^2}\right) \approx 0.1\left(\frac{n_e}{f}\right)^2 = 0.1\left(\frac{1}{fC_R t}\right)^2 = \frac{10^{23}}{(ft)^2}\,db/km$$

Figure D.1 shows the absorption as a function of time after a high-altitude explosion for frequencies of 0.5, 2, and 10 GHz. At PAR's 0.5 GHz frequency, absorption is over 1,000 db for short times. By t = 200 s, α drops to about 10 db/km. It would drop to about 0.4 db/km by 1,000 s, but that would still give two-way losses of about 2 x 0.4 db/km x 100 km = 80 db = 10^8.

The situation was more favorable at the threefold higher MSR frequency, which would cease to reflect from its n_c = 3 x 10^{10}/cc critical density contour at about $1/C_R n_c$ = 1/(3 x 10^{10}/cc x 10^{-12} cc/s) = 30 s. Its absorption would drop to about 1 db/km after 200 s. X-band radars have critical frequencies 20-fold higher than UHF, so their critical electron densities are about 1.2 x 10^{12}/cc, which would only be reached near explosions at about 150 km. Thus, they should not be reflected, and their losses would drop to 0.1 db/km after about 100s. The use of the nominal literature value of C_R of about 10^{-12} cc/s would lead to collision frequencies and absorptions about a factor of four higher than those estimated above. It would correspond to temperatures of about 0.1 ev, which is more appropriate for low-altitude fireballs. At high altitudes, electron temperatures are non-equilibrium and can remain higher than that of the surrounding air for times comparable to that of the engagement. The electron-ion collision

147. J. Jackson, *Classical Electrodynamics* (New York: John Wiley & Sons, Inc., 1975), pp. 284–289.

cross section scales inversely with energy, so its average rate scales as $1/T$. Thus, the value at 1 ev is about a factor of 3 lower than that for low-altitude bursts. Since absorption is quadratic in electron density, this factor of 3 produces a factor of 10-fold uncertainty in predicted absorption, which determines whether x-band radars could experience significant absorption.

Heave. The x-rays that escape explosions much higher than 100 km deposit a significant fraction of their energy at about 100 km, where the product of air density and absorption cross section is about the reciprocal of the scale height. At a distance R below a burst of energy Y, the energy deposited per unit mass e is about $Y/4\pi R^2 H_o \rho$, where H_o is the unperturbed scale height and the air density $\rho = mn$, where m is the mass of an air molecule and n is their number density. A 5 MT explosion 100 km above the deposition region gives a new internal energy per unit mass of about 5 x 10^6 J/kg, which is a 100-fold increase over ambient. That gives a sound speed $\sqrt{e}$ of about 2 km/s, a temperature T of about me/k, and an effective scale height H of about $kT/mg = e/g = 4 \times 10^6$ J/kg /10 m/s$^2 \approx 400$ km.

The air adjusts to this new equilibrium scale height by rising or "heaving" up at speeds $\sqrt{e}$ of about 2 km/s on a time scale of about $H_o/\sqrt{e} = 7$ km/2 km/s ≈ 5 s. In a few 10s of seconds, heave lofts air from about 100 km to much higher altitudes, which creates a new, denser atmosphere there that falls off exponentially on a length scale of about 400 km. Subsequent bursts take place in air of higher density than ambient and deposit their x-radiation in air of higher than ambient density and hence at higher altitudes. That air could in turn be heaved to still greater heights in an attack involving many detonations.

Refraction. At electron densities well below those required for reflection or excessive absorption, disturbed environments can cause significant degradation through refraction or bending of radar beams by irregularities in ionization distributions. The index of refraction at frequencies above the plasma frequency $\omega_p = 4\pi n_e e^2/m$, is $n = [1 - (\omega_p/\omega)^2]^{1/2}$. If a radar beam passes through regions of varying thickness that cause the two sides of the beam to experience a difference in physical path length d, that produces a phase difference that deflects the beam through an angle θ of about $(n - 1)d/\lambda$. For $\omega >> \omega_p$, n is about $1 - (\omega_p/\omega)^2/2$, so

Equation D.5

$$n - 1 \approx -\frac{\left(\frac{\omega_p}{\omega}\right)^2}{2} = \frac{-2\pi n_e e^2}{m\omega^2}$$

which produces a deflection of $-dn_e e^2/m\omega c$, which scales as dn_e/ω, i.e., it only falls off as $1/\omega t$, rather than the $1/(\omega t)^2$ of absorption. The previous section's predictions of electron density as a function of time gives a deflection

Equation D.6

$$\theta \approx \frac{de^2}{mcC_R t\omega}$$

Figure D.1 shows the amount of refraction expected at $f = 0.5$, 2, and 10 GHz as functions of time after a high-altitude burst. For 0.5 GHz, the refraction is initially about 2 x 10^4 rad/km, but drops to 200 rad/km after 1,000 s. A PAR or UEWR beam might be about 10 km across at 200–300 km altitude, so if the electron density contour was 1 km thicker for every 100 km horizontally, that would give a d = 10 km x (1 km/100 km) = 0.1 km, so the beam deflection would be θ = 200 rad/km x 0.1 km = 20 rad, which would completely disrupt the beam. At S-band, the radar beam width is about 3 km, so d = 0.03 and the late time deflection would be about 50 rad/km x 0.03 km = 1.5 rad = 90°. At x-band, the radar beam width is 1 km, so d = 0.01 km, and the late time beam deflection is θ = 10 rad/km x 0.01 km = 0.1 rad = 6°, which is still a problem, but would decrease in time.

Appendix E
Command Guidance and Hit-to-Kill Technology

The main progress in missile defense has been in accuracy. The meaning of "hitting a bullet with a bullet" has been refined from passing within a few kilometers with a nuclear bullet in 1960 to a few meters with a lethality enhancer in 1980 to a few centimeters with HTK today. Those increases have been supported by progress from command guidance to homing with on-board observations and computation. Both layers of Sentinel and Safeguard used radars to detect and track targets and command guide their interceptors. The radars tracked the incoming missile and the outgoing interceptor, measured the range and angle to each, computed their separation, calculated the optimal divert, directed the interceptor to execute it, and told it when to fuse. The few-degree accuracies of their radars produced miss distances comparable to the kill radii of the interceptors' nuclear explosives. This command guidance was conceptually simple, but introduced a communication channel that was susceptible to interference and nuclear effects. Proportional navigation (PN) gives the transverse acceleration

Equation E.1
$$y'' = KV_C\lambda'$$

needed for intercept as a function of closing velocity V_C, LOS rate of change λ', and the "navigation ratio" constant K, as determined by the interceptor kinematics, electronics, and its designer. The first integral of the acceleration gives the interceptor's relative velocity,

Equation E.2
$$V = y' = KV_C\lambda + V_M$$

where V_M is the missile's velocity perpendicular to the LOS. Velocity integrates the LOS rate, which is a smoothing process that reduces bandwidth requirements. At time t, if the interceptor's relative transverse displacement y is small compared to the distance to impact, λ is approximately $y/V_C(T - t)$, where T is the projected time of impact, so $y' = Ky/(T - t) + V_M$. For $t << T$, $y' = Ky/T + V_M$, whose solution is $y(t) = V_M T/K(1 - e^{-tK/T})$, which shows that the transverse separation decays on a time scale of T/K. Intercept requires a transverse velocity $y'(0) = V_M$. The exact solution,

Equation E.3
$$y'' = \left(\frac{V_M K}{T}\right)\left(1 - \frac{t}{T}\right)^{K-2}$$

has a maximum of $V_M K/T$ at $t = 0$ for $K > 2$ and falls monotonically to zero as t approaches T.

For a target with transverse acceleration A, the proportional acceleration is $y'' = KV_C\lambda' + A$, which for 0 initial LOS or velocity error integrates to $y' = KV_C\lambda + At$ and

Equation E.4
$$y'' = A\left(1 - \left(1 - \frac{t}{T}\right)^{K-2}\right)\frac{K}{K-2}$$

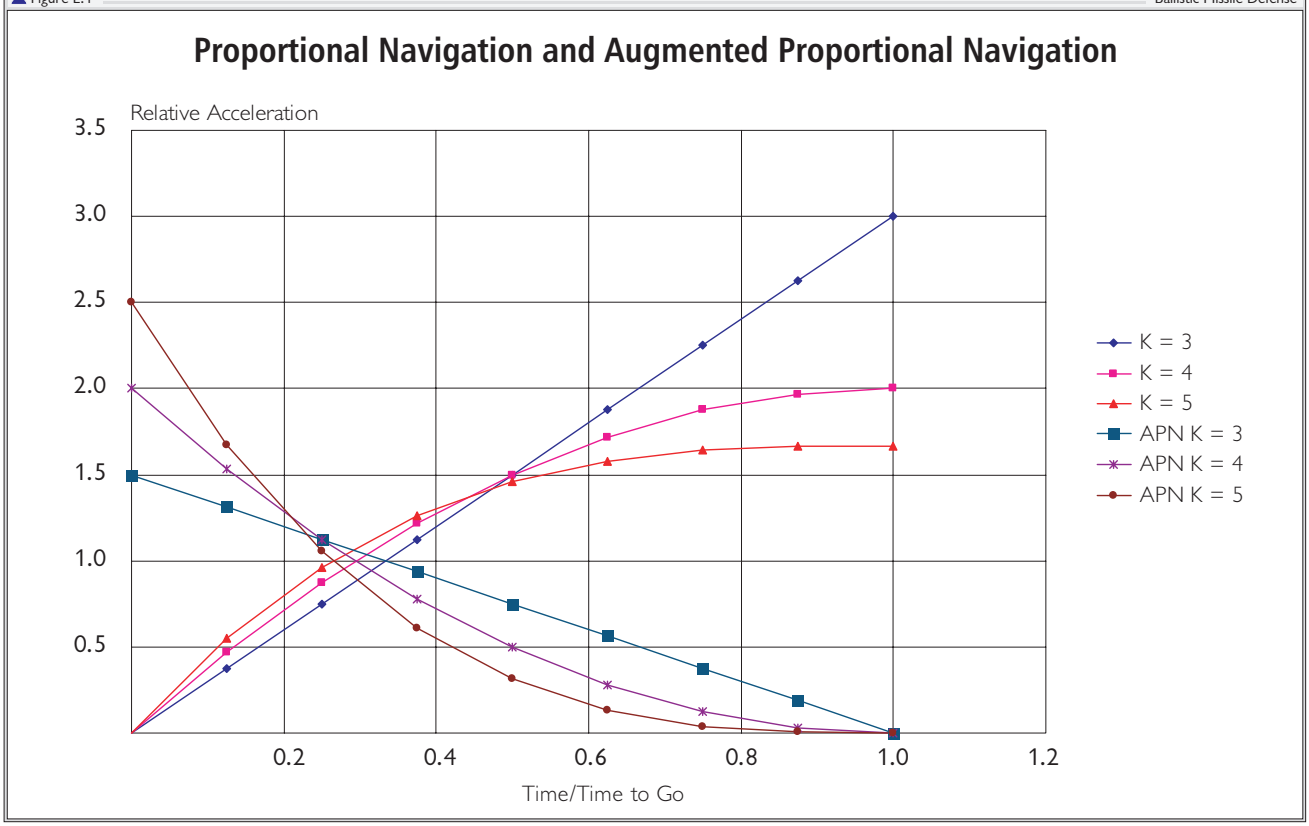
The relative acceleration is zero at $t = 0$ for all T and K, and has a maximum of $y''A = K/(K - 2)$ at $t = T$ for $K > 2$. Figure E.1 shows the PN relative acceleration $y''A$ as a function of t/T for various values of K by the curves that increase to the right, where the top curve is for $K = 3$. It has a maximum value of 3, which is the basis of the rule of thumb that an interceptor needs an acceleration about 3 times that of the missile to intercept accelerating targets. The second curve is for $K = 4$, whose maximum is 2; the third is for $K = 5$, whose maximum is 5/3. Their relative accelerations are about the same at $t = T/2$. For smaller t, large K has higher relative accelerations; for larger t, small K has higher relative accelerations.

For a missile that is accelerating transverse to the LOS, proportional navigation can be augmented to incorporate the missile's projected acceleration. Figure E.1 shows the accelerations required to intercept targets with acceleration A with augmented proportional navigation (APN), which has a maximum relative acceleration of $K/2$ at $t = 0$. At small t/T, the top curve is for $K = 5$, whose maximum acceleration is 2.5. For $K = 4$, it is 2, and for $K = 3$, it is 1.5. For typical values of K, the maximum accelerations for PN and APN, a_{PN} and a_{APN}, satisfy $a_{APN} = AK/2 - a_{PN}(K/2 - 1)$. Optimal guidance has smaller miss distances but requires similar accelerations.

The velocity increment needed to intercept accelerating targets using PN is $\Delta V_{PN} = KAT/(K - 1)$, so the relative velocity increment is $\Delta V_{PN}/AT = K/(K - 1)$, which is 2 for $K = 2$, 1.5 for $K = 3$, and 1.3 for $K = 4$. The velocity required for intercept with augmented proportional navigation is precisely half that.[148] For PN, higher navigation ratios decrease maximum relative acceleration and relative velocity increments. For APN, higher navigation ratios increase maximum relative acceleration but decrease velocity.

148. Zarchan, *Tactical and Strategic Missile Guidance*, p. 141.

Figure E.1 Ballistic Missile Defense



Appendix F
Directed Energy Weapon Scaling

A laser of optics diameter D and wavelength λ has an ideal beam divergence $\theta \approx \lambda/D$ = 3 microns/10 m = 0.3 microradians at a wavelength of 3 microns and a D = 10 m mirror. At range r = 3,000 km, that would produce a spot diameter $d = \theta r = 0.3 \times 10^{-6}$ rad x 3,000 km = 1 m. A laser power P = 3 MW would produce a flux F of about 3 MW/m^2. In 10 s, it would deposit a fluence J = 30 MJ/m^2, which is enough to melt unhardened structures. It could negate a missile of that hardness every 10 seconds, so a laser in space at that range could negate the roughly 1,000 missiles in the Soviet threat if launched over a period of two hours. The laser performance parameters can be collected together as

Equation F.1
$$\frac{J}{T} = F = \frac{P}{\left(\frac{r\lambda}{D}\right)^2} = P\frac{\left(\frac{D}{\lambda}\right)^2}{r^2} = \frac{B}{r^2}$$

where $B = P(D/\lambda)^2$ is the generalization of radar gain. The parameters above give an approximate brightness B of about 3 MW x (10 m/3 x 10^{-6} m)^2 = 3 x 10^{19} W/sr. By irradiating a target for time T, a laser deposits $J = BT$, which must be great enough to melt the target to structural failure. Thus, the brightness required for lethality at range r is

Equation F.2
$$B = \frac{Jr^2}{T}$$

For missiles launched from a point, the time to kill a missile at range r and switch to the next is $Jr^2/B + T_s$, where T_s is the switch time. This time must be averaged over missile range. The average density of satellites in orbit is $N'' = N/4\pi R_e^2$, so the average time is the integral of $N/4\pi R_e^2/(Jr^2/B + T_s)$ over the satellites in sight, which gives a kill rate $dM/dt \approx (N/R_e^2)B/J$. Equating it to the launch rate M/T gives

Equation F.3
$$\frac{M}{T} \approx \left(\frac{N}{R_e^{\,2}}\right)\frac{B}{J}$$

A constellation of 18 satellites of the above brightness and T_s = 1 s gives $M/T \approx 0.2$ kill/s, which could address 1,000 missiles launched over 500 s. Lasers also apply to rogue missiles. Their essential scaling is also $M/T \approx (N/R_e^2)B/J$, although it is more useful to rewrite it as

Equation F.4
$$BN \approx \frac{R_e^{\,2}JM}{T}$$

As for their launches, which are essentially from a point, it is the product of satellite number and brightness that matters, not their individual values. Rogue missiles are given credit for burn times T of about 250 s, hardness levels J of about 10^8 J/m^2, and the ability to launch about $M = 5$ missiles simultaneously. Meeting such an attack would require $BN = (6{,}400 \text{ km})^2 10^8 \text{ J/m}^2 \times 5/250 \text{ s} = 8 \times 10^{19}$ W/sr, which could be supplied by about 10 lasers with brightness about one-third of that estimated above, i.e., 3 MW, 5 m lasers. While the total costs for rogues would be reduced by about a factor of six from Soviet launches, their investment costs are high because the full constellation must be in place even for this small number of missiles, if they are fast and launched together. Rogue missiles could be hardened further, which would increase BN directly.

Appendix G
Space-Based Interceptor Scaling

Boost-phase interceptors maneuver in front of accelerating missiles before burnout so that the kinetic energy released as the missile runs over it is more than adequate to destroy the missile and cause its weapons and fragments to fall short of their targets. Space-based interceptors are deployed in orbit in advance of conflict to provide global coverage of launches and are minimally dependent on other sensors and systems.

Space-Based Interceptors. Space-based interceptors are small self-guided rockets pre-deployed in orbit in constellations large enough so that an adequate number would be within range of launch. They are assumed to have efficient, high velocity and acceleration engines and short release delays, so their constellations for strategic threats could be analyzed with geometric models. If a SBI can quickly reach its maximum speed V in a time short compared to the boost phase T of the missile that is its target, it can reach missiles within a distance $r = VT$ while they are still in boost. Each SBI can cover an area of πr^2, so the number of satellites required to cover the surface of the Earth, $4\pi R_e^2$, is

Equation G.1
$$N \approx \frac{4\pi R_e^{\,2}}{\pi r^2} = \left(\frac{2R_e}{r}\right)^2 = \left(\frac{2R_e}{VT}\right)^2$$

N is sensitive to reductions in T, although they could be offset by increases in V. Launches from distributed areas modify these calculations to take account of the SBIs over the launch area at the time of launch, which reduces the number of SBIs needed and the sensitivity of constellations to V and T. Soviet launch areas were distributed over much of the USSR, covering an area of an effective radius R of about 1,600 km. SBIs could fly in from a ring of width r around it, so the effective radius from which SBIs could participate in boost was $R + VT$. Soviet SS-9s had a boost-phase duration T of about 300 s, so for V = 6 km/s, r = 1,800 km and $R + r$ = 3,400 km, and the number of SBIs needed to address missiles in distributed launches during boost phase was

Equation G.2
$$N = \left(\frac{2R_e}{R+r}\right)^2 = \textit{14 per missile}$$

SS-9s accelerated for a boost time T of about 300 s and then deployed RVs for a bus time T_{bus} of about 300 s, during which they remained targets, although of decreasing value. The radius from which SBIs could fly in during busing was $R + VT_{bus}$ = 1,600 + 6 km/s x 600 s = 5,200 km. That gave an absentee ratio for intercepts of $N = (2 \times 6{,}400/5{,}200)^2 = 6$, which allowed 17 percent of the SBIs to engage. However, many of these intercepts would involve buses that had already offloaded some RVs. If each bus initially carried m_o RVs, every boost-phase intercept would kill m_o RVs, but between T and the end of bussing at T_{bus}, the number of RVs killed per successful intercept would fall with time. The SBIs added from T to T_{bus} from the ring between $R + VT$ and $R + VT_{bus}$ would number

Equation G.3

$$\Delta N_{bus} \approx N''\pi\left[(R+VT_{bus})^2 - (R+VT)^2\right]$$

where $N'' = N/4\pi R_e^2$ is the density of SBIs on orbit. For the conditions above, $R + VT_{bus} = 1.5(R + VT)$, so the number of engagements in the bus phase is about $1.5^2 - 1 = 1.35$ times that in boost. The number of RVs killed in the bus phase is the integral of $2\pi r dr N'' m(t)$, where m is the number of RVs left at time t, which is $m(t) = m_o(T_{bus} - t)/(T_{bus} - T)$

Equation G.4

$$K_{bus} = 2\pi N'' \sum_{T}^{T_{bus}} \frac{dtV(R+Vt)m_0(T_{bus}-t)}{dT}$$

$$= \frac{2\pi N'' m_0 V^2}{\Delta T}\left[\left(\frac{T_{bus}R}{V}\right)(T_{bus}-T) + \left(T_{bus} - \frac{R}{V}\right)\left(\frac{T_{bus}^2 - T^2}{2}\right) - \left(\frac{T_{bus}^3 - T^3}{3}\right)\right]$$

where $\Delta T = T_{bus} - T$. For the Soviet missile times and distributed launch area, $T = R/V = T_{bus}/2$, the number of SBIs from the ring is about 0.095 x N, and the number of RVs they kill is 0.44 x N. Their ratio $4.6 = m_o/2$ is the average number of RVs killed per intercept during busing, which indicates the kills during busing were roughly half as effective as kills during boost for the conditions of Soviet launches. Consistent with that, for the Soviet distributed launch area, the total number of RV kills in the boost and bus phases was ≈ 1.6 times the number of kills in boost phase alone.

Brilliant Pebbles. BP was a specific realization of SBI based on the interceptor technology available at a particular point in time. To be effective, BP had to get the maximum performance possible from a given level of available structure, propulsion, sensor, and guidance technology. The primary departure from the ideal scaling assumed for the SBIs above was in propulsion. The small, efficient, high *Isp* engines then available did not have the high velocities, accelerations, and mass fractions assumed by the SBI. Fortunately, for the distributed Soviet launch areas then of concern, these non-idealities did not compromise performance excessively, although they did have to be taken into account for accurate scaling estimates.

Figure G.1 shows the progression of sensors used by the BP during its detection, track, and intercept. BP detected missiles in boost with wide field of view (WFOV = 20°) mid wavelength infrared (MWIR = 3–5 microns) sensors, which from a typical range of r = 1,000 km gave it a ground footprint of 300 km, which it raster scanned over the area passing below. The MWIR sensors had 256 detectors, so their ground resolution was about 1 km, about that of DSP. BP then shifted to its 5° FOV UV-visible camera to guide it toward the plume. The UV-visible sensor's resolution was initially several hundred meters. It dropped to about 10 m by 10 s before impact. Then the BP shifted to a 1° FOV LWIR (10 microns) imager and a 1° FOV lidar to separate the missile hard body from its plume. At 10 s before impact, the LWIR sensor only had one pixel over the missile, so it could not resolve it, but by 1 s it had 10 pixels, so it could support imaging and aim point selection. It used a 0.1° FOV lidar to provide range information and backup for aim point selection for targets that were dim in the LWIR.

For a BP of average acceleration A, the time to reach speed V is $T_{accel} = V/A$, which was about 6 km/s/ 0.04 km/s^2 = 150 s. During this time, its average velocity was $V/2$, which reduced BP range by $T_{accel}V/2 = V^2/2A$ = (6 km/s)^2/2 x 0.04 km/s^2 = 450 km. For a release delay of 30 s, its effective range was about $V(T_{flight} - T_{accel}/2)$ = 6 km/s x (570 s – 75 s) = 3,000 km to a SS-18 bus and 6 km/s x (270 – 75 s) = 1,200 km to its booster. These range reductions of $(T_{accel}/2)/T_{flight}$ = 75 s / 540 s = 14 percent for the bus and 75 s/270 s = 28 percent for booster were significant but not debilitating for SS-18s.

Figure G.2 shows how BP range varied with V and A for a 30 s launch delay. For the initial V = 4 km/s BP, the maximum range to a SS-18 booster was ≈ 800–1,000 km, weakly dependent on acceleration. For the V = 6 km/s available with today's technology, ranges increase to 1,200 to 1,400 km for

accelerations of 4 to 8 g. For the $V \approx 8$ km/s possible with advanced technology, ranges increase to 1,300 to 1,750 km. Figure G.3 shows how V and A affect constellation size. For the distributed Soviet launch area, their impact on constellation size is less than that on range. At $V = 4$ km/s, the constellations require about 25 BPs per missile; at 6 km/s about 18–22; and at 8 km/s about 15 for 6 g and 8 g and 18 for 4 g. To negate 600 Soviet missiles with 4 km/s BP by engaging in boost would take about 600 x 25 = 15,000 BPs. It would take about 15,000/1.6 = 9,400 in boost and bus phases. With 6 km/s and 6 g it would take about 12,000 in boost or 7,500 in boost and bus.

BP liquid fuel engines had a specific impulse I_{sp} of about 300 s, i.e., an effective exhaust velocity c of $gI_{sp} = 3$ km/s, for which the rocket equation estimates the initial mass required to accelerate a $m = 4$ kg KV to 6 km/s at $M = me^{V/c} = me^2 = 7.4m = 30$ kg, although real engine performance could double it. KV cost C_K was estimated to be about $500,000. The cost to orbit C_V was about $20,000/kg, so the estimated total cost of a KV on orbit was $C = C_K + C_V M$. The total cost on orbit for a 6 km/s BP was about $500,000 + $20,000/kg x 30 kg ($500,000 + $600,000) = $1.1 million. A zero-speed BP would cost around $500,000 + $20,000/kg x 4 kg = $580,000, but would require many BPs. A 9 km/s BP would cost about $750,000 + $20,000/kg x 80 kg = $2.1 million, which is more than double the nominal design.

Figure G.4 shows constellation costs per ICBM in boost as functions of V and A. For large distributed launch areas, costs are insensitive to A and $V < 8$ km/s, although they would increase faster than these scaling models at large V. For $V = 4$ to 6 km/s and $a = 4$ to 8 g, the cost is about $0.02 billion/missile. Thus, engaging all 1,000 Soviet missiles in boost phase would cost about $0.02 billion/missile x 1,000 missiles = $20 billion. Engaging them in boost and bus would cost about $20 billion/1.6 = $12.5 billion.

BPs unable to reach the boost phase could engage in midcourse, where they would have lower absentee ratios but would face the same countermeasure and discrimination issues as other midcourse system. Boost-phase intercepts could be made by BP from radii less than $R_b = VT + R_{eff}$, where $R_{eff} = R_{Launch} - V(T_{delay} - V/2A)$ is the launch area radius, corrected for delays and finite acceleration. Bus-phase intercepts can be made by BP out to $R_{bus} = VT_{bus} + R_{eff}$. BP out to $R_{target} = V(T_{flight} - T_{delay} - V/2A)$ from the target can engage in midcourse or terminal intercepts. The fraction of the BP constellation in range is $f_{target} = [1 - \cos(R_{target}/R_e)]$, so there are about Nf_{target} BPs to engage the $m_o M_b - K_{boost} - K_{bus}$ RVs that leak through boost and bus-phase defenses. For intercontinental launches, T is about 1,800 s, so $VT_{flight} = 11{,}000$ km, and $f_{mid} = 56$ percent, and a large fraction of the BP constellation can contribute.

Figure G.5 shows the number of kills in each phase as a function of the number N of BPs for 300 Russian heavy SS-18 with 10 RV each in distributed launch areas for BP with 10 g acceleration and 6 km/s maximum speed. Boost-phase kills K_{boost} increase linearly with N to about 1,000 RVs at $N = 2{,}000$, for an absentee ratio of about 2,000/(1000 RV/10 RV/missile) = 20, in agreement with above estimates. The number of kills in the bus phase is about 70 percent larger. Their total is about 2,750 at $N = 2{,}000$. The curve slanting down and to the right is the number of RVs that leak through the boost and bus phase, which is large at $N = 250$, but approaches zero at $N = 2{,}200$. The bottom curve is the number of RV kills in midcourse, assuming BP have the same 0.9 kill probability in midcourse as in other phases. It lies on top of the K_{boost} curve for $N < 1{,}600$ and then turns over and falls with the leakage line as fewer RVs leak through boost and bus to be killed in midcourse or terminal. At $N = 1{,}600$ BPs, the boost, bus, and midcourse phases contribute roughly equal numbers of kills. The leakage through boost and bus is about 750 RVs, but enough BPs are available in midcourse to reduce the number of RVs leaking through all three layers to about 70. By $N = 2{,}000$, the number drops to zero.

Figure G.6 shows the number of RV kills for 20 missiles with 10 RVs apiece, which could represent a single SSBN on station, in bastion, or in port. Because the launch area is essentially a point, K_{boost} is only about 20 at $N = 180$. However, the number of kills in the boost and bus phases is still apprecia-

ble, approaching 180. Those two phases could kill all 200 RVs by $N = 200$, for an absentee ratio of 200/20 = 10, as reducing the boost phase produces more opportunities for kills in the bus phase. At $N = 120$, where K_{mid} peaks, the kills in bus phase are about 7.5 times those in boost. The kills in each layer are 14, 104, and 60. About 80 RVs leak through boost and bus, and 20 through all three layers. The leverage of boost-phase intercepts falls as the number of MIRVs or launch area falls.

Figure G.7 shows the variation of boost and midcourse kills for the launch of 100 single-weapon, 150 s burn time mobile missiles like Soviet SS-25s. Because their launch area is distributed, its boost-phase kills are strongly reduced. And because it is a single RV weapon with no bus, the only other contribution is midcourse. Boost phase contributes about 100 kills for an absentee ratio of about 4,500/100 = 45:1, twice that for distributed, MIRVed missiles. The 100 SS-25 missiles could be addressed in boost, but it would be expensive.

Figure G.1 Ballistic Missile Defense

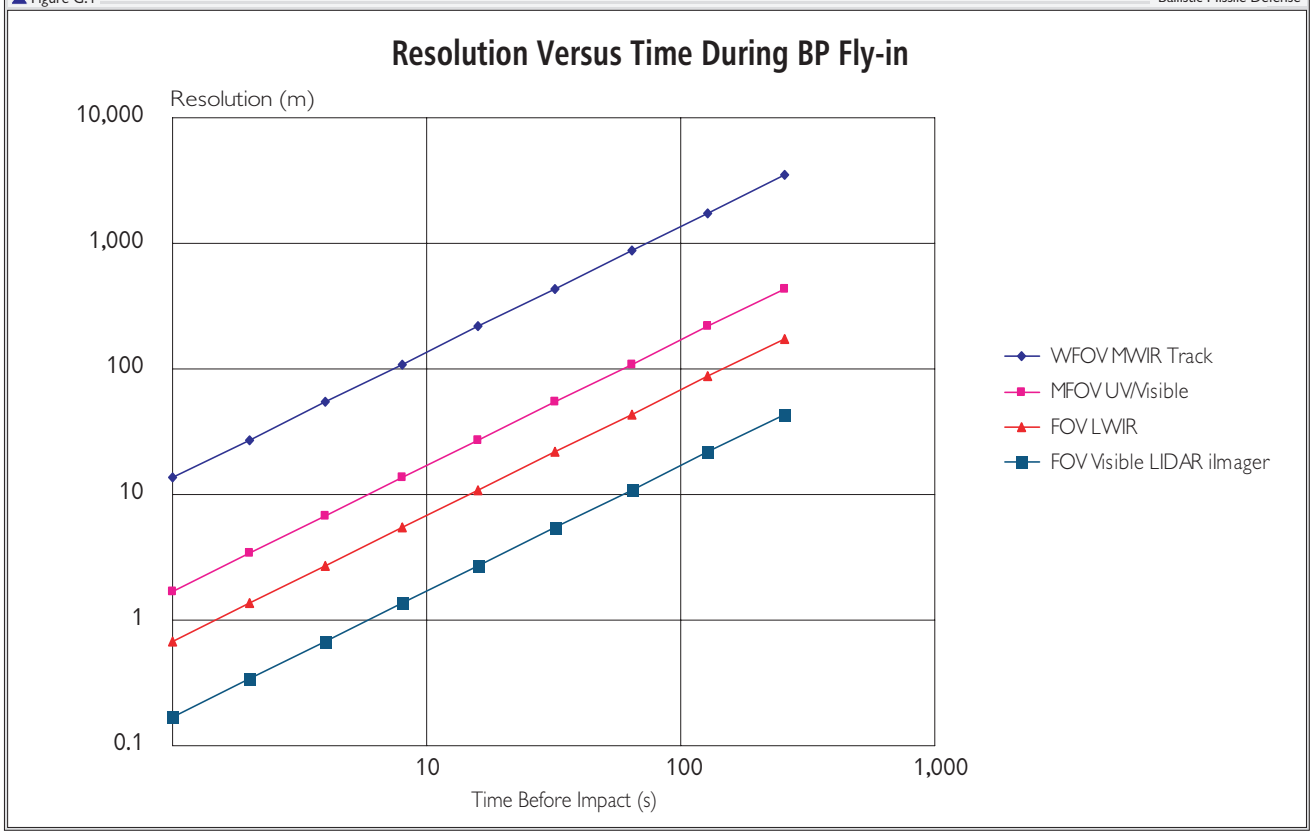


Figure G.2 Ballistic Missile Defense

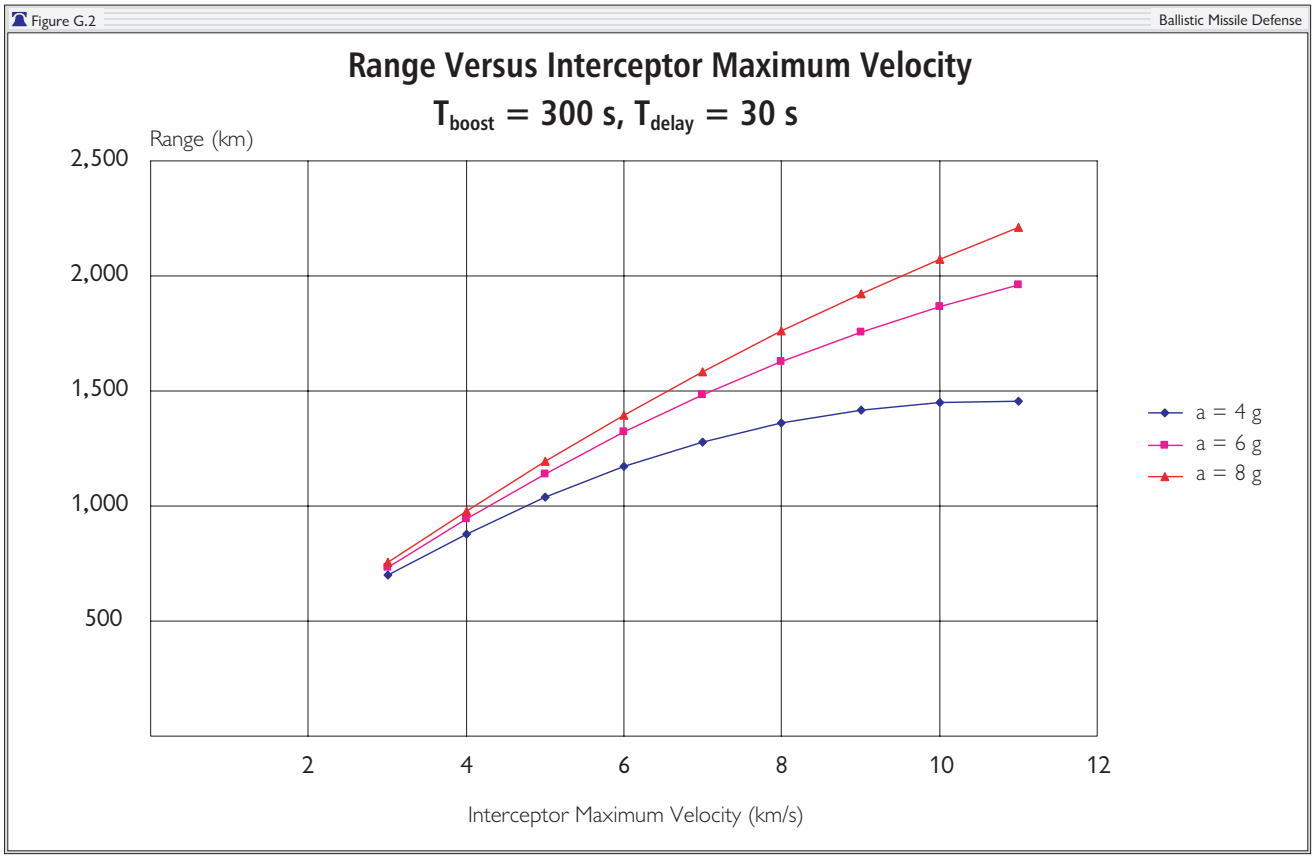


Figure G.3 Ballistic Missile Defense

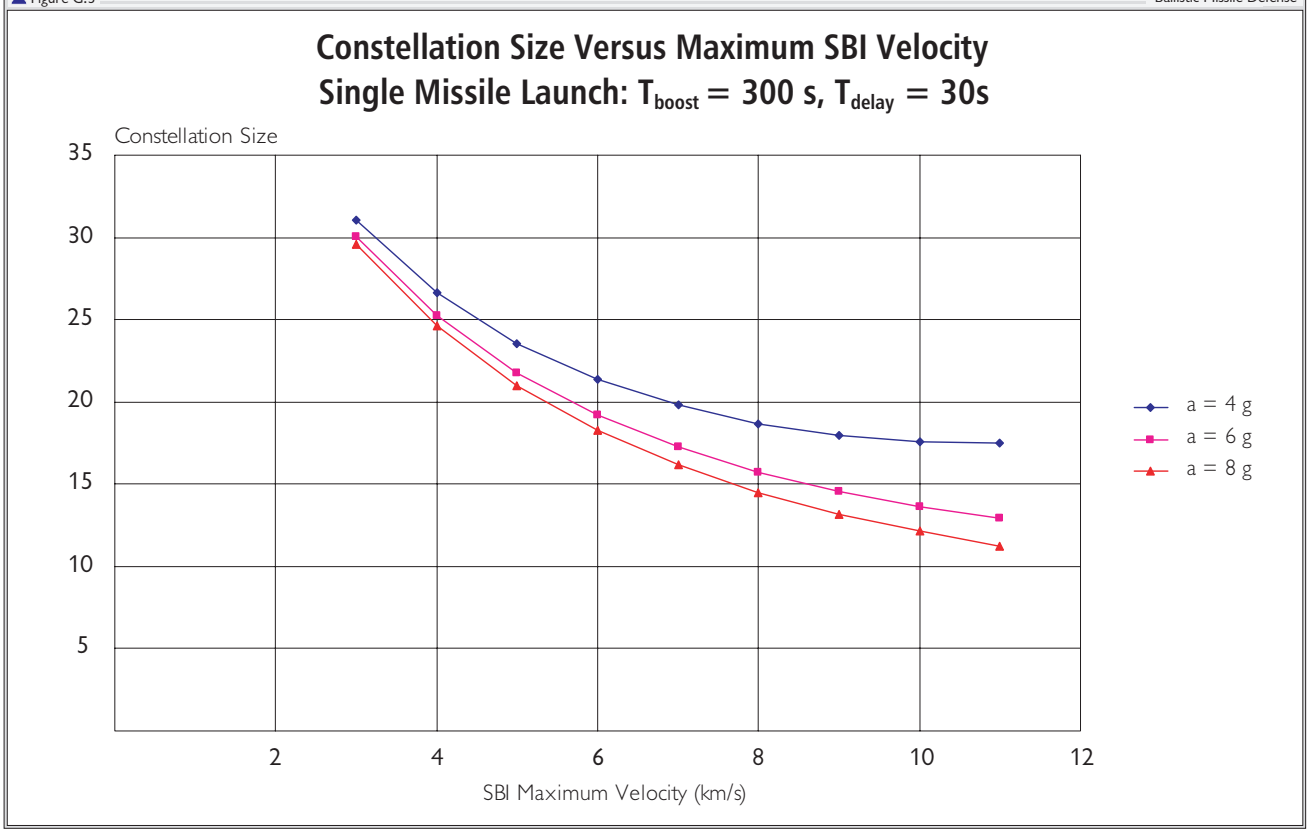


Figure G.4 Ballistic Missile Defense

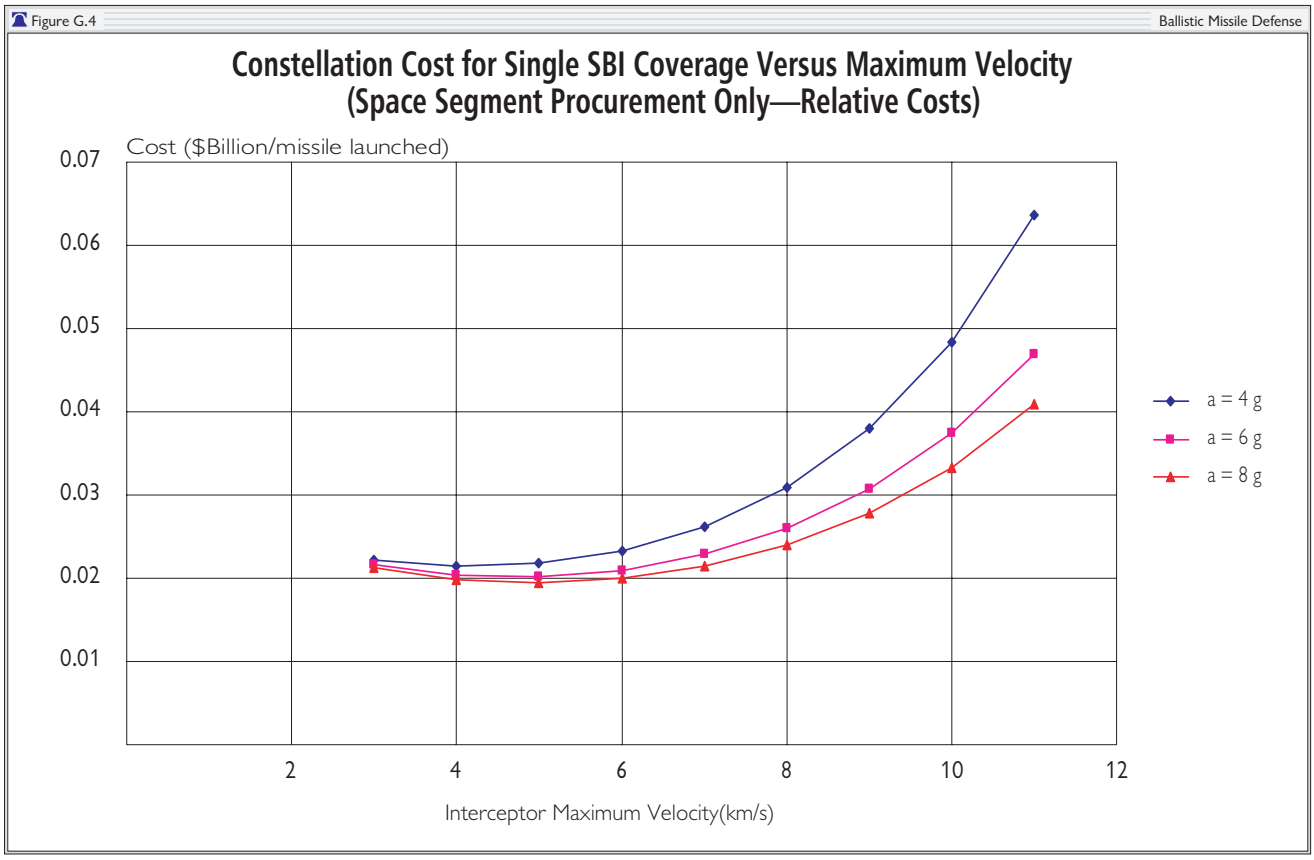


Figure G.5 Ballistic Missile Defense

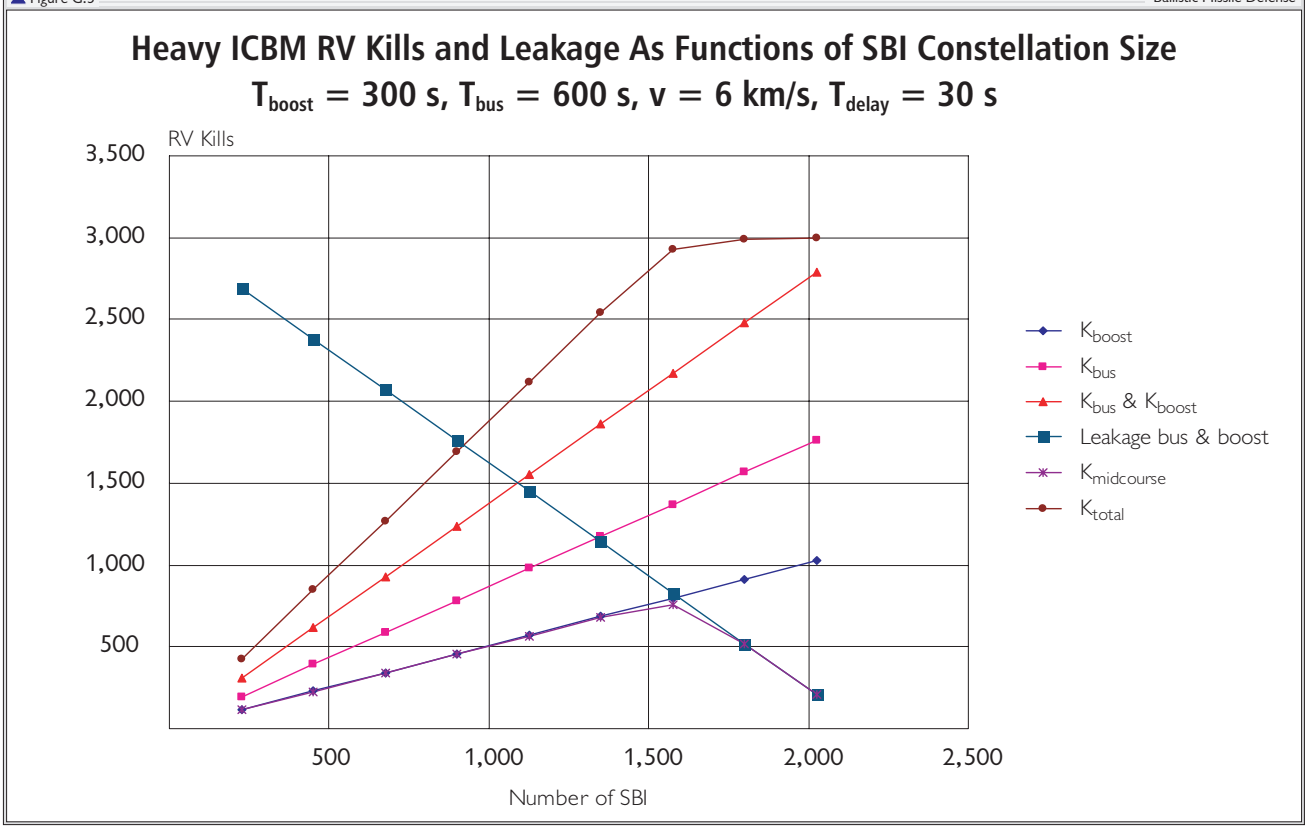


Figure G.6 Ballistic Missile Defense

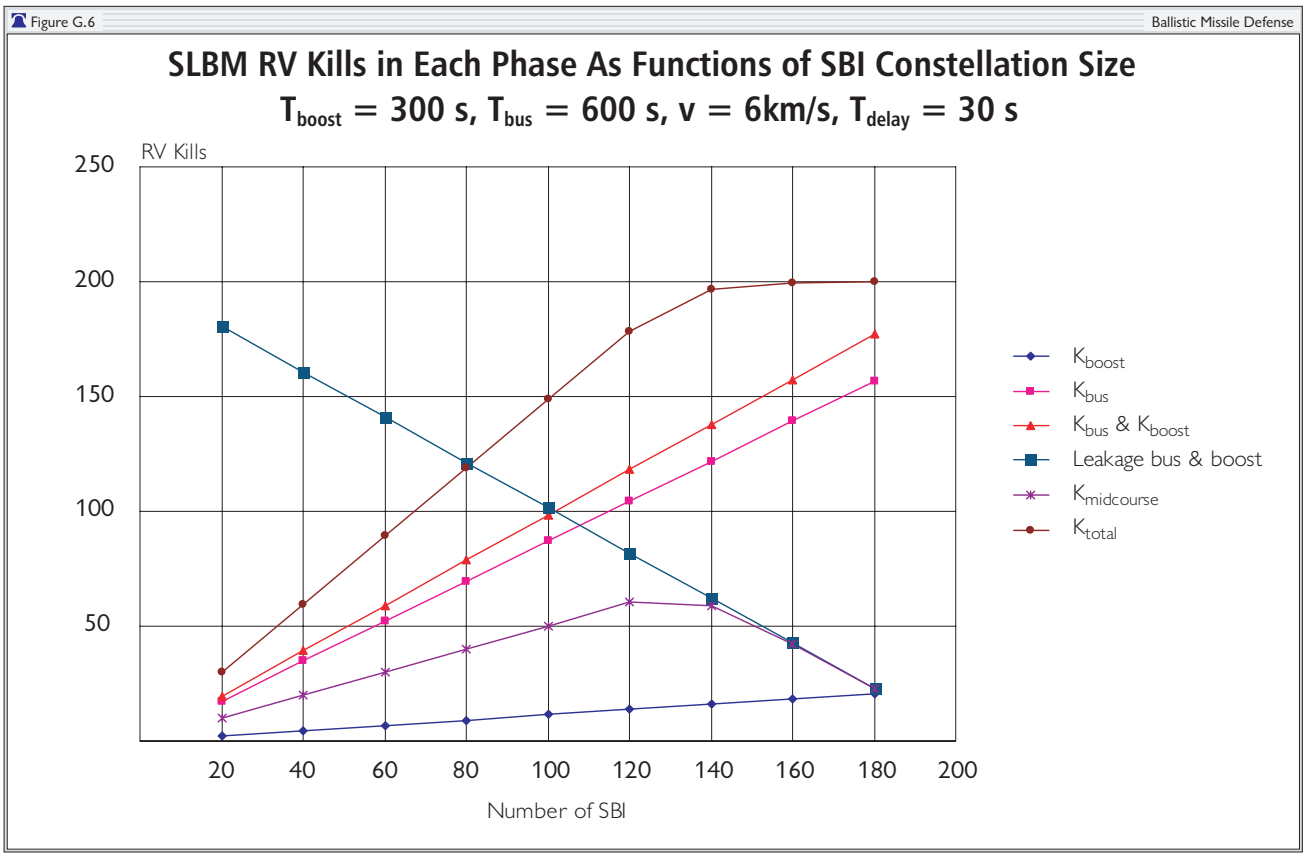
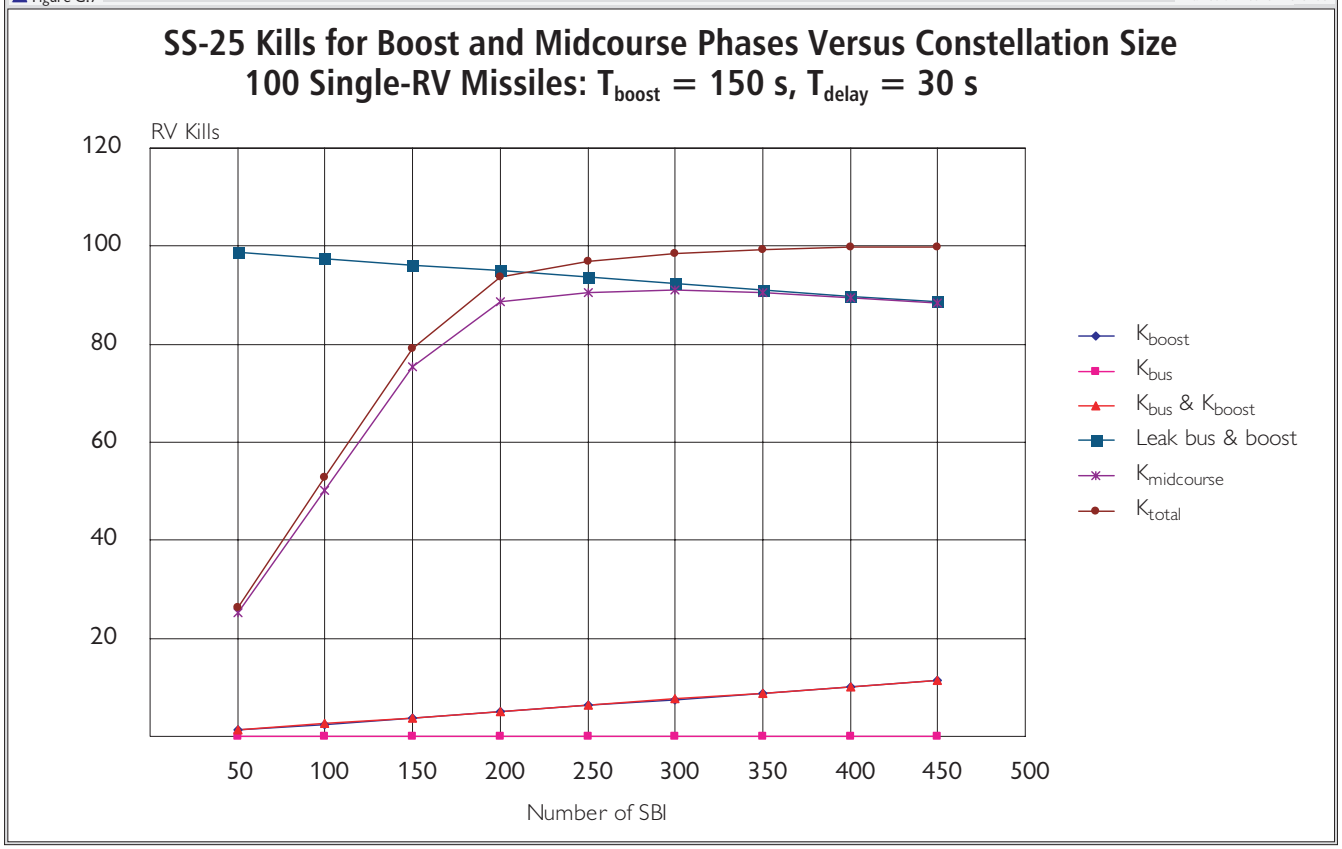


Figure G.7 Ballistic Missile Defense



Appendix H
Survivability

A nuclear explosion of yield Y at distance K produces a fluence $J = Y/4\pi K^2$; thus, to survive by hardening alone with material of J' energy withstood per unit area per unit thickness would require a thickness $\Delta \approx J/J' \approx Y/4\pi J'K^2$. A BP with fuel has a mass M of about 50 kg, a density ρ of 500 kg/m^3, and hence an effective length D of about $(M/\rho)^{1/3}$, which is typically about 0.5 m, and an area A of about D^2, which is typically about 0.3 m^2. If the hardening material is applied to the whole surface, the mass of hardening required is about

Equation H.1

$$M_{hard} = \rho D^2 \Delta = \frac{\rho^{\frac{1}{3}} M^{\frac{2}{3}} Y}{J'K^2} = \frac{CM^{\frac{2}{3}}}{K^2}$$

where $C = \rho^{1/3}Y/J'$ is about 10^9 J/m^3 for $Y = 20$ KT and $J' = 10^9$ J/m^3. The BP can divert to make K larger, but that requires fuel and hence mass. To miss the ASAT by distance K through a divert executed when the ASAT is at range R, the BP must deflect from its initial velocity V by an angle $dV/V = K/R$, which requires fuel of $dM/M = dV/c$, so the fuel for successful divert is $dM = M(dV/V)(V/c) = 2MK/R$. BP could reduce the fuel by maneuvering at larger R, but that would give the ASAT time to correct, which would reduce K.

The BP could deploy N decoys of mass m and maneuver enough to hide itself in them. Spacing the decoys about K apart would force the ASAT to treat each as a valid target and expend a nuclear ASAT on it. The overall diameter of the cloud is about $N^{1/3}K$, so the mass to maneuver the BP and decoys M_{man} is about $2N^{1/3}(M + mN)K/R$. The total mass for BP, decoys, hardening, and maneuver M_{def} is about

Equation H.2

$$M + mN + M_{hard} + M_{man} \approx M + mN + \frac{CM^{\frac{2}{3}}}{K^2} + \frac{N^{\frac{1}{3}}(M + mN)K}{R}$$

As all terms but the last involve the BP mass, it is clear that the defender's cost is minimized by reducing M. M_{def} also falls with N, but the attacker's mass is about $M_{att} \approx (1 + N)M_{wpn}/E$, which increases with N. The defender's goal is to make the ratio of M_{att} to M_{def} large, which occurs for large N. If the attacker must commit an ASAT to each credible object, that requires a total mass $(1 + N)M_{wpn}$. The effective attack mass is less by the absentee ratio E, because each time the attacker negates a BP overhead it also negates E more elsewhere in orbit, unless the attacker waits until new BPs rotate into view.

Figure H.1 shows the attack and defense masses and their ratio as functions of N for $K = 1$ km, $R = 300$ km, $m = 0.5$ kg, and $E = 10$. The bottom curve is M_{man}. For small N, the decoy mass M_{decoy} is slightly greater than M_{man}; for large N it is an order of magnitude greater. M_{hard} is an order of magnitude larger than them at small N, but it does not vary with N, so the decoy mass reaches it at $N = 30$. The mass of the wet BP M_{bp} is dominant at small N, but the decoy mass exceeds it by $N = 100$, so the total defense mass starts to increase there. The attack mass M_{att} increases with N. It is less that the total defense mass M_{def} for $N < 5$. The ratio of the attack mass to the defense mass reaches about 4 by

$N = 30$ and 10 by $M = 100$. For larger N, the ratio begins to saturate. For $1 << N << M/m$ light decoys, the ratio of attack to defense mass is $M_{att}/M_{def} \approx NM_{wpn}/E(M + mN)$, so the defender should use about M/m light decoys, maneuver at long range, and keep absentee ratios low. The attacker should attempt to discriminate, minimize miss distance, and use small yields.

Figure H.2 shows how the components of defense mass and attack mass M_{att} vary with satellite mass M for the conditions above and $N = 30$ decoys. The attack mass is about 300 kg for all M. The main defense mass is decoys M_{decoy}. Hardening M_{hard} is second at small M, but becomes the dominant element by $M = 50$ kg. Maneuver M_{man} is an order of magnitude smaller. The total defense mass M_{def} increases from 20 kg at $M = 1$ kg to about 1,400 kg at $M = 1{,}300$ kg. For small M, the attack/defense mass ratio favors the defense by over an order of magnitude. The ratio is unity at $M = 300$ kg and is about 0.2 by $M = 1{,}000$ km.

Figure H.1 Ballistic Missile Defense

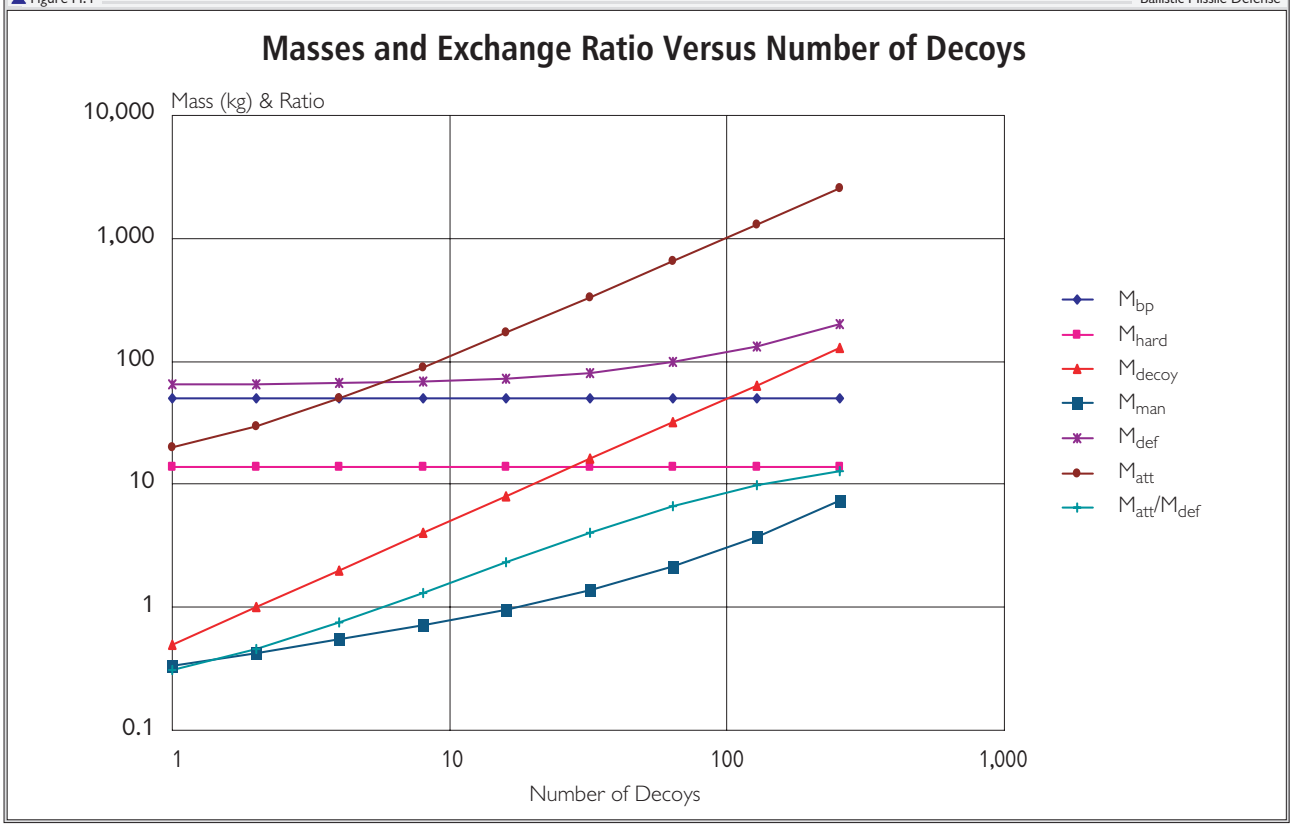
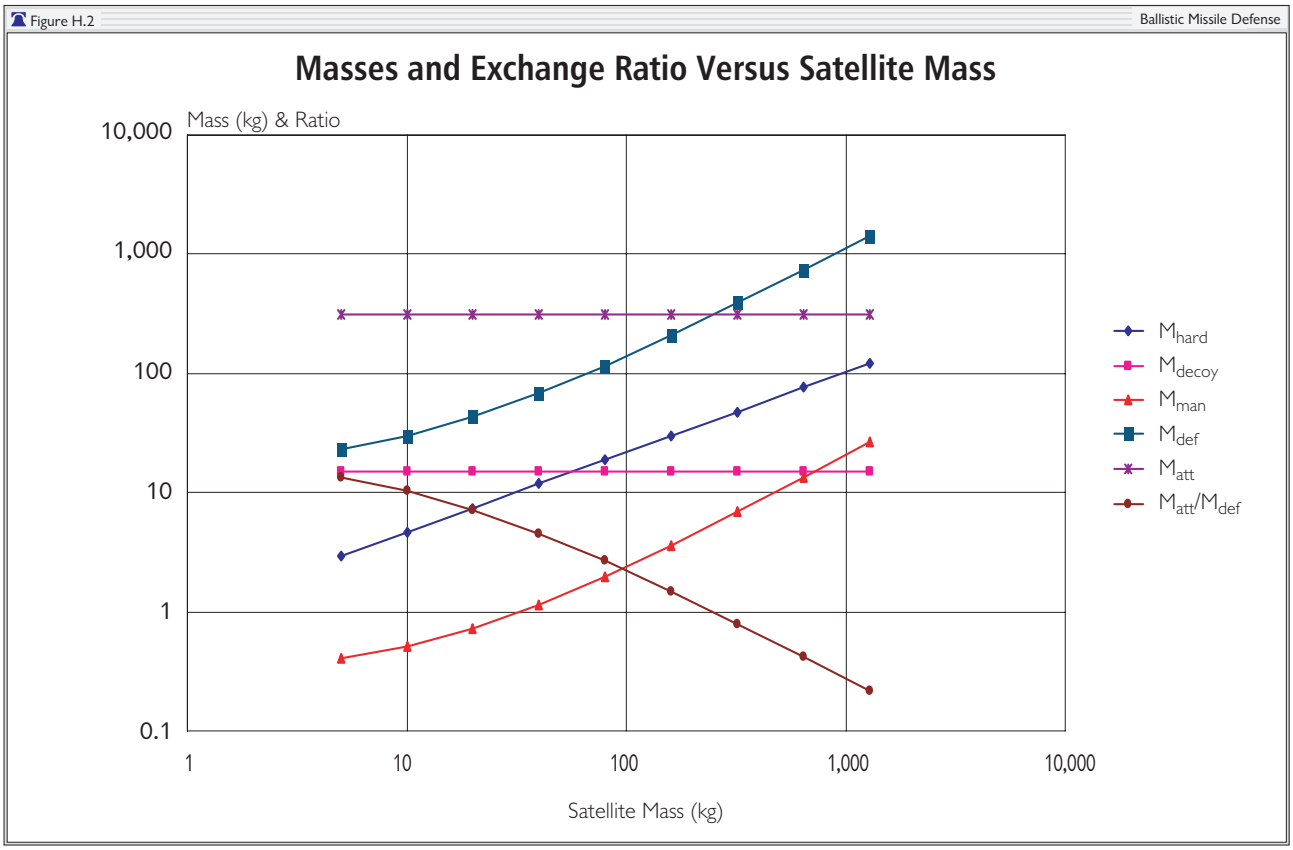


Figure H.2 Ballistic Missile Defense



Appendix I
Sensor Satellite Scaling

Surveillance satellites do not need to image; tracking satellites may need to image. This discussion provides a unified treatment of the main tradeoffs in their design. While the optics of such satellites are not necessarily designed to diffraction limits, they follow it sufficiently closely that the scaling arguments used in other sections give the approximate sizes and masses of elements well enough to discuss overall trends. An aperture of diameter D at wavelength λ produces beam divergence of about λ/D, which gives a resolution d of about $r\lambda/D$ at range r. If performing its function requires the sensor to produce resolution d, the aperture must have diameter $D = r\lambda/d$. To resolve 1 m RVs and decoys at λ = 10 microns from a range r of about 1,000 km would require $D = 10^6$ m x 10^{-5} m/1 m = 10 m, which would be prohibitively expensive. Conversely, a D = 0.3 m aperture would give $d = r\lambda/D = 10^6$ m x 10^{-5} m/0.3 m = 30 m at 1,000 km and about 100 m at the 3,000 km ranges that reduce constellations to "manageable" sizes and costs, which might be affordable but might not distinguish objects.

Figure I.1 illustrates those trades for sensors with D = 0.1, 0.3, and 1 m at λ = 10 micron for various ranges. A resolution of 1 m may be excessive, but 100 m cannot separate closely spaced objects, so an intermediate resolution of 10 m might be useful. That resolution could only be achieved by a 0.1 m sensor from a range of 100 m, but that is precisely where ground- and space-based intercepts occur. They do view clouds of objects from such distances. From about 300 km, a 0.3 m sensor could achieve 10 m resolution. That was one of the configurations studied by Brilliant Eyes. However, it would require several hundred satellites, so budget and judgmental factors drove it up and to the right while maintaining that aperture, which ultimately produced a SBIRS-Low with that aperture at 3,000 km range, which could only resolve 100 m. From that range it would take 1 m sensors to get back to a few meters resolution.

An aperture D has area D^2, to which the satellite area is proportional; thus, satellite volume is proportional to D^3. Satellite weight and cost are roughly proportional to volume, so its cost is approximately D^3. For the scaling discussed above, that implies cost $\propto (r\lambda/d)^3$, so that the cost of resolution scales as $\propto 1/d^3$, although the cost would remain fixed if range decreased proportionally. The number of satellites N needed to tile the Earth with average range r is about $(2R_e/r)^2$. The cost of a constellation is proportional to the product of the cost per satellite and the number of satellites in it, so

Equation I.1
$$Cost \propto \left(\frac{2R_e}{r}\right)^2 \left(\frac{r\lambda}{d}\right)^3 \propto \left(\frac{\lambda}{d}\right)^3 r$$

which is minimized for small r, i.e., large numbers of small satellites at short ranges, which was the approach followed in the Brilliant Eyes derived from BP. Figure I.2 illustrates these trades as functions of range for a required resolution of 10 m. The sensor diameter at the bottom runs from centimeters at ranges of 10 km to several meters at 3,000 km. The sensor mass increases from less than 1 kg to 10 tons. The BP sensor, which began imaging at about 100 km, had a roughly a 0.15 cm aperture and 1 kg mass. The number of satellites drops from 10,000 at 100 km to 30 at 3,000 km. The total BE mass, which is the product of the number of satellites and the sensor mass of each, increases from one ton at 10 km to about 500 tons at 3,000 km. This suggests that reducing the range from 3,000 to 300

km and increasing the number of satellites from 30 to 1,000 could reduce mass on orbit by a factor of about 30.

In its surveillance mode, the LEO BE could collect target signals more efficiently, so that it could get the same radiometrics as a 0.3 m aperture at GEO from a (1 mm/30 mm) x 0.3 m = 1 cm aperture at LEO, which is why BP could perform detection for itself using on-board sensors of roughly that size. In a tracking function, a BE constellation at ranges of 300 km should have a cost advantage of about a factor of 10 over a smaller constellation of larger satellites for any given resolution, which could be used to provide better resolution by about a factor of 2. Because of the large number of satellites, BE would have more opportunities for simultaneous observations than larger satellites. Iridium and Teledesic have demonstrated the practicality of constellations approaching this level.

Figure I.1 Ballistic Missile Defense

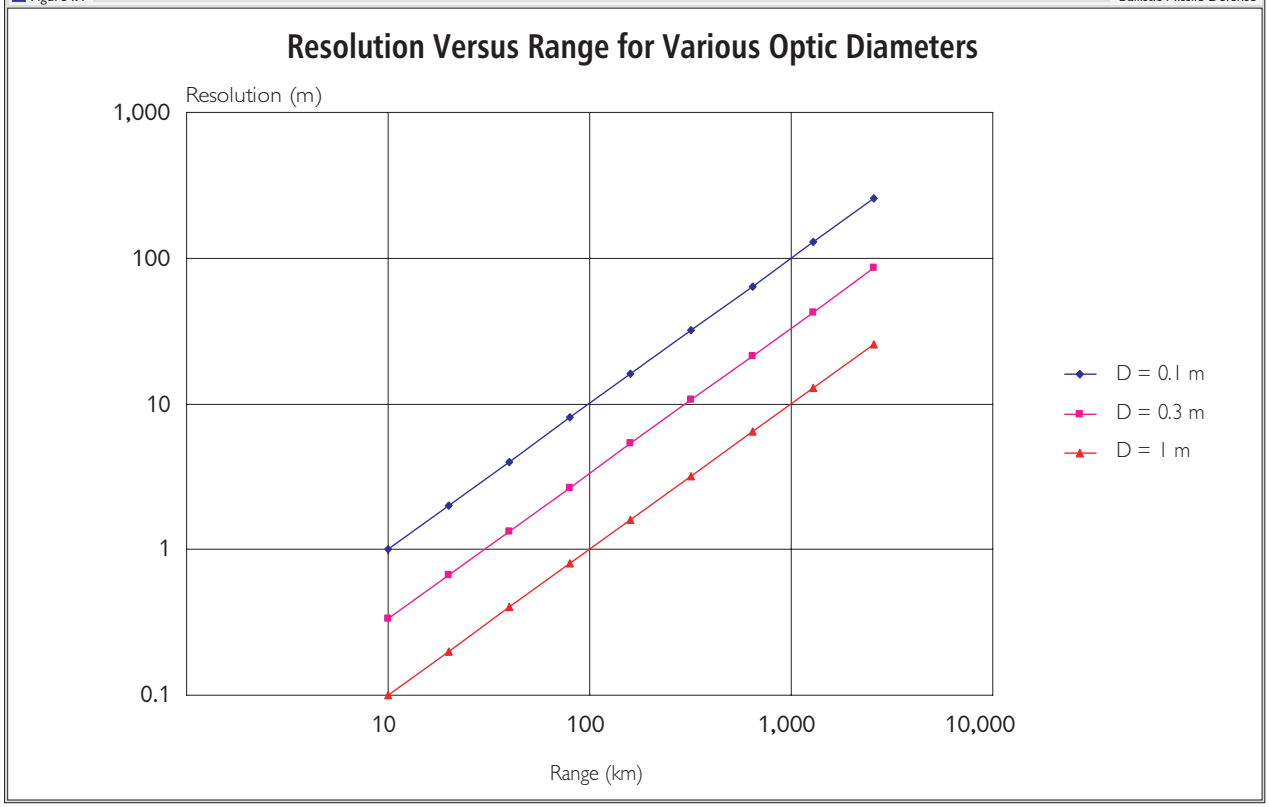
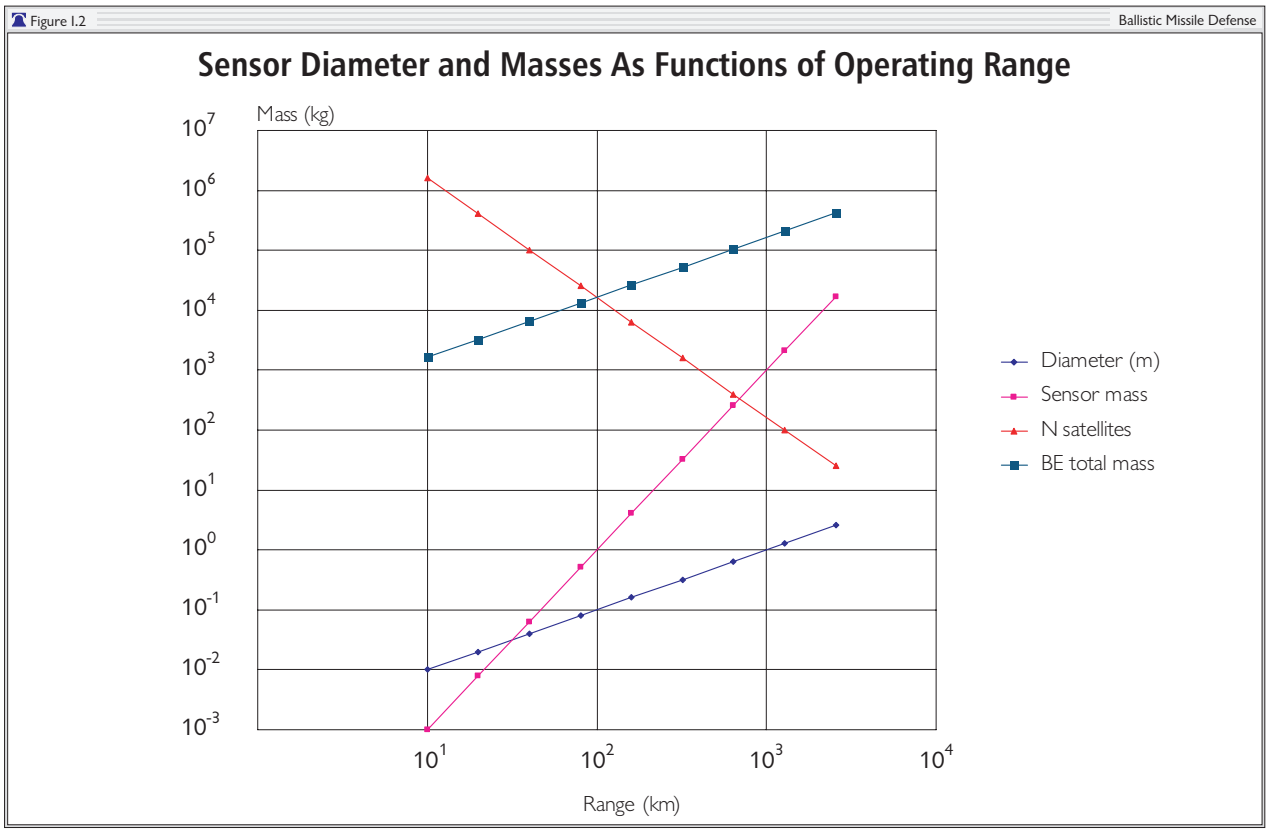


Figure I.2 Ballistic Missile Defense



Appendix J
Surface-Based Intercepts of Theater, Terminal, and Regional Missiles

Intercepting theater and regional missiles in boost requires quick reaction and favorable geometry. The interceptor velocities required increase with delay times. For typical missile and interceptor parameters, delay times over 60 s are unacceptable. Such release times would be difficult with current satellite sensors, but might be possible with organic or forward-deployed radars. Post-boost intercepts are less sensitive to delays, but less valuable and more sensitive to countermeasures. Without delays, nominal interceptors have very large footprints (See Figure J.1), but modest reductions in velocity (See Figure J.2) or increases in launch delays reduce footprints by similar factors, which would largely eliminate the advantages of more capable interceptors.

Theater missiles execute a gravity turn to roll over into an optimal flyout angle of about 45° with respect to the horizontal. Its reentry angle is about the same. If it is detected at a distance D from the interceptor launcher, its time to arrive is approximately $D/V \cos i$, where i is its reentry angle and its velocity is $v = gR$. At a similar level of approximation the interceptor can be treated as being launched at delay time T_{delay} after the missile, having an average acceleration a of about 6 g and reaching a maximum velocity v after an acceleration time T_A of $v/2a$. Thus, the maximum range the interceptor can reach in time T is

Equation J.1
$$r = \frac{v\left(T - T_{delay} - T_A\right)}{\sqrt{2}}$$

for $T > T_{delay} + T_A$, and $r = a(T - T_{delay})^2/2$ for shorter times. To reach maximum range, the interceptors are fired at an angle of 45° with respect to horizontal. The hit condition is that sum of the interceptor's maximum range and the distance the missile flies to land there must sum to the detection distance, or

Equation J.2
$$\frac{v\left(T - T_{delay} - T_A\right)}{\sqrt{2}} = \frac{D}{V \cos i}$$

for shorter delays and

Equation J.3
$$\frac{a\left(T - T_{delay}\right)^2}{2} = \frac{D}{V \cos i}$$

for longer ones. Whichever equation satisfies the condition on $T > T_{delay} + T_A$ determines the time of intercept and hence the interceptor range to the forward edge of its defended footprint.

For ICBMs it is only necessary to substitute the optimal reentry angle of $i = 22.4°$ to find the forward extent of the interceptor's footprint.

Sea-based systems can also use that analysis to determine their forward footprint, but it is not useful as the area covered is ocean. However, they can take advantage of the defended footprints they generate behind them to cover much larger areas. If the missile is detected at a distance D from the launcher and shore, the times for the interceptor to reach its maximum range inland, $r/(v/2)$ and that for the missile to reach there must be the same, which gives the relationship

Equation J.4

$$\frac{r}{\left(\frac{v}{\sqrt{2}}\right)} = \frac{D + r}{V \cos i}$$

where r is given by Equation J.2 for short delays and Equation J.3 for long delays. Either can be solve algebraically to determine the maximum backward extent of the interceptor coverage shown in Figures J.1 through J.7.

Figure J.1 Ballistic Missile Defense

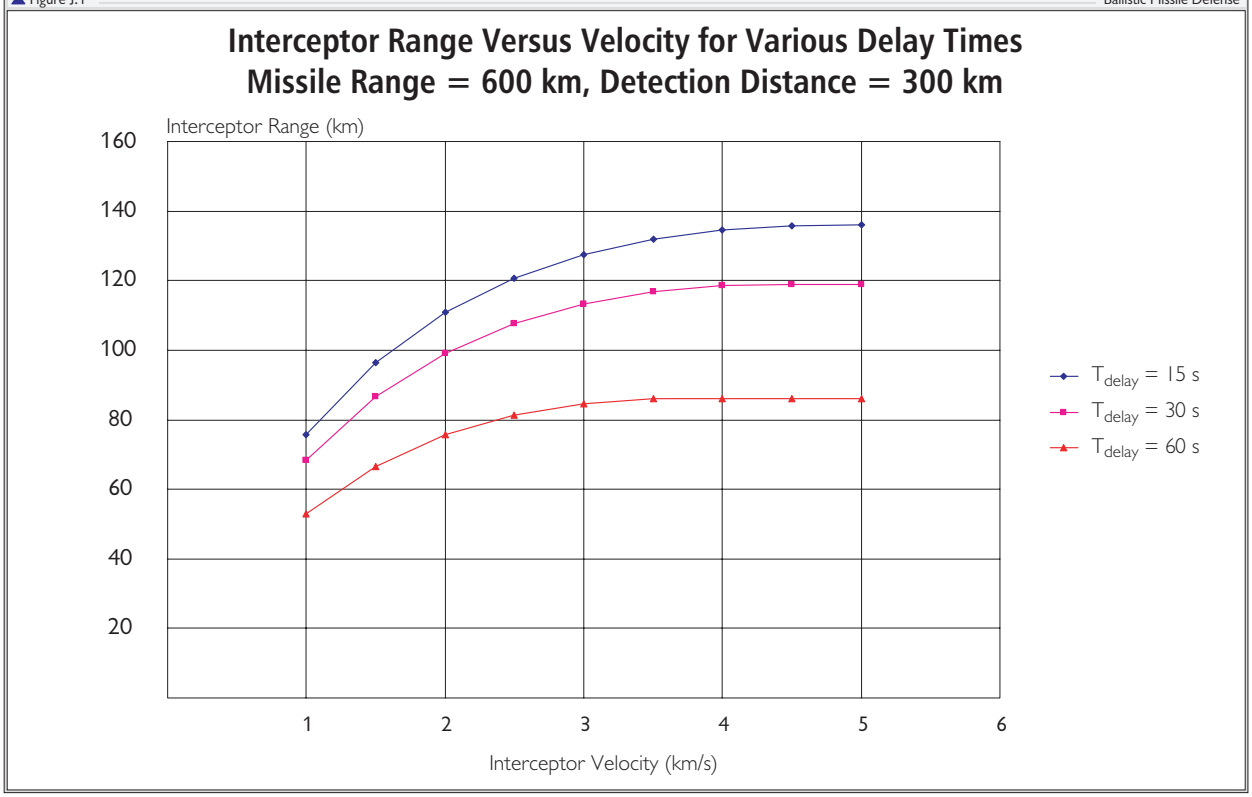


Figure J.2 Ballistic Missile Defense

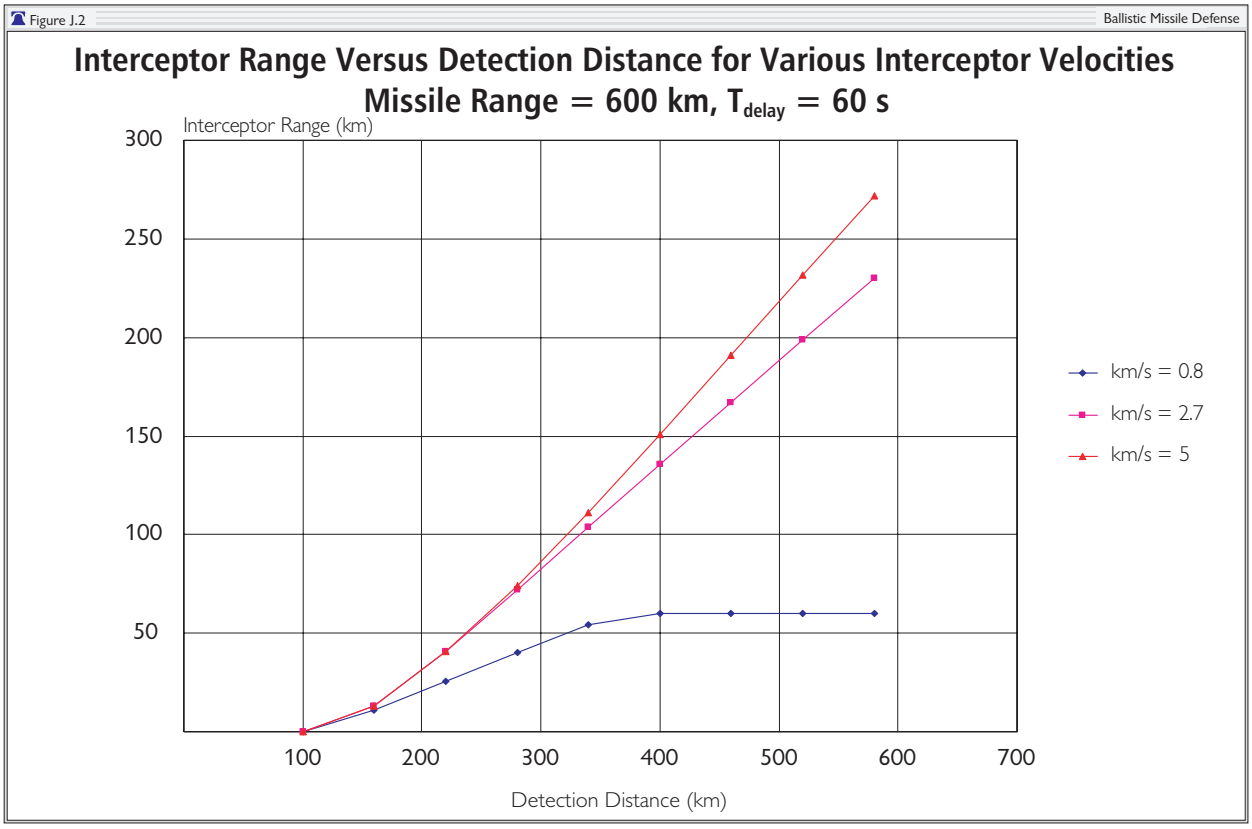


Figure J.3 Ballistic Missile Defense

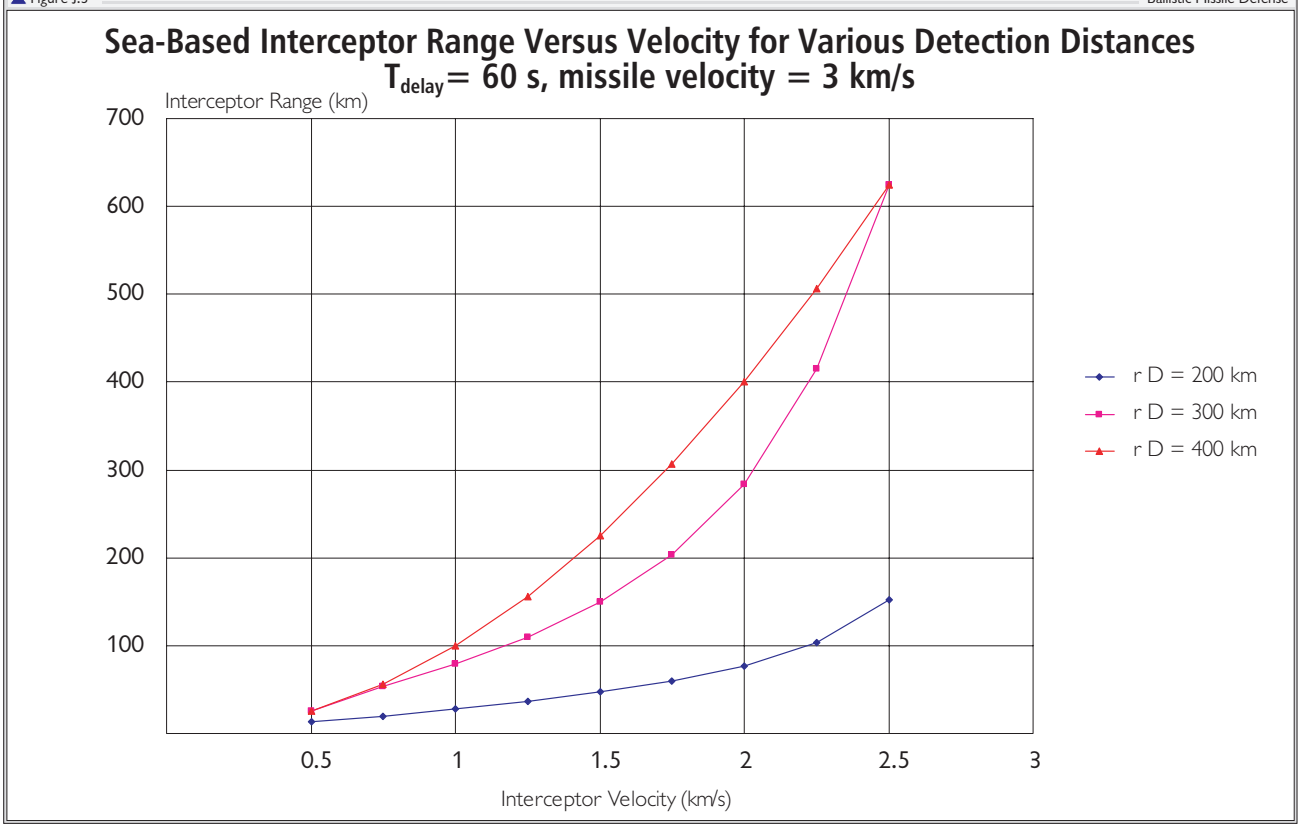


Figure J.4 Ballistic Missile Defense

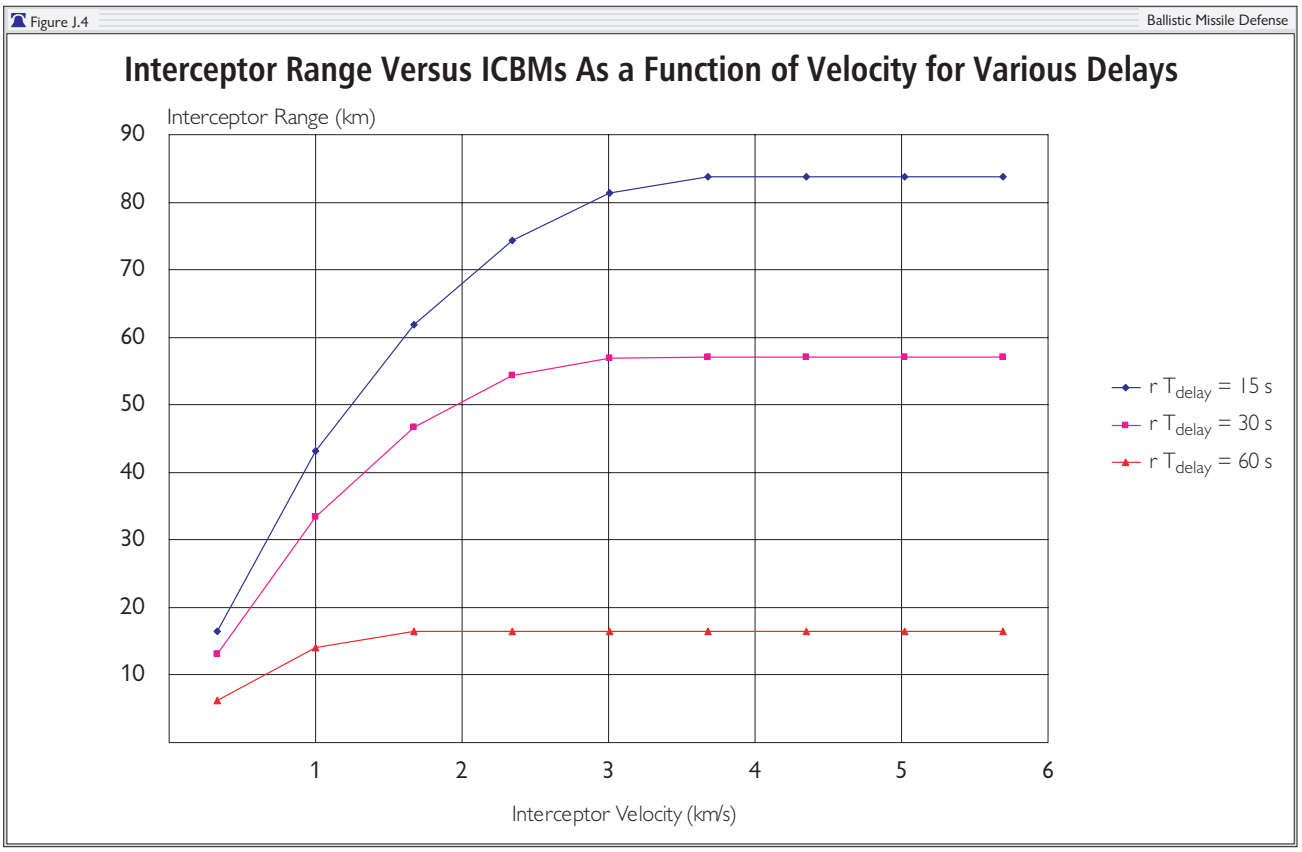


Figure J.5 Ballistic Missile Defense

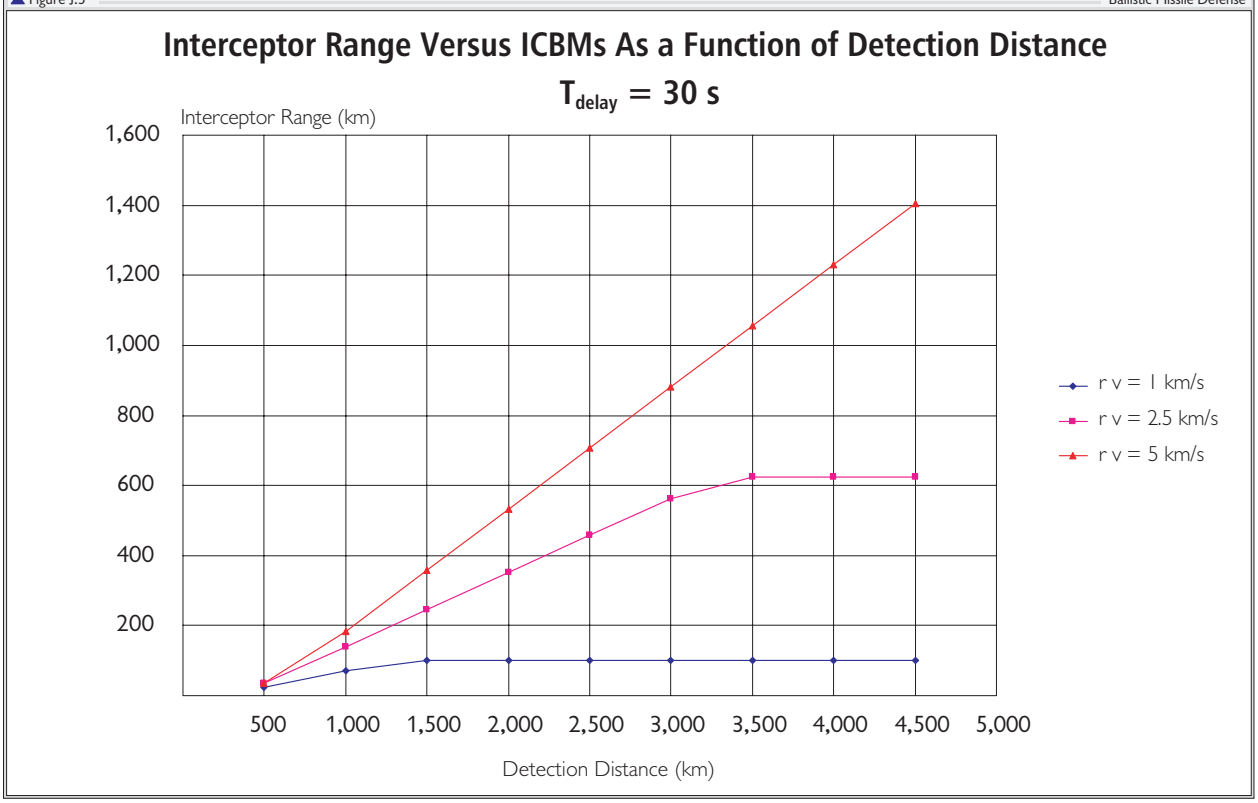


Figure J.6 Ballistic Missile Defense

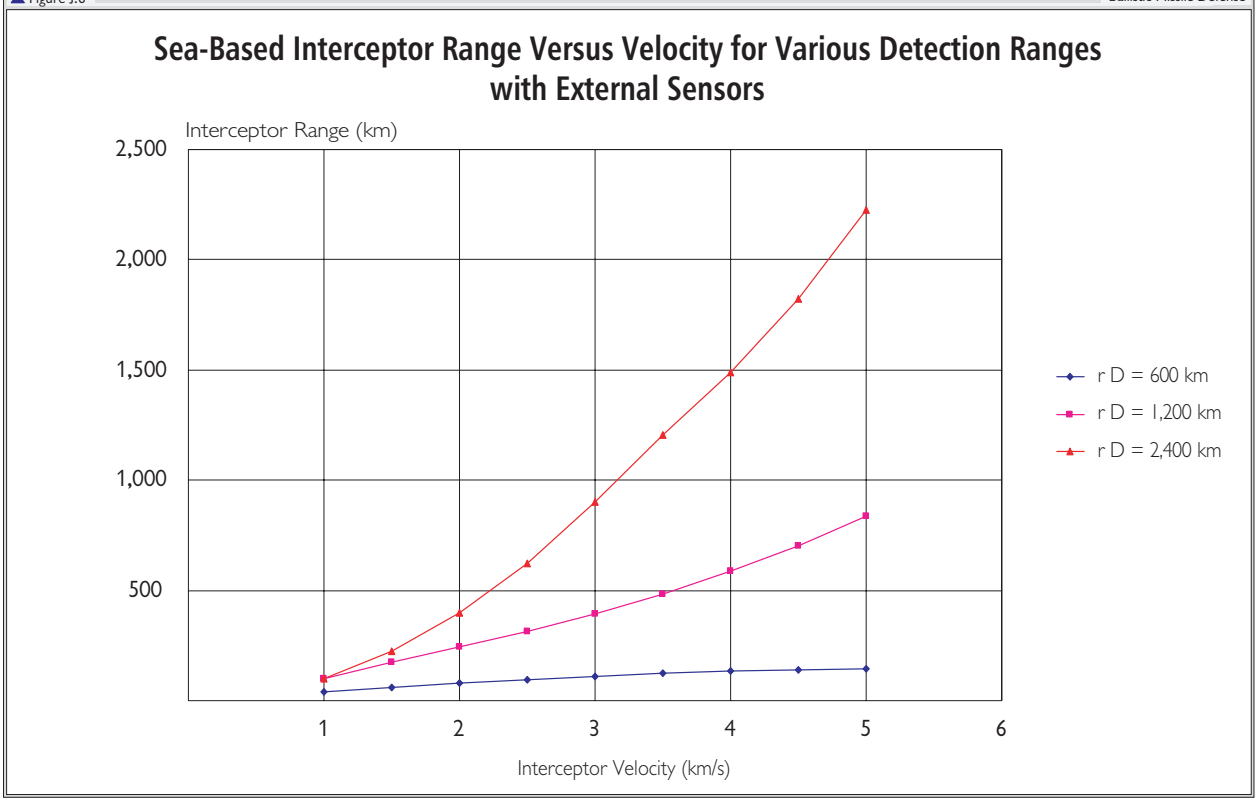
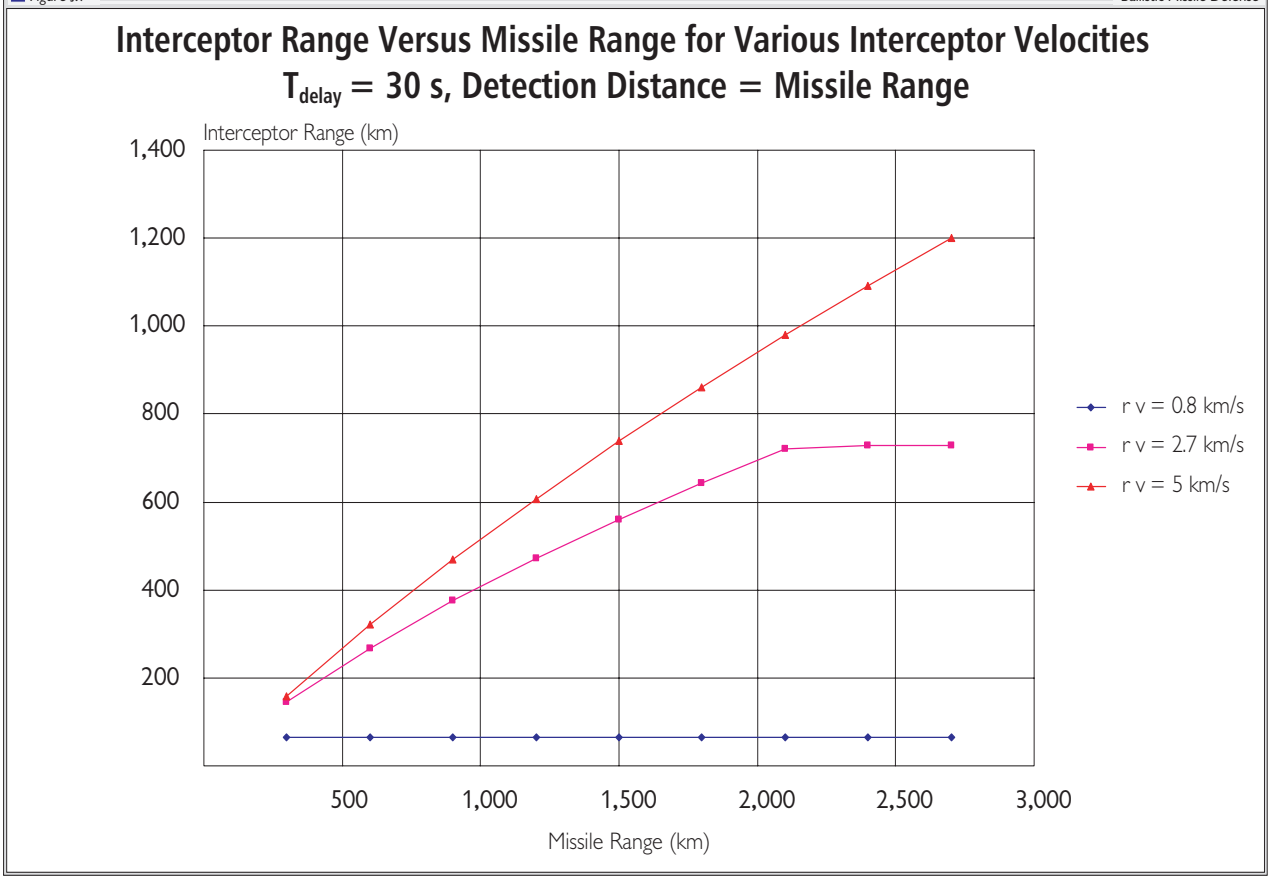


Figure J.7 Ballistic Missile Defense



Appendix K
Stability

Game theory can be used to discuss the strategic impact of the introduction of defenses into offensive balances.[149] The discussion below reviews the essentials of game theory and Nash optimal solutions. The stability of the current offensive configuration is shown to be high and insensitive to reductions in offensive forces, deployment of defenses, and the exchange of offenses for defenses. Damage and cost models are derived and discussed in earlier papers.[150] Strikes are estimated with conventional exchange models.[151] Costs are approximated by exponentials of damage to self and incomplete damage to the other. Thus, first and second strike costs are of the form Cost = damage to self + L(1 – damage to other), where L is a parameter that measures the attacker's relative preference for inflicting damage on the other side and preventing damage to itself.[152] If neither side strikes, the cost of inaction is L, which is a measure of the damage one side wants to be able to inflict on the other to deter untoward action or induce it to comply. During the Cold War, L was large because deterrence was the dominant role of strategic forces. As that conflict recedes, L should fall as the need for such deterrence diminishes. The side considering striking minimizes its first strike cost C_1 and decides whether to strike depending on whether C_1 is smaller than the cost of inaction L. That determines the optimal allocation of the first striker's weapons between missiles and value targets, as well as the cost of the side that strikes second, C_2.[153]

Figure K.1 defines the graph of play, the decision nodes, which side decides the next step at each node, and a set of payoffs to both sides for traversing each path.[154] The nodes represent decisions whether to strike first or strike back, so the first and second strike costs described above are the appropriate payoffs. The two sides are identified only as U (unprimed) and P (prime), corresponding to the unprimed and primed symbols used for their forces, strikes, and costs. The three nodes at the upper right corner illustrate the essential logic. U can strike or not at node 5. If it does, that leads to node 1, where the logical response to a first strike by U is a re-strike by P, which lacks any other use for its missiles, which has costs (C_1, C_2'). Inaction by U at node 5 would allow P to decide at node 2. A strike by P there followed by a re-strike by U would produce costs (C_2, C_1'). P not striking would result in inaction by both sides, which would have costs (L, L'), their damage objectives. P should strike at node 2 if $C_1' < L'$, but not if $C_1' > L'$. Thus, for $C_1 > L$, U chooses between C_1 from the top branch and L from the second, and would generally chose the latter. If $C_1' < L'$, U would also choose between C_1 and C_2, and would probably strike at node 5.

149. Canavan, "Analysis of Decisions in Bi- and Tri-Lateral Engagements."
150. G. Canavan, "Crisis Stability and Strategic Defense," *Proceedings of the Military Modeling and Management Session of the ORSA/TIMS National Meeting, November 12–14*, S. Erickson, ed. (Washington, D.C.: Operations Research Society of America, 1991).
151. G. Canavan, "Missile Defense and Proliferation," in A. Zichichi, ed, *25th Session of the International Seminars on Planetary Emergencies, 19–24 August 2000* (London: World Scientific, 2000).
152. G. Kent and R. DeValk, "Strategic Defenses and the Transition to Assured Survival," RAND Report R-3369-AF, October 1986.
153. A. Piontkovsky, "Global Defense and Strategic Stability," in A. Zichichi, ed., *16th Session of the International Seminars on Planetary Emergencies, 19–24 August 1992* (London: World Scientific, 1993).
154. Powell, *Nuclear Deterrence Theory.*

As C_1' fell to L', P would see an incentive to strike. Anticipating that transition, U would preempt P before C_1' reached L'. In this way a rational decision by P to reduce its costs slightly when L reaches C_1 would induce a rational decision by U to preempt that would impose larger costs on both sides than that of inaction, which would be the preferred path if either side controlled all decisions. Moreover, U's decision to preempt depends on its evaluation of P's cost C_1' and damage objective L', neither of which U knows with precision. Imperfect knowledge of P's decision parameters could cause U to strike by accident. The difference $C_1' - L'$, the margin of safety against such accidental exchanges, goes to zero as the decision approaches. The lower half of the decision tree contains the symmetric branch on which P decides first. The two halves are combined by a decision as to which of the two sides could strike first in a crisis that is conventionally modeled as a random decision by Nature, which was explored[155] in earlier notes and is represented here by the probability u that U can strike first in a crisis.

Figure K.2 shows the cost to U of node 7 in a bilateral interaction between two sides with START I forces as a function of U defenses and the probability that it can strike first in a crisis u.[156] Numbers of interceptors up to ≈ 600 do not change strike incentives because the decision variable, $C_1 - x$, remains positive. Larger numbers produce large U costs at small u because P has an incentive to preempt. At large u, U can strike first and use its defenses to negate P's suppressed second strike. At very large defenses, the cost to U is reduced below that of inaction for all u. However, that reduction is gained through reciprocal strikes that have large cost to P, and possibly U, through leakage not considered here.

Freedom to trade offenses for defenses makes it possible to deploy large defenses without inviting preemption. Figure K.3 shows the impact of U reducing offensive weapons W with and without defenses, starting from START III level U and P offenses, for $L = 0.5$, and $L' = 1$, i.e., moderately aggressive opponents.[157] The top and bottom curves are U's and P's first strike costs if U unilaterally reduces its offensive forces without deploying defenses and P maintains its offensive forces at 2,000 weapons. By $W = 500$, moving right to left, U's first strike cost increases to ≈ 2.2, while P's falls to ≈ 1.2. By $W \approx 100$, the discrepancy is a factor of ≈ 5, which could incentivize strikes or errors, given current uncertainties about damage preferences.

The two central curves are U's and P's first strike costs if U's offensive reduction is accompanied by a complementary increase in its defenses. The increase is not simply proportional to the decrease in U's offenses, but an increase to 1,600 defenses that is concentrated when U's offensive forces are reduced below 1,000 weapons makes C_1 and C_1' essentially equal at $W = 100$ and keeps them within a few percent of each other at all stages in the reduction. Thus, for these conditions it is possible to trade large offenses for defenses without significantly impacting the first strike costs of either side. As first strike costs are their primary decision variables, it is possible to trade offenses for defenses without impacting strike incentives, margins, or stability.

155. See T. Schelling, *The Strategy of Conflict,* and T. Schelling, *Arms and Influence* (New Haven: Yale University Press, 1966).
156. Canavan, "Analysis of Decisions in Bi- and Tri-Lateral Engagements."
157. Canavan, "Freedom to Mix Defenses in Modest Forces."

Figure K.1 Ballistic Missile Defense

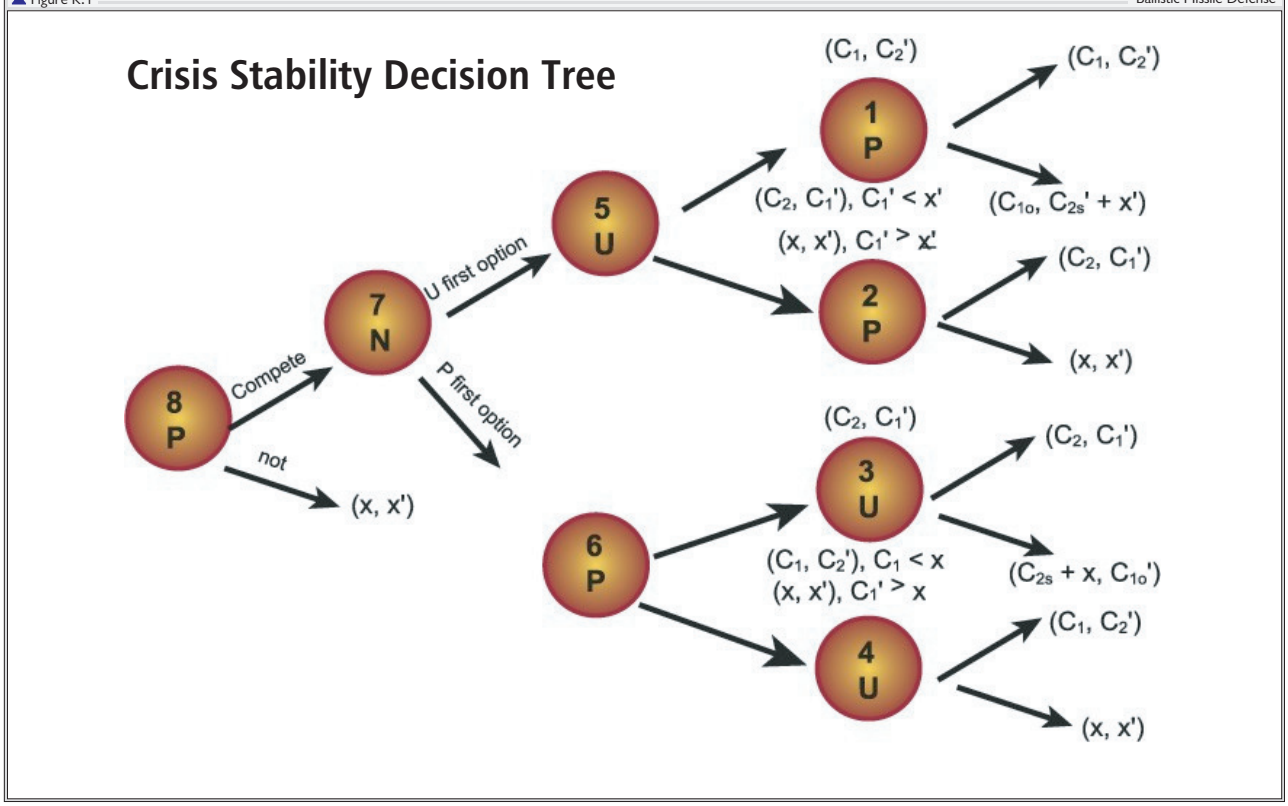


Figure K.2 Ballistic Missile Defense

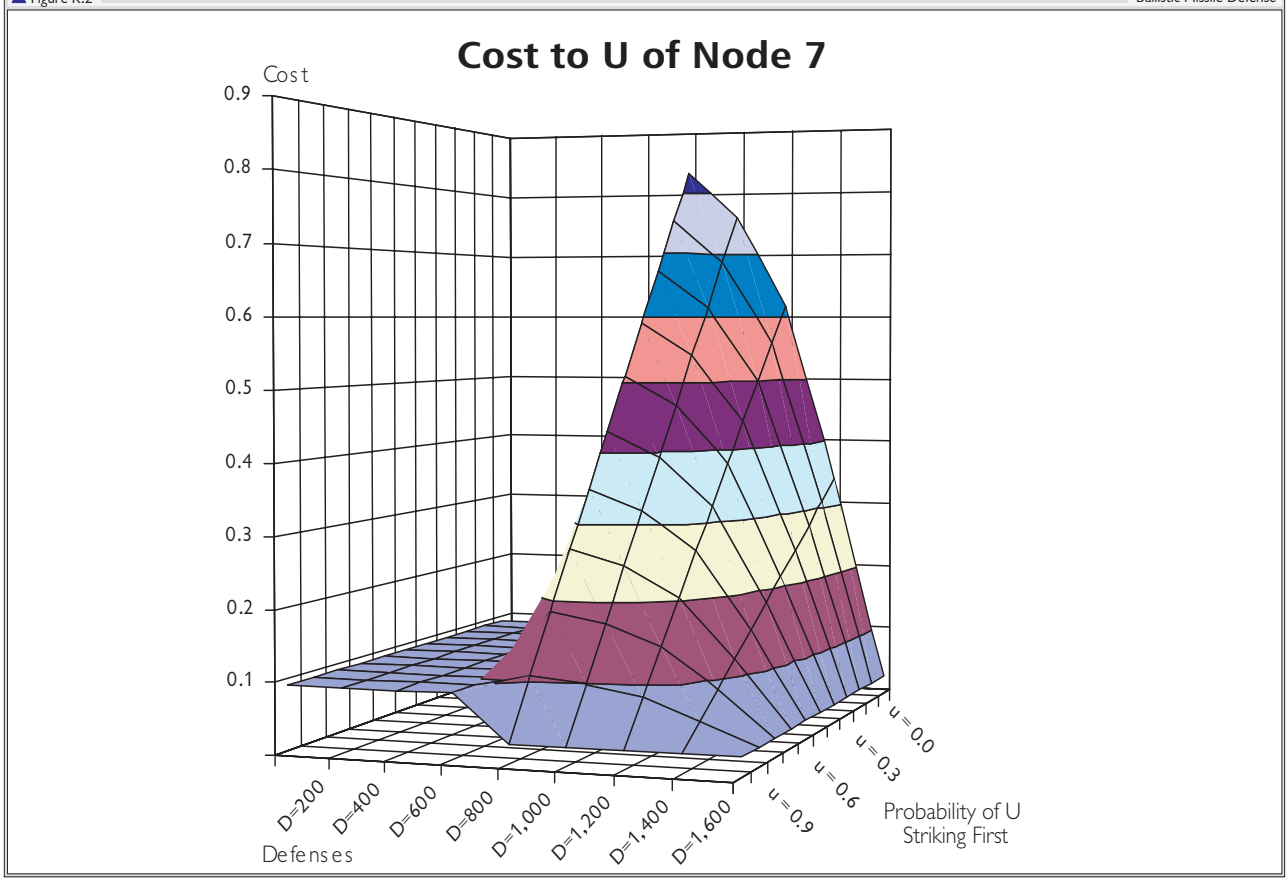
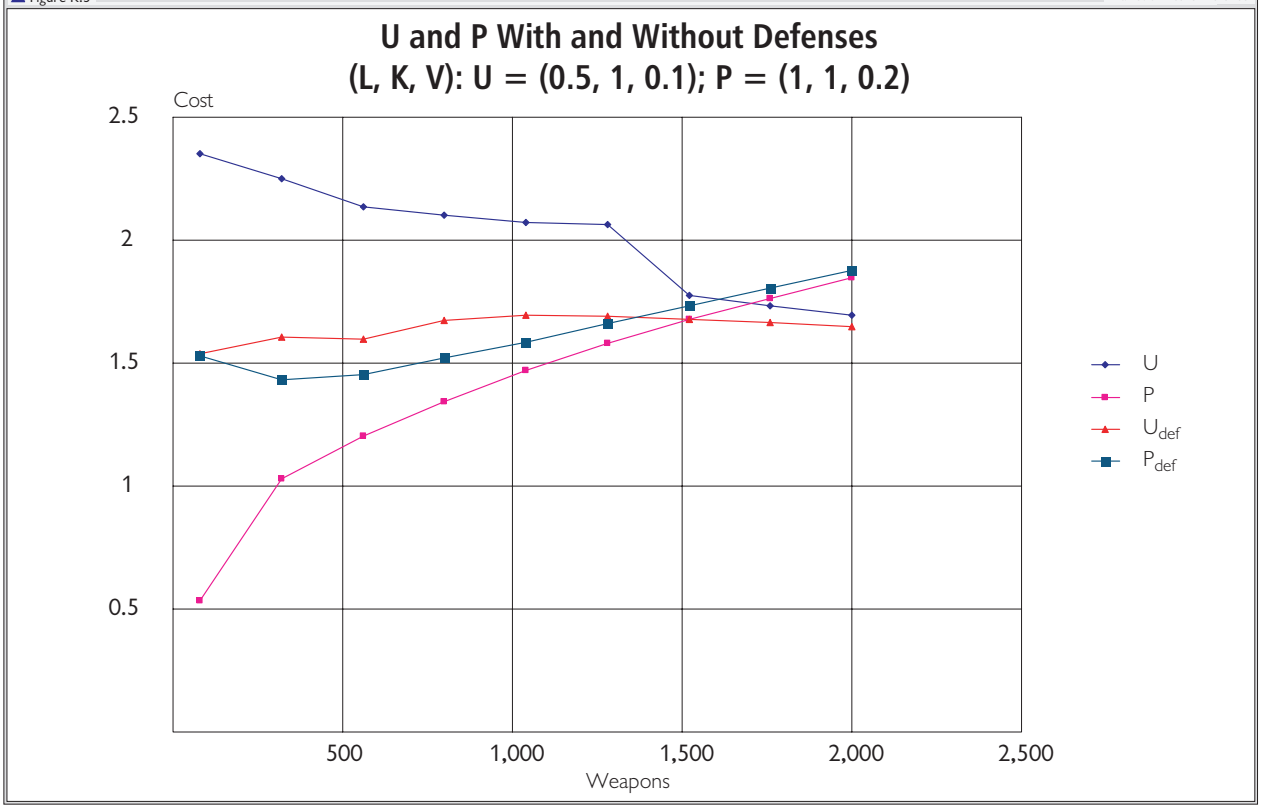


Figure K.3 Ballistic Missile Defense



Appendix L
Surface-Based Boost-Phase Intercept of Rogue Missiles

Ground-, sea-, and air-based interceptors can be used to intercept theater missiles on trajectories that provide access to them from nearby, secure basing. Figure L.1 shows approximate interceptor flyout distances and total ranges (i.e., interceptor and missile total flight distances), assuming head-on approach for a 4 g theater ICBM and 6 g interceptors with maximum speeds of 7 and 4 km/s. The top curve is the total range for the 7 km/s interceptor and the third curve is the distance it can fly out. The second curve is the total range for the 4 km/s interceptor and the bottom curve is the distance it can fly out. The 7 km/s could meet missiles head-on from a distance of 2,300 km with zero delay and about 1,800 km with 60 s delay. The interceptor would fly 1,500 to 1,100 km; the missile would fly the rest of the way. The 4 km/s interceptor could intercept from 1,700 to 1,400 km, flying 900 to 600 km to do so. While these ranges are large, they only apply to head-on intercepts. Trajectories from North Korea toward the eastern coast of the United States lie inland up to 700 km. Thus, the 7 km/s interceptor could engage them from ships 100 km off shore with 60 s delay, but the 4 km/s interceptor could not.

Analytic scaling models show that intercepting regional missiles in boost requires quick reaction and favorable geometry. Intercepting regional missiles in boost requires very quick reaction and favorable geometry. The interceptor velocities required increase with delay times. For typical missile and interceptor parameters, delay times greater than 60 s are unacceptable. Such detection and release times would be difficult with current sensors. Post-boost intercepts are less sensitive to delays, but more sensitive to countermeasures. Without delays, nominal interceptors have very large footprints. Modest reductions in velocity or increases in launch delays would reduce footprints by similar factors. Waiting for detection by organic sensors of few 100 km range would reduce footprints by orders of magnitude, which would largely eliminate the advantages of more capable interceptors.

It is possible to intercept a regional missile in its boost phase if the defenses react very quickly and its interceptors have a favorable geometry. A theater missile executes a gravity turn, rolling over into an optimal flyout angle of about 45° with respect to the horizontal. Its acceleration A can be treated as constant at its average value of about 4 g. Thus, it accelerates to its maximum velocity V in an acceleration time T_A of about V/A. At the end of acceleration, it is at a range $R_A = 1/2AT_A^2 = V^2/2A$ from its launch point. On an optimal trajectory, a regional missile of range R has $V = \sqrt{gR}$, so $R_A/R = g/2A = 1$ g/(2 x 4 g) = 0.12, so the boost phase is a small but non-negligible fraction of the total range. For R = 1,000 km, V = 3.2 km/s, so at burnout the missiles is at a range R_A = 125 km and height Z = down range X = 89 km. Since $Z \propto R_A \propto V^2/2A \propto Rg/A$, the burnout altitude increases linearly with range.

At a similar level of approximation the interceptor can be treated as having an average acceleration a of about 6 g to a maximum velocity v after a delay time T_{delay}. Thus, in the missile acceleration time, it can reach range

Equation L.1

$$r = v\left(T_A - T_{delay} - \frac{v}{2a}\right)$$

Interceptors are fired at the angle θ with respect to horizontal that minimizes the time to intercept; thus, they reach an altitude $r \sin\theta$ and ground range $x_{int} = r \cos\theta$. The hit condition for the interceptor to reach the altitude of the missile by burnout is that $r \sin\theta \geq Z = R/\sqrt{2}$ and that its ground range be greater than the separation between its initial position and the point below burnout. For head-on engagements, the interceptor can start from a distance $X + x_{int}$ from the launch point. For passing engagements, it must be within x_{int} of the burnout point.

As interceptor delay time T_{delay} lengthens, the interceptor's effective flyout time (missile boost time less interceptor delay and acceleration) drops from 43 to 23 s, which decreases its ground range x_{int} from 148 to 28 km and increases its flight angle with respect to horizontal from 31° to 73°. For T_{delay} = 24 s it must fire straight up. For longer delays, there is no solution for boost-phase intercept under these conditions. The missile's range at burnout is R_A = 125 km, and its distance down range is about 89 km, so the maximum range in head-on engagements is $X + x_{int}$ = 235 to 117 km. For small delays, x_{int} is about 146 km, which is the dominant contribution to the total ground range of about 235 km. For T_{delay} = 23 s, x_{int} is about 10 km, which would require a large number of interceptors. An interceptor with a longer delay could not reach the missile's altitude by burnout even if fired straight up. It has more than enough velocity to reach the missile later, but the missile would no longer be in boost. Once powered flight is terminated, various decoys and countermeasures can be deployed to complicate the intercept.

The short delays, or fast response times, needed to make the intercepts possible are a concern. For 1,000 km missiles, the response times needed are in the 10s of seconds, which are much shorter than those currently provided or projected for satellite systems. To achieve them, it would be necessary to use sensors that could observe the missiles as soon as they ignite, rather than waiting for them to penetrate clouds, and feed those observations to the missile launchers in near real time.

Regional missiles relax those timelines somewhat. The above equations apply, but as ranges increase, times lengthen and tolerances relax somewhat. Figure L.2 shows how times and distances vary with missile range from 1,000 to 5,000 km for T_{delay} = 30 s. The two top lines are the maximum interceptor range r and missile burnout altitude Z, which scales as $Z = R_A/\sqrt{2} = AT_A^2/2\sqrt{2}$ in terms of the missile acceleration time T_{msl} shown by the next line down. The effective interceptor burn time T_{eff} drops from 113 s at 5,000 km range to 16 s at 1,000 km. The interceptor's firing angle increases to about 85° by 1,500 km. At lower altitudes, there is no solution, for the reasons discussed above, which is reflected in the crossover between Z and r at about 1,500 km. For ranges of 5,000 to 2,000 km the interceptor ground range x_{int} is 100–150 km; for shorter ranges it drops to zero.

While $r > Z$ for R > 1,500 km, it appears that they could cross again for larger R. Both the short- and long-range crossings depend on missile and interceptor parameters, particularly the delay time. That dependence can be studied by equating missile altitude $Z = AT_A^2/2\sqrt{2}$ to the maximum interceptor range $r = v(T_A - T_{delay} - v/2a)$, collecting terms to produce the relationship

Equation L.2

$$\frac{AT_A^{\ 2}}{2\sqrt{2}} - vT_A + v\left(T_{delay} + \frac{v}{2a}\right) = 0$$

and solving it for the value of T_A that makes $r = Z$ for any T_{delay}, v, and a. Interceptor velocities of 3 and 4 km/s give a range of missile boost times for feasible intercepts of 100–180 s. The lower times are roughly the same for both velocities. The upper time is 60–80 s longer for 4 km/s. For 3 km/s, there is no feasible solution for T_{delay} > 30 s. For 4 km/s, there are solutions for T_{delay} > 40 s.

The missile boost time is related to its velocity by $V = AT_A$ and its velocity to range by $V = \sqrt{gR}$, so the times in Figure L.3 can be related to missile ranges by $R = (AT_A)^2/g$. For a 3 km/s interceptor at small T_{delay}, the range limits for feasible solutions are ≈ 135 and 5,200 km. By a 15 s delay, the lower limit

increases to about 460 km and the upper drops to 3,900 km. By T_{delay} = 30 s, the two converge at about 1,760 km. For a 4 km/s interceptor at small T_{delay}, the range limits are 240 and 9,300, which is larger than the range of validity of the model equations. By 20 s, they reduce to about 820 and 7,000 km. At T_{delay} = 40 s, they converge to about 3,140 km.

The upper range limit is not a serious constraint for 4 km/s; it is only a constraint for 3 km/s when T_{delay} approaches its maximum feasible value. The lower range limit is more of a constraint. For either velocity they eliminate intercepts of missiles with ranges under 1,000 km by interceptors with delays greater than 25 s. They eliminate boost-phase intercepts of missiles of any range by 3 km/s interceptors with delays greater than 30 s and by 4 km/s interceptors with delays greater than 40 s. All of these delays are short compared to current systems. The variation of allowable delays with interceptor and missile parameters can be evaluated with the discriminant, which determines whether the equation for T above has real solutions. The discriminant is $d = b^2 - 4ac$, where a, b, and c are the coefficients of T_A^2, T_A, and T_A^0 in Figure L.1. Thus,

Equation L.3

$$d = v^2 - 4\left(\frac{A}{2\sqrt{2}}\right)v\left(T_{delay} + \frac{v}{2a}\right)$$

so that $d = 0$ for the minimum interceptor velocity v_{min} that produces a feasible solution

Equation L.4

$$v_{\min} = \frac{\sqrt{2}AT_{delay}}{1 - \frac{A}{\sqrt{2}a}}$$

which is shown in Figure L.4 as a function of T_{delay} for interceptor accelerations of A = 4, 5, and 6 g. For T_{delay} = 0, there is no restriction on v, but 4 g interceptor with T_{delay} = 60 s would need a maximum velocity of about 12 km/s to overcome such delays. A 5 g interceptor would need about 8 km/s. A 6 g interceptor would need 6 km/s. Interceptors with under 3 g could not intercept at any velocity. Increasing acceleration to 5–6 g is clearly important; gains fall off for higher accelerations for A = 4 g. For higher missile accelerations, the interceptor acceleration needed increases in proportion to A. If v = 6 km/s is taken as a practical upper limit on interceptor velocity, a 4 g interceptor could only perform with a maximum delay of about 30 s; a 5 g interceptor with 45 s; and a 6 g interceptor with 55 s. However, at these maximum delays, the interceptor would have zero cross range and would have to fire vertically. Practical regional missile and interceptor combinations do not make T_{delay} over 60 s useful.

Figure L.1 Ballistic Missile Defense

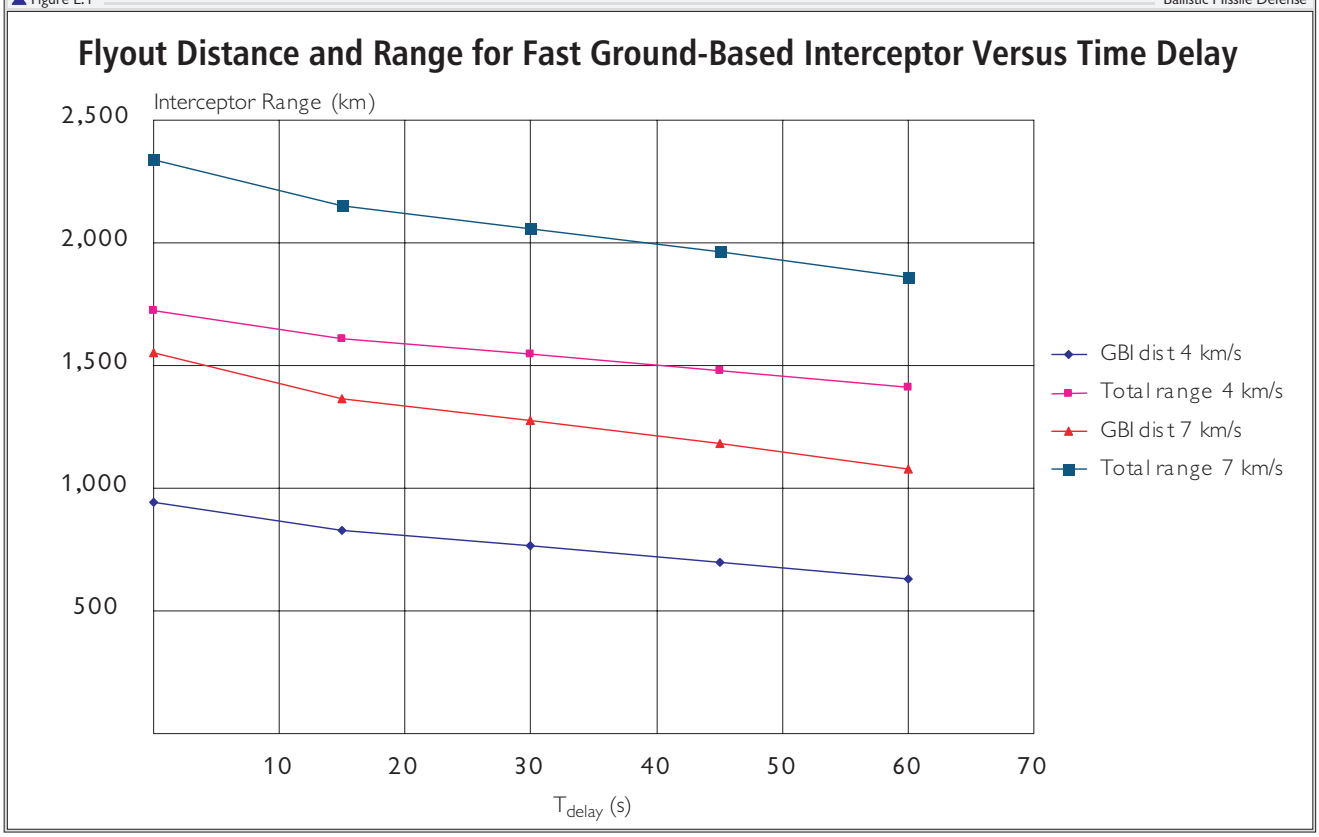


Figure L.2 Ballistic Missile Defense

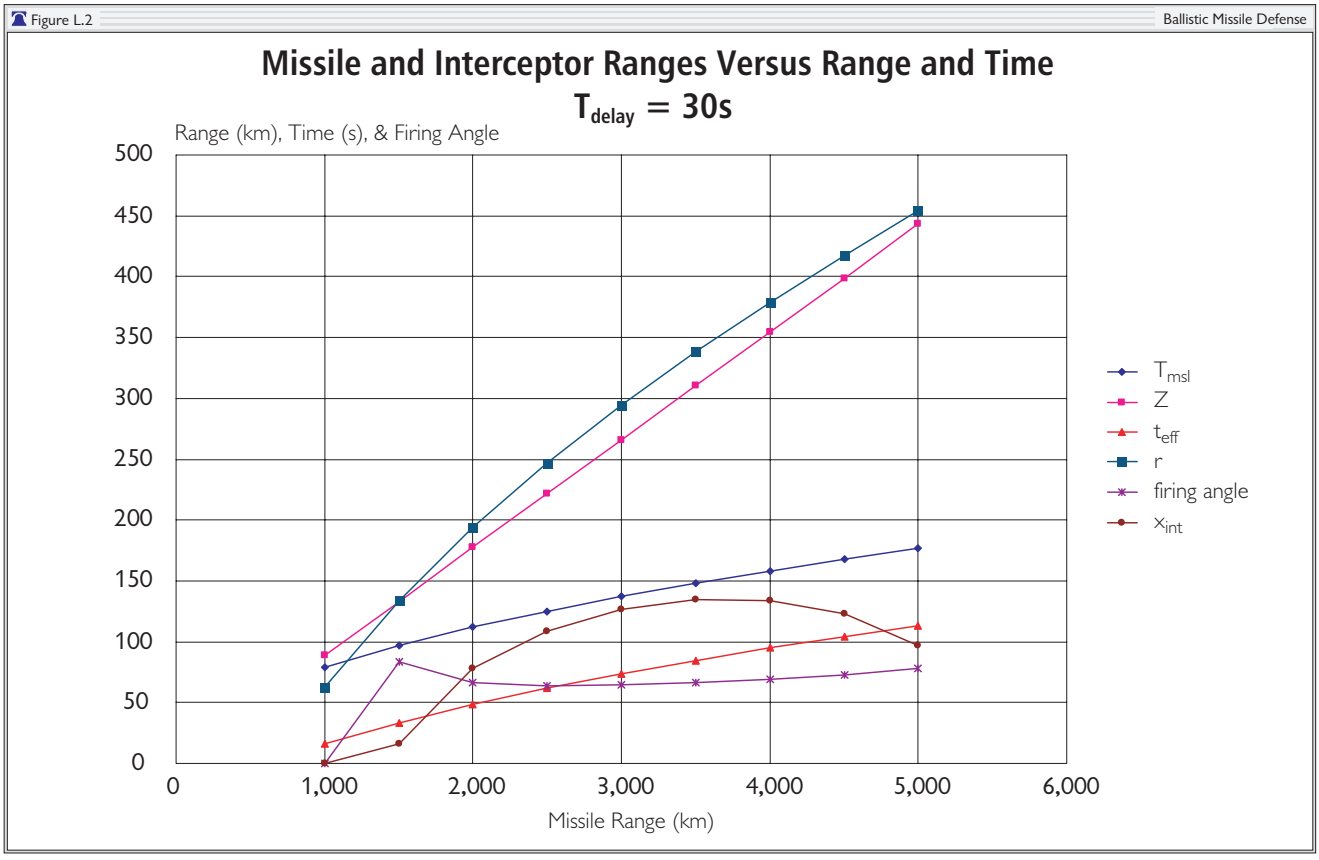


Figure L.3 Ballistic Missile Defense

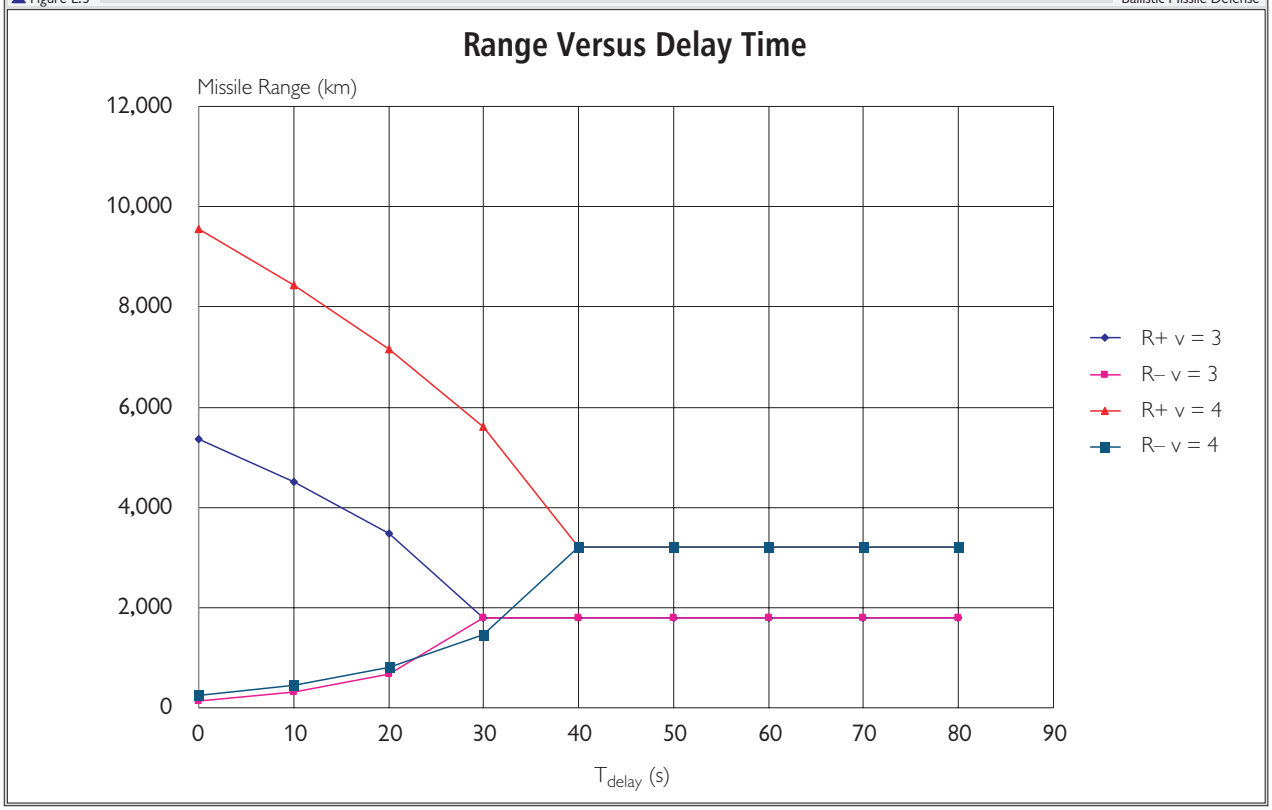


Figure L.4 Ballistic Missile Defense

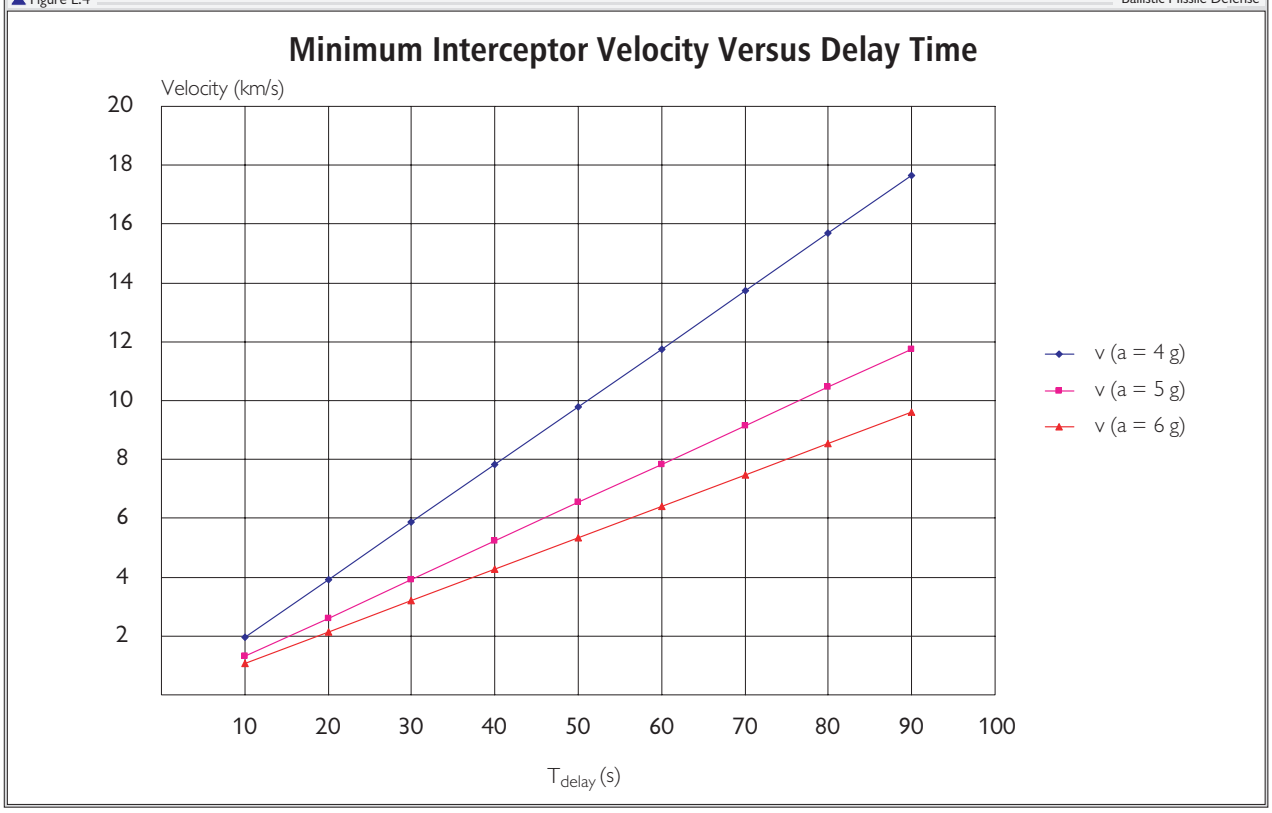
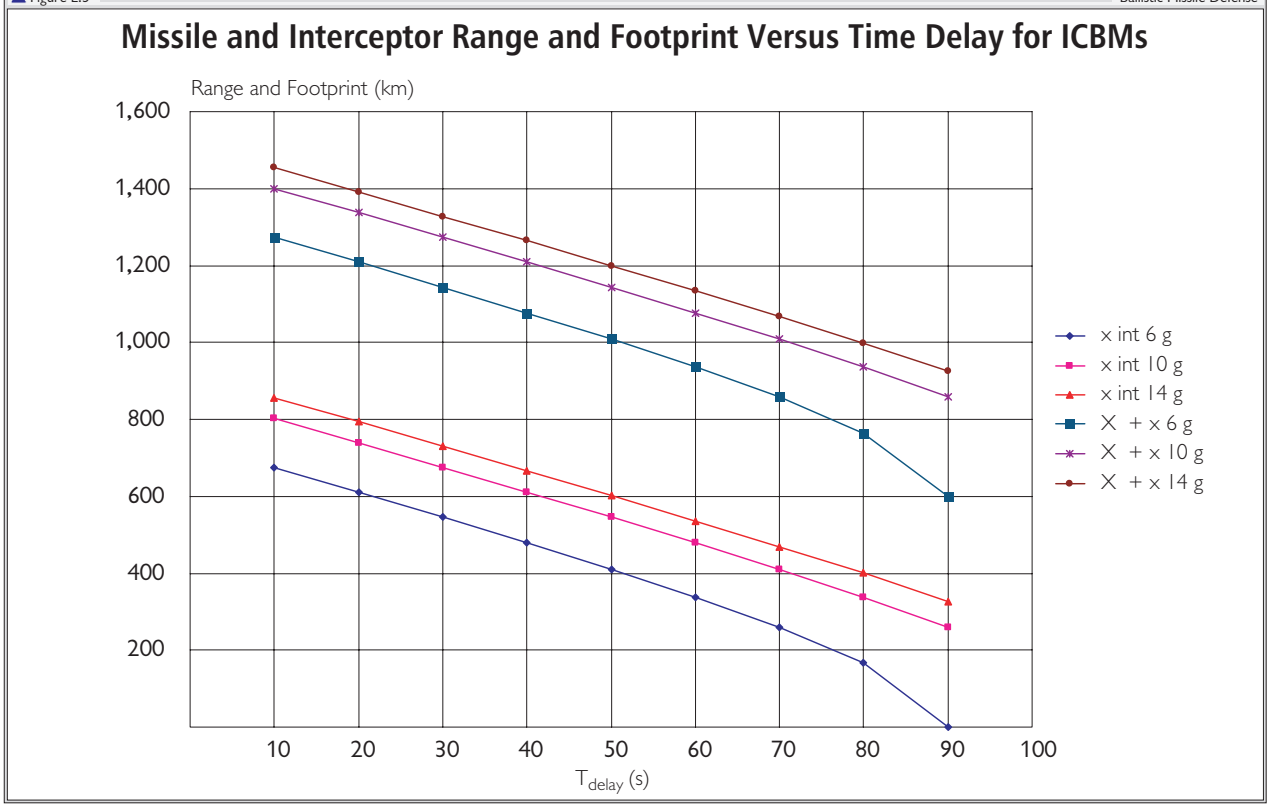


Figure L.5 Ballistic Missile Defense



Appendix M
Space-Based Boost-Phase Intercept of Rogue ICBMS

Figure M.1 shows the costs to intercept a 240 s rogue ICBM as a function of BP acceleration and maximum speed. The minimum cost of \$0.1 billion for a single point launched would increase by a factor of 2 if two BPs were allocated to it and another factor of 5 for that number of simultaneous launches to \$1 billion. The constellation costs are procurement costs for space hardware only, although they are assumed to include the BMC2 capability intrinsic to the GPALS BPs. Ground segment and 10–20 year operating costs could roughly double those costs to ≈ \$2 billion. Faster accelerating missiles in the 180 s range would increase costs by ≈ 1.8 to ≈ \$3.6 billion.

Figure M.2 shows the number of clusters that could be killed by intercepts in boost and cluster deployment phases. Rogue missiles are not expected to carry multiple weapons, but they are launched from a small area, can carry a large number of light decoys, and have relatively short burn times, all of which make them attractive but stressing targets for boost-phase intercepts. The end of boost at time T is the end of the opportunity to intercept their weapons free of decoys, but if the missile has a mechanism that continues to deploy decoys until time $T_{cluster}$, intercepting them before that time prevents some fraction of the clusters from being released, which avoids having to discriminate or commit an interceptor to them later. The number of clusters negated after boost can be several times larger than that intercepted in boost for typical conditions.

For point launch areas, the number of intercepts increases with the area in the BP constellation that can reach the missiles or dispenser. The range from which BP can reach the missiles in boost is ≈ VT. For N BPs the density overhead is $\approx N'' \approx N/4\pi R_e^2$, so the number of BPs that can engage in boost is $K_{boost} = N''\pi(VT)^2 R_e^2$. Each would kill the weapon and would incidentally kill all m_o decoys or clusters of decoys on board. Additional BPs would arrive after the end of boost, but while decoys were still being released. If their release was linear in time over T to $T_{cluster}$, the number negated at time t would be $m_o(T_{cluster} - t)/(T_{cluster} - T)$, which can be multiplied by the BP's rate of arrival, $2\pi r\, dr$, and integrated over T to $T_{cluster}$ to determine the cluster negations after boost

Equation M.1
$$K_{bus} = 2\pi N'' m V^2 \left(\frac{T_{cluster}\left(T_{cluster}^{\ 2} - T^2\right)}{2} - \frac{T_{cluster}^{\ 3} - T^2}{3} \right)$$

However, rogue ICBMs could use multiple RVs if the country had access to non-indigenous technology, as is suggested by the intelligence estimates cited earlier. If so, the time the bus takes to deploy the RVs effectively lengthens the boost phase, providing additional time for RV and cluster kills. Figure M.3 shows the boost- and bus-phase kills for $N = 30$ high-acceleration SBIs with $V = 6$ km/s intercepting a single missile with $m_o = 10$ RVs and a boost phase duration of $T_{boost} = 200$ s as functions of the time of the end of the bus phase T_{bus}. For $T_{bus} = T$, there is no contribution subsequent to boost, and $K_{boost} \approx 1.6$. For $T_{cluster} = 1.5T = 350$ s, $K_{boost} \approx K_{bus}$ and essentially all of the RVs are killed in one phase or the other by a modest number of SBIs. If the ICBM deployed clusters of decoys along

with each RV, proportional numbers of clusters would be killed in each phase. These numbers would be reduced as above by finite accelerations and long release delays.

Figure M.4 shows interceptor fly-in range $\approx V(T - T_{delay} - V^2/2A)$ for $T_{delay} = 30$ s as functions of missile range, which determines T through $V \approx \sqrt{(gR)}$. Accelerations of 6 g produce significant penalties, but 12 g is almost as good as 24 g except for $R < 1{,}000$ km where neither is useful. Figure M.5 shows the number of kills that result for these fly-in speeds for 30 s delay and 24 g BP. Increasing the number of BPs in the constellation doubles the boost-phase kills and halves the midcourse kills, shifting the crossover from ICBM ranges to ≈ 3,500 km.

Figure M.1 Ballistic Missile Defense

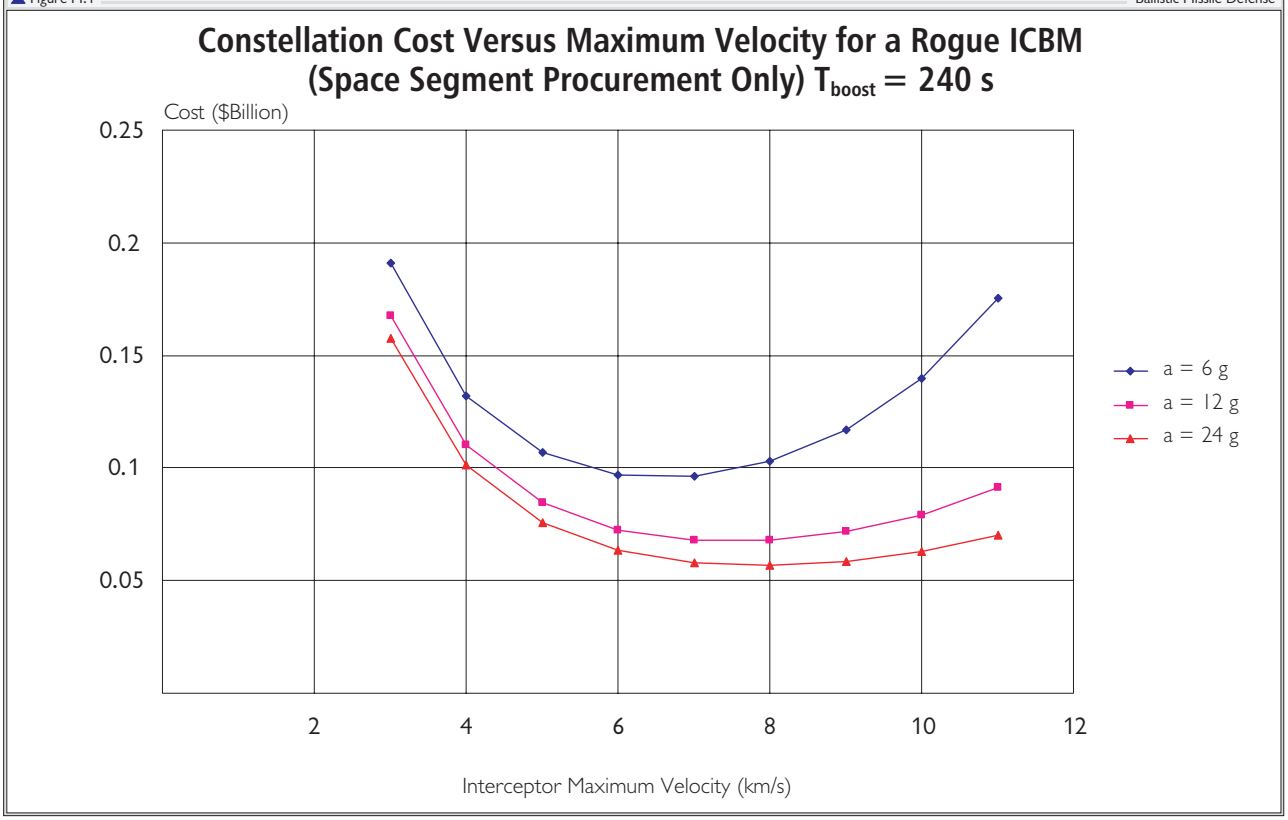


Figure M.2 Ballistic Missile Defense

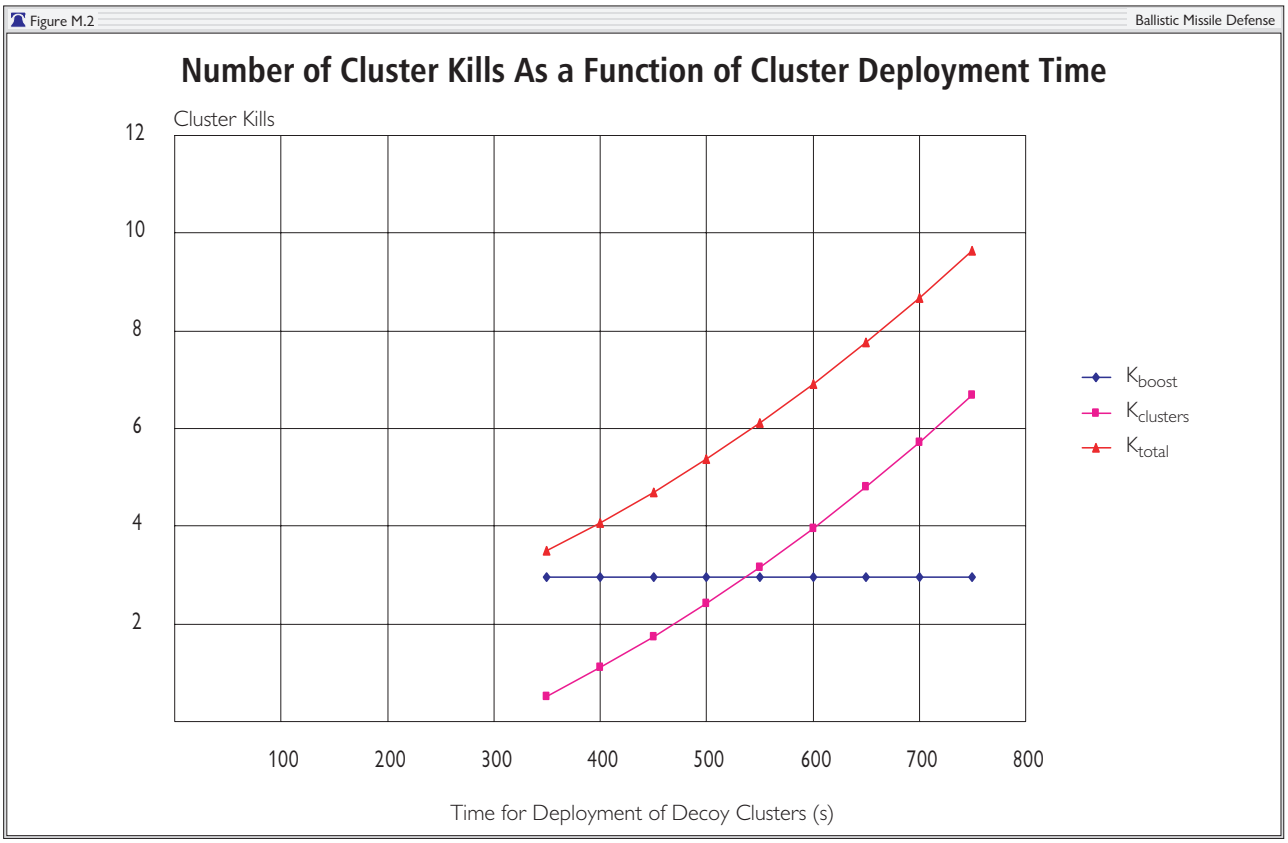


Figure M.3 Ballistic Missile Defense

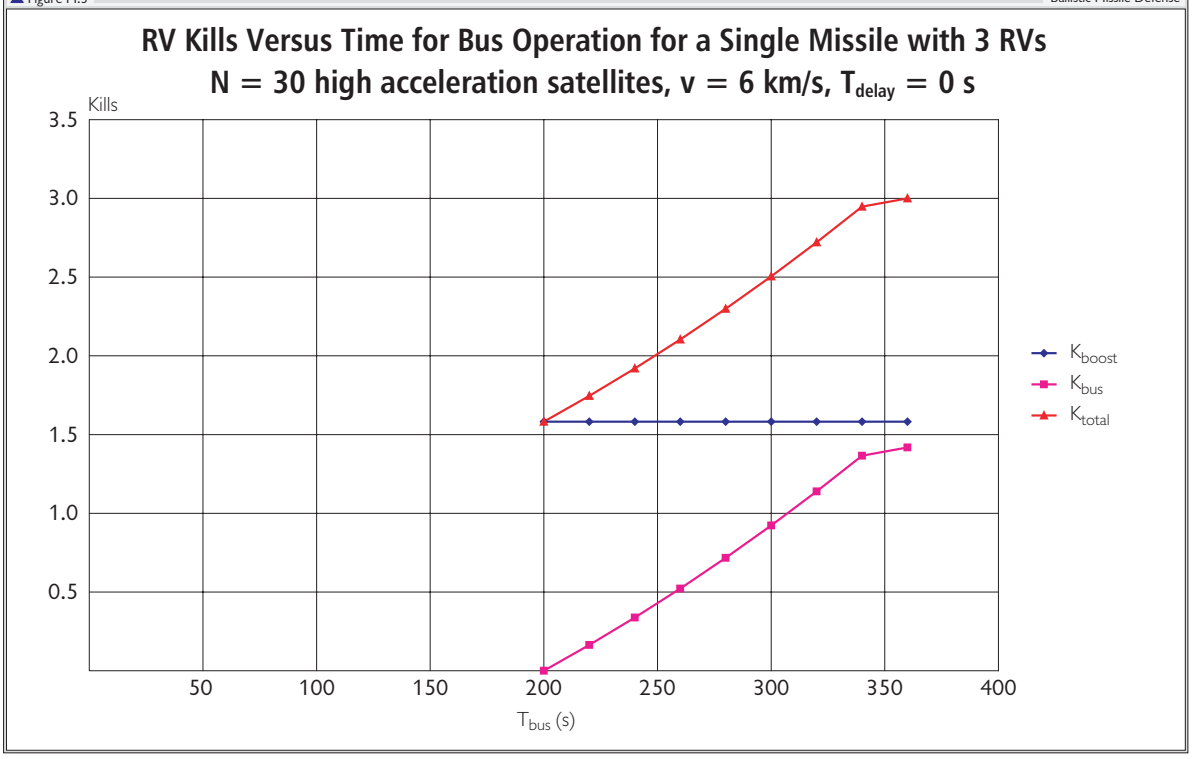


Figure M.4 Ballistic Missile Defense

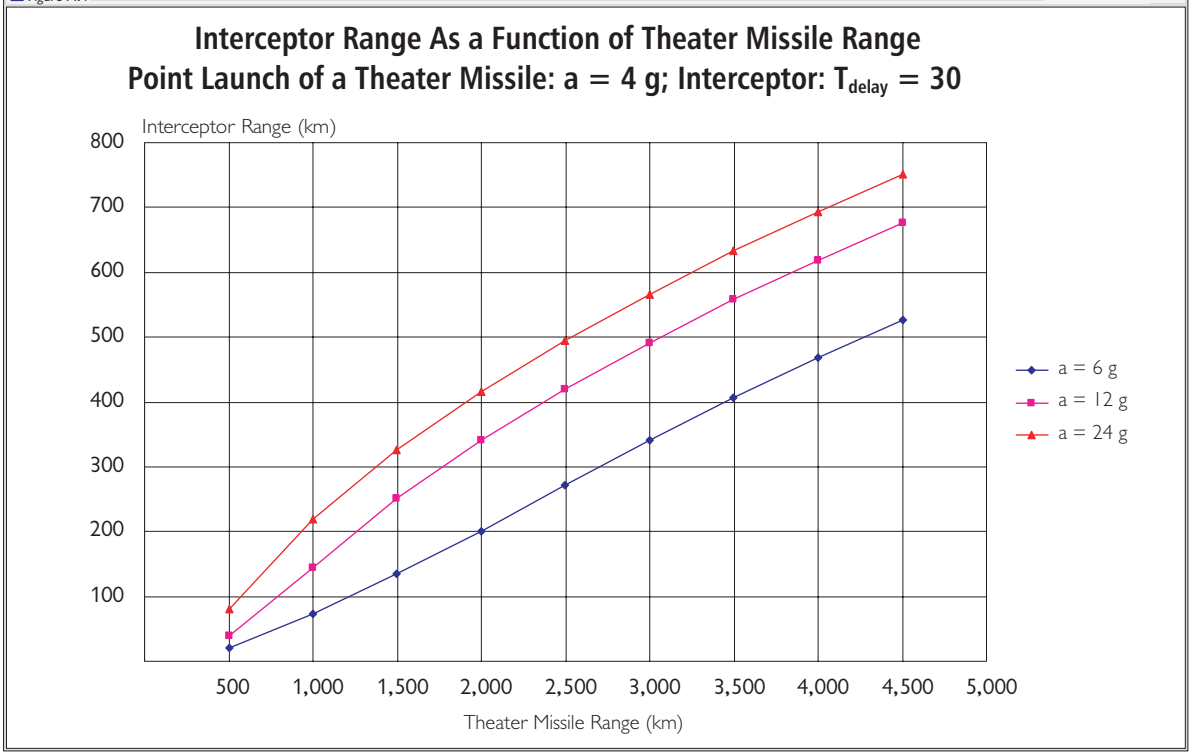


Figure M.5 Ballistic Missile Defense