Royal Netherlands Defence Academy

Military Operational Science

Introduction to air operations

Authors
Jan van Angeren, IJsselsteijn
Marcel de Goede, Dordrecht
Herman Koolstra, Best
Fred Oude Rikmanspoel, Deurne
Anne Tjepkema, Hoeven
Joop Voetelink, Breda

Printed at the Royal Military Academy Breda, The
Netherlands Proof reading and editing: Anne Tjepkema;
Wilma Koolstra Authors:
Jan van Angeren Chapter 11
Marcel de Goede Chapter 1
Herman Koolstra, Chapters 2-10, 12-14, 16, 17 and 19
Fred Oude Rikmanspoel, Chapter 18
Anne Tjepkema, Chapters 15 and 21
Joop Voetelink, Chapter 20

Table of Contents

Introduction

Some developments can be seen as a dividing line in the history of mankind. The development of agriculture, the industrial revolution and more recently the IT-revolution are typical examples of this. Such developments do not just affect one single aspect but have an impact on all areas of human life.

The invention of powered flight is also such a dividing line that has affected all areas of human behaviour. Of course this includes warfare, and after the introduction of the aeroplane the art of war has had to change dramatically. Aerial warfare is not simply the addition of a further dimension for man to fight in, nor does it only serve as an extension and support of land an sea forces. Now, this new dimension has made it necessary to completely rethink the art of warfare.

Therefore it is important that the study of airpower is not limited to those who will be employed in that particular field. The knowledge of airpower is essential for all who study warfare, because there is no separate land or sea war but only one joint operation.

Airpower has removed old constraints like terrain, has compressed time and disregarded frontlines, thereby turning whole countries into target areas. Airpower also came with its own set of constraints such as weather, range, payload and accuracy, but many of those constraints now no longer exist as a result of technological developments.

Whereas in the first decades of airpower the emphasis of technological development was on aerodynamic performance- more speed, altitude range and payload- the dramatic increase in aircraft capabilities came after the introduction of better sensors and computers. Now the point has been reached where almost every classical enemy unit can be detected and attacked under almost all weather conditions. This capability was achieved only recently but not by every air force. Not all constraints have however disappeared, at least two still exist. These are the lack of persistency and the enormous cost involved, which severely limit the number of aerial assets. Furthermore, the present emphasis on irregular opponents necessitates a re-evaluation of the proper use of airpower. The classical airpower assets are not always the most economical ones to use against an unconventional contender in an almost empty battlefield. Another challenge is the enhancement of the military network. It is technologically not difficult to connect a limited number of players, but is extremely challenging to connect an aircraft with every ground unit within its reach.

If technology developed over the years, so did airpower theory. Some theorists were way ahead of the technological capabilities and were envisioning a use of airpower that far exceeded its capabilities at that time. Others misjudged possible future effects by applying too linear an extrapolation. The classical example is the use of aerial bombardment against cities. The very limited attacks on London during the First World War wreaked havoc while the much heavier bombardments on London during the Second World War were not able to change the course of the war.

Finally, the way air operations are performed has changed and continues to change. Operations are constantly adapted based on lessons learned during actual battle and training missions.

The fundamentals of airpower and the theory of how to employ it have been the subject of many documents. However, most books only give a qualitative description of the fundamentals of airpower. The authors believe that a good understanding of airpower can be gained by balancing a qualitative and a numerical approach. Airpower is to some extent a technical achievement and understanding the theory behind it will result in a better understanding of its possibilities and limitations. Airpower is also an expensive resource and therefore a scarce commodity. To optimise scarce resources is the field of economics, the laws of which the application of airpower is also subjected to.

For this book a build-up approach is used. First, the reader is given a short historical overview to set the stage. The subsequent six chapters discuss the technology behind airpower and move from the properties of air to aerodynamic principles, to individual platforms, sensors and weapons. We conclude this section with the description of the maintenance involved. The next three chapters go on to describe the operation of the single aircraft and includes navigation, weapons employment and electronic warfare. After laying down these fundamentals the next eight chapters discuss complex air operations. The last section deals with command and control, legal issues, and strategy.

Reading the book from cover to cover is just one of the ways to use this book. It covers many issues and some readers may only be interested in certain aspects of airpower while others may wish to use the book as reference material. Furthermore, those acquainted with technology most likely don't need some of the technical chapters. This book is intended to serve all these different uses. We are confident that everybody who does study all these chapters will have a firm base to start more advanced airpower studies.

Chapter 1 – The evolution of airpower

A short introduction to airpower history

Preface

Air forces are the youngest of the military services. Unencumbered by existing traditions, they have created their own, starting with the pioneers of the Great War who developed their ideas and concepts while fighting, as the war raged and the technology of their planes and weapons evolved. At the end of that devastating war, the air weapon had taken great strides. Roland Garros escaped after three years of captivity and had to take flying lessons after returning to his country. The planes at that time were barely comparable to the Morane Saulniers L he had last flown.[1] Be it rudimentary, every form of air operation today can be recognized in that war that has been of such great importance to the development of airpower. Another factor that makes the war of 1914 to 1918 so important is the influence it has had on the major airpower theorists that today are considered to be the classics. Their ideas developed further in the interwar years and found application in air services of major world powers. The resulting concepts were put to the test when war erupted again some two decades after the Armistice. To what extent did the theory live up to expectations?

After the Second World War, the Cold War would determine international relations for some four decades. The character of the superpowers' armed forces and their respective allies was very much determined by a possible military confrontation. Considerations with respect to the opposing power block were leading for the composition, posture and concepts of operations. For the air forces, the strategic (in this case almost equal to nuclear) aspect of the confrontation had the highest priority. Whenever another adversary had to be countered, both blocks had difficulty adapting.

After the Cold War had ended, it took considerable reorientation to come to grips with a no less volatile world. Difficulties encountered in all kinds of internal conflicts and civil strife such as in the former Yugoslavia would become the rule. The Gulf war of 1990/91 proved to be an exception. In all conflicts however, airpower played a very distinctive role. A role it continues to play as the world turned another corner after the terrorist attacks of September 11th 2001.

This chapter gives an overview of airpower history in its first century. It is a selection of highlights to illustrate the development of airpower. It goes without saying that this summary can be expanded upon almost endlessly.

Origins

In December 1903 the Wright brothers were the first to make a heavier-than-air powered flight. This historic flight was to be the start of a new line of developments next to the other, already much older, variant of flight by means of lighter-than-air platforms.

The Wright brothers were the first to achieve success, but were by no means the only ones on the right track to developing an aeroplane. In many countries pioneers had built aircraft, laying the foundations for aircraft manufacturing. Even before

1 Lee Kennett, The first air war, 1914-1918 (New York 1999) 93.

aircraft flew at all, the phenomenon had been drawing the public's attention and featured in books and magazines. With the actual realization of powered flight, this attention grew even further. In the spirit of the years between the *fin de siècle* and the First World War, developments surrounding flight were very popular. It became one of the arenas in which nationalistic rivalry found an outlet.

The first military aeroplane was a Wright Flyer bought by the US Army in 1909. The first military use of aeroplanes in war is ascribed to the Italians in their campaign against the Ottomans in the Italo-Turkish war in present-day Libya in 1911. The Italians had fielded a mixed bag of some 10 planes, which provided reconnaissance and out of which several grenades were tossed, making the Italians the first to have executed an aerial bombardment. All in all, these efforts are more of importance because of their historical noteworthiness rather than that they have any military significance. On the eve of the war of 1914-1918 however, both major powers, and also the smaller ones, all possessed a service for military aviation, albeit one embedded in the existing services. The war really got things going and gave military aviation a truly flying start.

The Great War

Trial and error: a first time for everything

The importance of the First World War to the development of airpower can be seen as twofold. Firstly, every air operation known today has its origins in the Great War.[2] Secondly, the war had a profound influence on the airpower theorists of the interwar years, whose influence stretches to the present day.

At the outbreak of the war, all the major belligerents acknowledged the potential usefulness of aircraft in warfare. All of them had an air service in place. The aircraft initially fielded were used for reconnaissance and observation. It was to be during the first weeks of the war that these unarmed planes would already make contributions that carried great consequences. German aviators provided crucial information on their enemy at Tannenberg, enabling the Germans to check the Russian advance westward. French airmen discovered the German right flank turning east of Paris during their advance, thereby exposing it to attack. This precipitated the battle of the Marne, which meant the final failure of the German execution of the Schlieffen plan.

The peculiarly static nature of the Western front enabled lighter-than-air craft to play a significant role in the First World War. Both sides used cable balloons in large numbers for artillery observation. Furthermore the Germans used large dirigibles ("Zeppelins") as bombers, which bombed London in 1915. With a complete absence at that time of any air defences, they could do so in broad daylight. The British soon corrected this, chasing the Zeppelins into the darkness of the night. By 1917, the Germans had new large bomber planes at their disposal. The numbers of both Zeppelins and bombers were always low and the effect of this early strategic bombing campaign was mainly psychological. It did lead to

2 Nowadays, the Royal Netherlands Air Force like most other NATO-countries, as well as the Alliance as a whole, distinguishes four main categories of air operations, notably counter air, strategic air, anti-surface force air and supporting air operations. Counter air ops are aimed at combating the enemy's air assets. Strategic air ops aim at achieving strategic results. Anti-surface air force ops combat the adversary's surface forces. Supporting air ops comprise all other operations, some necessary to facilitate the first three categories, others as operations in their own right. Examples of the latter are intelligence surveillance and reconnaissance (ISR), air transportation and air-to-air refuelling.

serious investigation into the quality of air defence, which eventually led to the creation of the world's first independent air force, the Royal Air Force, in April 1918. Also, the results of the German bombing effort were used in calculations that were to become part of the interwar British doctrine on strategic bombing.

Back to the original task of aircraft: observation. In order to enable military operations, observation was something that had to be denied the enemy and secured for oneself. Now the fight for superiority in the air was on. The first aircraft used as fighters were armed observation aircraft. One way of gaining an advantage over other aircraft was by reducing the weight through the use of only one pilot, as opposed to a pilot and an observer. Among the first planes to be used as fighters, the British used scout planes, long-range reconnaissance planes with a crew of one, that were comparatively light and somewhat faster than other planes; hence the designation "scout" for a fighter plane. Apart from having an aircraft capable of the performance necessary to take on an adversary in the air, the most important consideration was armament. The earliest aviators of the war improvised with whatever they could take into the air. This comprised not only firearms. The advent of lightweight machine guns like the Lewis made it possible to take more serious firepower up into the air. Initially, the machineguns were most often fired by gunners, firing in front of the plane when it was a pusher, and backwards when the engine was in front. In both cases the gun was manoeuvrable and had to be aimed by the gunner. An early innovation in this field was made by the Frenchman Roland Garros. He was the first to invent and use a device that enabled the pilot to fire the - now fixed - gun forward, aiming the gun by aiming the plane. This was realized by the crude but simple expedient of armouring the propeller blades where they crossed the line of fire of the gun. The Morane Saulnier N scout monoplane thus fitted was arguably the first real fighter plane.

The next leap forward was made by Anthony Fokker, the Dutch manufacturer who had his factory in Germany. In search of an answer to Garros' invention, he invented an interrupter device, which withheld the machinegun's fire every time a propeller blade passed in front of the barrel. This was first implemented on a Fokker Eindecker plane, "an evil-looking monoplane scout".[3] These Eindeckers, as flown by famous airmen like Boelcke and Immelmann, ensured the Germans had the upper hand in the skies over the front from summer 1915 to spring 1916. They only operated behind their own lines, so downed aircraft would not fall into enemy hands. By remaining enigmatic the Fokkers gained an enormous reputation in allied circles, far exceeding their actual performance and numbers. It is estimated that there were never more than fifty Eindeckers at any one time on the whole of the Western front.[4] The Fokker mystery lasted until finally one did fall into allied hands. When it was tried and tested against allied planes it turned out to be an averagely performing plane.[5] Soon the allies would also be equipped with interrupter gear. The practices laid down by these early German airmen form the basis of fighter doctrine up to this day. Men like Boelcke and Immelmann, actually young subalterns, laid the foundations for both fighter tactics and the organization of the fighter arm. A start was made with organizing single-type units, the *Kampfeinsitzerkommando's (KEK)*, as opposed to the existing mixed units of observation and escorting scouts. The Allies had their pioneering counterparts like Albert Ball, Jean Navarre, Georges Guynemer and Charles Rickenbacker. While

[3] Cecil Lewis, Sagittarius rising (Harmondsworth 1977, first ed. 1936) 51.

[4] Kennett, The first air war, 110.

[5] Lewis, Sagittarius rising, 52. Lewis describes the Fokker as "perfectly orthodox". It was flown against a Morane Saulnier Bullet (the British name for the N type), which was "all over the Fokker" and "had everything its own way. A cheer went up from the ground. The bogey was laid."

fighting they developed their ideas and laid the foundation for new traditions for this novel weapon.

1916 was an important year in the development of airpower. Solitary flying by lone huntsmen became the exception. Tactics now prescribed the flying in pairs, or multiple pairs. The French, who had been the first to organize uniform aeroplane units, were quick to react to the German scheme of the Luftsperre, where they tried to seal off the air above the battleground at Verdun. They were able to swiftly transfer sufficient fighter units to deny the Germans superiority in the air. French major Tricornot de Rose is credited with being the first to have fighter units actually fight as units as opposed to being a group of individual pilots, however apt at flying they may have been. For the remainder of the war the belligerents would organize themselves in the same manner, the Germans perfecting their *Keks* into Jagdstaffeln and Jagdgeschwader. Turned into mobile units these were later dubbed circuses, which travelled along the front to be employed at hotspots.

The arms race in the air continued. The allies fielded new planes like the Nieuport Bebe's and the DH 2's that could outfight the Eindeckers. The Germans gained the upper hand again with the advent of the Albatros line of fighters, whose success lay in their powerful engine. The Staffel thus equipped brought about Bloody April, the month in 1917 that saw such high losses for the Royal Flying Corps. The answer was the arrival of the Hispano Suiza line of engines, which resulted in fighters like the SE 5a and the Spad XIII. Arguably, at the end of the war the Germans fielded the best fighter with their Fokker D VII. The numerical advantage of the Allies however ensured they kept the upper hand in the skies. As a result of the wartime competition between the belligerents, warplanes had taken great strides. The fighter planes of 1918 bore hardly any resemblance to the aircraft that were taken to war in 1914.

Next to several variants of observation (which by eliminating surprise and being the eyes for the artillery- the greatest killer of the war- contributed to a large extent to the characteristic stalemate of trench warfare[6]), air-to-air fighting and strategic air operations, the Great War also saw the development of the fourth category of air operations, nowadays called anti-surface force air operations (ASFAO). It comprised fighting the enemy on the ground, either in direct contact with friendly forces (close air support), or at more or less distance from the front line (interdiction). During the course of the war, both sides developed specialized aircraft for this task. Some scouts were modified in order to survive in the lethal environment at low altitude, like the Sopwith Snipe that became the Salamander. These types were designated TF, for trench fighters. Some planes were especially designed for this task, like the German Junkers J 1. The German army organized dedicated units for support of ground forces, named Schutzstaffel, protection squadrons. The growing importance attached to air support for ground action is mirrored in the renaming of these units to Schlachtstaffel, battle squadrons, in the final year of the war.

All in all, compared to the total efforts during the war, the part played by the air forces was small, and could not change the course of events.[7] For airpower however, the war meant an enormous development across-the-board, from technology to organization and operations.[8] Firm foundations were laid for the

[6] Denis Winter, The first of the few. Fighter pilots of the First World War (London 1982) 14.

[7] In 1918, in an army of more than 2 million men, the Germans employed some 5,000 pilots. (Kennett, The first air war, 121) The Commonwealth forces lost some 1 million men during the war, of which some 9,000 were airmen killed in action. (Winter, The first of the few, 153.)

[8] Kennett, The first air war, 226.

century to follow. For some the new weapon promised an alternative to the massive slaughter the war had seen.

The interwar years

Founding airpower theory

As mentioned above, the Great War had a profound impact on the airpower theorists of the interwar years, whose influence stretches to the present day. Airpower seemed to promise a way of avoiding another stalemate with accompanying massacres such as seen in the Great War. People like General Smuts, under whose chairmanship a committee had studied the use of airpower for the defence of Great Britain after the German bomber attacks on British cities in 1917, envisioned great things for airpower. The Smuts report stated that "there is absolutely no limit to the scale of its future independent war use."

Today three theorists rank as classics in airpower theory: the Italian Giulio Douhet, the Englishman Hugh Trenchard and the American William Mitchell. Each from their own position they advocated the use of airpower as the way of the future. To them, war meant war as they had experienced it from 1914 to 1918. It was total war in the sense that it took a nation's total spectrum of resources. To defeat your enemy, you had to look beyond his armed forces and focus instead on his capacities as a whole. By attacking a nation's will and power to make war, a long and bloody stalemate like that seen on the Western front could be avoided, they believed. This could be achieved by airpower, which they saw as being of decisive strategic importance. Strategic bombardment of the enemy's heartland, while bypassing its armies through the air, would cut wars short. To this end, in their opinion, air forces should be organized as co-equal to and independent of the other services. To a varying degree, this position brought them into conflict with the military establishment of their day. Trenchard fared most successfully and managed to keep the young Royal Air Force in existence, at that time the only independent air force in the world. The other two were not as successful, both clashed with their superiors to the point of being court-martialled.

The bottom line of their ideas was that air forces should be employed offensively, at the outbreak of hostilities, aimed at the "destruction of a hostile nation's power to make war" in Mitchell's words. This power was viewed in a broad sense and also included the morale of the population, which was seen as embodying the enemy's will. Douhet thought that through bombardment the enemy's population could be driven to force their government to sue for peace. In his view, the bomber was the most important weapon. Notably Trenchard viewed aircraft as ill-suited for defence; bombers would "always get through". In the interwar years Bomber Command was the most prestigious arm of the RAF.

In the wars of the future, aircraft would be the central weapon. Military analyst JFC Fuller envisioned in 1920: "Fleets of aeroplanes will attack the enemy's great industrial and governing centres. All these attacks will be made against the civil population in order to compel it to accept the will of the attacker…" Words like these, read with a knowledge of World War II strategic bombardment, conjure up pictures of death and destruction on an immense scale. During the interwar years however, the theory was that bombardments would prevent long wars, thereby justifying civilian casualties. It must be noted there was still a huge gap between the theory and the possibilities of the existing aeroplanes. Accuracy, payload, reach and available aircraft were nowhere near sufficient to execute a campaign as would be necessary to achieve strategic results. At the same time radar was non-existent. This new invention would render it impossible for bombers to enter an airspace largely undetected, a premise on which the whole theory was based. Bombers were seen as unstoppable. Radar however, made air defence possible. Still, the beliefs were firmly held and solidly embedded in the doctrines prevalent at the beginning of the Second World War.

The Second World War

The theories put to the test

Blitzkrieg

The Second World War commenced with the German invasion of Poland on September 1st, 1939. The new Blitzkrieg concept the Germans applied was successful as the backbone of the Polish defence was defeated within little more than a week. Before the month was out, Poland was defeated and carved up. The Soviets administered the final blow from the east, while the Luftwaffe bombed Warsaw. Airpower played an important role in this success for Germany. The combination of Junkers Ju 87 Stuka dive bombers used in close coordination with the armoured spearhead of the Wehrmacht, proved very successful in forcing a swift operational decisions. It was a first, both for armour and air support on this scale. In May 1940 it was used again with great success in France and the Low Countries. The Stuka had become the archetypical symbol of the Blitzkrieg.

The Battle of Britain

The Battle of Britain in the summer and autumn of 1940 was another milestone in the history of air warfare. Here, Germany tried in vain to subdue Great Britain with airpower alone. In theory the German air effort intended to pave the way for an invasion of England. The preparations for operation Seelöwe however, cannot be qualified as having been serious. Germany lacked the means to execute a sea-borne invasion; Hitler thought the British would choose to seek some kind of understanding with Germany after the defeat of France. He did not want to fight Great Britain. He set the Luftwaffe loose however, Hermann Goering boasting that his pilots would bring England to its knees all by themselves. Initially, the Germans attacked the British integrated air defence system in order to establish air superiority. A host of factors contributed to the lack of success in the time spent on attacking this system. The Germans had overconfidence in their own capabilities, reinforced by the successes against Poland, Norway, France and the Low Countries, and underestimated British strength and aircraft production. At the same time they failed to accurately assess the British air defence system.[9] Furthermore, the German aircraft lacked the capabilities necessary for the job in hand. The Luftwaffe only had medium bombers while the Stuka proved too vulnerable to use. The fighters lacked range, which left most of England out of reach for escorted bombers.

In frustration, the Germans turned their attention to British cities. The British however were not to be coerced. In fact, the shift of the German efforts away from the hard-pressed air defence was a welcome breather for Fighter Command. The German daylight raids on English cities soon proved too costly, resulting in the next shift in operational strategy: the bombers would fly at night. This period is

[9] Richard Overy, Goering (pb London 2000) 170.

known as the Blitz. Stubbornly the British withstood the pressure. The British themselves came a long way in countering the theory that strategic bombing of the civilian population would cause enough pressure to force an adversary to submit.

The Battle of Britain is a perfect example of a counter-air operation; the German campaign being the offensive (OCA), while at the same time the British efforts constituting the defensive variant (DCA). The Germans had lost this battle when they turned their attention away from the British fighter arm; the offensive counter air campaign was stopped short. Whether it could have been successful is a matter of conjecture.

Strategic bombing

For all the interwar theorizing on the strategic use of airpower, it was initially largely used tactically, as supportive to ground combat, primarily and with great success by the Germans. With the exception of the bombings of Warsaw and Rotterdam, the first year of the Second World War did not see any use for the air arm in the strategic sense. This can be seen as particularly odd when it concerns Great Britain, whose Trenchard doctrine was the most elaborate doctrine on the strategic use of airpower in the world. When the war broke out, Bomber Command was hardly used for well over a year. When London did start to employ the bomber force, it turned out they were far too vulnerable to survive during daytime. For the remainder of the war, Bomber Command would have the bellies of their bombers painted black for night operations. This caused huge problems as it turned out to be exceedingly difficult to find the targets. Herein lay the origin of the practice of large-scale carpet bombing of German cities. It was simply not possible to hit smaller targets. During the first years of the war, navigating to the right city was already considered accuracy. In practice, the American contribution to the bomber offensive against Germany from 1942 onwards, which was flown by day and aimed at more accurately hitting the enemy's industrial capacity, hit German cities as hard as Bomber Command did. The Combined Bomber Offensive was now a round the clock affair. A very important sideline to the bombing offensive was the battle of attrition in the skies over occupied Europe. In this battle, the escorting fighter planes that from the fall of 1943 were able to accompany the bombers to Berlin and back, established and maintained air superiority for the remainder of the war.

Warsaw was the first in a long line of cities to follow. The strategic bombing of cities by both Axis and Allies resulted in civilian death and destruction on a scale never seen before. The Allied Combined Bomber Offensive against Germany and the fire bombing of and nuclear destruction wreaked upon Japanese cities, are sources of controversy to this day.

To be sure, the immense effort put into strategic bombardments contributed to the defeat of Germany and Japan. Much material and human effort that could otherwise have been employed at the fronts had been put into air defence. In the end it was the accumulation of punishment from the air combined with the advances on the ground that did the Axis in.

Airpower applied on a grand scale did not prove to be a shortcut to victory. It could not prevent a long drawn out conflict from having to be fought on the ground and the high seas, as was hoped and believed by airpower theorists like Douhet and Trenchard. It added one more dimension of total-ness to war. In the First World War civilian populations had suffered the consequences of their countries being at war. In the second, they were in the frontline themselves.

The Cold War

After the Second World War, the (western) allies demobilized on a huge scale. For the US this was made easy as it enjoyed the monopoly on the atomic bomb. At the same time the Cold War started to develop. The atomic weapon constituted the strategic deterrent that was the safeguard in a confrontation with the USSR. The USAF was the arm that wielded this so-called big stick. The bomber force of Strategic Air Command was the nation's most important weapon. Later the bombers were supplemented by intercontinental ballistic missiles, first in underground silos, later also on submarines. During the decades of the Cold War, the confrontation between East and West, between NATO and the Warsaw Pact, was the leading factor in security and defence matters for both sides. The armed forces were geared to this. The competing other side, or the perception thereof, was the yardstick applied when moulding the armed forces .

The confrontation between NATO and the Warsaw pact never turned into a shooting war. Wars were waged in other theatres however, some closely connected to the worldwide ideological confrontation. The wars in Korea, Vietnam, and the Middle East as well as the Falklands war are touched upon below.

Korea

When war broke out on the Korean peninsula on June 25, 1950, it turned out that the US had demobilized to an almost dangerous extent. The Americans were barely able to help the South Koreans hold on to the peninsula in the face of the North Korean onslaught. Only five years after World War II, the US armed forces almost had to start from scratch again. Having to fight a limited conventional war in a time when the nation's military strategy was directed towards a worldwide nuclear confrontation with the USSR and its allies caused considerable difficulty. After the first year that saw the front line moving up and down the peninsula several times, the war lasted two more years until a ceasefire was reached, which left the front line almost exactly in the same position as before the war. The outbreak of the war marked the starting point of the irreversible militarization of the Cold War.

In the context of the early Cold War years, the USAF had its focus on the competition with the US Navy over the question of which service was to wield the nation's atomic sword in the struggle with the Soviet bloc. In the standing war plans, air force strategic bombers were to destroy the USSR's capacity to wage war by bombing industrial cities. Doctrine had not changed since the last war. In the event, this strategic role could not be played in Korea. Other tasks, most notably the tactical use of airpower now in demand at the Korean front, had been completely neglected. Like the other services, the air force had to improvise during the first months and make do mostly with aircraft of World War II vintage rushed to the front. Until the Chinese intervened in the war when America and its allies came too close to their borders in November 1950, this was not a great problem, as the North Koreans did not have much of an air force. B-29 bombers that could be spared from their main task of deterring the USSR could fly and bomb unmolested by day.

The war took on wholly new dynamics with the Chinese intervention in November 1950. On land it meant the total reversal of the advance into North Korea up to the Yalu, the border river with China. This would only be redressed again after half a year of heavy fighting, which saw the frontline around the old inner Korean border again. In the air, the allies were surprised by the MiG-15, the new cutting edge Soviet fighter. The B-29s were soon driven into the night. The response to the MiG was the F-86 Sabre, of which eventually some 150 were sufficient to gain and maintain air superiority. This must however be viewed as a 'political air superiority'[10] as both sides restrained themselves, having no interest in widening the war. The Chinese stayed on the defensive in order to deter an American incursion into Chinese airspace. Their planes never embarked on a sustained offensive campaign into Korean airspace. Thus, this relatively small force of Sabres was able to keep the upper hand in Korean airspace against a force of MiG-15s that was to grow to a strength of 830 by the end of the war. Over the years, the Sabres scored 792 MiGs against a loss of 78 of their own. This huge difference is credited to the American pilots' superior experience and the armament of the Sabre, which was better suited to combating fighters.

The success of the fighter war, on which most media attention was focused, was counter-balanced by the lack of success of strategic bombardment and interdiction. This was cause of much concern among air force leadership, as their newly-won independent position[11] was based on a theory that now seemed to buckle under the stress of reality. Strategic bombing of the sources of North Korean war-waging capacity was out of the question as these were in China and the USSR. Bombing of the Korean population, which was done on a large scale, could never be decisive in any way as it meant little pressure on China and the USSR, who kept the communists upright. As for interdiction, the battlefield was never isolated despite immense efforts from the air. The Chinese and North Koreans proved amazingly effective in battle damage repair. They furthermore needed far less supplies to be able to keep on operating than a western army. In the event, the interdiction effort meant "trading coolies for aircraft at an uneconomic rate".[12] In the end, the country was ravaged by the war, with an unclear number of victims among the civilian population, exceeding one million. The Chinese lost some 900,000 men. The war was limited as far as the battle space was concerned. The Korean population however experienced very little limitation in death and destruction. The country was to suffer even more than Japan during the Second World War. As for the price the air component had to pay, this was high as well. Allied losses amounted to 1886 allied aircraft; 147 by MiGs, the rest through ground fire or accident. To be sure, the air forces contributed to not being defeated, playing a significant role in the precarious first weeks, in which a foothold on the peninsula was maintained. There was however a gap between raised expectations with respect to airpower and the reality of the war where in the end it failed to succeed in dislodging the enemy from his position, either operationally or strategically.

Vietnam

Another limited war, in a geographical sense at least, was fought in Indochina, where the US fought the insurgency against the South Vietnamese government by the Vietnamese communists, the Viet Cong (VC). The insurgents, backed by North Vietnam, fought patiently while the US tried in vain to get a grip, in the process expanding their military presence to more than half a million men in 1968. In the end the VC had more staying power, fighting in its own country against unpopular governments. The US efforts could not uphold this regime, the massive US presence in its totality increasingly doing more harm than good in a situation that could not be solved by military means.

The air war in Vietnam was waged in two theatres, each with its own characteristics. In the south air assets supported US ground troops, depending in a large part on Army aviation for their mobility. Air manoeuvre operations, executed with a

---

[10] Callum A. MacDonald, Korea. The war before Vietnam,(New York, 1987) 246.
[11] The USAF became an independent service in 1947.
[12] MacDonald, Korea. The war before Vietnam, 239.

combination of transport helicopters and helicopter gun ships, was a new phenomenon of the war. In the north where no US ground forces were present, the operations had a strategic character, aimed at putting pressure on the North Vietnamese in order to stop the support of the communist efforts on the ground in South Korea. This was complemented by air interdiction in an effort to stop the flow of supplies from north to south, which included bombing of the so-called Ho Chi Minh trail that ran through neighbouring Laos and Cambodia. As in Korea, both the strategic and interdiction efforts cannot be said to have been successful.[13] The US pilots faced a tough job in the skies over North Vietnam, having to contend with MiG-17, -19 and -21 air defence fighters, complemented by the densest concentration of ground-based air defence among which were the most modern Soviet SAM systems ever fielded up until then. Throughout the years the air war took a toll. American aircraft and crew were lost, aircrew were taken prisoner, some of them ending up being a POW for years.

The air war in Vietnam was an enormous effort for the US. Most assets in the inventory were used, from the B-52 designed as a strategic nuclear bomber to obsolete planes of World War II design like the A-1 Skyraider and B-26 Invader taken out of storage for use in the counter-insurgency role. Even C-47 Dakotas were modified into gun ships. Aircraft flew off US Navy carriers in the South China Sea, over North Vietnam from Yankee Station, in the south from Dixie Station. To support the war in the south with their fighter bombers, the US Air Force and US Marine Corps had several bases inside the country, whilst the USAF also flew against the North from Thailand. Next to this there were B-52s flying off Guam in the Indian Ocean. Tankers and other support aircraft supplemented the fighter and bomber planes from places sometimes as far off as Okinawa.

In order to keep the war limited, the air efforts against North Vietnam and outside South Vietnamese borders were strictly regulated. The centralized direction went up to the president himself, a subject of controversy ever since. The US eventually managed to extract itself from Vietnam in 1973, after having increasingly handed over responsibility on the ground to South Vietnamese forces in a process called Vietnamisation. The last phase of the war saw a success for airpower when the strategic bombing campaign Linebacker II succeeded in coercing the North Vietnamese back to the Paris negotiating table. These negotiations bought just enough time for the US to withdraw with a face-saving delay before the North Vietnamese were triumphant in 1975 in what was basically a Vietnamese civil war.

The Middle East

The Middle East qualifies as one of the more volatile regions of the world. Since its inception in 1948, the state of Israel has had to fight for its survival against its neighbours. Wars were fought in 1948 (War of Independence), 1956 (Suez), 1967 (the Six Day War) and 1973 (The Yom Kippur War). In 1982 Israel invaded southern Lebanon in an attempt to attain security for northern Israel by chasing the Palestine Liberation Organization out of Lebanon. Although Israel firmly established itself after the militarily victory in the first four wars (those against the neighbouring states), disputes over the exact delineation of borders continue to this day. Likewise, hostility from several extremist organizations like Hezbollah in southern Lebanon and Hamas in the occupied territories, Gaza and the West Bank, proxies being used to exert pressure by countries like Syria and Iran, remain a problem. Airpower has played a significant role in the wars between Israel and its

[13] James J. Wirtz, "Vietnam war", in: Richard Holmes, ed., The Oxford companion to military history (pb. Oxford 2003) 956: "Bombing the jungle is probably the least cost-effective form of warfare ever devised by man."

neighbours. The Israeli opening of the 1967 Six Day War surely ranks as one of the best OCA campaigns ever. In a matter of hours the entire Egyptian and Syrian air forces were destroyed on the ground, after which the Israeli ground forces could swiftly take the Sinai Peninsula, the Golan Heights and the West Bank, including Jerusalem. In 1973 the Egyptians surprised the Israelis by their use of new Soviet SAM systems, which constituted an effective umbrella under which they could successfully dislodge the Israelis from the entire east bank of the Suez Canal.

Another remarkable example of the use of airpower was the destruction of the Iraqi nuclear reactor at Osirak by an Israeli precision strike in 1981. The strike set the Iraqi's quest for nuclear weapons back by years.

In the symmetrical conventional wars with neighbouring states, Israel's superior air force has played an indispensable role. In the confrontation with the extremist organizations Hamas and Hezbollah however, it is much more difficult to employ aircraft in a useful way. In the occupied territories Gaza and the West bank, much intelligence gathering, surveillance and reconnaissance work is done by air assets. Occasionally high value targets are hit by either helicopters or fighter bombers. Assets designed to combat tanks, like the Apache or Cobra helicopters, now take out single civilian vehicles or houses in order to eliminate extremist leaders. The fight with southern Lebanon-based Hezbollah, flared up dramatically in the summer of 2006, after the organization kidnapped two Israeli soldiers. In an effort to punish Hezbollah, or even to coerce it into returning the soldiers and at the same time to stop firing Katyusha missiles into northern Israel, the southern half of Lebanon was hit heavily from the air. To what extent these actions, either pinpoint or large-scale, contribute to longer-term success is unclear, however skilfully they were executed on a tactical level.

The Falklands war

The Falklands war of 1982 is interesting in many respects. It has been called "a freak of history"[14], being the last colonial war Great Britain has fought. For the British it was fought at an astonishing distance from their homeland, mainly with and off ships. The Falklands are situated in the South Atlantic, 8,000 miles from Great Britain. The nearest base was Ascension Island, roughly halfway to the islands. Sea power therefore was at a premium during this conflict. The Royal Navy had to establish sufficient sea control to be able to conduct operations. In the event the Argentines challenged this by airpower.[15] The only way to counter this was by also employing airpower. The first British move in this respect was Operation Black Buck on 1st May. In this memorable operation, an RAF Vulcan, a venerable strategic bomber about to be retired, bombed the runway of Port Stanley airfield on East Falkland. To reach its target it had to be refuelled by 15 RAF Victor tanker planes, which also had to transfer fuel among them in order to enable the Vulcan to cover the distance of the Ascension – Falklands round trip, a 16-hour flight with 17 fuel transfers in all. In the event, one of the 21 bombs the Vulcan dropped pierced the 4,100 ft runway. Although the length of the runway was hardly suitable for the Argentine Skyhawks, Daggers and Super Etendards, the British wanted to exclude this possibility altogether.[16] Perhaps even more important than this counter air

[14] Max Hastings and Simon Jenkins, The battle for the Falklands (pb London 1983) 7.

[15] The Argentine Navy contributed with its air assets from land bases. After the loss of the Belgrano on 2 May, they chose not to engage their surface fleet, including their one aircraft carrier. The cruiser was torpedoed early in the conflict by the submarine HMS Conqueror, costing the lives of 368 Argentines . It was a severe blow.

[16] Martin Middlebrook, Operation Corporate. The Falklands War, 1982 (London 1985) 115. Seven Black Buck missions were mounted, of which three could not be carried through. Two missions were flown

aspect, was the signal the operation sent to the Argentines. The British meant business and had demonstrated its reach.

The capital ships of the task force were the aircraft carriers Hermes and Invincible. The 20 Sea Harriers on board were in for a very tough job.[17] As the airbases in Argentina were beyond reach, the confrontation would have to be in the Falklands theatre. From the moment the task force reached the area of operations at the end of April, to about a week after the amphibious landing at San Carlos on 21 May, fierce battles were fought between Argentine aircraft and British ships and fighters. The British tried to establish air superiority in order to facilitate an amphibious landing on the islands. In the event, this was not achieved. The landing on May 21 did proceed however, weather conditions and a good deal of luck contributing to its success. Both sides paid heavily in this war of attrition in which ultimately the British prevailed as the Argentines could not prevent the amphibious operation and later failed to spoil the landing itself or the subsequent march on the capital Stanley.

From the air warfare angle, the war in the South Atlantic has many interesting aspects. Both sides fielded western weapon systems, in the time of the Cold War when British attention was fully focused on the Warsaw Pact adversary. The dangerous Exocet anti-ship missile, which accounted for the sinking of HMS Sheffield and Atlantic Conveyor, was a French product. Asked later why the British were so vulnerable to this threat, a high-ranking officer said: "The Russians have no Exocet."[18] Even though the Argentines only had literally a handful of the missiles at their disposal, the threat of air attack caused the British carriers to operate, after the sinking of Atlantic Conveyor, at a considerable distance east of the islands, as much as 200 miles. This seriously restricted the availability of the Sea Harrier combat air patrols (CAP's) over the amphibious task force. It was one of the reasons why the British proved unable to establish air superiority before the amphibious landing took place. This in turn led to a severe loss of ships for the task force.

The war in the Falklands illustrated the importance of airborne force multipliers, which today are taken for granted in air warfare but were absent in this conflict. The British lacked airborne early warning, "the single most critical British deficiency of the war".[19] It was another reason why they failed to establish air superiority. The Argentines applied the tactic of flying the last 50 miles' distance hugging the sea, so that they approached their targets mostly undetected until seconds before dropping their bombs. By staying below the range of the radar, the British ship-borne radar-guided surface-to-air missiles had little time to react. The Harriers, which ultimately accounted for 31 kills, engaged most of the Argentine aircraft after they had made their attacks. The Argentines on the other hand were severely hampered by the lack of sufficient air-to-air refuelling capability. The fighter-bombers had to operate from mainland bases, giving their pilots not much time to manoeuvre once they were over the area of operations. In the event, the Argentines did not engage in dogfights with the Harriers, they concentrated on attacking the ships and then had to return to the mainland. Confrontations proved "an entirely one-sided affair" in favour of the Sea Harriers, which made most kills with the very reliable and deadly heat-seeking AIM 9L Sidewinder missile, although several kills were made with the gun.[20] Because of the scarce resources and having to operate without early warning,

with Shrike anti-radiation missiles, while the final one dropped bombs fused for air bursts on the airfield, in conjunction with the final stages of the battle for Stanley.

17 Later on these were supplemented by 8 more Sea Harriers and 8 Harriers GR3 of the RAF. Eight of their number made a record length, single seat aeroplane, air-to-air refuelled flight to the Falklands.

18 Hastings and Jenkins, The battle for the Falklands, 189.

19 Hastings and Jenkins, The battle for the Falklands, 168.

20 Ibidem, 239.

the numerically superior Argentine air force was able to "swamp" the British air defences,[21] ultimately resulting in the loss of four warships, and some ten others being put out of action. These losses could have been even worse, as several bombs that were on target failed to explode.

The impact of the Argentine air threat on British operations, in other words: the consequences for the British of having to operate without air superiority, shaped the further course of the war considerably. When Atlantic Conveyor sank, three Chinooks and six Wessex helicopters went down with it. The plan for the advance on Stanley centred on these Chinooks. Now it had to be done on foot. The air threat also seriously hampered the logistic build-up of the bridgehead around San Carlos. The supply ships had to unload and shuttle between San Carlos Water and the safety of the Task Force ships east of the islands, all in the available hours of darkness. It resulted in the build-up being too slow for policy makers in London, who then forced Brigadier Thompson to attack Goose Green, an Argentine position he wished to leave aside in his own plan, as it was operationally of no importance.

The events of the Falklands war illustrate the importance of air superiority, a maxim in airpower thinking ever since its inception. Not having it seriously hampers a campaign. Brigadier Thompson later said "If the air threat had been properly appreciated, I don't think this whole venture would ever haven been undertaken."[22]

From the end of the cold war to the post 9/11 world

New developments and appreciation boost airpower usefulness and practicality

Desert Storm

Someone not as well known as, say, H. Norman Schwarzkopf, played a conspicuous behind the scenes role during the conflict in the Persian Gulf, which started little more than half a year after the fall of the Berlin wall.[23] His name was John A. Warden III, then a colonel in the USAF. At the time he was a planner at a Pentagon section called Checkmate. He and his staff provided General Schwarzkopf, commander of Central Command[24], with a strategic air campaign plan for use against Iraq. Air assets being the first to be deployed in reaction to Iraq's invasion of Kuwait in August 1990, they were to deter any further advance by the Iraqis, which was initially feared. Eventually, the Coalition troops build-up amounted to some 600,000 ground troops. Checkmate's initial air plan was to develop the first part of the overall air campaign, which constituted the first phase of Operation Desert Storm. This air campaign in preparation of the ground advance into Kuwait and southern Iraq in order to liberate Kuwait, lasted 38 days. The counter air part of the air campaign gained air supremacy in a matter of days, resulting in a near complete freedom of movement in the air, which enabled the coalition air assets to concentrate on preparing the battlefield for the ground advance. The "thunder and lightning of Desert Storm" weakened Iraq as a strategic entity, and wore down the Iraqi armed forces so severely that the ground war lasted only some 100 hours.

---

21 Ibidem, 242.

22 Hastings and Jenkins, The battle for the Falklands 212.

23 David Halberstam, War in a time of peace. Bush, Clinton and the generals (New York, 2001) 47.

24 The US Combatant Command that comprises the Middle East and the Horn of Africa that was to lead the military campaign to oust Saddam Hussein from Kuwait.

Several key technological developments had reached the stage where operational use was possible. For the first time F-117 Nighthawk stealth fighter bombers were used on a large scale. Their stealth characteristics enabled them to put their payload of two laser-guided precision bombs to good use in environments that would have proved deadly to conventional fighter bombers, in this case downtown Baghdad. Among the first targets hit were key nodes in the Iraqi integrated air defence system, thereby blinding it and enabling the Coalition air forces to secure air supremacy in a matter of days. Other airborne assets of great value were the two E-8A JSTARS air-to-ground surveillance planes, which were rushed to the front while actually still in development. The Iraqi desert landscape was perfectly suited for detection of ground movement by the plane's synthetic aperture radars.

A noteworthy action in the offensive counter air campaign was the attack of Task Force Normandy. This force consisted of eight US Army Apache attack helicopters, guided by 2 USAF CH-53J Pave Low special operations helicopters. They executed the first combat action of Desert Storm when they destroyed two early warning radars, thereby making room for the first planes to enter Iraqi airspace.

The air campaign, dubbed "Instant Thunder" by Warden (a variation on the Vietnam war air campaign called Rolling Thunder) can be seen as the pinnacle of what could be achieved by airpower in a symmetrical conflict. It can, at least partly, be seen as a vindication of airpower theory as it originated in the interwar years. Technology had finally made it possible to really employ airpower in a strategic sense. The end of the Cold War and the confrontation with a large but manageable foe provided the opportunity to do so. Although still relatively small scale, the employment of stealth aircraft, precision-guided munitions (PGMs comprised some 5% of air delivered ordnance) and JSTARS had already made a difference on the battlefield. It pointed the way to the coming years in which developments further accelerated.

The war liberated not only Kuwait, in a sense it also freed airpower from the fairly confined roles it had grown into during the preceding decades. In the USAF, strategic airpower comprised the assets of the Strategic Air Command (SAC), with its prerogative to deliver the nation's nuclear arsenal to the enemy's homeland. Outside this Cold War setting, this apparatus was pretty much useless. In the wars where SAC bombers were used, this was always in a different manner. The Air Force's other major combatant command, Tactical Air Command, very much focused on combating the enemy's armed forces. In the post Cold War setting, faced with adversaries of a different order, the dichotomy of strategic versus tactical had to be revised from an operational point of view. Strategic operations were no longer synonymous with nuclear operations; strategic ends could be attained with "tactical" means, and vice versa. Small fighter bombers, like the F-117s equipped with only two laser-guided, conventional, bombs (LGBs), attacked strategic targets. Cruise missiles developed to deliver nuclear bomb loads, hit strategic targets armed with conventional explosives. SAC B-52s carpet bombed dug-in Iraqi troops in preparation of the coalition's ground offensive. F-111s designed to penetrate Soviet airspace at low level for the delivery of nuclear bombs, proved the greatest tank killer when equipped with PGMs.

Warden's strategic thinking made a very important contribution to the evolution of airpower, by arguing that the military must not mainly (in practice often only) think about combating an adversary's armed forces. Stated simply, one should start thinking about a conflict from the strategic level, by considering an adversary in its totality as a strategic entity that can be analysed as a system. When analysed correctly, it is possible to assess fairly accurately what is required to make the system do whatever it is you want it to do to reach your strategic objectives. If this means using military force, military planners should also consider that system as a whole, and not only its military power. By using this idea as a starting point, he paved the way for a new appreciation of the bulk of the air force, formerly mainly envisaged in a tactical role influencing the surface fight between opposing armed forces. The USAF organization reflected this change in perception, when in 1992 SAC and TAC were merged into one Air Combat Command; a telling illustration.

Warden was not the first to have come up with ideas regarding the strategic use of airpower. These ideas were rooted in the theory of the classics of the interwar years, and in the concepts that were further developed by the Air Corps Tactical School during the Second World War. It took someone like Warden however, to revitalize the old ideas and adapt them to his time, when technological developments enabled the use of airpower in a precise manner. Air forces across the globe recognize Wardens thinking as very useful. His five ring theory is both incorporated and works as catalyst in much airpower and wider strategic thinking since its formulation.[25] In this way he became even more influential than as a planner at Checkmate, where he laid the ground work for the successful air campaign of Desert Storm.

As mentioned, the Gulf War turned out to be the exception rather than the rule. Civil wars like those in former Yugoslavia, or in many countries in Africa like Somalia and Rwanda, proved to be much more the rule in the post Cold War era. Taking a position with respect to these internal conflicts proved much tougher than determining a position regarding Saddam's invasion of Kuwait.

Former Yugoslavia

When the world's attention was focused on the Gulf region and America's military victory over Iraq in early 1991, the Balkans erupted into yet another of its conflicts. After several rounds of secession starting in 1991, Yugoslavia collapsed, resulting in years of civil war. With the collapse of communism in the USSR and among its allies, the essentially communist government of Yugoslavia could also no longer muster enough ideological power to keep the country united. Hindsight even shows the bankruptcy had already been in the making for more than a decade. Marshall Tito, with his considerable political virtuosity and prestige built up during World War II, in which his partisan movement had won the civil war that raged during the years of axis occupation, had been able to keep the republics constituting Yugoslavia[26] together. After his death in 1980, no successor ever came close to being able to do the same. When communist ideology no longer formed the basis for power, politicians turned to nationalism as a substitute. Populations of differing ethnic origins - Serb, Croat, Bosnian Muslim and Kosovo Albanian - were divided to the point of taking up arms against each other.

Deny Flight and Deliberate Force

Civil war erupted in Croatia in the summer of 1991 after the Croatian government declared its independence from Yugoslavia. The same happened in Bosnia-Herzegovina a year later. In response, the international community sent in the UN peace keeping force UNPROFOR (UN Protection Force). Added to this, NATO conducted air operations in support, with several NATO nations contributing to both military operations. From 1993 to 1995 the air operation Deny Flight was executed. Initially this was aimed at maintaining the no-fly zone the UN had

[25] Col. John A. Warden III, USAF, "The enemy as a system", in: Airpower Journal (Spring 1995) 41-55.
[26] The six republics were Slovenia, Croatia, Bosnia-Herzegovina, Serbia, Montenegro, and Macedonia.

declared over Bosnia. To this end, NATO air defence fighters flew combat air patrols over Bosnian airspace, supported by NATO E-3 Sentry AWACS aircraft, airborne warning and control stations. The Sentries monitored every air movement and directed the fighters to any bogey they detected. Later on, with UNPROFOR peacekeepers increasingly encountering difficulties on the ground, and in an attempt to gain more leverage over the warring factions, it was arranged that the peacekeepers could call on the fighters for support. Most participating fighters like the F-16, F-18 or Harrier, because they actually were fighter-bombers, could fairly easily be armed with bombs, next to the air-to-air missiles they needed for their original Deny Flight task. Several encounters took place during this operation. On 21 November 1994 , after a violation of the no-fly zone, a Serbian airbase, actually situated in Serbian-held territory in Croatia, was bombed. A group of planes forming a COMAO package (composite air operation) struck the base Udbina at several key points, disabling it for some time. On another occasion, four former Yugoslav air force Galebs were shot down after violating the zone. On several occasions Deny Flight aircraft gave air support to UNPROFOR soldiers. Sometimes to the dismay of the airmen involved in the NATO air effort, very little could be done in this respect, given the precarious situation on the ground. Until the summer of 1995, UNPROFOR soldiers were vulnerable because they were scattered across the country, some military observers were deployed almost on an individual basis. In short, they were dependent on the whims of the warring factions. Any NATO air support prompted action against soldiers on the ground, as the hostage taking of UN personnel after air strikes on Bosnian Serb ammunition storage sites near Pale in May 1995 showed.

Things changed when the situation on the ground started moving in the spring of 1995. In Croatia, the Croatian army had reached a stage where it could stand up to the Croatian Serbs who held about a third of Croatian territory, land that by the way had been inhabited predominantly by Serbs for more than a century, the so-called Krajina. They had evicted non-Serbs, in a process of ethnic cleansing that started sometime in 1990 when the Croatian government in Zagreb began to voice its desire to secede from Yugoslavia because the Croat-Serb population felt increasingly threatened. Zagreb had wanted to retake the Krajina ever since. In the first years of the civil war in the former Yugoslavia the Serbs had the upper hand however, as they had enjoyed the support of the former federal government, and with it the support of the better part of the Yugoslav armed forces. Now, four years on, both the Croatians and the Bosnians had become stronger, with international opinion (mainly American) favouring their predicament over the Serbian side of the story. Furthermore, the two ethnic entities had come to an agreement to stop fighting each other and concentrate their efforts against the Serbs. In May 1995 this resulted in a combined offensive. In a matter of days, the Croat Operation Storm blitzed the Serbs out of Krajina, evicting the civilian population in the process. In Bosnia the tables were turned as well, the continuing pressure pushing back the Serbian lines. On the other side of Bosnia, the Serbs took the enclaves of Srebrenica and Zepa and threatened Gorazde, places the UN had declared safe areas. UNPROFOR troops were deployed in these areas in an effort to deter Serbian attack, causing a certain fragmentation of UN troops, one of the factors hampering air efforts. It soon became clear that the taking of the safe areas was accompanied by indiscriminate murdering on a large scale. This total disregard for human life and defiance of the international community, endangering its credibility, proved too much for the governments who had looked on for years, unable to change the deplorable course of events in any significant way. The mustering of political will to act combined with circumstances on the ground to bring military force to bear: Serbs were on the defensive and in retreat, UNPROFOR was no longer scattered all over the country, in fact was even being reinforced in order to also combat Serbian forces on the ground. In August 1995 operation Deliberate Force was launched, an air campaign aimed against the Bosnian Serbs in order to force them to take part in talks to end the war in Bosnia. The campaign was to last some three weeks, short but significant. At the same time diplomatic pressure was directed against the Croats and Bosnians to halt their offensive and join the talks. Now was the time. Going on would endanger their gains of the last months and lose them the international support they had. Finally the Bosnian Serbs gave in, after even having been threatened by Serbian president Slobodan Milosevic that they would lose his support if they continued defying the world. Their stubbornness had become a danger to the Serbian cause (and Milosevic's position) in the whole of former Yugoslavia. So, in the summer of 1995, a host of factors contributed to the war in Bosnia coming to an end, firstly by cease fire, later in the year by agreement reached in Dayton, Ohio. In this complex context, airpower had played an important part. Both Deny Flight and Deliberate Force were significant for NATO, as they were the first military operations executed by the Alliance since its establishment in 1949. War on a larger scale was to follow before the end of the decade was out. Again it was to involve airpower.

Allied Force

A further round of war in former Yugoslavia took place in Kosovo, a province of Serbia. The fight between the Kosovo Albanian insurgents and the Serb government had in 1998 turned into a situation with severe implications. NATO deemed the situation as potentially dangerous if left un-addressed. The alliance moved itself into a position in which it was finally forced to respond militarily. While increasing political pressure on the Serbian government, it had threatened with military force. America and its allies wanted to avoid another long drawn out conflict like the one in Bosnia earlier. After the success of Dayton, which was perceived as being for a large part brought about by Deliberate Force, military intervention by air seemed to offer a possibility to stop such an escalation short. By so doing the Alliance put its credibility on the line, in the end a consideration even more important than the concerns regarding international instability on the Balkans.

In the night of 23 March 1999, the Alliance started operation Allied Force, an air campaign that was to coerce the Serbian government into submitting to NATO's demands regarding Kosovo, basically saying Serbia had to leave the province, to be replaced by an international military presence. It took 78 days before the total pressure on Serbia forced it to do so. Allied Force was (and remains) a unique phenomenon in that it was an air-only operation. NATO did not want to put ground troops at risk, could in fact not even afford casualties. This was a part of the assignment SACEUR had received. NATO's air forces were to succeed without casualties on its own side.

NATO had initially planned for a short series of air strikes, expecting the Serb government to be more forthcoming once the threat proved serious. The Serbs would not budge however. In fact they even responded by stepping up the campaign in Kosovo, causing large numbers of refugees to pour into neighbouring Macedonia and Albania and increasing the pressure on the Alliance. The question now was who could sit out this period of tension the longest: Serbia or the Alliance with its internal diversity. The air campaign can be seen as having consisted of two parts. On the one hand it was a strategic air campaign against Serbia. Within limitations, in order to prove to the Serbian population that not they, but only the regime was targeted, targets were selected in order to exert growing pressure on the regime. Targets comprised for instance bridges, refineries and electricity plants, which were temporarily put out of order by causing them to short circuit. A TV transmission tower was also hit in an effort to curb the government's influence. On the other hand, NATO tried desperately to engage Serbian forces on the ground in Kosovo, in an effort to halt their actions against the Kosovo Albanian population. This proved impossible. The air campaign was seriously hampered by the way the Serbs used their ground-based air defence (GBAD) assets. They operated in an active mode only very sparingly. In combination with good camouflage this prevented detection. As a result, the GBAD remained a threat to NATO aircraft, even forcing them to operate above 15,000 feet (5 km). In combination with not having own forces on the ground, this made it all but impossible to seriously counter Serbian ground action. Unfavourable weather conditions were also a hampering factor, considering the dependence on electro-optical means for observation and targeting.

In the end, after 78 days of campaigning, the Serbs gave in. The air campaign must be seen as one of several reasons they did so. The Serbian government had become politically isolated. Even Russia no longer supported it. The military force that was to enter Kosovo after the Serbian withdrawal had been made more acceptable by placing it under UN auspices, and by adding a substantial Russian component. Still, in the end Allied Force was seen as proof that airpower can achieve strategic success by itself. Whether it is the exception to the rule or the vindication of airpower theory is still debatable.

The airpower decade

The 1990s have been called "the airpower decade".[27] In this time, several successes could be chalked up. The decade started with the Gulf War in which the impressive air campaign all but defeated the Iraqis, or so it seemed, catching the public's eye during the air campaign weeks. It appeared to be the factor that after years of muddling around in the former Yugoslavia finally put a stop to things, first with Deliberate Force, later with Allied Force. With the latter, airpower even seemed to have been the only deciding factor. Finally airpower had succeeded in achieving the strategic results the classical theorists had foreseen. The means had caught up with the theory. Now, through precision guidance, bombs could be put on target. The percentage of precision-guided ammunition, first employed in Vietnam, had risen from some 5% during Desert Storm to 40% during Allied Force. Consequently, airpower proved an appealing weapon for politicians, as it exposed few own personnel to danger. Own losses, which nowadays are almost unacceptable to a critical public, can be avoided by minimizing the employment of ground forces. The 20th century thus ended on a note of optimism for airpower enthusiasts. The future proved to be even more promising because of the increasingly swift development of information technology. Weapon delivery is precise and increasingly less dependent on weather conditions when adding satellite guidance to electro-optical guidance. Sensors are improving in quality, enabling ever better target acquisition. The next level is reached by networking all sensor and shooter platforms, as coined in the concept of network centric warfare (NCW), or network enabled capabilities (NEC). Real-time shared situational awareness combined with very short (minutes maximum) sensor-to-shooter time -the technology to deliver ordnance on target- has never been better, resulting in unprecedented tactical superiority. The first years of the new century were to show however, that this is no

[27] Sir Timothy Garden, cited in: Frans Osinga, *"Airpower in het postmoderne tijdperk. Revolutie in de lucht"* in: <u>Militaire spectator</u> jg. 172, nr. 6, 2003, 338.

guarantee for success on higher strategic levels, least of all in today's asymmetrical conflicts.

After 9/11: Beyond the airpower decade

Operations in Afghanistan: OEF and ISAF

The post-Cold War era entered a new phase after the terrorist attacks on the World Trade Centre in New York and the Pentagon in Washington on September 11, 2001. Responsible for the attacks was the terrorist organization Al Qaida, led by Osama bin Laden. In retaliation, America declared war on terror, designating the concept rather than specific perpetrators as the enemy. The first military action in this fight was to be Operation Enduring Freedom (OEF), the war against Al Qaida and its Taliban government hosts in Afghanistan. In October 2001, after the Taliban had refused to hand over the Al Qaida operatives who had their bases in Afghanistan, the US started OEF. The first strikes came from the air. Every fixed target that could be attributed to Al Qaida was bombed, as were the scarce air defence assets the Taliban had. After this, the strategy was aimed at regime change in Kabul through the support of anti-Taliban forces in the country. To this end small-scale US and allied ground forces, mainly Special Forces teams, joined up with the anti-Taliban forces, known as the Northern Alliance. These small units could direct air strikes on Taliban positions. This air support turned the balance of power in Afghanistan in favour of the anti-Taliban forces. In November 2001 the Alliance took Kabul, after which the Taliban and their Al Qaida allies withdrew to the mountainous terrain bordering Pakistan. Since 2002, besides efforts to combat remaining Taliban and Al Qaida operatives, military operations in Afghanistan have focused on contributing to the international community's efforts in support of the government of Afghanistan, in order to enable it to bear responsibility for the country on its own. At the time of writing, the NATO-led International Security Assistance Force (ISAF) is executing this mission, increasingly replacing the American-led OEF mission as the primary military operation in the country. Air assets play an important supporting role in the assistance efforts on the ground. ISAF forces can call on fighter-bombers like the RNLAF F-16s and RAF Harriers for air support, ranging from air presence to the delivery of precision-guided bombs or other ordnance. Attack helicopters like the AH-64 Apache are able to follow ground forces during hazardous operations. Transport helicopters provide operational and logistic support, in addition to medical and/or casualty evacuation. A prerequisite for this mission in land-locked, mountainous Afghanistan with its challenging accessibility is fixed-wing air transport, intra-theatre with tactical air transport means like the C-130, as well as into Afghanistan with strategic air assets.

Operation Iraqi Freedom

The most recent major military confrontation again took place in Iraq. As part of the broad strategy aimed against the so-called "axis of evil", of which Saddam Hussein's Iraq seemed to be a considerable part, George W. Bush's government decided to invade the country and oust the dictator in order to replace him with a new government. This invasion, named Operation Iraqi Freedom (OIF) started on March 23, 2003. It took only some four weeks to reach the capital Baghdad, seize it and most other major cities. The opposition as mounted by organized military units was crushed in the process. The symmetrical confrontation with Iraq's armed forces was so "extremely lopsided"[28] in favour of the attacking coalition of

---

28 Williamson Murray and Robert H. Scales, The Iraq war. A military history (Cambridge, Mass./ London, 2003) 183.

Americans and British that the Iraqis never stood a chance. A quick advance to Baghdad from the narrow staging area in Kuwait was at the heart of the operational plan. The military reforms Defence Secretary Rumsfeld had fervently endorsed had resulted in an invasion force of only some 150,000 troops pulling this off. Reliance on advanced technology enabled the military to shift their attention from old-fashioned massing of troops to massing of effect. Airpower played a major part in this new so-called "effects-based approach to operations". Airborne assets provided an unprecedented level of situational awareness, while weapons platforms were better equipped than ever to exploit this advantage. All these sensor and shooter platforms were connected with the command and control (C2) structure in a networked system in accordance with the concept of network centric warfare. Airborne sensor platforms ranging from JSTARS (detecting ground movements in all weather conditions) and AWACS (providing an air picture) to a variety of unmanned aerial vehicles made it all but impossible for an adversary to stand a fighting chance in a symmetrical encounter. The total available range of fighters, fighter-bombers and bombers provided the coalition with every shooter platform imaginable. Control of the air was absolute, enabling optimal use of both sensor and shooter platforms, providing near continuous intelligence and fire support for ground units. Fighters operated from Kuwait and carriers in the Arabian Sea, while bombers flew from bases in Europe or from Diego Garcia. The B-2 Spirit stealth bombers even flew from their home base in Missouri. The huge number of aircraft operating over the Iraqi theatre of operations placed high demands on the availability of tankers. This was exacerbated by the limitation on the use of air bases in countries bordering Iraq, as international political support for the invasion was not widespread. The ordnance these aeroplanes delivered was mainly of the precision-guided kind: 70 % of the ordnance used was "smart". GPS guidance was added to the electro-optical guidance, significantly reducing the limitations imposed by weather conditions.

The level of interaction, the joint-ness between the services, was higher than ever before in this conflict with its incredibly high operational tempo. "Speed of ground movement flushed the enemy; airpower killed him while he was exposed, massed and in the open."[29] Furthermore, with the size of ground forces getting smaller, the relative part played by air assets expanded. These assets in turn could only be used in this way because of the total control over the airspace. Surveillance planes and large bombers can only loiter above a battle space when there is no threat to these vulnerable planes. In the case of OIF, air supremacy was a given from the outset. Since Desert Storm, Iraq had already been subjected to severe restrictions on the use of its airspace that were strictly enforced by America and the United Kingdom. In fact, it can be said Iraq had been subjected to a counter air campaign for twelve years.[30] Thus, the simultaneous start of the ground and air campaigns on 23 March 2003 was in appearance only. The importance of air superiority cannot be overstated nor taken for granted.

Regarding OIF it should be noted that the striking operational success of the initial four weeks of the campaign is only part of the story. The American presence in Iraq since then has still not resulted in anything resembling lasting meaningful progress in the process of replacing the regime with a new government, contributing to security in a broader sense. At the time of writing (2006) the situation in Iraq, if not a civil war then closely resembling one, is still highly unstable, with American losses increasing on a daily basis. The Iraqi dead cannot

[29] Murray and Scales, The Iraq war, 245.
[30] Ibidem, 162.

even be accurately assessed. It is too early to tell to what extent the decision to invade Iraq has been a wise one.

Summing up

A century after its appearance on the stage, airpower has made its distinct mark on the history of warfare. The First World War saw the first large-scale use of aircraft, growing in variety and importance at a fast rate. Every air operation known today is rooted in the Great War; the air had become a theatre of war in its own right. Another important legacy of the experience of this war is summarised in the ideas developed by the classical airpower theorists Douhet, Trenchard and Mitchell as a result. During the interwar years they formulated theories, which remain at the base of airpower doctrine to this day. Most important among these ideas are those regarding the strategic use of the new air weapon. Airpower promised an alternative to the way war was fought on the Western front with all its stalemate and slaughter. The Second World War saw the new concepts applied, with weapons and within contexts that had come a long way since the last war. Bombers had incomparably more payload, which was to increase even more as the war went on. Getting the bombs on target remained a problem however, resulting in the civilian population bearing the brunt of the applied strategic pressure. The strategic use of airpower however did not prove to be the sought after shortcut to victory, the means to prevent a long drawn out conflict. Total war had taken on a new dimension as the inhabitants of Warsaw, Rotterdam, Coventry, Dresden and Hiroshima could attest to. In the confrontations in the various theatres of battle on the other hand, airpower was shown to be of major importance. In order to enable surface forces to be operationally successful, one had to prevent an adversary from interfering from the air. Simultaneously, air superiority meant being able to support ground action, or action at sea. This battle for air superiority, the counter-air operation, has been at the core of air warfare ever since the advent of the aeroplane.

The Cold War can be viewed as an era with its own characteristics. Despite the "limited" wars actually being fought, the confrontation between NATO and the Warsaw Pact countries determined the composition of the armed forces to a large extent. Airpower grew into distinct roles, the strategic one dominated by the nuclear weapon, while "tactical air" focused on the confrontation with the adversary's armed forces on the battlefield. The end of the Cold War provided opportunities for new ideas regarding strategic and tactical use of military power, reshuffling the means in the process.

From an air warfare angle, the examples mentioned each merit attention. The wars in Korea and Vietnam, although part of the ideological struggle of the Cold War, were long drawn out conflicts, limited wars that in a sense can be seen as sideshows to the main stage. In both cases the US, as the lead actor, had difficulty in adapting the means and mindset to the circumstances at hand.[31] The wars in the Middle East emphasized the importance of airpower. Especially Israel's use of the air arm has been impressive, with the air force forming Israel's first line of defence since its establishment. More recent history however also shows the limitations of superior conventional forces when confronted by an irregular adversary. The war in the Falklands saw western hardware pitted against each other in very inhospitable circumstances in the South Atlantic. It illustrated once more the importance of air superiority as well as the limitations of airpower in the absence of supporting means

[31] Something comparable happened to the USSR in the 1980's when it was confronted by an irregular foe in Afghanistan.

like air-to-air refuelling and airborne early warning. The Gulf War of 1990/1 was waged at a moment in time when the Cold War had just come to an end. It saw the US-led coalition confronted by a symmetric enemy with regular armed forces for a change. The outcome was accordingly. It was to remain exceptional.

The 1990s were further characterized by the necessity for intervention in civil wars, of which those in former Yugoslavia dominated the larger part of the decade. This chapter dealt with the complexities of those Balkan wars at some length in order to illustrate and emphasize the context in which airpower was used, which can only be ignored at peril. The circumstances have become even more complex after 9/11. America's war on terror has upset existing balances in Afghanistan and Iraq. After achieving swift operational success in the first phases of operations Enduring Freedom and Iraqi Freedom, the difficult follow-up period of providing security assistance while building new structures, places totally different demands on the armed forces in theatre. They have to contend with an elusive opposition that uses irregular and terrorist methods.

This chapter showed the forms of air operations developing from their humble beginnings more than a century ago. The introduction of the aeroplane in warfare added a dimension in many senses. The basic underlying ideas regarding counter air operations, strategic operations and operations aimed against the enemy's surface forces have not changed much. Discussions on where to apply pressure with airpower, what kind and to what extent, have always been lively however. The means with which these operations are executed have taken great strides, influencing the debate and keeping it alive. The search for a useful application of this so powerful tool will remain a challenge.

Suggestions for further reading

Budiansky, Stephen, Airpower. The men, machines and ideas that revolutionized war, from Kitty Hawk to Gulf War II (New York, 2004) Readable, recent broad overview of the century-long history of airpower.

Deighton, Len, Fighter (London 1977) Highly readable book on the Battle of Britain by British author best known for his best-selling fiction, also of a high quality.

Hastings, Max, Bomber Command (London 1979). Excellent study on British World War II strategic bombing theory and practice; treats both the higher levels of politics and strategy and the quality of the airmen executing the missions. Combines high scholarly standards with outstanding readability.

Johnson, J.E. "Johnnie", The story of air fighting (London 1985) A history of fighter combat from World war I to the Falklands by Canadian World War II ace. Originally published in 1964 as *Full circle*.

Kennett, Lee, The First Air War, 1914 – 1918 (New York 1999) Very useful starting point for study on First World War aviation. This readable book touches upon most aspects and provides further leads.

Loo, Erwin van, Crossing the border. De Koninklijke Luchtmacht na de val van de Berlijnse Muur (Den Haag 2003) The story of the Royal Netherlands Air Force after the Cold War comprehensively told. Volume in the official history.

Murray, Williamson, Luftwaffe (London 1985) Admirable history of the Luftwaffe from 1933 to 1945, meaningfully placed in the contexts of German military strategy and the air war over Europe as a whole. Contains a very useful appendix on the "Pre-war development of British and American doctrine and airpower."

Nordeen, Lon, Fighters over Israel (London 1991). A history of the Israeli Air Force from the War of Independence to the Eighties.

Olsen, John Andreas, Strategic airpower in Desert Storm (London 2003) Comprehensive study on the use of post Cold War strategic airpower by Norwegian Air Force officer-scholar. Contains very useful analysis of Wardens theory.

Overy, Richard, The battle (London 2000). Concise treatment of the Battle of Britain by eminent historian of the Second World War.

Chapter 2 – The Physics

Introduction

All aircraft share the same atmosphere and thus the same physical constraints. Understanding the physics makes it easier to comprehend the possibilities and restrictions of airpower platforms, from the largest bombers to the smallest unmanned air vehicle. Therefore the physics are discussed first before proceeding to the next chapter on different types of airpower airframes.

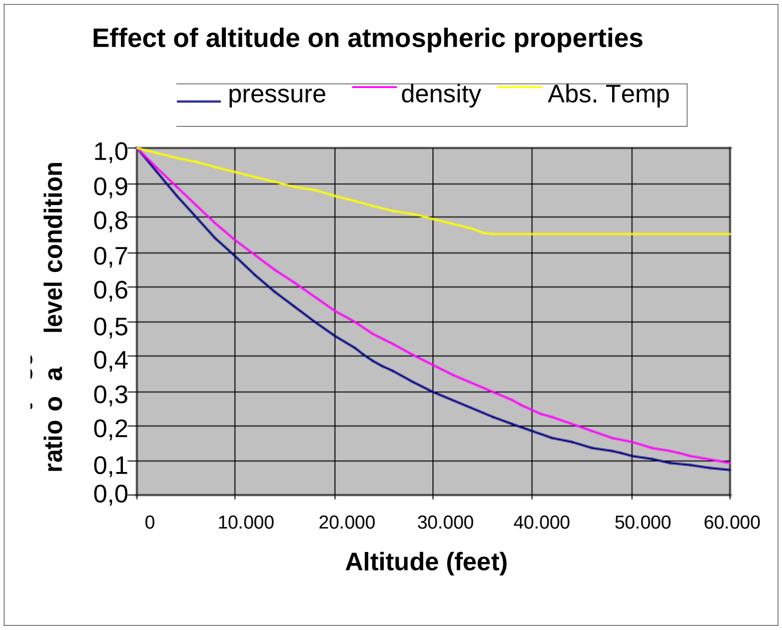
Properties of air

The usable portion of the atmosphere consists of a relatively thin layer. Man can live without oxygen augmentation in the lower 3 to 4 kilometres' region, but above this altitude the partial oxygen pressure is too low. As altitude increases, pressure decreases rapidly, halving approximately every 18,000 feet. At 80,000 ft (the approximate altitude where the TR-1 can still fly) pressure is less than 3% of the pressure at sea level. The rapid decrease in pressure limits the altitude at which most aircraft can fly. Low pressure limits the lift produced by the wings and the thrust generated by the engine. Aircraft and engine design will determine which of the two factors will be the limiting factor for the individual aircraft ceiling.

Temperature also decreases with altitude at a constant rate of 2º Celsius per 1,000 ft until the upper limit of the troposphere is reached. At this altitude (45º latitude around 36,000 ft) the temperature is minus 56°C. The graph below shows the decrease of pressure, density and temperature with altitude, expressed as a ratio compared to sea level values.

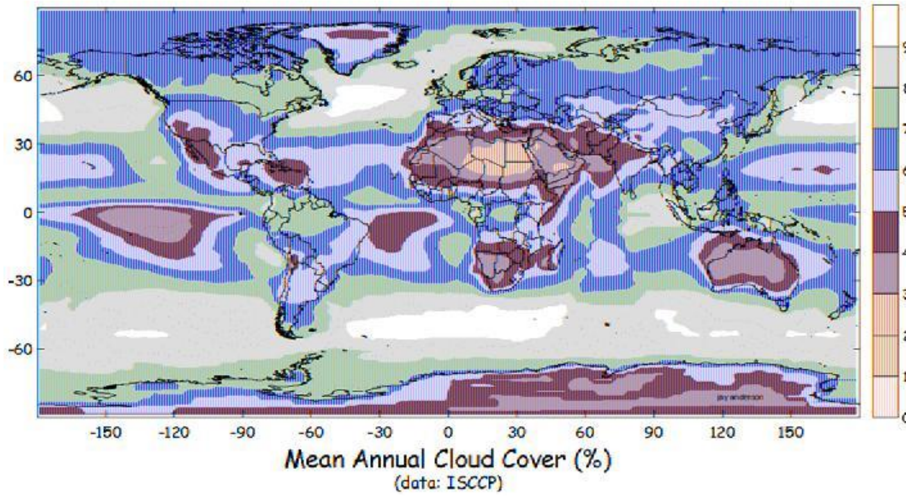
Another important property of air is sonic velocity, which is dependent on temperature. At 15°Celsius (the average sea level temperature), this velocity is 660 knots, at -56°Celsius (the average temperature at 36,000 ft) it is 575 knots. In aviation, airspeed is expressed in relation to the speed of sound; this ratio is called the Mach number (named after Ernst Mach). Assuming that the speed of sound is about 600 knots, the Mach number can easily be converted to nautical miles a minute (0.8 M = 8 nm/min).

Sea level conditions standard atmosphere		
	Metric	US standard
Pressure	101320 N/m²	2116.22 lbs/ft²
Temperature	288.15°Kelvin (15º Celsius)	518.7°Rankin (60º Fahrenheit)
Density	1.225 kg/m³	0.00238 slugs/ft³
Sonic velocity	340.43 m/sec 1225.km/u)	1116.88 ft/sec (661.7 Knots)



Air is invisible, not only to the visual part of the electromagnetic spectrum but also to radar frequencies and large parts of the infrared spectrum. Meteorological conditions like rain, snow and clouds will affect visibility and infrared radiation but this effect is limited on (most) radar frequencies. Adverse meteorological conditions have prevented the employment of airpower in the past and they can still be a limiting factor. Only aircraft equipped with advanced radar and weapons are capable of all-weather precision attacks but they are unable to identify and attack their targets with the same precision as under favourable meteorological conditions. The precise limitations will be covered in chapters 4 and 5.

Given the right meteorological conditions, air allows the aviator an unrestricted view of the earth when using a combination of eyesight, night vision goggles, infrared thermal imaging and radar. The same holds true for the air and ground-based air defender.



The fact that it is difficult to hide in air has necessitated the use of technical measures such as jamming, chaff and stealth by aviators, whereas those combating them have had to employ electronic countermeasures and multi-spectral sensors.

Since air covers the whole planet one could assume that aircraft can fly everywhere. However, thunderstorms and hail are avoided by aircrew because they can severely damage and even destroy aircraft. On take-off and landing, fixed-wing aircraft in particular are limited further and require that crosswind, ceiling and visibility plus runway conditions are within the limits of aircraft and crew.

Aerodynamic principles

Lift

The art of flying is the art of generating lift. Lift is the force required for sustaining flight but also the force required for manoeuvring aircraft and missiles. The process of lift generation can be explained in several ways. The easiest way to understand lift is by applying the physics principle of Action and Reaction and the famous law of Sir Isaac Newton: Force is Mass times Acceleration (F=M x a). To generate lift a certain mass of air must be accelerated downward to give a reaction force that is large enough to counter the gravitational force on the aircraft.

One can imagine that the air mass that passes around the wing of the aircraft is deflected downward by a few degrees[32]. To answer the question of how much air mass is to be deflected we change the formula from mass to mass flow. Force = Mass per second times velocity change per second or:

F= Mass flow x Δ V

Mass flow is determined by three variables. Firstly the size of the wings determine what area of air is affected (in technical terms the S or surface area of the wing), secondly the density (σ) of the air and thirdly the speed (V) of the aircraft: the faster the aircraft flies, the more air mass passes the wing. In other words: Mass flow is proportional with V x S x σ.

The angle at which air is deflected is dependent on two variables: the angle between the wing and the free air stream (angle of attack), and the shape (design of the wing). However, rather than the deflection angle value, we need to know the downward velocity change. If an airfoil deflects the passing air by 6 degrees, the downward speed of the air will be 10% of the forward speed[33], making the downward acceleration velocity **times** deflection percentage.

Δ V =V x deflection percentage

Because the lift force is mass flow **times** downward acceleration, lift will be proportional to velocity squared (V^2) and also to S and **σ**. To express lift in an exact formula we need a coefficient, the so-called lift coefficient or C_L. The value of C_L depends on the shape of the wing and the angle of attack. With this coefficient the Lift formula is complete.

32 In reality the air close to the wing surface will be deflected the most and the amount of deflection will decrease with increasing distance to the wing surface. In actual fact the deflection angle is an average deflection angle.

33 Downward acceleration is velocity times the sinus (deflection angle).

The formula used for lift is:

$L=\frac{1}{2}\sigma V^2 SC_L$	
L= Lift	(Lbs or Newton)
σ = Air density	(slugs or Kg/m^3)
V=Airspeed	(ft/sec or M/sec)
S= Wing area	(ft^2 or m^2)

Lift and speed limits

The lift coefficient (C_L) is dependent on the shape of the wing and the angle of attack. An increase in angle of attack will increase the C_L, but to a limit. Above a certain angle the flow behind the wing will become turbulent and lift will decrease. This phenomenon is called the stall. On most wings the stall will occur around a 10-15 degree angle of attack. However, some special wing designs (F-16, F-18) are capable of reaching very high angles of attack of more than 25 degrees. The lowest velocity at which an aircraft can fly is just above the stall speed. To make takeoff and landing feasible, a low stall speed is required. To reach that capability most aircraft are equipped with flaps, which increase the surface area and the maximum value of C_L.

There is not only a minimum airspeed an aircraft can fly at, but also a maximum airspeed. The maximum airspeed can be limited by the available power but also by the wing's structural limit.[34] To fly at high speeds requires a strong and relatively short wing, to fly at a low airspeed requires a large wing with preferably a large wingspan. So, every aircraft designer has to compromise. If the aircraft needs to fly at high speeds it requires flaps to be able to fly at a lower speed and one has to accept higher takeoff and landing speeds that require long runways. On the other hand, a design for slow speed implies a wing with a large wingspan, but this will result in structural limitations that dictate a lower maximum speed.

Drag

It is common experience that drag increases with velocity. This is valid for cars, trains bikes and also for aircraft. In aviation this 'normal' drag is called profile drag. The profile drag is proportional with the density, the size and shape of the body and the velocity squared. Profile drag can be reduced by streamlining the shape of the aircraft and by using a smooth surface. The increase in drag with velocity squared holds true until the speed of sound is approached. Starting above .85 Mach, shockwaves start to form around the airframe in places where the air is accelerated above the sound velocity (typically around the nose and the top of the wing). Shockwaves dramatically increase the drag. This is the reason why almost all commercial aircraft tend to stay well below this speed.

Contrary to cars and trains, aircraft produce lift. Everything comes at a price and this is also true for lift. The generation of lift will cause a special type of drag called induced drag. But here a strange phenomenon takes place: induced drag **decreases** with the velocity squared. The logic behind this can be understood if one realizes that at slower speeds less air is deflected, so a much larger deflection angle is required.[35] This large deflection angle tilts the reaction force more backwards. This

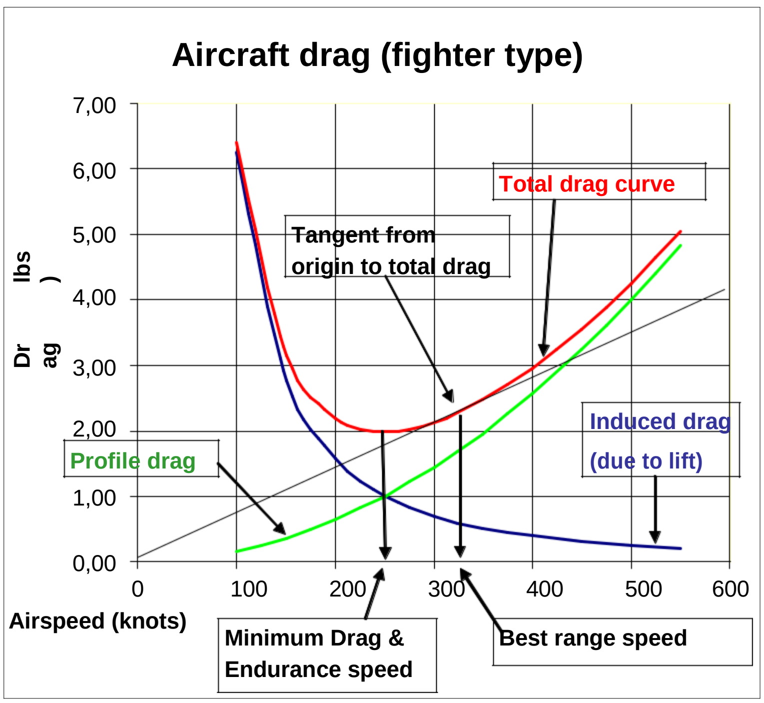
---

34 The maximum speed is called the wing divergence speed; at this speed the aerodynamic forces will tear the wing from the fuselage. To increase the wing divergence speed the wing must be swept or short.
35 If the speed decreases by 50%, the mass flow decreases by 50% and the downward acceleration decreases by 50% as well. Consequently the lift decreases to 25% of the original value. To restore the lift

feature was initially not well understood and it took the Wright brothers years of experimentation before they discovered that a thin wing with a small curvature (which deflects the air over a small angle) was the way to generate lift without too much drag.

This combination of profile drag and induced drag gives an aircraft a unique drag curve that is depicted below. This 'bathtub'-like drag curve differs considerably from that of other types of transport vehicles. For most means of transportation the drag only increases with speed. Consequently the best speed range is normally achieved at the lowest practical speed, with cars and trains achieving maximum endurance by stopping the vehicle and turning the engine off.

For aircraft the drag curve is quite different and to understand the typical drag curve of an aircraft means to understand aircraft performance. In this curve the lowest drag is well above the lowest speed the aircraft can fly at. In this example it is around 250 knots. This is automatically the speed for best endurance and also the best glide speed. Military aircraft that are optimised for endurance (e.g. reconnaissance UAV) will fly at this speed when loitering above the target area. Aircraft in orbit, for example tankers, surveillance aircraft or aircraft holding for landing will also fly at this speed to minimize fuel consumption.



The best range speed is by definition the speed where the fuel used per nautical mile is the lowest. Fuel used per nautical mile is equivalent to fuel flow divided by velocity.[36] For jet aircraft the fuel flow (lbs/hour) is proportional to the drag, so the best range speed is found where the ratio between drag and speed is the lowest. Graphically this point can be found by drawing from the origin the tangent to the total drag curves. The point where the tangent touches the drag curve is the best range speed. This drag is of course higher than the minimum drag, so the required

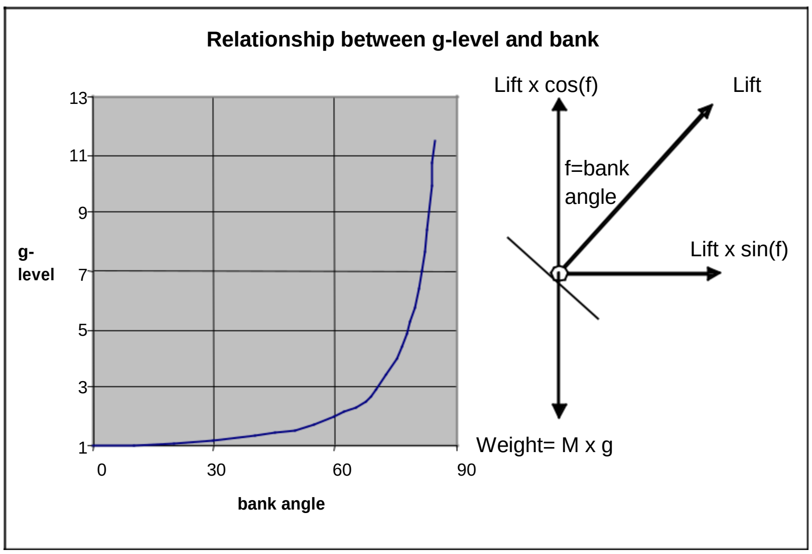
to its original value the deflection angle (and angle of attack) has to become 4 times as large. At small deflection angles the induced drag will be proportional to the deflection angle and be 4 times as large.

[36] Dividing both fuel and distance by time gives this new relationship.

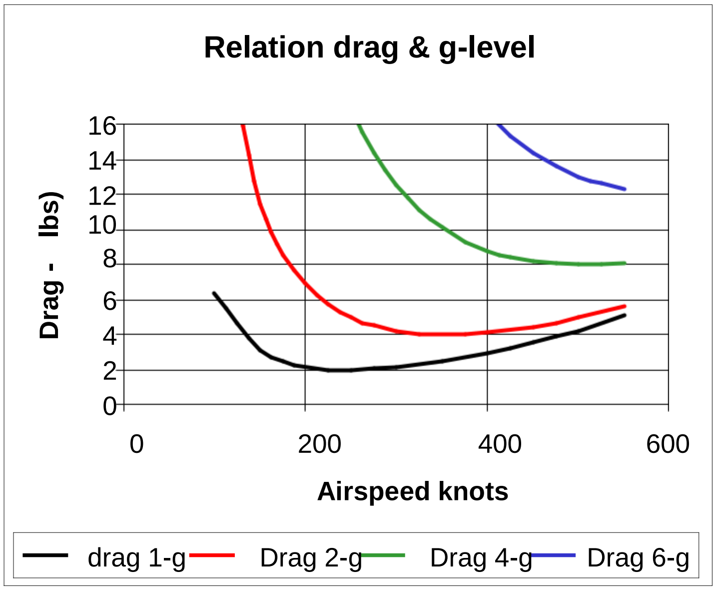
thrust and fuel consumption will be higher as well, but the fuel per covered distance will be the lowest. The aircraft designer tries to optimise the design in such a way that the best range speed is as high as possible, but he also has to stay away from Mach effects. Therefore most turbojet aircraft (from fighters to wide bodies) have a best range speed between .78 Mach and .84 Mach. Aircraft cannot slow down to save fuel like other transport vehicles; it might result in not reaching the destination at all. Nor can aircraft, (except for specialized military aircraft) accelerate much beyond their best range speed.

Manoeuvring

Lift has more applications than countering gravity. Lift also supplies the force needed to change direction. By banking the aircraft the lift vector is tilted and part of the lift force becomes available to turn the aircraft. Most aircraft turn with small bank angles up to 30°. These small angles require an even smaller increase in lift (a 30º bank only requires a 15% increase in lift). But fighters need more agility. To make tight turns, the F-16's lift can be increased nine-fold. Because force and acceleration can be exchanged it is common practice to call this 9-g; which means that the acceleration generated by the wings is nine times the gravity acceleration of the earth. The wings must be extremely strong to withstand a load of nine times this weight. The maximum C_L of a wing is fixed, so to be able to generate enough lift the aircraft speed must be high enough. Because lift is proportional to airspeed squared, a speed of three times the stall speed will give the capability to generate a lift of 9 times the weight.



Generating more lift to turn the aircraft comes at a price because the induced drag will also increase. Again there is a non-linear relationship. As the induced drag increases with the g-level squared, the 9-g turn has a penalty of an 81-fold increase in induced drag. This changes the drag curve considerably. The only way to counter this enormous drag is to use very powerful engines or the pilot will sacrifice energy (loss of airspeed or altitude). Note also that the minimum drag point now occurs at a much higher speed!



From the graph it is easy to imagine that at high g-levels and low airspeeds the drag can easily be higher than the available thrust.

Turning is more than g-level

Fighter pilots in air combat are not just interested in g-level. As will be explained in chapter 9 they are interested in turn circle and turn rate. If g-level and true airspeed are known, turn radius and turn rate can be determined with the centripetal acceleration formula. (V^2 / R = C).
At high g-levels the centripetal acceleration and the g-level are very close and the following rules of thumb are valid:

-The turn rate is proportional to the true airspeed divided by the g-level. Consequently the best turn rate is achieved when high g-levels at low speeds are attainable. This requires a relatively large wing area.

- The turn circle is proportional to the speed squared divided by the g-level. Consequently, if two aircraft have the same turn rate but one has a higher speed (and consequently a higher g-level), it will have a much larger turn radius.

V = velocity	(ft/sec or m/sec)
R = turn radius	(ft or meter)
C = centripetal acceleration	(ft/sec ^{2} or m/sec ^{2})

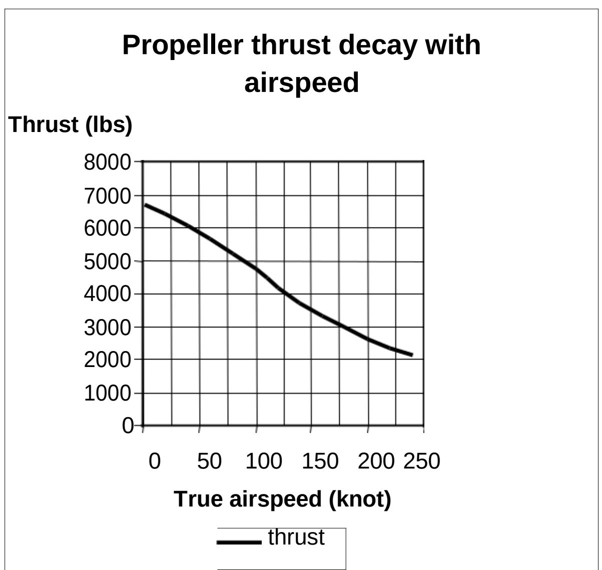
Propulsion

Propellers

Lift is generated by pushing air downwards and thrust is generated by pushing air backwards. Traditionally this was done by propellers. Propellers are airfoils just like wings. Everything that has been said about lift also applies to propellers, so the larger the propeller, the more air is affected. An increase in angle of attack of the propeller blade will increase the thrust. The higher the speed of the aircraft the more air mass is flowing through its propeller and consequently less air acceleration is required to give the same amount of thrust.

Propellers are limited to low-speed operation due to two contributing factors. Firstly, to prevent excessive drag the tip speed of the propellers (which is the vector sum of the rotational and forward velocity) must be well below Mach. Secondly, as the airspeed is increased, the propeller blade has to adjust its pitch more forward to account for the change in relative wind direction. However, when the blade is pitched forward the reaction forces on the propeller blade change direction as well and will point more sideways. The result will be that the thrust portion becomes smaller and the propeller drag increases. Normally propeller aircraft will have a maximum operating speed of around 0.65 Mach. On the positive side, the propeller can generate large amounts of thrust at slow speeds while being very fuel-efficient.

A typical propeller performance is depicted in the next graph.



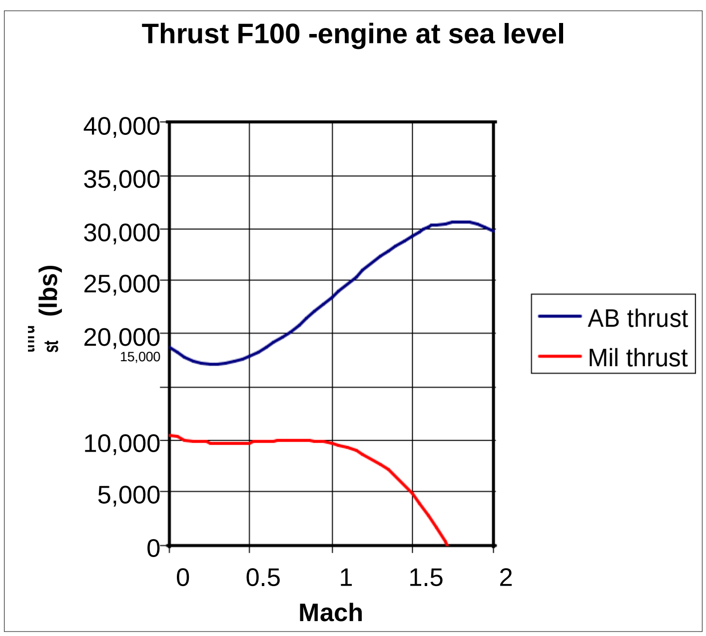
Typical thrust for propeller

Jet engine

The jet engine has been developed to overcome the limitations of a propeller. Instead of accelerating air like a propeller, the air is initially slowed down in the inlet (except at low airspeeds where the air is sucked in and accelerated). When the air is slowed down, pressure increases. The air is then further compressed by the compressor, which is driven by the turbine. After the compressor, fuel is added in the burner chambers and part of the oxygen is burned. It is not feasible to burn all oxygen; this would create such a high temperature in the engine that the turbine would melt. So, most of the air is used for cooling. After passing part of the energy to the turbine, the air passes through a nozzle where the pressure is finally converted into speed again. Optionally, some of the remaining oxygen can be burned after passing the turbine in the afterburner section. This afterburner section, which is made of titanium, is able to withstand much higher temperatures than the turbine. As the titanium wall needs some cooling air as well, even in afterburner not all oxygen is burned.

Contrary to the propeller, the jet engine can also develop thrust at high Mach speeds. Even at supersonic speed however, the air entering the compressor must still be at low subsonic velocity. Therefore the air must be decelerated in the intake. This deceleration will cause drag. The aircraft designer will take great pains to design the intake in such a way that most of the speed loss of the entering air is converted into pressure, which in itself increases engine performance. Supersonic velocities present the biggest problem to the intake designer. Supersonic air can be slowed down through a single shockwave or through a number of weak (oblique) shockwaves. It can be proven that a large number of weak shock waves gives a better pressure recovery than a single (large) shock wave. To generate a large number of weak shockwaves, the air inlet can be fitted with a cone (F-104) or a moveable inlet ramp (F-4, F-15). Fighters without such a cone or ramp (F-16) can not generate a large number of oblique shock waves and are less efficient at high supersonic Mach numbers.

Thus a good pressure recovery improves the efficiency of the engine, which will compensate for most of the momentum drag in the intake caused by the deceleration of entering air. Furthermore, the larger airflow means that more air mass can be heated and accelerated. A pilot sees this from looking at the fuel flow in maximum power, which will increase at higher Mach numbers. This large amount of thrust and especially this large amount of power at high speed is the key to supersonic flight.



There are also limitations to the jet engine. Its fuel consumption is considerably higher when compared to a propeller engine. In afterburner fuel consumption will be very high. The use of afterburner might increase the power by 40-80%, but will increase fuel consumption by more than 400%. However, for some flight conditions, e.g. supersonic flight and high turn rates this extra power is essential.

There are two variations to the normal jet engine. The first is the bypass engine, which can be best described as a combination of a propeller and a straight jet. This type of engine has two compressors driven by two turbines. Most of the air flows through the first compressor but bypasses the second compressor, is not burned and does not flow through the turbines. The ratio between the air bypassing the second turbine and the air entering the core is called the bypass ratio. Modern engines powering commercial aircraft have a bypass ratio of 8 and higher. Bypass engines have a much better fuel economy than straight jet engines and are able to operate at higher airspeeds than propeller aircraft. However, if the airspeed increases above .85 Mach these engines become less effective. The F-100 engine, which powers the F-15 and F-16 has a very low bypass ratio: only 40%. A large bypass ratio means quite a large intake that is not suited for supersonic flight at all.

The other variation is the ramjet. If the engine is operating at very high Mach numbers it is possible to build an engine without a compressor and a turbine. The air will be compressed by the shock waves and no turbine is needed because there is no compressor to be driven. It is a very simple design, but the aircraft or missile must be accelerated to a high Mach number before the engine can operate. Therefore this type of engine is more appropriate for missiles that can be accelerated to high Mach numbers with a booster.

All air-breathing propulsion systems will have decreased power at higher altitude. Generally speaking, thrust reduces with air density. So while the friction is decreasing (which allows for higher airspeeds), the available thrust also decreases. Most aircraft operate below 40,000 feet. Operating at higher altitudes requires powerful engines or very special designs. The TR-1 and the SR-71 are perfect examples of this.

Chapter 3 - Platforms

Introduction

The aerodynamic properties described in the previous chapter make it clear that each aircraft is limited by lift, drag and available thrust. But, depending on the role and mission range, payload and speed requirements are different. That is the reason for the variety in platforms that will be discussed in more detail in this chapter.

Fighters

To design a fighter means to compromise. The fighter pilot's wishes are clear: high supersonic speed and fast acceleration, good manoeuvrability, long range and low observability. But all these requirements lead to different designs. High supersonic speed and the ability to intercept will dictate relatively small thin wings, a small frontal area and a large engine. Furthermore, to limit the weight, the fuel fraction should be low. Preferably the aircraft should be single-engine in order to limit the frontal area. The F-104 was an aircraft designed to those specifications. The drawbacks of this design were: limited range, bad manoeuvrability, especially at low speeds and hardly enough space to install a sizeable radar antenna.

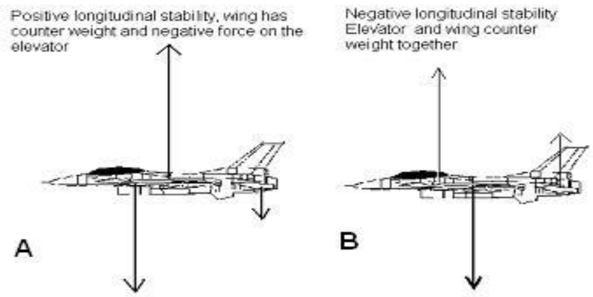
To give an aircraft better manoeuvrability, more lift is required at low speed. This means larger wings, but this will generate more drag at high supersonic speeds. A solution is to give the aircraft a variable sweep wing. Theoretically this solves the problem: the aircraft will have a small and nicely swept wing at high speed and a larger straight wing when manoeuvring at low speeds. The drawback of a variable sweep wing is more complexity and weight. The weight penalty will reduce manoeuvrability. Thus, most modern fighters are designed without a variable sweep wing and with a reasonable wing area to support manoeuvrability.

Range is another problem area for fighters. If the fighter carries too much fuel, the weight will reduce its manoeuvrability and acceleration capability. To give a fighter an in-flight refuelling capability helps, but is of little use during a supersonic intercept or scramble. Another well-tried method is the use of pylon tanks. These tanks extend range and can be jettisoned if conditions so dictate. However, the drawback of pylon tanks is that the aircraft profile drag increases considerably, far more than if the same amount of fuel was carried inside. (Room inside the aircraft could, for example, be generated by stretching the airframe.) The present tendency is to design fighters with more range and a higher internal fuel capacity(e.g. F-35). The need for more range is, of course, related to the experience with operations abroad. During these operations, well equipped airfields close to the target area are scarce. This range increase will however affect the manoeuvrability of the aircraft.

Another way to increase range and manoeuvrability is to make the aircraft unstable. In conventional aircraft the centre of gravity is situated in front of the aerodynamic centre (the point where the lift is supposed to act). A forward centre of gravity makes for a stable trajectory - an arrow is a good illustration of this principle[37]. To stabilize a conventional aircraft the horizontal tail must generate a downward force.

37 When the aircraft is disturbed nose up, the lift on the wing increases, the negative tail lift decreases, and both forces generate a moment around the centre of gravity to rotate the nose back. For an unstable aircraft the situation is reversed and a disturbance will be increased by the lift of the wing and counteracted by the tail.

The consequence is that the lift generated by the wing should compensate for the weight of the aircraft and for the negative force of the tail.



When an aircraft's centre of gravity is behind the aerodynamic centre (B), the horizontal tail will have to generate positive lift, so the wing has to carry less than the full weight of the aircraft. This means that with the same wing design the aircraft can pull more g. The consequence of this design is that the aircraft will be unstable; instead of correcting any disturbance the aircraft will increase it. Such an aircraft would be very demanding to fly, but with the help of computers the aircraft's behaviour can be changed in such a way that it feels like a 'normal' stable aircraft.

Comparing fighter performance should be done carefully. It is dependent on many variables: internal and external fuel load, weapons load, temperature and altitude. Normally the aircraft manufacturer will list maximum speed and ceiling for a clean aircraft using afterburner. Turn performance is calculated for an aircraft with only half internal fuel and is different at each altitude. Range however, is calculated at optimum altitude, without burner. To make a good comparison between different fighters it is best to use Energy & Manoeuvrability diagrams.[38] These show excess energy for every flight condition. A complete fighter comparison analysis is beyond the scope of this book. Instead we will give some typical fighter parameters.

Typical fighter performance	
Thrust to weight ratio (clean)	1- 1.3
Max Indicated airspeed	750-850 knots
Maximum Mach number	1.6 - 2.2 M
Ceiling	>50,000 ft
Maximum time in afterburner (fuel limited)	15 - 20 minutes
Maximum g-level	8-9 g
Sustained turn rate low level	20-25 degrees/second
Endurance at optimum power, without external fuel	1.5- 2 hours

38 An example of a EM diagram is given in chapter 9.

Fighter-bomber

It is not surprising that most fighters can be transformed relatively easily into fighter-bombers. The key to this versatility is the external carriage of weapons and fuel. As we have seen, the wing of a fighter has to be strong in order to withstand high g-forces. If extra weight is added to the wing, some of its manoeuvrability is lost. It is not uncommon for the g-limit to be half the value of the clean configuration. Drag increases, the extra weight raises the induced drag and the stores add a considerable amount of profile drag. So, maximum speed will be reduced significantly and only low supersonic numbers are attainable. But it takes more to transform a fighter into a fighter-bomber, although most of the changes are software modifications in the fire control computers. These modifications make the aircraft capable of delivering bombs, enable it to operate special equipment like targeting pods and add air-to-ground modes to its radar. An important feature of fighter-bombers is that by jettisoning stores they can instantly regain their full manoeuvring capability.

For the air battle commander, a fighter-bomber is a safe choice. It can be used for air defence and ground attack, moreover it can rapidly alternate between these roles. This compromise is cost-effective as well. The price of a fighter-bomber is slightly higher than that of a straight fighter, but far lower than that of specialized aircraft for ground attack and air defence. Given these advantages then, it is no wonder that most new fixed-wing combat aircraft are fighter-bombers.

Typical fighter-bomber performance with maximum load of fuel & weapons	
Thrust to weight ratio (clean)	0.60 - 0.80
Max indicated airspeed	500-600 knots
Maximum Mach number	0.85 - 1.2 M
Ceiling (military power)	20,000-30,000 ft
Combat range medium level	350-550 nm
Maximum g-level	4.5 – 6 g
Sustained turn rate low-level	10-15 degrees/second
Weapon load	2000 - 8000 lbs
Endurance at optimum power, without external fuel	2.- 3 hours

Transport

Military transport aircraft differ somewhat in construction from commercial airliners. The specific military requirements necessitate the ability to land at remote places with less than optimum runways. The need to quickly load and unload the aircraft (think of a load of combat vehicles) leads to a design with a low cargo floor and a ramp. This automatically means a high wing. Short runway operations require a low take-off speed and fast acceleration, so a large wing is combined with propeller propulsion. At low speed the large wing gives sufficient lift and the propeller delivers optimum thrust. Special care should be given to the undercarriage; tire pressure should be lower than that of commercial aircraft to enable soft field landings. Because of the weight of the aircraft under those conditions, more or larger tires are required to carry the load. Typical aircraft designed according to these specifications are the C-130 Hercules and the C-160 Transall. Because these aircraft are able to deliver their cargo very close to the troops they are classified as tactical transport aircraft. They can be equipped with self-protection equipment (chaff, flares and jammers).

Typical parameters of the C-130H are:

C-130H Tactical transport	
Takeoff ground run maximum weight standard day	4700 ft
Maximum landing roll	2400 ft
Maximum Mach number	0.6M
Ceiling	25,000 ft
Maximum ferry range	4200 nm
Range with max payload	1000 nm
Empty weight	80,000 lbs
Maximum payload	22,000 lbs
Maximum take-off weight	155,700 lbs

The most important parameters of a cargo aircraft are payload and range. These parameters are interrelated. It is possible to give up payload and carry more fuel to increase the range. Therefore the maximum ferry range (with zero payload) is presented and the range with maximum payload is given separately. There is a more or less linear relationship between the two, so at half the maximum payload the range is about halfway between ferry range and maximum payload range.

Another type is the strategic transport aircraft. This type is not required to land on rough fields but must take its cargo to an airfield close to the combat zone. For landing on normal runways, it is advantageous to design it like a (wide-body) airliner with turbo engines, so it can fly faster, higher and cover intercontinental distances. For easy loading and unloading a high-wing design with a ramp is required. Aircraft like the C-141 and the C-5 fall into this category. A special aircraft is the C-17. With this design the USAF has built a strategic airlift aircraft (considering range, speed and load) but with the landing capability of a tactical transport aircraft. The benefit of this design is that no time is lost in the transfer of cargo from a strategic transport aircraft to a tactical transport aircraft.

C-17 strategic & tactical transport	
Take off ground run at maximum weight standard day	7600 ft
Landing roll (possible on 3000 ft austere runway of 90 ft wide)	1880 ft
Maximum Mach number	0.77M
Ceiling	45,000 ft
Ferry range	5290 nm
Un-refuelled range max payload	2400 nm
Max payload	170,000 lbs
Empty weight	236,600 lbs
Maximum takeoff weight	523,000 lbs

Bombers

In a historical sense it is difficult to give a proper definition of a bomber. Strategic bombers from the Second World War had about the same range and payload as a modern fighter-bomber. Maybe the best definition is not payload or range-determined but more defined by the capability to attack deep into enemy territory from the own homeland. In the Second World War therefore, a range from England to Berlin was sufficient for a strategic bomber but in the Cold War it became important to reach the USSR from the USA and vice versa. Intercontinental strategic bombing requires highly specialized aircraft. The end of the Cold War has not changed these requirements.

Strategically, it is of great importance to have the capability to strike from the homeland, without having to depend on foreign forward bases. Technically it is impossible to build a large aircraft and make it highly manoeuvrable at the same time in order to be able to outperform fighters and ground-based air defence. To increase the survivability of the bomber, several approaches can be chosen, all resulting in a different design.

The first is to protect the bomber with long range fighters as was done from 1943 onwards. If that approach is chosen, the bomber will look more like an airliner in design. Equipped with a subsonic large wing, it will fly economically at high level and high subsonic Mach numbers. This altitude will at the same time minimize the effect of small arms. Furthermore, a wide variety of Electronic Warfare equipment is needed to counter attacks by surface-to-air missiles. The B-52 and the Russian Bear fall into this category. If fighter protection is not available, the bomber can still be employed if it is equipped with stand-off weapons. These minimize its exposure to threats or make it still usable in a low-threat environment as was seen during many recent conflicts.

Another approach is to make the aircraft capable of operating at high speed and low level. Low-level flying will delay detection and the high speed will minimize the chance of a successful interception. To combine cruise capability at high altitude with high dash capability requires a special wing design. A favourite solution is the variable swept wing as used on the B-1B, the F-111B and the Backfire. However, with the present look-down shoot-down capability of fighters, fast and low is no longer a safe solution (as will be further explained in chapters 4, 5 and 10).

The last and more recent approach to the survivability problem is to make the aircraft extremely stealthy. The B-2 is the only present example. Here a very special wing design is used as well: the flying wing. Chosen for its stealthy characteristics, the flying wing is also aerodynamically very effective and has excellent range capability.

Despite their large fuel tanks, all bombers depend heavily on air refuelling to complete their mission. It was therefore not surprising that in the past, Strategic Air Command possessed about the same number of tanker aircraft as bombers.

These B-52 figures give an impression of bomber capabilities:

B-52 Strategic bomber	
Empty weight	185,000 lbs
Maximum takeoff weight	488,000 lbs
Maximum Mach number	0.86 M
Ceiling	50,000 ft
Un-refuelled range	7650 nm
Typical bomb load	27 bombs internal & 18 external (MK –82)
Thrust (sea level)	8 x 17,600 lbs

Surveillance & reconnaissance

The first military use of an aircraft was gathering data on enemy positions and this still takes place. Aircraft that are used for information gathering are very different in design. First we should distinguish between operations outside and those inside the threat area. A second criterion is mission related: surveillance over an extended period or short-haul reconnaissance. Both criteria will determine what type of airframe should be used. The four possibilities are presented in a matrix:

	No threat	Threat
Surveillance	Optimise the aircraft for low speed; this will give the best endurance. Unmanned aircraft are particularly effective in this role. Examples of this class are AWACS, Predator, and Maritime Patrol Aircraft.	Aircraft needs endurance and protection. This can be achieved with a combination of extremely high altitude flight and stealth. Aircraft with extremely low wing loading are required. Examples Global Hawk and TR-1.
Cover area	Optimise the aircraft for range. Platforms will look like airliners. An example is the KC-135R but satellites can be used as well.	Speed, self protection and stealth are required. Examples are tactical reconnaissance aircraft, but satellites can be used as well because they can operate above the threat.

As can be concluded from the matrix, there is no single surveillance or reconnaissance platform. The platforms include modified commercial aircraft, specialized fighters-bombers, UAVs, some very specialized aircraft and satellites. The platform characteristics of reconnaissance fighter aircraft are, of course, similar to those of the regular fighter-bomber. There is also a similarity between modified airliners and strategic transport aircraft in terms of range and endurance. Two platforms that need further discussion are the UAV and the satellite.

Unmanned Aerial Vehicles

The UAV is the preferred choice for surveillance missions[39]. The crucial parameter for surveillance missions is very long endurance. To give a platform a long endurance certain requirements must be met and drag must be low. This leads to a design for a wing that generates a lot of lift at low speed. The design resembles that of a sailplane. Weight must be as low as possible. This consideration favours the use of an unmanned design as this saves on the weight of the crew, life support systems and pilot instruments.

There are several types of UAVs with endurance ranging from a few hours to several days. Interesting to note is that most UAVs lack self-protection capability (less weight and cost), which means that they can operate only when there is a favourable threat level. Below the figures for the GNAT-750 are depicted. This airframe, developed in 1989, was the basis for the later development of the Predator.

GNAT-750	
Weight	1140 lbs
Payload	330 lbs
Endurance	48 hours
Ceiling	25,000 ft
Wingspan	35 ft
Length	16 ft
Cruising speed	±75 knots

Satellites

Satellites have several applications from which airmen benefit. Communications, reconnaissance and navigation are the most widely used. Each application requires a special kind of satellite and orbit. A satellite in orbit is not free of gravity. There is no end to the gravitational field of the earth, but the strength of this field decreases with the square power of the distance. The speed of the satellite in orbit should be so high that its centrifugal force balances the gravitational force of the earth. This force is always directed towards the earth's centre. To balance this force the centrifugal force of the satellite should be exactly opposite the gravity force. This is only possible when the centre of the orbit coincides with the centre of the earth.[40]

[39] UAVs are also applicable for other mission types (e.g. dirty and dangerous missions) but these mission types might generate different design considerations.
[40] The result is that the satellite always follows a large circle around the earth.

This can be an orbit over the poles, or an orbit with some angle (declination) with the earth's axis. It is impossible for a satellite to orbit for example around the 30°N parallel[41]. The higher the orbit, the longer it will take for a satellite to complete it. In the lower layer a complete orbit takes approximately one hour and 40 minutes. GPS satellites have a much higher orbit taking approximately 18 hours.

A special case is the geo-stationary satellite, which always orbits around the equator at approximately 36,000 km. Its velocity exactly matches the turn rate of the earth, therefore its position relative to the earth is always the same. Geo-stationary satellites are extremely useful for communications and weather data collection. But the space available is limited: all these satellites in the same orbit over the equator should be well separated to minimize interference.

For reconnaissance satellites the orbit altitude should be as low as possible to give the best possible view. Normally, the orbit altitude is between 200-400km and orbit time is approximately 100 minutes. The consequence of the orbit is that a particular location is only spotted for a few minutes after which the satellite is beyond the horizon again. A relevant term used for satellites is revisit time. If the (low) orbit is exactly over a meridian, the satellite will return over the same spot every 100 minutes. This is the shortest time possible. The consequence of course is that only the locations along that meridian are reconnoitred. If the orbit has some declination with the meridian, the path over the earth is continuously changing. Consequently a larger part of the earth can be covered. It can require many orbits (depending on the inclination angle) until the satellite is over the same spot again. This time is called the revisit time and may be as long as two weeks.

Helicopters

The limitations of fixed-wing aircraft are that they need a runway and are limited in low-speed capability. Helicopters or rotary-wing aircraft don't need runways and combine terrain-independency with a relative freedom in speed. To create a hover capability by rapidly moving the wings is as old as Mother Nature. Most flying insects and many birds are capable of doing this. A practicable engineering solution proved to be more difficult in the past. The idea of flapping was disbanded early in the development and a rotary design proved to be the easiest solution. But it came with a price.

If we compare the surface area of a helicopter rotor with the wing area of a fixed-wing aircraft of the same weight, it is obvious that the helicopter wing area is much smaller. The relatively small wing area suffices because the rotor blade always turns at high speed. As discussed in chapter 2, at high speed it is easier to generate enough lift and the lift-dependent drag is much smaller. But there are limits: the first constraint is that the tip speed should be below Mach to stay away from shock waves, which would produce noise and drag. Normally the design tip speed is around .7M.

The consequence of the first constraint is that in forward flight the retreating blade must continue to generate enough lift. In forward flight, the speed of the retreating blade in relation to the air is, of course, much lower. The retreating blade will need a higher angle of attack than the advancing blade in order to balance the lift forces on both sides of the helicopter. But at a certain forward speed the retreating blade might stall. This constitutes the maximum speed of a helicopter. Most helicopters

[41] In that case the centre of the orbit is not the centre of the earth's centre but a point on the earth's axis half way between the Equator and the North Pole.

will not fly much faster than 140 knots because of this constraint. A related problem is the fact that the advancing blade can encounter shockwaves. The shockwaves will be most pronounced at high speed and high angle of attack[42] (high collective settings)[43]. Mach waves will significantly increase the drag and increase noise.

Rotor-generated lift also constrains flight at altitude. At high altitude, a fixed-wing aircraft can compensate for the decrease in density (and lift) by flying faster, which will not cost more energy because the drag is also decreased. As long as the engine can generate enough thrust and the speed is still below Mach, flying high is not a problem. However, the helicopter rotor speed is already at its maximum and can not be increased. The only way to cope with the decrease in density is to increase the angle of attack of the rotor blades. An increase in the angle of attack gives more drag while at the same time engine power is decreasing with altitude. So the helicopter is doubly affected by the decrease in density and consequently the ceilings of most helicopters are considerably lower than those of transport aircraft. There are some other peculiarities. With helicopters the power required is normally high in hover and at maximum speed. The minimum drag is normally at around 60 knots. The consequence of this is that a helicopter is able to fly at a certain altitude without being able to land. In short, altitude has a considerable effect on payload and hover capability.

A typical helicopter feature is the vortex ring state (also called settling with power). A vortex ring state can be induced when both the helicopter descent rate and the collective settings are high . The high sink rate allows the air that is pushed down by the helicopter rotor to move around the blade edge and to re-enter the rotor on the top side. The re-entering air will decrease the rotor lift[44] and an increase in collective setting will only aggravate the situation. For recovery, the pilot has to decrease the collective setting and increase the forward speed. If this action is not taken in time the helicopter may crash.

Another consequence of rotor-generated lift are high structural loads. The centrifugal forces on the rotor are several times higher than the lift force and the rotor generates a high vibration level. These structural loads lead to severe design restrictions and make the helicopter more maintenance-intensive than a fixed-wing aircraft carrying the same load.

Driving the rotor generates a moment in the horizontal plane. The standard way to counter this moment is by the force of a tail rotor. The required tail rotor force is high in the hover state and lower when the helicopter has some air speed. The reason for this is that the tail itself generates some directional stability in forward flight. A particular situation arises when the helicopter is hovering in crosswind. If the crosswind comes from the wrong side[45] the tail rotor has to compensate for the wind force on the tail and for the rotor moment, making hovering in crosswind conditions one of the most severe flying conditions for a helicopter. It can however be necessary for attack helicopters when they have to point their weapons in the right direction. The tail rotor has a few disadvantages:

---

42 High angle of attack will cause a significant increase in the airflow speed on the upper side of the rotor blade.

43 The collective regulates the pitch setting (and so the angle of attack) of all the blades collectively. When more power is required (e.g. turns & climbs) the collective setting is increased. With the stick, called cyclic, the pitch difference between the blades is changed to create a momentum in order to roll or to pitch.

44 If air is pulled from below and pushed downwards again there is no resulting upward force.

45 If the tail rotor is pushing the air to the left and the wind is from the right the wind is coming from the wrong side.

Takes away power (approximately 10%);

Creates a slight side force that must be balanced by the main rotor;

The high speed tail rotor is vulnerable, especially in confined areas and dangerous for people boarding or disembarking.[46]

Several design solution have been found to overcome some of the disadvantages of the tail rotor. Interesting to mention are the shrouded tail rotor and the NOTAR. The NOTAR uses the exhaust stream of the engine to counter rotor momentum. [47]

There are other ways to cope with the inherent disadvantages of a rotor system. The first is to use two counter-rotating rotor systems in tandem. The second is to use two counter-rotating rotor systems mounted on one shaft. The third is to use two intermeshing rotor systems on two laterally mounted shafts. These solutions overcome the above mentioned drawbacks of the tail rotor. A technically complex solution is the tilt rotor design as used in the Osprey.

CHINOOK CH-47	
Empty weight	24,000 lbs
Maximum gross weight	50,000 lbs
Range	385 nm
Ceiling (no pay load)	20,000 feet
Hover ceiling max payload	5,000 feet
Centre cargo hook	26,000 lbs
Forward and aft cargo hooks:	17,000 each or 25,000 lbs tandem
Cruise speed	150 Knots
Useful load	26,000 lbs

Missiles

There is a wide variety of missiles, ranging from short-range air-to-air missiles to intercontinental ballistic missiles. The larger the missile the longer its range. The logic behind it is simple: the missile drag is proportional to the frontal area and the available energy is proportional to the weight of the fuel. If all dimensions (length and width) are doubled the frontal area will be four times as large but the fuel capacity will be increased eight times.

There are some features that influence range that are common to all missiles and worth understanding. First, most missiles are only powered during the small initial portion of their trajectory and deplete their kinetic energy during the remaining portion. The missile has two ways to control its flight path: a moveable nozzle and aerodynamic controls (similar to an aeroplane). The moveable nozzle is only effective during the propulsive phase while the aerodynamic controls are effective as long as the missile has enough speed. Consequently, if the missile does not have a moveable nozzle but only aerodynamic controls, it will not be controllable at low speed. This is a typical problem in the initial phase of every ground-launched missile; therefore the initial phase should be as short as practicable. Normally, ground-launched missiles are designed with a booster engine for fast initial

[46] According to U.S. Army safety reports, 20 percent of all peace time helicopter accidents are caused by tail rotors.

[47] The tail boom is designed in such a way that the rotor down wash mixes with air which is relieved from slots in the boom. This creates an aerodynamic force, which counters the momentum as well. This is called the Coanda effect.

acceleration and a sustainer for the remaining part of the powered flight. Missiles fired from aeroplanes already have an initial velocity and normally don't require a booster. Almost all missiles are designed for high supersonic speed, therefore the wings and controls have a typical supersonic design: thin and small to decrease the drag. But small wings result in low lift capability at lower speed. Given these design considerations the following observations on missile performance can be made.

Missile range is dependent on how much energy can be given to the missile. The missile engine just adds a velocity component to its initial velocity. Because kinetic energy is related to the square power of the velocity, the effect of the speed of the launching vehicle is profound. E.g. compare two aircraft, one flying at 1000 knots and the other at 500 knots. If both fire the same missile with a 1000 knots' delta velocity capability, the missile from the faster aircraft has 78% more kinetic energy and about the same increase in range. Missile range is also affected by the potential energy increase or decrease required. E.g. a missile at Mach 2, descending to a target 10,000 ft lower receives a 16% energy increase. Shooting downwards or upwards makes for a considerable difference in range.

Missile manoeuvrability is normally high in terms of g-load. It is not uncommon to see missiles turn with 30-g. This does mean that such a missile will always outperform a fighter turning at 9g because turning performance is dependent on g-level and velocity. A typical comparison is made in the following table, showing that the missile is able to exceed the turn rate of the fighter, but that its radius is more than twice as large.

	Fighter	Missile
g-level	9	30
Mach	.7	2
TAS at 15000'	439 kts	1253 kts
Turn rate (deg/sec)	22	26
Turn radius	1903 ft	4635 ft

Strategic missiles

The strategic bombing mission is no longer limited to bombers. A wide variety of missiles possess a strategic attack capability. The list includes intercontinental ballistic missiles, medium-range ballistic missiles and cruise missiles. These weapons, although expensive, can be more cost-effective than strategic bombers and all have the capability to destroy static targets in the enemy heartland. The only bomber mission missiles can't replace is loitering above the battlefield and giving on-call air support to ground operations. This tactical bomber mission is only feasible when air superiority is achieved.

Sea basing

It was not long after the invention of aircraft that aeroplanes were launched from ships. The obvious advantage of sea basing airpower is that the force can be brought close to the enemy without the need to obtain overflight rights and staging permissions from other countries. Another reason for sea basing is that other fleets were fought more and more with aircraft and less with ships engaging each other. But sea basing comes with a price. First of all, equipment placed on board ships is exposed to salt water, which is highly corrosive. Secondly storage space, even on large carriers, is limited and aircraft wings and rotor blades have to be made foldable. Thirdly, the limited deck space requires alternate procedures for take-off and landing. These will be discussed in more detail below.

- **Starting from a carrier.** Carrier-based fixed-wing aircraft have only a few hundred feet to accelerate to takeoff speed. The required longitudinal acceleration is approximately 2-3g depending on takeoff speed and wind speed over the deck. This acceleration can not be accomplished by the aircraft engines alone, so assistance is needed from a steam catapult, which is normally connected to the nose gear of the aircraft. Consequently, the nose gear has to be extremely strong (and heavy). It is good to note that the catapult not only accelerates the aircraft but also restrains it from forward motion before the aircraft is launched. (The aircraft is in Full AB before the catapult is fired.) Another way to minimize the takeoff length is to have a 'ski' jump installed on the carrier. The principle behind the ski jump is to convert forward speed to vertical speed. Because the aircraft is still well below its takeoff speed when it 'jumps' in the air, the pilot can not maintain level flight and it follows more or less a ballistic trajectory. The trick is that the aircraft will accelerate to takeoff speed before the trajectory brings it too close to the water. This type of start requires an aircraft with a lot of thrust and also a strong landing gear to withstand the forces of the jump.
- **Arresting.** During normal landings, a pilot will flare and reduce vertical velocity in order to touch down gently. This method is not practical when landing on a carrier. The plane must touch down in a very narrow spot so the hook can grab one of the cables. So, the landing itself is firm without a flare. This puts a heavy load on the landing gear and the aircraft structure, which must be strong and heavy. The same is of course true for the hook, which has to decelerate the aircraft within a few hundred feet, which requires a deceleration of approximately 3-g.
- **Landing from the hover**. Another landing method is the vertical landing from a hover such as employed by helicopters but also by the GR-1 and the Harrier. For fixed-wing aircraft this requires a strong engine with nozzles that can be turned vertically. The Marine version of the F-35 uses an extra lift fan as well, in order to generate sufficient lift at low speed. A first impression might be that landing on a ship does not mean any additional requirements for helicopters. However, landing on a moving (rolling and pitching) platform does create problems. A particular problem is that the rotor blades may become unstable during the deceleration on a moving platform. For this reason a rotor brake must be installed.

The bottom line is that navy aircraft need to be stronger, will be heavier and consequently their performance is lower than their land-based counterparts, and the cost of the airframe will be considerably higher.

Chapter 4 – The theory behind sensors

Introduction

All modern military aircraft are equipped with a variety of sensors such as targeting pods, lasers, radars and missile warning equipment, all using the Electro Magnetic (EM) spectrum. Understanding the characteristics of the EM spectrum helps to understand the possibilities and limitations of these sensors. Therefore, some fundamentals of the EM spectrum will be discussed before proceeding to the actual sensors in the next chapter.

The electro magnetic spectrum

General

Electro Magnetic (EM) radiation is not a human invention that started with the development of the first radio, but a natural phenomenon. Visible light is a form of EM energy everyone is acquainted with, but the sun also radiates at other frequencies that cannot be observed by the human eye. For these frequencies we need special sensors. The fact that (most) radiation cannot be seen and that some EM frequencies are dangerous have led many to believe that all radiation is life-threatening. However, the opposite is closer to the truth. Without EM radiation, no form of life would be possible. In fact only the very high frequencies (e.g. X-rays) have the ability to destroy tissue or trigger DNA mutations. Most frequencies are harmless, but as with sunlight, too high a dose can give unwanted side (heat) effects.

The EM spectrum	
Frequency-band	Name
30 KHz-300 KHz	Low Frequency (LF)
300 kHz – 3 MHz	Medium Frequency (MF)
3 MHz -30 MHz	High Frequency (HF)
30 MHz- 300 MHz	Very High Frequency (VHF)
300 MHz – 3 GHz	Ultra High Frequency (UHF)
3 GHZ –300 GHz	Super High Frequency (SHF) (Micro Waves)
10,000 – 400,000 GHz	Infra Red radiation (IR)
384,000GHz -690,000 GHz	Visible Light

EM energy sources

There are several ways to generate EM energy:

- Alternating electric current passing through a conductor (antenna) causes radio waves.
- The movements of molecules and electrons cause light and infrared.

- The movements of the atom nucleus cause EM radiation in the form of X-rays.

Basic parameters

EM waves, as the name suggests, consist of a changing magnetic field and a changing electric field perpendicular to each other. EM waves can be classified by the following basic parameters:

- Intensity, normally expressed in Watts per square meter;
- Frequency, the number of cycles per second expressed in Hertz (Hz);
- The wave length or λ is directly related to frequency as shown in following formula:

$$\lambda = c/f$$

(c= speed of light & f= frequency in Hz)

- Polarization, which is the direction of the electric field of the EM wave. A vertical radio antenna will have a vertical polarization (the alternating electric current flows through the antenna) and the magnetic waves form a (horizontal) circular pattern around it. Polarization can be vertical, horizontal or circular, which is a mix of both.

> During aircraft manoeuvres the orientation of the aircraft changes and with it the orientation of its antennas. Consequently the polarization of the signal may also change. For successful communication both transmitter and receiver need to have the same frequency and polarization. By applying circular polarization this becomes independent of the orientation of the aircraft.

- Phase. Because EM is a wave, the magnetic and electronic fields are constantly changing in magnitude and direction. The phase, which is expressed in degrees, indicates where in the cycle the wave is. (180º is halfway and after 360º the cycle is complete.)

> Phase is the key to understanding modern phased array radar antennas. If two elements on a radar antenna are in phase, their signals will add up and form a wave front straight ahead. In other directions they dampen each other. By changing the phase of antenna elements, a phased array antenna can be steered.

Propagation

All EM waves propagate at the speed of light, which is approximately one billion feet per second in vacuum. The speed of EM through a medium is lower. When EM energy passes from one medium to another and encounters a speed change, diffraction will occur. The speed difference between layers in the atmosphere is the reason that EM waves don't move exactly in a straight line but bend slightly. The lower frequencies bend more than the higher frequencies, as can easily be detected when watching light passing through a prism (blue has a higher frequency than red).

This is also the reason for medium wave radio transmissions to reach further than the higher FM frequency. The Line of Sight (LOS) range of an EM source can be approximated by the formula:

$$R = 1.3 \times \sqrt{H}$$

(R is LOS range in nautical miles, H in feet)

Line of sight table	
Altitude (feet)	Line of sight
1000	41 nm
10,000	130 nm
20,000	184 nm
30,000	225 nm

For instance, an AWACS aircraft flying at 20,000 ft will be able to detect a low-flying aircraft at a maximum distance of 184 nm. However, if the target has some altitude, say 1000 ft, the LOS range of the target -41 nm- has to be added, which will give a maximum detection range of (184 +41 =) 225 nm.

An exception has to be made for the short-wave radio band (7 – 30 MHz). Short-wave EM radiation reflects against ionised layers in the upper atmosphere and 'hops' around the earth. Before satellite communication existed, this was the only frequency band that could be used for global communications.

Energy decreases with distance at the same rate as the surface area of the radiation increases, namely the square power of the distance. Radar energy is a special case. Because radar works with reflected energy, the signal strength decreases on the way out and on the way in. Therefore the received energy decreases by the fourth power of the distance.

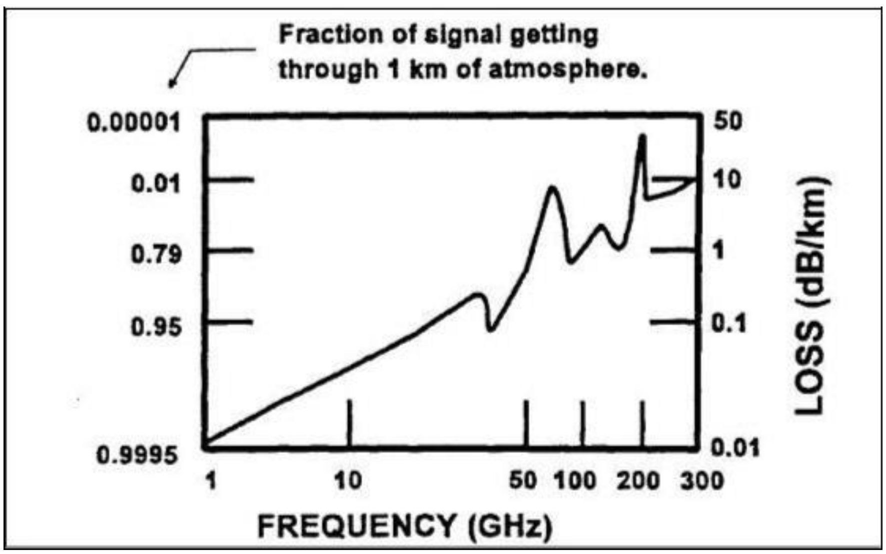
Absorption, reflection and scattering

In theory EM energy can propagate an infinite distance, losing intensity with each mile travelled, but the signal can get distorted when passing through a medium. When EM energy passes through a medium it can be absorbed, scattered, reflected or pass through without any disturbance. The way it behaves is dependent on the combination of frequency and medium. Some examples are given below:

- Visual light is scattered by dust, smoke and small particles in the atmosphere. At a low sun angle the light passes through more air and is consequently scattered more. Blue light is scattered more than red, which explains the red sunset. (If blue is filtered out, the red part of the spectrum is more dominant.) It also explains the blue sky, which is scattered blue light from small particles in the atmosphere. At very high altitude the blue colour of the sky will disappear and the sky will turn black.
- Radio waves will pass through stone buildings and glass but will not penetrate metal structures.

- Radar energy is absorbed and reflected by water droplets so radar range is limited in rain showers. That is also why radar is well suited to detect showers (like weather radars do). However, detecting an aircraft in a shower will be harder than in clear weather. Radars working in the lower frequency band (3 GHz) are less affected than the higher frequency radars. Radars working at 10 GHz (the most commonly used band in aviation) are moderately affected, while millimetre wave radars (30-40 GHz) are strongly affected by rain.
- Water vapour and carbon dioxide absorb a large portion of IR energy. Within the IR spectrum two bands are usable: the far IR band 8-12 μ, and the mid IR band 3-4μ. IR energy, like visual light, will not pass through clouds, sandstorms, snow and heavy rain, nor will it pass through normal glass. Therefore an IR detector needs a lens made from special material.

The absorption characteristics of the earth's atmosphere are of special interest to the aviator. If EM frequencies are absorbed by the earth's atmosphere, the signal strength will be heavily attenuated and the frequency may not be usable at all.



Bandwidth

EM energy sources do not emit on a single frequency. In reality the signal occupies a certain bandwidth. Normal white light for example has a large bandwidth (384,000GHz - 690,000 GHz) while a laser will have a very small bandwidth. EM waves can be modulated in amplitude[48] or in frequency[49] or be used in pulses (as for radar). By this modulation, data is added to the signal. This modulation frequency will increase the bandwidth of the original signal.

48 Amplitude modulation is a change in the carrier signal strength. This modulation carries the information ,which can be speech, video or data.

49 Frequency modulation changes the carrier frequency. The frequency change can be the frequency of a speech signal but also of a data signal.

If CF is the main or carrier frequency and MF is the modulation frequency, the signal bandwidth will be from CF-MF to CF+MF. E.g. if the carrier frequency is 100 Mhz and the modulation frequency is 20 kHz the bandwidth is from 99,990 kHz to 100,010 kHz. For a pulsed signal the bandwidth is 2/pulse time.

Most airborne radars work around 10 GHz. A common pulse length is 1 microsecond. This pulse length will generate a bandwidth of $2/10^{-6}$ second =2 MHz. The bandwidth is from 9,999 MHz to 1,001 MHz.

Available bandwidth is dependent on the main frequency (or carrier frequency) of the signal. A one percent bandwidth of a 1 MHz carrier frequency (medium wave radio) is only 10 kHz. A one percent bandwidth in the colour red (375,000 GHz), which is used in fibre optics communication, yields a bandwidth of 3,750 GHz. The available bandwidth of an EM signal is a very important parameter and determines how much data can be added to a signal. This explains why a medium wave radio can only transmit one programme while a fibre optic line can handle many thousands of communication lines simultaneously.

The standard UHF radio of an aircraft has only a 25 KHz bandwidth; this is more than enough for radio communications, but not enough for the high speed data link required on the modern battlefield. The standard air data link (LINK-16) uses a bandwidth of 2MHz.

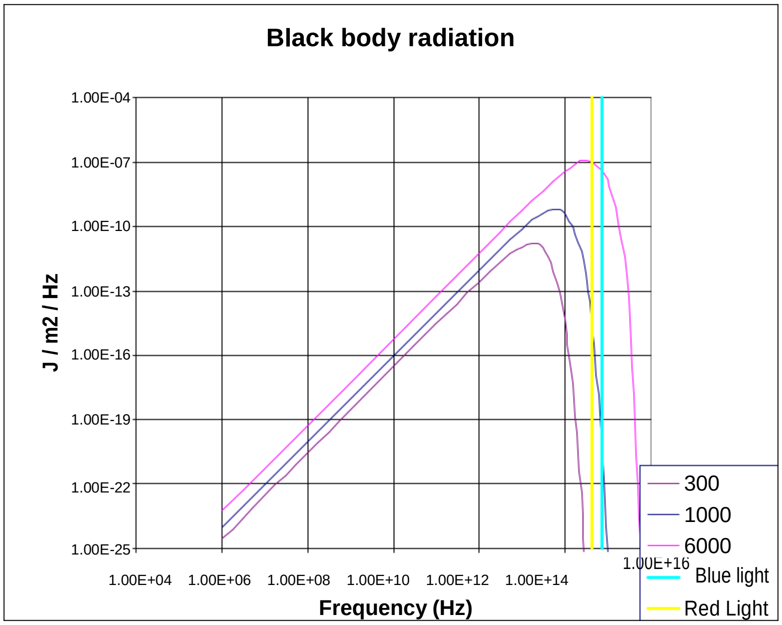
Heat and signature

Part of EM radiation is generated by the movement of molecules. The movement energy is directly related to the absolute temperature. If the temperature of an object increases, the amount of radiated energy will increase rapidly.[50] The peak frequency of the radiation also increases with temperature. At room temperature the maximum radiation is in the far infrared band (between 8 and 12 Micron). At 5500° Kelvin (the temperature of the sun's surface) the peak energy is in the visible spectrum (between 0.4-0.8 micron).

Fortunately 8-12 micron radiation is not completely absorbed by water vapour as can be seen in the picture of the IR windows in the previous paragraph. The result is that this IR energy (which every object at room temperature emits) will not be absorbed and can be detected during day and night operations. This is exactly what a thermal imaging system does.

The next graph depicts the radiation spectrum at different temperatures. Note that the spectrum of the higher temperature completely encompasses the lower temperature, showing that higher temperature objects radiate more energy at every frequency. The energy values depicted on the scale might seem low but the value is Joule per square meter per spherical radian per Hertz. Over the complete visual band (3.46×10^{14} Hertz) the power of the sun is immense, more than 32 MJ per square meter per spherical radian (at the surface). On the other hand, radiation levels at room temperature are much lower. At 300°K the total radiation in the 8-12 micron band is about 100 watt per square meter. This might seem enough but a thermal imaging sensor does not measure the temperature of an object but the minute temperature differences between different parts of an object, making an IR sensor extremely sensitive.

[50] It has been proven that total radiated energy increases with the fourth power of the absolute temperature.(Planck's Law)



The energy graph shows the theoretical radiation of a black body, which can radiate at each frequency equally well. In real life every material has frequencies at which it will neither radiate nor absorb. This is a welcome feature, for the lens of a thermal imaging sensor should neither emit nor absorb IR energy nor should the air through which the energy must pass, emit or absorb energy.

Sensor theory

General

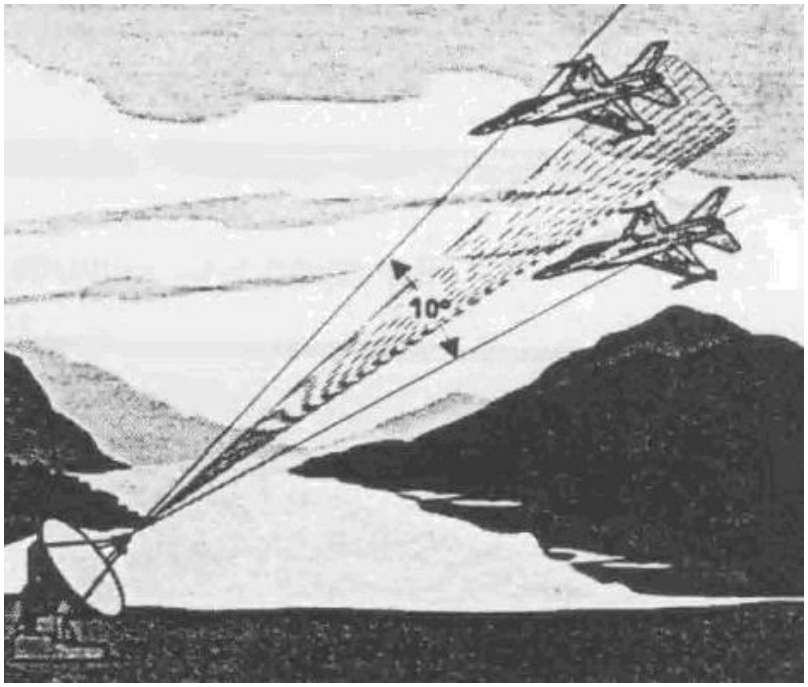
All EM sensors may differ in appearance but face the same challenge. They have to collect as much energy as necessary in the correct frequency band to measure at high resolution. Radio and radar equipment collect EM energy with antennas, IR and visual sensors use optics. In this paragraph we will use the more general term aperture instead for every device that collects EM energy. Resolution is a combination of accuracy and sensitivity. The accuracy can be an angular accuracy that depends on the size of the aperture, but accuracy may also be necessary in measuring range or frequency. Sensitivity is not only the capability to measure very small quantities of EM energy but also minute differences in intensity and frequency. As with the human eye, a small change in contrast may reveal the target better than the absolute amount of energy that it emits.

Aperture and resolution

The size of the aperture should be large enough in relation to the receiving wavelength, normally a quarter of the wavelength is a minimum size. Secondly, the larger the aperture, the more energy it can collect, which will be favourable for its range. Thirdly, the size of the aperture in relation to the wavelength (λ) is also the determining factor in the angular resolution. Angular resolution can be approximated by the following formula:

> Resolution (radials) = λ / aperture diameter
>
> (Wavelength and diameter should be expressed in the same units!)

The energy distribution around the antenna axis can have many different shapes depending on the kind of aperture used, but by definition the beam width[51] is the area between the 3 dB points. This simply means that 50 percent of the received or emitted energy is outside or inside that area. Consequently 50% is outside that area and some of the emitted energy can be focused in side lobes. Especially in radars these side lobes can give unwanted returns. The next picture shows that if both targets are within the same beam the radar will detect only one target and can not resolve between the two targets (assuming both targets have the same range).



Beam width effect of a radar antenna

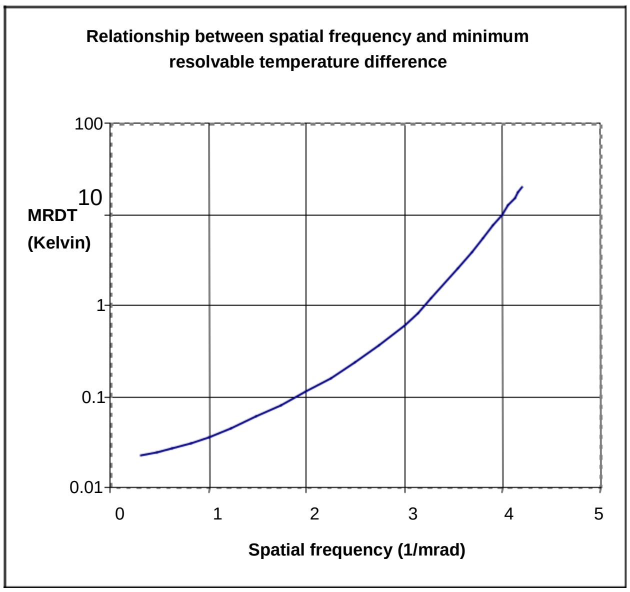
Space in an aircraft is a scarce commodity; therefore apertures from sensors can be much smaller than operationally required. Sometimes there is a way around this. The received signals are integrated over time and correlated. This is done with ground observation radars and is called Synthetic Aperture Radar (SAR).

A radar warning receiver with a 15 cm aperture and searching for 3 cm radar waves has an angular resolution of (60 x 3)/15 = 12 degrees. A radar antenna on a fighter may have a diameter of 90 cm. When it uses a 3 cm wavelength signal the angular resolution is 2°.

This also works for an IR sensor. Let's assume a targeting pod is equipped with a far IR sensor (wavelength average 10 μ). If the designer wants a resolution of 1 foot at 10 nm (60,000 ft) the required angular resolution is 0.001°. This would mean that the aperture should be at least 60,000 times the λ (10 micron) which gives 60 cm. However, for an optical sensor, working between 0.4 and 0.8 micron, a 3 cm aperture would give the same result. This simple rule explains why IR sensors on a targeting pod are bulkier than visual sensors.

51 There is a difference in beam width for planar array antennas and parabolic dishes. The first has a beam width of $0.88 * \lambda / D$ (in radians); the second has a beam width of $1.02 * \lambda / D$ in radians).

The size of the aperture is not the only feature that determines the angular resolution, signal strength variation in relation to sensor sensitivity is also important. For visual sensors the contrast and light level should be sufficient in relation to the sensitivity of the sensor. For IR sensors the temperature difference should be large enough and for radars the RCS variation of the ground features should be large enough. In general, there is a non-linear relationship between the measured angular resolution of a sensor and the signal strength. Below, an example of this relationship is given for an IR sensor. In more general terms, the size of the aperture determines the maximum achievable angular resolution while the signal strength variation in combination with the sensitivity of the detector will determine how much of the angular resolution can actual be realized.



Radar theory

The basic principle of radar is easy. Emit a burst of EM energy in a given direction and measure the time until the echo arrives. The time interval is a measure for the distance and by rotating the radar antenna a radar picture can be built. This principle works fine in simple sea radars that are used for navigation. These radars have to detect ships with large radar cross sections, navigating in water that hardly gives any radar returns[52]. For airborne radars the situation is more complex. The heart of the problem is that the radar returns of aircraft are almost insignificant compared to the ground returns. It is literally like finding a needle in a haystack. Fortunately aircraft move, and this movement gives a frequency change to the radar frequency. This frequency shift can be detected by *coherent pulse Doppler radars*. The techniques to filter the right frequencies will be explained later, but this technique is not foolproof and pilots exploit the shortcomings of airborne radar during aerial combat. Before explaining the coherent pulse Doppler radar, the standard radar and radar parameters will be discussed.

52 Except in rough sea; waves will reflect radar energy especially if the radar is looking into the wind.

Measuring distance and scan times

Radar waves travel at 1 billion feet per second. If we want to look at objects as far away as 250 nm, the radar energy needs approximately 3 milliseconds to cover that distance there and back. As a single return will not stand out between the clutter, multiple hits are required. Assuming that the radar needs approximately 10 replies to distinguish the target from the clutter; the radar has to dwell on the target for at least 0.03 second. This might seem short, but if the radar scans 360° with a 1° accuracy, the time for a complete scan has increased to more than 10 seconds! Within 10 seconds a fighter aircraft could have made a 180° turn. If the radar beam weren't so narrow, a faster but less accurate scan would have been possible. The first limitations of radar are therefore:

If the radar range increases, scan time increases.

A smaller beam gives a more accurate target position but requires more scan time.

Pulse repetition frequency

The pulse repetition frequency (PRF) is one of the key parameters of radar. In principle low PRF is used for long-range observations and high PRF for short distances. If the PRF is 1000 Hz, the radar energy has 0.001 second to travel to the target and back. This gives a maximum range of 500,000 feet or 83 nm. So, PRF limited radar range is:

Radar range = c / (2 x PRF)

(c is the speed of light).

The radar warning receiver uses PRF as one of the features to identify a threat. Some radar warning receivers will transform the received PRF into an audio signal that the pilot can hear in his headset. Most airborne radars use a variety of PRF, firstly because the radar has to be capable of working at different ranges. There are other reasons as well, one of them being range ambiguity. The radar can not distinguish whether an echo is from the last pulse or from the last-but-one pulse[53]. By working at different PRFs this ambiguity can be resolved.

Detection range

The range at which a target can be detected is dependent on the following set of parameters:

- **Emitted energy**. The emitted energy depends on the average power (P_{avg}) and the amount of time over which returns are integrated (T_{int}). In mathematical terms $\mathbf{P_{avg} \times T_{int}}$.
- **Antenna gain and effective antenna area.** Antenna gain is the parameter that describes how well the radar transmits its energy in one direction, and is dependent on the antenna size. The symbol for antenna gain is **G.** The effective antenna area determines how much energy is captured. The symbol is $\mathbf{A_c}$.

53 If a large target exists just outside the PRF limited range, the echo will still be large enough to be detected.

- **Attenuation by atmospheric conditions**. The primary cause of radar attenuation is rain, for simplicity reasons attenuation is not included in the formula below.
- **Radar cross section of the target**. The radar cross section determines how much energy is reflected back to the antenna. The radar cross section is dependent on the physical size of the object, its reflectivity and the directivity of the reflection. The radar cross section will differ considerably depending on the angle at which the target is seen. The symbol used is: **σ**.
- **The minimum discernable energy.** The receiver needs to receive a minimum amount of energy for detection. We use the term S_{min}.

Emitted EM energy can be imagined as a sphere: its surface area increases with the square of the distance, consequently the energy per square meter decreases with the square power of the distance. Because radar energy has to travel out and back, the received energy decreases with the fourth power of the distance. An approximation[54] of detection range is given in the following formula:

$$\boxed{R_{\max} = \sqrt[4]{\frac{P_{avg.}\mathrm{T}_{\mathrm{int}}.G.\,A_C.\sigma.}{(4\pi)^2.S_{\min}}}}$$

In reality detection is a probability that increases if signal-to-noise ratio increases. A critical choice for the radar design engineer is to set the detection level at a threshold that optimises radar detection range without introducing a lot of false alarms in the radar[55].

Radar resolution

Angular resolution has been discussed in the previous paragraphs, but radar discriminates in distance as well. The combination of angular resolution and range resolution is called the resolution cell. The practical application is that two targets that are separated from each other by less than a resolution cell will be seen by the radar as a single target. The detection resolution is half the length of the emitted pulse. A typical pulse time is one microsecond, which gives a pulse length of 1000 ft and a resolution of 500 ft. If targets are closer together the echo from the second target will overlap the response of the first target.

Pilots can exploit the size of a resolution cell by flying a formation within one resolution cell, thereby disguising the number of aircraft in the formation. At a certain moment they can split up fast, confronting the enemy with an unexpected number of targets with and even having a chance for one of the targets to split off undetected.

A range resolution of 500 ft may be acceptable for search radars but is not good enough for accurate tracking and accurate ground mapping. One solution is to use a shorter pulse, but there is a limit to that. Shorter pulses increase the bandwidth and decrease the amount of energy received. In most mapping radars pulse

[54] The signal attenuation is not taken into account in this formula.
[55] Peter Swerling has developed special relationships for required signal to noise ratio's and detection probability.

compression techniques are used.[56]. In most Synthetic Aperture Radar (SAR) applications a resolution of one foot is attainable.

Doppler measurement general

As stated earlier, the key to airborne radar is that not only range and direction are measured, but also the frequency shift of the signal due to its movement. This measurement is crucial to discriminating the target from among the massive ground clutter. Nowadays, this filtering is done digitally by using the so-called Fast Fourier Transformation (FFT). The advances in digital computer power have made it possible to perform the massive number of calculations needed in real time. The theory behind the filtering is not easy. In the remaining part of the chapter this theory is explained with some simplifications, in order to give the reader an understanding of the possibilities and limitations of pulse Doppler radars. The remaining part of the chapter may be skipped by the less technically interested reader, but the reader should realize the following limitations:

- It takes several different PRF's to measure target velocity (for example 7-8 different PRFs);
- It takes several different PRFs to measure target distance;
- Each target is only seen by in a limited number of PRFs, for example 3 out of 8. This Doppler visibility will influence the detection range.

Doppler measurement in detail

If a radar pulse is reflected by an object that is moving relative to the radar, the frequency of the returning echo changes. In case of airborne radar the frequency is changes by the speed of the transmitter and the receiver. In formula:

$$\text{Freq. Doppler} = 2 \text{ x } (V_r + V_t) / \lambda$$

Where V_r is the speed of the radar relative to the target, V_t is the speed of the target relative to the radar and λ is the wavelength of the radar.

For airborne radars, working at around 10 GHz, or with a λ = 3 cm the Doppler shift is approximately 35 Hz per knot closing velocity. So even at a high closing velocity of thousands of knots, the Doppler shift is only 35,000 Hz, which is low compared to the bandwidth. As was calculated earlier, the bandwidth of an airborne radar working at 10 GHz and with a one microsecond pulse is 2 MHz. It is intuitively obvious that with a bandwidth of that magnitude it is impossible to measure a shift of only 2 % of that bandwidth.

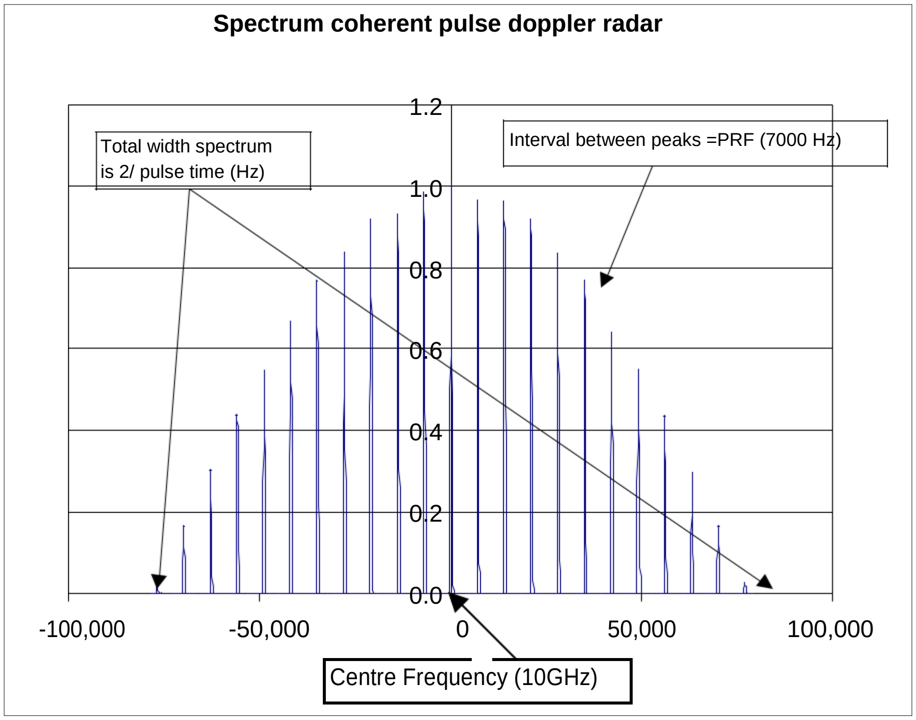
It would have been different if the pulse width were wider. If the pulse width were infinite (the ultimate case) every frequency could be measured. However, with a continuous signal it is easy to measure a frequency shift but hard to measure distance.

The solution to the problem is using a coherent pulse Doppler signal. Coherency means that each subsequent pulse is in phase with the previous signal. Or, to put it differently, the radar signal can be considered as one continuous signal that is interrupted during some periods.

The spectrum of this coherent pulse Doppler radar is completely different from that of non-coherent (standard) radars and is depicted in the following graph. The

[56] One common technique is 'chirp'. The chirp technique adds a linear frequency modulation to the emitted pulse. In the receiver the radar filters the different frequencies[56] and rebuilds it to a short pulse.

example is given for 10 GHz radar with a PRF of 7000 and a pulse width of 1 microsecond.



The wide spectrum is now changed to a collection of small frequency bands. The total spectrum width is still the same. But within this bandwidth there are different peaks at frequencies that are spaced around the centre frequency and with an interval that is the same as the PRF. We call this main lobe and side lobe frequencies.

Measuring small Doppler shifts is now possible, because the bandwidth of each lobe is much smaller than the initial bandwidth. But a new problem arises: ambiguity. If a frequency difference is measured, it is impossible to determine if this frequency change arose from a main lobe or one of the side lobes. In other words, the Doppler filter can only determine a frequency shift within the interval of the PRF.

The following example might help clarify this ambiguity. Imagine looking at the pendulum of a clock every second (PRF = 1 Hz). You see the pendulum on the left side the first time and on the right the second time and so on. You might conclude that the pendulum swings from right to left in 1 second, or at a frequency of 0.5 Hz. However, if the pendulum has a frequency of 1.5 Hz (0.5 + PRF) it would have swung from left to right to left to right within one interval and you would have made exactly the same observation. The conclusion should therefore be that the measured frequency of F1 can be generated by all the frequencies that are equal to: F1 + n x PRF.

Another example: If the PRF is 7000 Hz, only a Doppler frequency of between 0 and 7000 Hz can be measured. Assume 3500 Hz is measured. This 3500 Hz may represent any combination of 3500 Hz + N x 7000, so 3500, 10500, 17500 etc. Assuming that every 35 Hz represents 1 knot closing velocity (which is true for 10 GHz radars), the measured Doppler shift of 3500 can represent a closing velocity of 100 knots, 300 knots, 500 knots etc.

The radar is faced with another problem as well. It has to cancel all ground returns and discriminate them based on their Doppler shift. But by cancelling certain frequencies, all frequencies that are N x PRF higher or lower will also be cancelled. Let's assume that the speed relative to the radar beam (the closing velocity of the ground) is 250 knots and that a 50-knot speed interval is cancelled out (225 to 275 knots). If the PRF were 7000 Hz and the carrier frequency 10 GHz (as in the previous example) the result would be that all speeds with a 200-knot interval are cancelled out as well (25-75, 225-275, 425–475 knots etc.). But not only the ground return is cancelled, all aircraft with closing velocities in these speed ranges will not be detected. The speed region that is cancelled out is called the notch. On some radars this notch is fixed while on other radars the pilot can adjust the notch.

Similar problems exist in the range channel. Taking again that 7000 Hz PRF; this PRF is unambiguous to 12 nm. But if the target beyond that range has enough radar cross section it can show as well. In other words, a measured range of 10 nm can in reality mean 10, 22, 34, 46 nm etc.

The problem goes further; the radar is not able to measure when it is transmitting, so in the first part of the range the radar is blind. When the radar is blind in the first .2 nm it is also blind from 12-12.2 nm, 24-24.2 nm, 36-36.2 nm etc.

Solving the ambiguity problem

Solving these ambiguities is not difficult; the trick is to use more than one PRF. Typically a radar will use a set of 7 or 8 PRFs. The complete set is called a look. In the first example (PRF of 7000) the ambiguity was: 100, 300, 500 knots etc. If a second set was used with a PRF of 10,000 Hz the following set of velocities could have been generated: 215, 500, 785, etc. In this example, 500 knots closing velocity is the velocity that shows a match in both. The other reason for multiple PRFs is that it can compensate for holes in the detection speeds and distances. Every PRF has its own blind speeds, therefore a target will not show in every PRF set. The question is in how many PRF sets does a target have to be visible, e.g. 3 out of 7, 4 out of 7. If the number is too high, targets will be missed, if the number is too low there will be many false alarms. Typical correlations used are 3 out of 7 or 4 out of 8.

Using multiple PRFs will not solve all holes in detection speeds. Aircraft that have the same closing velocity as the ground will be filtered out in each PRF. This feature will be employed by pilots as a defensive manoeuvre against a radar threat. If the course is changed in such a way that the opposing aircraft is at the beam, the closing velocity is exactly the same as the closing velocity of the ground return and the aircraft will not be detected.

Resolution in velocity and range

A digital pulse Doppler radar will divide the incoming signal into range gates; normally the number of range gates is a power of two (2^n). Typical numbers are 128 or 256. Every interval is processed in all frequency filters. The resolution of a coherent pulse Doppler radar is determined by the number of filters in the filter bank. The number of frequency filters in a digital filter bank is also a power of two; typical numbers are 64 or 128. One other physical feature is important. The Fourier Transformations can only be calculated if there are enough samples (individual returns) in the filter bank. The minimum number of samples is the same as the number of filters. So, when 64 Doppler filters are used the radar should transmit 64 times with one PRF. Taking again the example of a 7000 Hz PRF, 64 filters will give a resolution of 7000/64 or 109 Hz, which equals 3 knots (for 10 GHz radar). The same calculation can be made for range filters. If there are 128 range gates and the PRF is 7000 (which equates to an unambiguous range of 12 nm), every range gate is 12/128= 0.09 nm or 560 ft.

Scan time for pulse Doppler radars

The following issue is scan time. Consider the fact that one look consists of 7 PRF and each PRF needs 64 samples. It follows that a target should be illuminated by 7 x 64 pulses = 448 pulses to complete a look. To calculate the required time the PRF set must be known. A rough estimate can based on an average PRF of 7000 Hz. It will take 448/7000 = 0.064 seconds to complete the look. It is imperative that the radar scan is slow enough; otherwise the dwell time on a target will be shorter than the look time, which would hamper detection. The required dwell time follows directly out of the beam width.

Example: Consider a radar with a beam width of 2 degrees employing 3 bar scans over 120° (from 60° left to 60° right). The total scan area is 3 x 120°= 360°. The radar has to dwell at least 0.064 seconds on every two degrees. Therefore the minimum time to complete the scan is not less than 360/2 x 0.064 = 11.52 seconds.

The basic conclusion from these calculations must be that there is no such thing as a "free lunch": everything comes at a price. To measure the frequency takes time, the more accurate it has to be, the more time is required (more filters, more samples) and the same is true for angle tracking where a more accurate angle requires the radar to scan more slowly.

Detection range revisited

The detection range of a coherent pulse Doppler radar is more complicated than that of a normal radar. Initial detection takes place in the Doppler filter. Detection basically implies that the value of one filter is significantly higher than the average value of the surrounding filters. Detection is therefore dependent on clutter level, jamming, side lobe returns and other distracting factors. The possibility for detection will be different with each PRF. It is possible that one aircraft flying at a 400 knots closing speed will be detected at 40 nm, while the same aircraft at the same position would go undetected if the closing speed was 350 knots. The crux of the matter is that, depending on the PRF set chosen, one aircraft can be visible in five out of seven PRFs while the other is barely visible in three. It is the task of the radar engineer to choose a PRF set that has no holes in the operationally important speed region.

Clutter and false targets

Pulse Doppler radars have far less clutter and false targets than non-Doppler radars but they are definitely not immune to clutter and false targets. There are basically two types of false targets: false main beam returns and side lobe returns.

False main beam returns can be caused by fast moving objects on the ground that present legitimate Doppler. Not only fast trains and cars qualify but also wind turbine blades. It is impossible to prevent those targets from entering the system. If the radar has a track-while-scan capability, the speed of those targets is calculated and their relatively low speed can be an indication for the pilot to discard them.

Side lobe returns are different. The problem is that ground targets in the side lobe have a different Doppler. This can easily be illustrated from the next example. If the main beam of the radar is looking straight ahead, relative movement on the ground is the same as aircraft speed, say 400 knots. So, all speeds between 350 and 450 knots are cancelled by the Doppler filters. But if a side lobe exists at a 30-degree angle, the relative closing speed in the side lobe is 400 x cosine (30) = 346 knots, so outside the 'notch'. Normally the power in the side lobes is significantly lower but ground features with large radar cross sections (corrugated iron roofs, steel fences etc.) can easily present a strong return.

Synthetic aperture radar (SAR)

Radar is the only ground sensor that can see through dust, smoke and clouds. The resolution of radar is normally not good enough to distinguish ground targets with enough detail.

For example, a beam width of 1 degree and a pulse length of 1 microsecond (common parameters) the range resolution will be 500 ft and at 20 nm the angular resolution will be approximately 2000 ft.

However, SAR is a radar application that makes high-resolution ground mapping possible. Two features are important for SAR. A short pulse should be used for accurate range resolution and a narrow beam is required for accurate angle resolution. Neither of the two can be achieved directly. A very short pulse causes a very large bandwidth, which is not useful and a good angular resolution requires an extremely large antenna, which does not fit inside an aircraft.

Both problems are solved in the frequency domain. Pulse compression can be achieved by changing the transmitting frequency during the pulse time. This frequency shift is also present in the return signal. In the receiver the different frequencies are delayed in such a way that a short pulse is achieved artificially.

The angular resolution improvement is achieved by integrating pulses over a long period of time. Because of the aircraft's movement, every subsequent pulse is from a different position. This way an almost similar situation is achieved as if the radar antenna were really the size of the distance travelled during the integration time. Prerequisites are that the radar is a coherent pulse Doppler radar and that the phase of the different pulses is adjusted to emulate a large antenna. This phase adjustment is called focusing[57]. It can be proven that the angular resolution of a SAR antenna is half the resolution of a normal antenna of the same size.

Example: If an aircraft flies at 480 knots, approximately 800 ft/sec, all pulses during one second are integrated and the wavelength used is 0.1 ft, the angular resolution is (2 x λ) / diameter = (2 x 0,1) / 800= 0.00025 radial or 0.0143 º. At a distance of 20 nm this equates to a resolution of 30 ft.

It is not possible to endlessly increase the synthetic antenna. The target should be illuminated during the complete integration period. The unexpected result is that a small physical antenna that has a large beam will allow a large integration period. Mathematically it can be proven that the smallest possible resolution is half the width of the physical antenna.

The choice of PRF. The PRF of SAR should be low enough to prevent ambiguous range measurement[58]. The PRF can also be adjusted for aircraft speed. Another consideration is side lobes; in a focused beam all the target returns out of the main beam are in phase and add up. However, at other angles than the main beam, returns might also be in phase. These directions are the SAR side lobes. In a SAR side lobe, angles are dependent on the PRF. The PRF should be high enough so that the 'synthetic side lobes' are well away from the main beam.

---

[57] Instead of considering the antenna as one large antenna, it is also possible to see SAR as a frequency problem. In different parts of the radar beam targets will have a different Doppler. By measuring this small Doppler shift and arranging returns based on their Doppler a fine beam can be emulated. This approach has a lot of similarities with the normal Doppler solving in the air-to-air modes of the radar.

[58] Unambiguous range = c/(2 * PRF)

To enable accurate phase and frequency measurements it is imperative that airspeed and aircraft acceleration are measured with great precision[59]. It is not uncommon for the antenna to be fitted with its own inertia measurement unit.

Another important feature of SAR is that it can normally identify moving targets. To recognize and correctly position these targets is not easy. The problem is that they have a different Doppler than the ground. Because a frequency difference is also a phase difference, the SAR might see moving targets at the wrong angle. With some advanced calculations however these errors can be corrected.

SAR summary

From the above description of SAR it should become clear that to build a SAR picture requires many calculations and takes time. Typically 1 to 10 seconds are used just to integrate over one physical beam width (so 1 -2 º). SAR pictures can improve even further if the same area is mapped several times. Another characteristic is that moving targets can be spotted and sometimes even classified (e.g. tracked vehicle). There is one operational disadvantage with the use of SAR: you have to transmit, so the enemy can notice your presence.

[59] Consider a SAR with λ=3cm, a position difference of 1 cm already equates to a 120º phase angle difference. At a 180º phase angle difference the signal will be subtracted instead of added!

Chapter 5– Sensors

Introduction

The combat value of a modern weapon system is the product of its aerodynamic performance, the quality of its sensors and the effectiveness of its weapons. Over the last decades aerodynamic performance has not changed significantly, but the quality improvement of the aircraft sensors has been staggering. Modern avionics enable an aircraft to operate, navigate and locate targets in almost all weather conditions and with a high accuracy.

The other side of the coin is that the quantity of information available to the pilot can be overwhelming. Therefore much attention is given to cockpit design, data presentation, data fusion and automation. A modern cockpit is equipped with two or three Multi Functional Displays (MFD), which present a myriad of data using quite a number of symbols. In a single display, data from different sensors can be integrated. For example, instead of a classic radar picture, the 'radar' display might now show navigational data, targets that are received by data link, target identification based on secondary radar information and, of course, the 'standard' radar data. But nice displays are only part of the solution. If a pilot has to listen to two different radios, to the audio warnings from the radar warning receivers and he also has to deal with messages from the aircraft's audio warning system, he can easily get overwhelmed in the audio channel. A possible solution is to cut down on audio transmissions in favour of a digital data exchange that is presented on the MFD. So, instead of getting verbal information about a target position the data is sent in a digital format an presented as a symbol on the display. Clicking on the symbol is the only action required to turn aircraft sensors to the target and the appropriate steering information is calculated automatically. The quality of a modern fighter pilot therefore is not only determined by how well he is able to fly the aircraft but also by how skilled he is at operating the systems.

Cockpit F-16-Block-15 after Mid Life Update improvement programme

Sensor classification

Aircraft sensors fall into four categories:

- Navigational sensors. These sensors tell the pilot the position of his aircraft. Navigation and navigational sensors will be covered in more detail in chapter 8.
- Communications. Standard radios like Ultra High Frequency (UHF) radio and Very High Frequency (VHF) radios fall into this category. There are new communication systems as well, for example Link 16, which is the new standard data link for military aircraft.
- Target location, identification and tracking. The most important sensors for target locator systems are radar, secondary radar and targeting pods equipped with IR and EO systems.
- Self-defence systems. These are radar warning receivers, missile launch warning systems and jammers. These systems are covered in detail in chapter 10.

Another division can be made between active and passive sensors. Military operators favour passive sensors because transmissions can give away the position of the aircraft. However, active sensors, like radar, are the only sensors that can operate in all weather conditions. Sensors also differ in resolution and range and will have different limitations based on the part of the EM spectrum and technique that are used. We will discuss these in more detail but will first present an overview of the most common sensors.

Sensors overview

The following table depicts EM energy types and their military application.

Name	Frequency	Application
High Frequency (HF) radio	7 MHz- 30 MHz	Used for long range communication, increasingly replaced by satellite communication
VHF Omni Range (VOR)	108 MHz- 117.95 MHz	Civil navigational beacon, gives a direction to a ground beacon based on a phase difference between two signals
Instrument Landing System (ILS) – Localiser	108 MHz- 111.975 MHz	Landing system, gives a position relative to the runway centre line, based on the modulation difference between signals
Very High Frequency (VHF) radio	118 MHz-136 MHz	Civil air traffic communication band
Ultra High Frequency (UHF) radio	225 MHz-400 MHz	Military air communication band
Instrument Landing System (ILS)- Glide slope	328.6 MHz- 335.4 MHz	Landing system, gives a position relative to the designed glide slope, frequencies are paired with localizer
Tactical Air Navigation (TACAN)	960 MHz- 1215 MHZ	Military navigation beacon, gives direction and distance to a ground beacon. The distance measurement part of a TACAN is called DME by civilian users.
Link 16	960 MHz- 1215 MHz	Military tactical data link
Radar Transponder	1030 MHz & 1090 MHz	Radar system for location and identification of aircraft. The aircraft transponder will respond with a coded message when a valid interrogation is received. Different modes are used for ATC and military identification.
Global Positioning System (GPS)	1567 MHz- 1587 MHz	Global Satellite Navigation System
Radar	3 GHz 30 GHz	Multiple applications: Weather, air surveillance, ground surveillance
Far Infra Red	Wave length 8- 12 micron (30,000 GHz)	Thermal imaging, used in targeting pods. Objects with temperatures of 300ºKelvin have their maximum energy in this band.
Mid Infra Red	4-6 Micron	Thermal imaging used for (older) IR missiles. Hot exhausts have their maximum energy in this band.
Visual	384,000 GHz- 690,000 GHz	Human eye Optical sensors
Ultra Violet	>700,000 GHz	Missile launch detector

Communication

Airborne communication systems traditionally employ the VHF (civil) and UHF (military) frequency band. The advantage of these high frequency bands is that there are many channels available. Many channels are also required to enable different control facilities to operate without interference. A restriction is that these frequency bands are limited to line-of-sight range. This may not hinder a high flying aircraft talking to an air traffic controller two hundred miles away, but a low flying aircraft will not be able to communicate with another aircraft or ground controller on the other side of a hill. Also, over the ocean there is no complete coverage with VHF ground control stations. There are two ways around this problem. The first solution is HF radio. HF radio has a much larger range because HF energy is reflected by the ionosphere layers in the higher atmosphere. By multiple reflections between these layers and the ground the signal can 'hop' over a considerable distance. Limitations are: the available bandwidth- which is smaller- and the possible fluctuation of the propagation quality due to ionosphere variations. A more modern way to communicate worldwide is satellite communication, which employs extremely high frequencies that enable the use of many channels. Another advantage of satellite communication is that it is much harder to jam.

A military requirement for all communications is that they are secure and hard to jam. Security is normally achieved by scrambling the signal. To prevent jamming, the signal can be spread over a wide frequency band. An example of this technique is the Have Quick UHF radio, which hops over a number of frequencies. A prerequisite for a hopping radio is that all radios use the exact same time to synchronise frequency changes.

The quantity of data transmitted in aviation has traditionally been low. Voice communication and simple data transmission generally do not require a large bandwidth. But when there is a requirement to send pictures (reconnaissance) or even streaming video (surveillance) the required bandwidth multiplies. There is a fast increase in demand for bandwidth both for civil (for example mobile phones and satellite TV) and for military applications. For example, the number of reconnaissance UAV that are employed in a certain theatre is limited by the available bandwidth, so while on the one hand the available bandwidth is increased by new systems such as satellite communications, the required bandwidth has increased even more. A possible way around this problem is the use of data compression. Present computers enable real-time compression with virtually no data loss.

Identification

The ultimate military requirement is to identify own and hostile ground and air units. This requirement is far from being met. Presently the only requirement that is being met satisfactorily is the identification of own airborne assets. This is primarily done with a transponder system, a system based on interrogation. The interrogation radar transmits a coded signal. When this is received by the aircraft transponder it will transmit a coded reply, which is received by the interrogator radar. Similar to normal radar, the distance is determined based on the time delay between interrogation and reply. The direction of the responding aircraft is measured by the antenna of the interrogator radar and the code can reveal the aircraft's identity or altitude.

The transponder is used in civil and military applications. The civil modes are mode 2, Mode 3, Mode C and mode S while the military modes are mode 1 and mode 4. Mode 2 is a fixed code assigned to an aircraft. Mode 3 is the code used by Air Traffic Control (ATC) and is set by the pilot. Mode C code is dependent on the aircraft's altitude. By reading the mode C code, ATC has a much better presentation of the aircraft's altitude than with the surveillance radar. Mode S is the newest civil code and the coded reply has even more capabilities, for instance the incorporation of exact position and heading.

The military modes are mode 1 and 4. Mode 1 is the oldest mode that is not encrypted and is close to obsolete. Mode 4 is an encrypted mode. A transponder will only answer to a correct interrogation. This is the prime identification method for own aircraft. Ground radars are almost always equipped with a mode 4 interrogator while the more advanced interceptor aircraft have an interrogator as well.

The primary limitation is that only friendly squawking aircraft can be positively identified. It is not possible to differentiate between an enemy aircraft and an aircraft with a broken transponder. Through analysis of the radar return it is possible to determine the radar cross section and when the aircraft inlets are exposed it may be possible to measure the Doppler shift caused by a turning engine. This Non Cooperative Target Recognition (NCTR) is a feature that is installed in some radars. It requires that the target is tracked for several seconds to enable an accurate Doppler measurement. This system is not foolproof but in combination with other systems it can help establish the identity of an enemy aircraft.

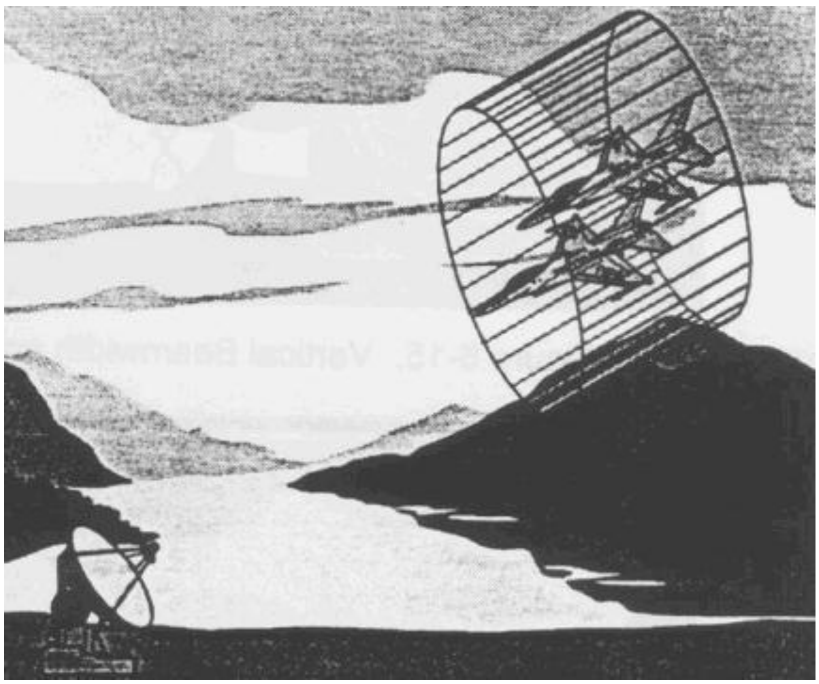
Target location and tracking

There are only three types of sensors that have been proven useful in target detection and tracking for aviation. These are radar, (electro) optical sensors and infrared sensors. The oldest sensor is without a doubt the optical sensor (the human eye), which still combines an excellent resolution (but only in a very small portion of the field of view) with the ability to measure frequency (we see colour) and the fact that it is passive. The limitations are that it has a small range, does not work well at night or in bad visibility and is not capable of measuring distance. The range problem and the range measurement problem can be overcome by using magnification and adding a laser range finder. This is exactly what a targeting pod does. The electro-optical sensor in a targeting pod magnifies the image and the computer adds some automatic tracking capabilities. Because of the collocation of the laser and the sensor, the distance can be calculated by measuring the time interval between the emitted laser pulse and the measured reflection. The newest sensor is the Imaging Infrared sensor, which senses variations in radiated IR energy. As explained in chapter 4, everything that has a temperature above absolute zero emits energy. At moderate temperatures most of the emitted energy is in the IR band, so consequently an IR sensor is the best sensor to use under conditions of darkness. Normally the resolution is slightly lower than an optical sensor. Because radar, with its unique capabilities, is the most important sensor it will be covered first.

Radar general

Introduction

Radar is the most important aircraft sensor. It is the only sensor that is all-weather capable and works over a large range. Both features are essential when employing aircraft. Radar also has some restrictions. Radar is active, will give away the aircraft's position, is susceptible to electronic jamming and has a limited resolution compared to EO and UV sensors. As explained in the previous chapter, resolution has two components, the angular resolution, which is typically around one degree and the range resolution that may be as large as several hundred feet. The combination of range resolution and angular resolution determines the resolution cell of the radar.



Two aircraft within the same resolution cell will be seen as a single target

Because angular resolution is - as the name says - an angle, the resolution cell of the radar increases with distance. An angular resolution of one degree is only 1000 ft at ten nm but 10,000 ft at one hundred nm. The practical significance is that radar can not discriminate between different targets flying within one resolution cell.

Radar modes

There are two basic air–to-air radar modes, search mode and track mode. In the search mode a radar does not dwell on a target longer than absolutely necessary to detect it, while in the track mode the radar dwells much longer on a target (sometimes continuously) to accurately determine its position. This accurate target position and speed are required to deploy a weapon with a high degree of success. It is important to note here that the angular tracking accuracy of radar is better than its angular resolution. Typically tracking accuracy is around one mil[60] radian.

Search modes are used to scan large areas. For an AWACS type of radar the search scan is a 360° scan, which can take approximately 12 seconds. The antenna beam of an AWACS is small in azimuth but wide in elevation. Consequently, the

[60] In the military two types of 'mils' are used. The first is the mili-radian, 1/1000 of a radial; there are 17.45 of these mils in one degree. The other is the old artillery mil, which is 1/6400 of a circle, with this mil there are 17.78 in one degree.

altitude accuracy is small, but adequate for scanning a large altitude band. For a fighter aircraft the situation is different. The radar beam is small in azimuth and elevation (enabling accurate positioning in altitude and azimuth). The fighter radar scans at different elevation bars, which can be a 4 bar from 6° up to 6° down and from 60° left to 60° right. The pilot can adjust the search volume to the tactical situation and scan volumes will be coordinated between flight members. The search mode in a fighter aircraft will also take up several seconds. This slow update cycle is normally not adequate for finding an accurate firing solution for the fire control computer, especially if the target is manoeuvring.

Search radar (Patriot)

The track mode is used for the accurate measuring of target position and velocity. Therefore an accurate target azimuth, range and velocity must be measured. The most simple and also most accurate mode is single target track. The radar is constantly aiming at the same target. To aim the radar antenna accurately in azimuth, it measures the target signal strength difference between different parts of the antenna. Older type radars use circular scan patterns for their antenna around the target to accurately locate it. One of the undesired consequences of the single target track mode is that all other targets of interest are not seen.

To alleviate this problem, radar engineers have developed combinations of track and search modes. The following combinations are possible:

Single target track interlaced with small search scans;

Two or more tracked targets with and without some search

volume; Track-while-scan.

The last mode is in essence just a search mode, but based on the search data the radar computer builds a track file that can be used by the fire control computer for a firing solution. This mode is normally only usable to track non-manoeuvring targets.

As described in the previous chapter, the performance of radar is dependent on many parameters and is therefore not a fixed value. In the next paragraph typical performance figures are given for the most common radars.

- Long Range Search radars. These radars are employed by air defence sites and air traffic control. Ranges of these radars are normally in excess of 200 nm and the scan rate is around 5 scans per minute or less. The horizontal angular resolution is around one degree and the vertical angular resolution is typically 3 degrees or more. Range resolution is around 250-1000 ft. These radars are perfectly suitable for early warning but lack accuracy especially in determining the altitude of a target. Furthermore, all ground-based early warning radars are limited by line-of-sight restrictions that prevent them from discovering low flying aircraft at a large range. This was the most important reason for developing airborne early warning aircraft. The problem with the altitude accuracy can sometimes be solved by adding special height finding radars to radar early warning sites. Long range search radars normally have no capability for tracking a target, but based on the scan information they will still build a track file. This is called track-while-scan; in the 'track' mode it has a very low accuracy due to its slow scan rate. The most important limitations of this type of radar are its slow update rate, its low accuracy and the fact that it will not detect close-in targets that fly above (or below) the scan volume[61]. A further limitation is that these radars (except AWACS) are not mobile, which makes them vulnerable to pre-planned air attack.

- Medium range search radars. These radars are typically used by ground-based surface-to-air missile systems. The angular and range accuracies are marginally better than long range search radars but the scan rate is twice as high: ten or more scans per minute. The faster update rate of these radars will enhance their track accuracy. Ranges are normally from 60 to 120 nm. These radars do not always have a separate track function. However, modern SAM systems such as the Patriot have radars with track and scan capability. The phased array antenna of the Patriot radar even enables rapid changes between the track and scan modes. One important advantage of these radars over long range radars is that they are generally mobile, so less vulnerable to a pre-planned attack. Furthermore, their tracking accuracy is much higher. The downside however is limited range.
- Tracking radars. Ground-based air defence systems have often dedicated radars to perform the tracking of aircraft. These radars have antennas with a horizontal and vertical resolution of one degree. The tracking accuracy is normally around one mil radian; the range resolution is around 100 ft. The range is normally based on the maximum range of the missile. Typical ranges are 40-80nm.

Fansong (SA-2) tracking radar

- Fighter airborne radars. The angular horizontal and vertical resolution for these radars is approximately one degree and the range resolution is 100-200 ft. The typical detection range against fighters is 25-60nm. Fighter radars are always multi-mode, which can include: search, track,

[61] If an attacking aircraft wants to prevent early detection by AWACS it can come close to the AWACS platform but well below or above the radar beam, so that it will not be detected. For this reason these valuable assets are normally accompanied by fighters.

combinations of search and track and ground mapping. The most important limitations are the scan volume, which is always limited to the forward quarter, and the range. The size of the radar and the available power are the main factors limiting the detection range. An advantage of these radars is the accurate tracking capability. However, the more targets are tracked the smaller the remaining track volume will be.

- Synthetic Aperture Radars (SAR). Synthetic Aperture radars are airborne radars that integrate the ground returns over a time interval (=distance). In this way they emulate a large antenna with a very good angular resolution. With the SAR principle a high accuracy ground map can be made. Typically the resolution is one foot or better. The low power systems are normally capable of mapping ranges of up to 20-50nm while the high power systems can map up to a 200 nm range. The path that is mapped during one scan is relatively small. Standard limitations apply here as well. You can either see a large area at a low resolution or a small area at a large resolution.

Targeting pods

Most combat fixed-wing aircraft and helicopters today have at least provisions for targeting pods. The targeting pod enables pilots to accurately aim passively at a ground or air target and guide weapons to those targets. Targeting pods combine an IR and EO sensor with a laser designator for weapons aiming and ranging. The IR sensor normally uses the 8-12μ wave length but the newer systems also use the mid-IR band (3-5μ) . The pods can be steered in different ways:

The pod can be stabilized to the body axis of the aircraft (Bore);

The pod can be steered towards a certain geographical point based on its coordinates and the aircraft position (geo-stabilized);

The pod can be slaved to another sensor, for instance the radar;

The pod can autonomously track a position on the ground based on contrast (Area track);

The pod can be steered to remain pointed in a certain direction. It will be adjusted to compensate for aircraft turn and pitch rates (rate track). This mode is normally a back-up mode when the pod can not maintain an area track;

The pod can sometimes track a point target because of its signature (point track).

The fact that the pod will normally stabilize on the target environment and is not affected by aircraft motion enables the pilot to acquire the Desired Mean Point of Impact (DMPI). It can be confusing for the pilot when the orientation of the picture as seen through the targeting pod is completely different from the direct view from the cockpit. For this reason the pod display can indicate with a symbol where the pod is looking relative to the aircraft's nose. Another reason for showing this position is that the pilot must ensure that the pod's field of view is not obstructed by parts of the aircraft. This is especially important when the aircraft is turning.

Each sensor normally has a wide and a narrow field of view. The wide field of view is necessary to find the target environment and the narrow field of view is used to accurately point the pod to the DMPI.

The pod designer should take special care that target track is not lost during changes in field of view or sensor changes. Not every manufacturer has been able to do this effectively.

Sensors and lasers are all collocated in the pod and theoretically looking at the same spot. To correct for minor alignment errors some targeting pods display the spot the laser is aiming at.

The tracking of the pod is based on contrast. There are several occasions when contrast tracking for a targeting pod is difficult:

When clouds suddenly obscure the target;

When a mobile target is tracked and it passes under a bridge;

When a target does not have sufficient contrast; a good example is an empty road. The pod will clearly identify the edges of the road but it is much harder to find contrast along the road.

IR sensors have more trouble-causing characteristics. The sensor measures temperature differences, so during daytime it can easily spot a hot steel bridge over a cold river. However, when night falls the bridge will cool more rapidly than the water and at a certain moment their temperatures are the same. This is called the crossover temperature. When the target and the environment are at the crossover temperature it will be invisible to an IR sensor. The duration of the crossover temperature is dependent on meteorological conditions, target and environmental characteristics and 30 minutes is not uncommon.

The IR visibility of a target can be predicted. There are software programmes that predict IR visibility based on target and meteorological conditions. With these tools the pilot can optimise his attack direction and attack time.

The laser of the targeting pod has two functions. It will emit a special code to guide the laser bomb. Also, the laser being a pulsed sensor, it can be used to accurately measure the slant range to the target.

Laser energy is concentrated in extremely short pulses. It is so intense that it can easily destroy the eyesight. A compounding problem is that the eye will not notice the laser because of its wavelength (around 1μ) but the laser will destroy the retina. During training the laser is normally tuned to a different frequency (1.54μ), which is eye-safe.

Lantirn targeting pod

Sensor overview						
	Resolution		Speed Measure-ment	All weather	Range	Active/ passive
	Range	Angle				
Electro-optical with magnification (targeting pod)	No range measurem ent	Better than 0.003°.	No	Day clear weather only	Same as visibility	Passive
Infra red with magnification (targeting pod)	No range measurem ent	Better than 0.006°	No	Day and night clear weather	Same as visibility	Passive
Search radar	200-500 ft	± 1-2°	Yes	All weather	>200 nm	Active
Airborne radar	100-200 ft	± 1-2°	Yes	All weather	20-60 nm	Active
SAR	1 ft	± 1-2ft	Yes	All weather	20-200nm	Active

Self defence systems

Ultra Violet (UV) sensors

UV sensors are used today to detect missile firing. The firing of a missile generates much heat that it is clearly visible in the UV band. The sensors will detect it and give the pilot the direction of the threat. Missiles and rockets are not the only sources of UV energy. The sun is a powerful UV source and direct sunlight as well as sun reflections must be cancelled out. Electric welding and open fires will also generate UV energy. The designers of UV sensors will try to filter out false alarms based on their signal characteristics.

RWR and jammers

Radar warning receivers detect radar signals and identify them based on frequency, pulse rate and scan type. RWR typically have several antennas to cover the whole area around the aircraft. Based on the relative signal strength an estimate can be given of the distance of the threat and based on the signal strength difference between the antennas, the threat azimuth can be calculated[62]. To prevent too many false alarms most RWRs only trigger the pilot when the aircraft is tracked and not when only a search radar is received.

To prevent accurate tracking by enemy radar, electronic jammers can be used. A jammer emits a signal at the same frequency as the enemy radar. This will be discussed in chapter 10.

Sensor integration

Because a single sensor is always limited, it is advantageous to combine sensors. There are several forms of data integration. Images from different sensors can be fused on a platform to get a higher definition target image. Another possibility is to combine similar sensors from different platforms. This for instance is done with the radar data to make a recognized air picture. The fused data from several radars,

[62] More advanced system measure phase difference to calculate the azimuth.

disseminated over a network, have a higher resolution and redundancy when compared to a single radar covering the same area.

Sensor integration with different platforms requires high speed and secure data links. The standard data link for airborne application is LINK-16.

Chapter 6 - Weapons

Introduction

The purpose of a weapon is to destroy or immobilize a target without causing any harm to the delivering aircraft and minimal collateral damage. To fulfil these requirements the weapon should ideally have just enough destructive force, be very accurate and posses enough standoff to prevent any harm to the delivering aircraft. However, it is difficult to meet all these requirements. Just enough destructive force for a target would lead to a large variety in weapons, each optimised for their specific targets. But an aircraft can only carry a small number of weapons and flexibility is increased if a multi-purpose weapon can be used. Large standoff range seems attractive, however for a lot of weapons accuracy will decrease with increased standoff. So weapons are full of compromises. In this chapter we will focus more on the unique characteristics of air-delivered weapons and less on the general characteristics of each weapon. First the general characteristics of air-deployed weapons are discussed followed by the specific types of weapons, bombs, missiles and guns.

General requirements

Platform and weapon

All three types of attack aircraft (bombers, fighter-bombers and attack helicopters) have different weapon employment capabilities. Bombers are the only aircraft capable of carrying heavy weapons such as the Air Launched Cruise Missile (ALCM), fighter-bombers can normally carry weapons up to 2000 – 3000 lbs class and most attack helicopters carry weapons that weigh less than 500 lbs. Attack helicopters are not capable of carrying bombs, they prefer light missiles and/or rockets and often carry a light-weight gun.

Airborne weapon special requirements

A weapon carried on an aircraft must meet many additional requirements when compared to weapons employed from the ground. It should be able to withstand the g-forces, temperature and pressure changes, cope with high vibration levels, separate safely from the aircraft and not detonate unless at a safe distance from the aircraft. For missiles there is the additional requirement that the smoke of the missile engine must not be too close to the aircraft engine intake otherwise the engine can stall. Preferably the weapon should be as light as possible, firstly because payload is always limited, secondly because heavy bombs on the wing are prone to cause wing flutter[63] and thirdly because they limit the aerodynamic performance of the aircraft. During weapon certification tests, the carriage and deployment limitations of each configuration are determined.

Technically, safe separation of a bombs is achieved by rods that push the bomb away at the release moment, giving it an extra downward velocity. This prevents collision with the tail of the aircraft[64]. To prevent the bomb from exploding when it

[63] Flutter is an oscillating mode of the wing. Wing stiffness, aerodynamic forces and weight distribution determine frequency and amplitude. Wing flutter is normally a divergent mode which can cause the wing to break. Some aircraft can experience flutter modes that are not diverging. These are called limited cycle oscillations or LCO.

[64] This is of particular importance under low g-conditions. Gravity pull is then limited and aerodynamic forces may blow the store against the tail plane.

is still in the vicinity of the aircraft, the bomb fuse is not armed before the preset time period (normally 3 - 9 seconds) has elapsed.

Weapon effectiveness

Most of the destructive power of a weapons is caused by the energy it possesses. A weapon on an aircraft has both kinetic and potential energy by virtue of its speed and altitude. When a weapon is dropped the potential energy is converted into kinetic energy.[65] A rocket or missile engine may increase kinetic energy even further. A second source of potential energy is the explosive compound in the weapon. This explosive material has heat and blast effects. The blast can accelerate the warhead fragments thereby increasing their lethality and working area. Typically, a bomb that explodes into large fragments has a larger lethal area than a bomb exploding into smaller fragments. So thanks to its speed, altitude and its carriage capability an aircraft can put considerably more kinetic energy on a target than could be achieved with conventional artillery.

With some special weapons the kinetic effect is of less importance. The purpose of such weapons may be to set fire (napalm), to short-circuit high tension electricity systems, or for area denial (mine laying), or to cause non-conventional effects such as produced by nuclear, chemical and biological weapons. International laws forbid the use of chemical and biological weapons because of their indiscriminate nature. The present focus on anti-insurgency warfare has revived the interest in non-lethal weapons. Presently the number of non-lethal weapons that can be deployed from an aircraft is very limited.

Fusing and penetration

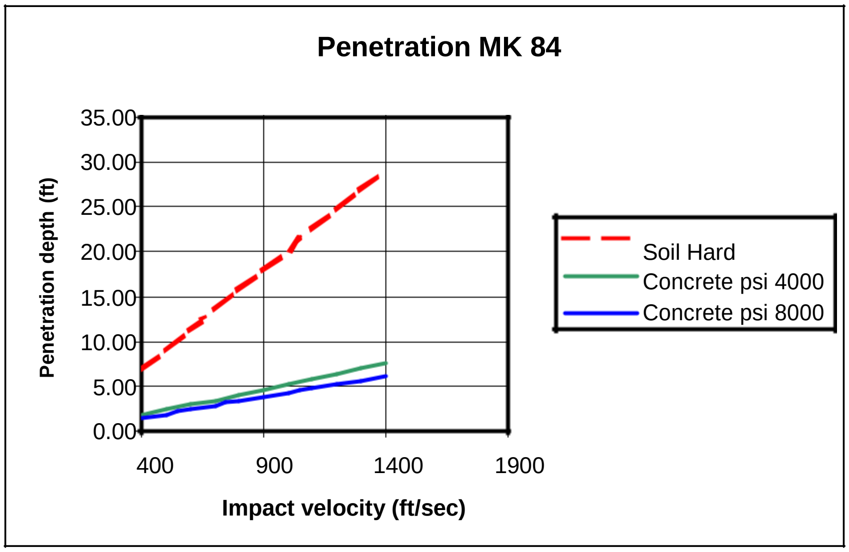
The type of fuse and fuse setting determines where the weapon will detonate in relation to the target. Most weapons explode on impact but that is not always the optimum moment. For some warheads it is advantageous that they explode before hitting or passing the target. Under those conditions proximity fusing (used for air-to-air missiles) or radar altitude fusing (used on cluster weapons and sometimes for general purpose bombs) will be used. The other extreme is when a large penetration is required and the bomb must explode a fraction of a second after contact. To achieve this, the fuse is given a detonation delay. Under these extreme circumstances the fuse must be strong enough to survive the impact and it will be positioned in the tail of the weapon. Normally a bomb is equipped with two fuses. When the delay settings of the two fuses are different, the weapon can be deployed in different roles.

If the tail fuse has a delay and the nose fuse is instantaneous the pilot has the following options: either arming only the nose fuse, which will give an explosion on contact; or arming only the tail fuse, which will give the bomb penetration capability. Arming both fuses will cause an explosion on contact and the bomb will still function if the nose fuse fails.

Penetration depth is dependent on many parameters including: impact velocity and angle, weight and shape of the weapon and the type of material the weapon must penetrate. In the following graph penetration is given for a 2000lbs MK-84 General Purpose bomb for different impact velocities and materials[66]. The impact angle is not only important for penetration depth, but will determine if the weapon will penetrate at all. At a small impact angle the likelihood of a ricochet is high, at moderate angles the weapon might broach. (When a weapon broaches it will initially enter the ground but curve back up.) Only at steep impact angles is

[65] Some energy is lost due to the friction in the atmosphere.
[66] Calculations are based on Sandia report SAND 97-2426.

penetration ensured. To increase the impact angle the pilot can increase the delivery angle by making a steeper dive to the target, or delivering at a higher altitude.

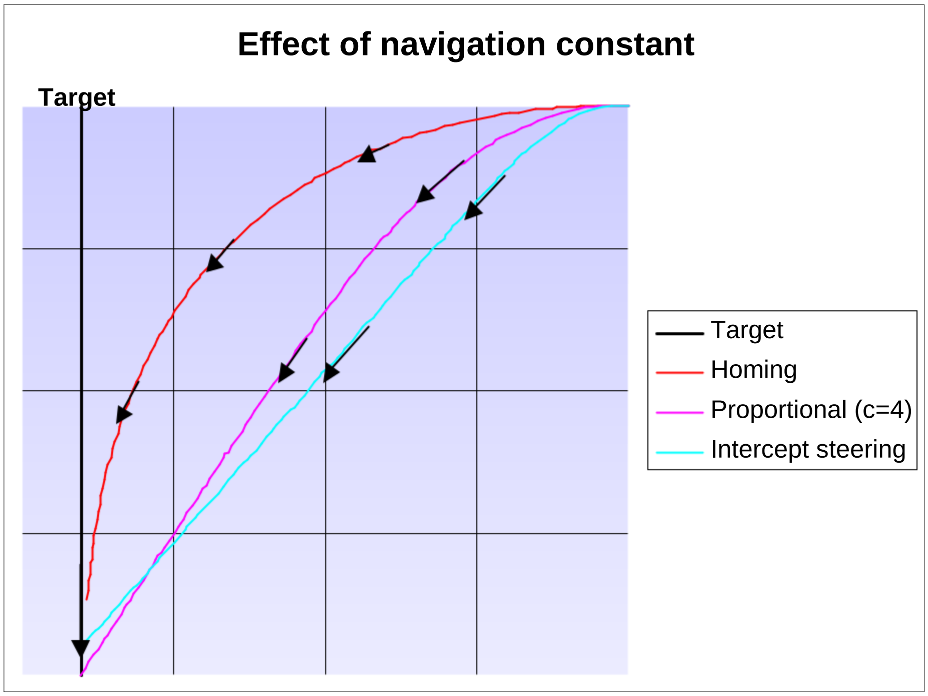


Guidance

There are several parameters for the classification of a weapon guidance system. In this paragraph we will consider the navigation principle used, the flight path control and the autonomy of the weapon.

Navigation. The path via which the weapon is steered to the target differs per system. There are three basic techniques, the oldest being the so-called homing. With homing the nose of the weapon is always pointed in the direction of the threat. This type of guidance was common for weapons that homed in on the reflected radar energy of the target. It is easy to design but the disadvantage is that the weapon has to be able to manoeuvre very aggressively when it comes close to the target. The second technique is intercept course steering, which can be based on a ground radar tracking the target and the weapon, in which case it is called command guidance. However, an active (transmitting) missile is able to calculate the intercept course by itself. In this technique the weapon is directly steered to the pre-calculated impact point. The manoeuvring requirements are now the lowest but the computing and tracking accuracy required will be very high[67]. The third guidance technique is proportional navigation. The weapon measures the change in angle of the target and changes its direction based on this angular change. If you divide the direction change by the tracking angle change you get the navigation constant. Experiments have shown that a navigation constant of around four gives an optimum trajectory. Proportional navigation is for example used in the Sidewinder air-to-air missile.[68]

67 Computing an impact point is only easy when the target and the missile are at constant speed, but in reality the missile speed is constantly changing and the target may manoeuvre as well.
68 Homing and intercept steering can be considered special forms of proportional navigation. The navigation constant for homing is 1 and for intercept steering ∞.



Flight path control The flight path is normally adjusted by aerodynamic control surfaces on the weapon, much the same as in aircraft. A second option for a missile is to equip it with a moveable exhaust nozzle. The moveable nozzle is extremely effective at low speeds where the aerodynamic controls are less effective. The combination of a moveable nozzle and aerodynamic control surfaces will give the missile a high manoeuvring capability.

Aerodynamic controls are operated in different ways. The less advanced control surfaces only have the option of switching from one extreme position to the other. This system is often called a 'bang-bang' control system. The advantage of this system is its low price but it has a negative impact on the flight path. The flight path is less accurate and this system generates a lot of drag, which will negatively impact on weapon range. A more sophisticated way of adjusting the control surfaces is called proportional control, where the control surface movement is proportional to the required change. This results in better accuracy and improved range but is also more expensive.

Autonomy Several levels of autonomy are employed in weapons. Each different level of autonomy has its advantages and disadvantages. The most common levels and techniques are presented in the following table.

AUTONOMY		
Technique	Pro	Con
Command Guided. Weapon is steered by the weapon platform through a data link. Used in several SAM-systems	Weapon is cheap because it does not need a sensor. Weapon can be steered on intercept course, this is less demanding for manoeuvring	Weapon platform has to track target and missile during the whole engagement. Requires a very accurate sensor to guide the weapon
Semi-active. Weapon platform illuminates the target and the weapon guides itself based on the reflection. Used in laser-guided bombs and air-to-air missiles (e.g. Sparrow)	Still relatively cheap. Relatively easy to achieve accurate guidance	Weapon platform has to track the target during the whole engagement. Weapon needs a passive sensor. Manoeuvring requirements are higher than command guided
Autonomously guided. Weapon sensor tracks the target and steers autonomously to the target. Tracking can be done by a passive or active sensor. Passive example: Sidewinder, Maverick. Active example: Exocet	Fire and forget weapon. The weapon platform only has to track the target until weapon launch. Easy to achieve accurate guidance. Manoeuvring demands can be less stringent	Weapon is expensive. Range can be limited to weapon sensor limits
Mixed guidance. Many weapons are command-guided during the initial phase and autonomous during the last phase. Examples: AMRAAM	Partial fire and forget. Easy to achieve high accuracy. Manoeuvring demands are less stringent	Very expensive

Missiles

General

Missiles are several orders of magnitude more expensive than bombs, but there are two important reasons for using missiles. They generate stand-off and are the most effective weapons to be used against a (fast) moving target.[69] If the target

[69] If the target speed is low (ships and slow moving cars), a guided bomb may be able to hit as well.

manoeuvres as well, the missile must have superior manoeuvring capability in order to be able to hit the target.

Warhead

The allowed weight for a warhead in a missile is limited, especially if it needs a high manoeuvring capability. Other considerations are that manoeuvring can cause the missile to pass the target at short distances and also that the closing speed can be extremely high (particularly for a ballistic missile intercept or a head-on air-to-air engagement). A solution often used for air-to-air missiles is a warhead containing fragments that are pushed outwards by an explosive compound. This increases the working area, which is important if the missile does not hit the target. Thanks to the high closing speed and the fact that aerial targets are not armoured, the fragments usually have enough kinetic energy to destroy the target. It is of course necessary for the warhead not only to have a contact fuse but also a proximity fuse.

If the missile's target is large and armoured, for instance a ship, the warhead needs to be much larger and normally a contact fuse is used[70]. Because of the required warhead size the missile will also be much heavier.

Sensors

Missiles basically employ one of each type of sensors as described in chapter 5 (Radar, laser, EO and IR). The missile sensor can be cheaper than a normal sensor because it will only have to operate for a limited amount of time. On the other hand, it should be capable of withstanding large acceleration forces. A limitation for some sensors is the available power and the size of the aperture. This is particularly relevant for radar sensors, which consume a lot of energy and need a large aperture to achieve accurate angular resolution.[71]

Range

Missiles differ considerably in range. The range record for a ground-launched missile is held by the ICBM, which has global reach. The maximum range for an air-launched missile is around 2000 nm, which can be achieved by air-launched cruise missiles. Air-to-air missiles have a much shorter range, typically below 40 nm.

Bombs

General

Bombs are cheap, easy to assemble and can carry a lot of destructive power. To deliver a bomb accurately from an aircraft has proven difficult[72]. The effectiveness of bombs was greatly improved when they became guided. The guidance potential is normally limited due to the weight of the bomb and the relatively small control surfaces, making the bomb inefficient against (fast) moving targets. With heavy and accurate bombs it has become possible to destroy bunkers, aircraft shelters, bridges and other heavy structures.

When there are many targets, not all of them equally well protected, an area weapon is needed. The best known area weapon is the cluster bomb. Each bomb is a canister containing a number of small sub-munitions. At a certain altitude or after a certain amount of time the canister will open and scatter the sub-munitions over a wide area. The sub-munitions can be expelled by spring force or by rotational movement of the canister. Cluster weapons were designed particularly to fight

---

70 Examples of these types of missiles are the Maverick (USA), the Exocet (France) and the Penguin (Norway).

71 See chapter 3, Aperture

72 This will be discussed in more detail in chapter 9.

armour. An undesired side effect of these sub-munitions is that they do not explode when they hit a soft surface. Unexploded sub-munitions must be destroyed by qualified personnel and are brightly coloured to ease detection. However, if an unaware person tries to move them it can be lethal. This problem could also be prevented if sub-munitions were equipped with self-destruct devices. Without these devices the unexploded sub-munitions resemble anti-personnel mines, which are politically unacceptable. Added to this, the present shift to more precision has made the cluster weapon a less attractive weapon, except for the new generation that contains sub- munitions with sensors that guide each one to a separate target.

Warhead

Whereas a missile has both propulsion and a warhead, the bomb itself *is* more or less the warhead. The guidance system of a bomb accounts for less than 10% of the weight. The destructive capability is therefore much larger. The most common bombs are the General Purpose (GP) bombs. The MK-82 (500 lbs), the MK-83 (1000 lbs) and the MK-84 (2000 lbs) GP bombs are used by many air forces around the world. The prime destructive effect of the bomb is its blast and kinetic energy. The blast effect can be effective up to ca 2000 feet[73]. So, even with an accurately guided weapon the collateral damage could be substantial if the target is in the proximity of the civilian population. Present developments are aimed at making bombs smaller, with more precise effects. Cluster bombs can be equipped with different sub-munitions. Mines are an example of a weapon to be spread out over an area and are therefore among the most common sub-munitions. Other sub-munitions are designed to penetrate armour. They employ the hollow charge principle and burn a hole through armour[74].

Incendiary bombs have been used frequently in the past, the best-known being the phosphor bomb used during the Second World War and the napalm bomb used during the Vietnam War. While incendiary bombs can be quite effective against some military targets, their use is associated with bombing campaigns against the population. Therefore the use of these bombs is presently considered politically incorrect.

Sensor

The most common sensor on a bomb is a laser seeker, which was first employed during the Vietnam War. The laser sensor measures the relative angle of the reflected laser from the target. It recognizes the pulse repetition frequency and compares it with a preset value, which prevents the bomb from heading to the wrong target and protects it against laser jammers. The laser seeker is built as a strap-on kit that can be put on a GP weapon. Because the laser light (similar to normal light) will not pass through clouds, the laser bomb is only usable in clear weather conditions.

Today GPS-guided bombs have become an essential part of the inventory. A common type is the JDAM strap-on kit for GP bombs. The heart of the system is not the GPS receiver but the inertial navigation unit.[75] The GPS might not lock on to the satellites before launch, but the INU works independently from outside sources. Even when the bomb is not able to track satellites it will still have an increased accuracy when compared to non-guided bombs. A good GPS position enhances the accuracy even further. A GPS weapon can be used in all weather but

---

[73] This of course depends on bomb weight and type of penetration. An air bust will give the largest fragmentation envelope and a deep penetrating bomb the smallest.

[74] The hollow charge is a nozzle through which hot metal particles accelerate to hypersonic speed. This stream can disintegrate almost a foot of armour.

[75] INU will be explained in the next chapter.

only against stationary targets. The accuracy is slightly less than that of laser-guided bombs.

Some bombs are currently equipped with dual sensors. They have a GPS and a laser. This gives the pilot flexibility to use the more precise laser mode when weather permits and use the GPS sensor under adverse weather conditions.

Range

The bomb normally follows a ballistic trajectory. If the delivery is low level the bomb trajectory will be around one nautical mile. If the bomb is delivered medium level or tossed, the range can increase to around 5 nm. However, a bomb deployed from a medium altitude has a considerable amount of potential and kinetic energy, which can be converted to range provided that the bomb has some wing area. Equipped with wings (which will deploy after delivery) the range can increase well beyond 10 nm. It should however be remembered that the increased range of a glide bomb will negatively affect its penetrating capability, because it is converting potential energy into range and not into impact velocity.

Rockets

Rockets are normally carried in pods and can be fired single or ripple. Rockets, like missiles, have a high kinetic energy and a low weight but lack the accuracy of the missile. The lack of accuracy makes it an area weapon, much like a cluster bomb. The forward velocity gives it a good separation from the aircraft, which is advantageous for slow flying aircraft like attack helicopters. The present trend to move from area weapons to precision weapons makes the rocket a less suitable weapon. The range of rockets is normally less than 10,000 feet and the warhead is small (a few pounds) and only effective against soft-skin targets.

Guns

General

The gun is an old but still effective weapon. The main disadvantages are that it is relatively heavy, it only works over a short range and the effect of the bullets is limited. But it has a few advantages as well. High muscle velocity and short time of flight enable accurate aiming even against a moving target.[76] The aircraft normally carries several hundreds of bullets and this gives a multiple target engagement capability. The small calibre decreases collateral damage. The gun is probably the weapon with the lowest collateral damage. Another advantage is that it is possible to house the gun in a turret and shoot in all directions. In the early days this capability was employed by fighters and bombers, in the Second World War it was employed by bombers only and presently it is the accepted way to house a gun in an attack helicopter.

Warhead

The size and weight of bullets might be limited, but the high velocity and design of the nose (for example depleted uranium) may still give them an excellent armour penetrating capability. A good example is the GAU-15, which is installed in the A-10. Not all bullets are designed to penetrate armour; some will have an incendiary warhead while others have a high explosive warhead that is more suitable in the engagement of aerial targets.

[76] Aiming accuracy is increased by laser range finding, rate measurement etc but basically the impact is always close to the bore sight of the barrel.

Range

Range is primarily determined by muzzle velocity and bullet weight. Traditionally effective ranges were between 1000 and 1500 ft for air-to-air, while for air-to-ground gunnery (strafing) the range is 1500-2000 ft. However, state of the art bullets have a concave nose that reduces aerodynamic drag. This has almost doubled the effective range of the gun. It is self-explanatory that a large gun such as the GAU-15, with a very high muzzle velocity (around Mach 3) has an even larger effective range.

Sensor

Although the most commonly used sensor for the gun is the human eye, its effectiveness is greatly enhanced by the use of the gun sight. The gun sight corrects for aircraft and target movement based on the measured target distance with laser or radar and the relative angle change of the target. The usability of the gun sight is greatly enhanced when it is projected on a helmet-mounted visor. This is the only practical way to employ a turret-mounted gun, which is how it is used in attack helicopters.

Weaponeering

Weaponeering is the art and science of matching weapons to targets in order to achieve the desired results with the minimum effort. Weaponeering has to take into account the following parameters: target vulnerability, weapon effects and the accuracy of the weapon delivery. The best known source manual for weaponeering is the NATO Joint Munitions Effectiveness Manual (JMEM). This publication gives the weapon effects for a variety of weapon-target combinations.

Target vulnerability is dependent on its structure and size. Not only the horizontal size matters but vertical extension of the target will also increase its vulnerability. Based on the weapon used, it must be delivered from within a certain distance from the target to be effective.

The weapon accuracy is dependent on several parameters. Wind prediction errors, ballistic errors, aiming errors and system errors all contribute to the total weapon accuracy. Accuracy of a weapon can then be expressed as a Circular Error Probable (CEP). The CEP is the radius within which 50% of weapons will fall. E.g. a CEP of 15' will mean that 50% of weapons will fall closer than 15' to the target. CEP is a proper measurement for guided weapons[77].

Non-guided weapons are not evenly distributed around the target. Experiments have shown that these weapons are dispersed around the bomb fall line (the line between the aircraft and the target). This dispersion is not measured in feet but in angle; the unit used is mill-radians or Mils[1]. This circular dispersion around the bomb fall line is called CEP_N *and the pattern on the ground resembles an ellipse. The range error will always be larger than the deflection angle. The difference is the greatest when the bomb fall line is shallow and decreases when the bomb fall line is steep.*

Typical figures for CEP_N are 40 Mils for non-computed deliveries and around 10 Mils for computed deliveries. The pattern on the ground depends on the dive angle and the delivery altitude as well as the weapon used. Some examples are given

[77] For guided weapons the distribution around the target is in general not normal. If the guidance system does not function correctly, errors may become quite large. However for non-guided munitions the dispersion around the bomb fall line is close to a normal distribution.

below for low and medium-level deliveries. A CEP_N of 10 Mils is assumed for the calculation of the CEP in feet on the ground.

MK - 84 GP Bomb Delivery speed 480 KTAS					
Dive angle (degrees)	Release height (ft)	Bomb slant range (ft)	CEP Range error over (ft)	CEP Range error under (ft)	CEP Deflection error (ft)
30	2200	3865	69	67	39
30	10000	14567	214	210	146
45	12000	14614	179	177	146

The table shows that a 10 Mil error at low-level deliveries still gives an acceptable error. But errors at medium altitude are quite large, even when the attack angle is very steep. Also interesting to note is that the pattern on the ground for a steep 45-degree delivery becomes more circular. The primary question to answer is how many bombs are needed to destroy a target. First we have to realize that we are working with probabilities. Therefore we have to multiply. The probability of target destruction is the probability that the aircraft reaches its target x the probability that the pilot detects the targets x the probability that the bomb functions x the probability that the bomb falls within the lethal area of the target. Only the last two probabilities- if the bomb functions and if it falls within the lethal area- can be calculated with JMEM; the other two have to be estimated based on experience.

To visualise the effect of accuracy on the number of weapons needed, we can compare weapon accuracy with the vulnerable area of the target. In the next table the first column shows how large the target vulnerable area is compared to the weapon CEP. So, if the target vulnerable radius is 10 ft and the weapon CEP is 100 ft, the ratio of target vulnerable area/weapon CEP is 0.1. The result is that only 1% of all bombs will destroy the target. As a consequence 229 bombs are needed to destroy the target with a 90% probability.

Target vulnerable area / weapon CEP	Percentage of weapons that destroy the target.	Weapons needed for a 50% kill	Weapons needed for a 90% kill
3	99.8%	1	1
2	93%	1	1
1	50%	1	3
0.5	18%	4	12
0.25	3%	23	76
0.1	1%	69	229

From this example it becomes clear what tremendous force multipliers precision munitions are. From these types of examples it is often argued that a modern F-16 has the same fire power as a complete bomber force in World War II. Some caution should be exercised: on some area targets (refineries, shipyards etc.) the vulnerability is quite large and even a World War II bomber could have achieved good results against those targets. But without a doubt, this increased accuracy was one of the contributing factors to the success of operations Desert Storm, Enduring Freedom and Allied Force.

Weaponeering can also be applied to calculating possible damage to nearby civilian targets. By doing so, the collateral damage can be minimized.

Chapter 7 – Maintenance and logistics

Introduction

The required maintenance man hours needed to keep an aircraft in the air quite often exceed the number of flying hours. A typical ratio for a fighter-type aircraft is seven maintenance hours for every flight hour. For a car the typical figure is just one maintenance hour per 20 hours of driving. In other words, flying is very maintenance-intensive. The obvious reasons are that aircraft are complex machines and that aircraft maintenance requires a bigger safety margin than, for example, motor vehicle maintenance. A broken down car can wait at the roadside, a broken down aircraft can crash.

There is another reason why we should be interested in airpower maintenance. The air battle is won with serviceable and mission-capable aircraft only. When comparing countries and air forces there is a general tendency to look at the aircraft in inventory only. However, a 100-aircraft air force with 60% serviceability is as strong as a 150-aircraft air force with 40% serviceability!

> *A classic example where good maintenance was instrumental in winning the 'war' is the Berlin Blockade (1948). The number of aircraft available was limited, and all aircraft needed to fly almost constantly. So, aircraft maintenance was done 24 hours a day seven days a week. Most aircraft technicians worked for 12 hours and were off duty for the next 24. The wear and tear was very high, particularly on the landing gear and a constant flow of spare parts from the USA to Germany was needed.*

In this chapter the factors affecting aircraft maintenance will be explained further.

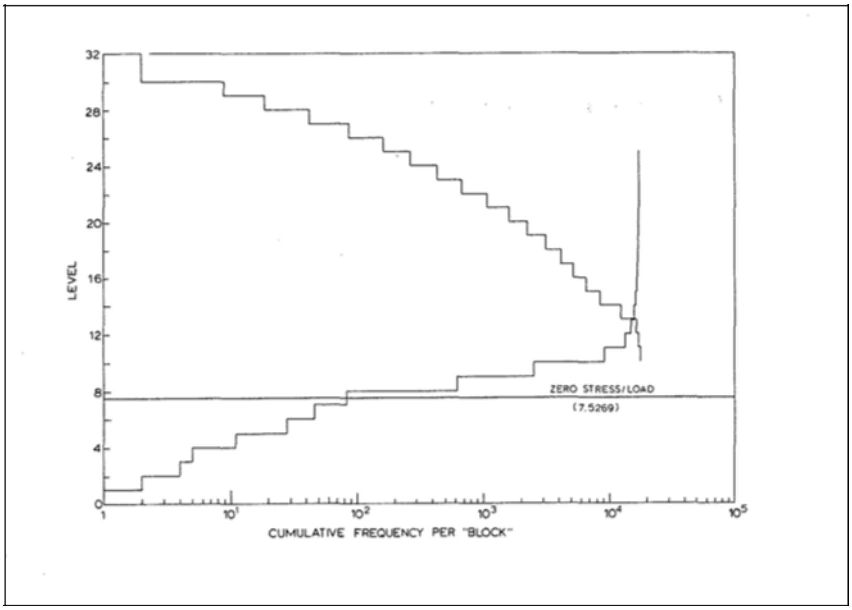
Wear and tear

There are a many factors that cause wear and tear on an aircraft. Before we start discussing those we have to realise that aircraft are normally built with small safety margins. Safety normally equates to extra material, extra material means extra weight that adds to drag and decreases range and performance. So the design is always a compromise between weight and safety versus performance. The following are the most common reasons for wear and tear:

Metal fatigue

A common fact is that metal breaks after it has been bent several times. The two most important factors that determine the level of fatigue are the load level and the number of cycles. The combination of load level and number of occurrences is known as the load spectrum. Before an aircraft is designed the load spectrum must be determined. For a fighter aircraft the load spectrum may look like the Fighter Aircraft Loading Standard for Fatigue (FALSTAFF) reference. The vertical axis shows the normalized wing load and the horizontal axis the occurrences per block of 200 flights[78].

[78] Normalized here means that 0 equates to the largest negative g-load for that airframe and 32 for the highest positive g-load. For an F-5 this 0 FALSTAFF equates to -3g and 32 to 7.33 g.



FALSTAFF Load spectrum

The aircraft prototype is not only designed against this spectrum but will also be tested against these loads. One of the early prototypes will not be fitted with avionics and engines but will spend the rest of its life in a test rig where it will be exposed to this load spectrum until the structure fails. Metal fatigue starts with small, still invisible cracks that will grow under the influence of load applications. It is important to know how fast these cracks will grow. If a crack increase is rapid, the inspection intervals must be short, but slow crack propagation allows for larger inspection intervals. There are three types of loads that are the most dominant factors for metal fatigue in aircraft.

The g-load, which causes metal fatigue in the wing and the wing spars;

The landings and take-offs cause heavy loads on the landing gear and on the spars and bulkheads where the landing gear is connected;

The aircraft fuselage is pressurized at altitude. This pressurization causes stress cycles on the fuselage and bulkheads.

Temperature cycles

All parts of an aircraft are subjected to heat and cold cycles, but the the cycle differs considerably per aircraft part. The most extreme heat cycle occurs of course in the hot section of the engine, where temperatures can easily exceed 1000ºC. At the same time the airframe can get as cold as -56ºC when it flies slowly at high altitude to as hot as 100ºC when it flies Mach 2. Temperature cycles change the molecular structure of the material causing it to lose its strength. There is no standard way a material reacts to heat cycles. Some material, like titanium, can withstand large heat cycles with ease while other materials will degrade rapidly. Heat cycles also cause material to expand and shrink. The problem with a jet engine is that the casing and the core expand and shrink at different rates. The core holds more mass, consequently it reacts more slowly than the much lighter casing. When the casing shrinks more rapidly than the engine core, the casing will chafe the core. To prevent damage by chafing the inside of the core is covered with abrasives. It is obvious that a larger space between core and casing will cause the engine to lose power.

This is the reason for the abrasives to be replaced after a certain number of heat cycles. However, if a jet engine runs continuously and is not subjected to heat cycles, it will hardly experience any wear. An often forgotten heat cycle occurs in all electrical equipment, notably when turning it on and off. What is true for a jet engine also applies to electronic equipment, when it is operating at a constant temperature it will wear more slowly. This is one of the reasons why electronic equipment has to be cooled or some electronic appliances have internal heaters to prevent them from operating at too low a temperature.

Friction

Friction is another important cause of wear and tear. Because a jet engine has rotating parts and no reciprocating parts the basic friction level is low. Well lubricated roll bearings are essential to keeping the friction as low as possible. If the bearings start to wear out, the number of metal particles in the lubrication oil will increase. Consequently, the status of the bearings can be inspected by checking the quantity of metal particles in the engine oil. Many aircraft types, e.g. the F-16, require such a check after each trip. This quantity is determined by burning a small sample of the engine oil and checking it for metal particles with spectrographic techniques.

Friction is not only a problem in the engine. Sand, dust and water drops will chafe off the paint on the airframe and particularly on the leading edges of the wing and rotor. Sometimes the leading edges of wings and rotors are fitted with easily replaceable tape to protect them. Some parts are designed to cause high friction-brakes and tires for example. Consequently these parts need to be changed regularly.

Ageing

All previously discussed forms of wear and tear are caused by use. Ageing however, is a process that is mostly independent of aircraft use. Parts can age by oxidation or under the influence of UV light or through contact with fuel or oil. Ageing can affect metal parts, seals, wiring etc. Traditionally ageing was not a dominant factor in aircraft maintenance. In the old days the usable life time of military aircraft was short, ten years at the most. But presently most aircraft types have a life expectancy of thirty years or more. As a result ageing has become an important factor in aircraft maintenance.

Maintenance general

There are two reasons for maintenance. Firstly, the repair of a broken aircraft, which is called corrective or non-scheduled maintenance and secondly inspections of the aircraft in order to prevent it from failure during the next service interval, which is called preventive or scheduled maintenance. Modern developments are aimed at limiting scheduled maintenance. This limiting can partly be achieved by equipping the aircraft with sufficient health monitoring devices. If all systems are continuously monitored, periodic inspections are superfluous and corrective maintenance can be carried out just before the system fails. Theoretically this concept is very attractive, but a prerequisite is that all failure mechanisms are known and that all relevant processes are monitored. While this might hold for a lot of systems, this is definitely not the case for all materials and subsystems. Particularly difficult is the prediction of the failure mechanism and modes of new materials and equipment.

Preventive maintenance

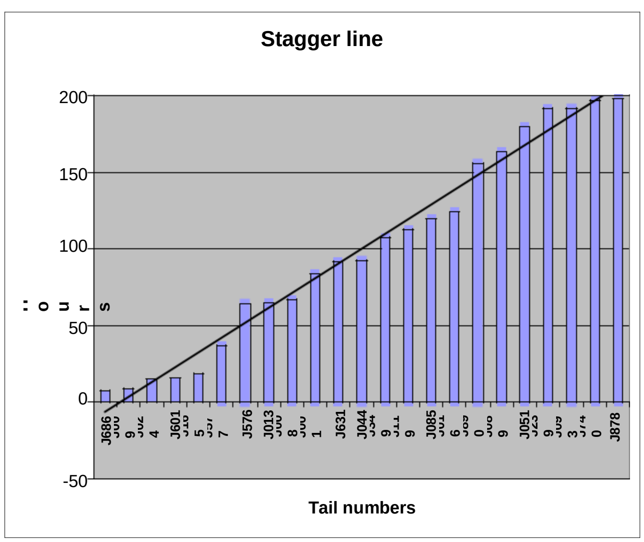
The prime method of preventive maintenance is inspection. Not all parts require the same inspection intervals. The engine needs an inspection after certain cycles, the landing gear after a number of landings and the airframe itself after a certain number of load cycles. Furthermore, washing, corrosion inspections and painting are required after certain time intervals. For practical reasons most inspections are grouped together in small and large inspections, in order to limit the number of periodic inspections and to standardize them. A possible schedule could be: a small inspection every 50 hours, a major inspection every 200 hours and an extensive inspection every 1000 hours

To guarantee that the aircraft structure does not fail between two inspections the interval should be short enough, meaning that the growth from a just discernable crack to a structural failure is less than half the inspection period. In this way there are always two opportunities to detect the crack before the structure fails.

Inspecting components can be done by removing them and scrutinizing them in the shop with special test equipment. Tests on the structure of the aircraft are more difficult. As it is impractical to take the aircraft apart completely, non-destructive methods are used, such as applying sonic, X-ray or UV-sensors.

It is important that periodic inspections are planned in such a way that a minimum number of docks and maintenance personnel is required. To achieve this, the fleet should be 'staggered'. That means that each tail number has a different number of flying hours left until the next inspection. If the aircraft are not well divided over the stagger line, more aircraft might need inspections than there are docks available and consequently, during other periods the docks will be empty and personnel idle because no aircraft has yet reached the required flying hours.

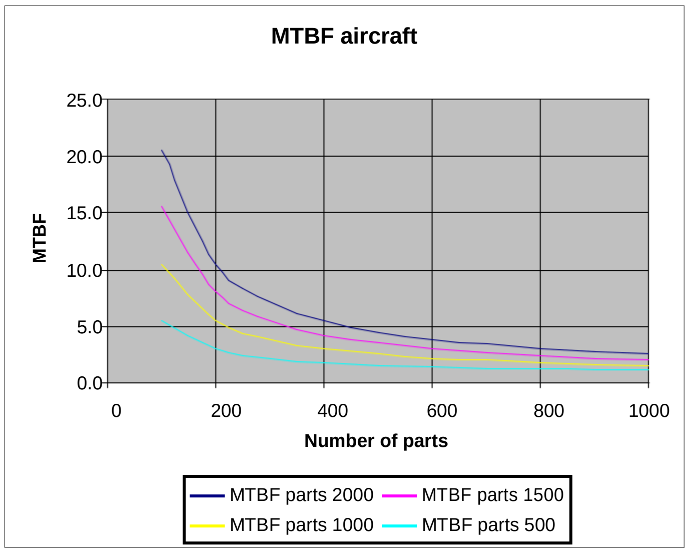
After the required inspections are carried out and all removed subsystems are re-installed, the aircraft needs 'follow-on maintenance' (FOM). FOM is required to ensure that all reassembling is carried out correctly. Not everything can be tested sufficiently on the ground. Therefore, it is often necessary to perform a maint-enance test flight also. Helicopters in particular need many maintenance test hours to trim the rotor system[79]



[79] This is called track and balance. With a correct weight distribution the vibration level can be lowered, furthermore the blades are trimmed in order to fly in the same plane.

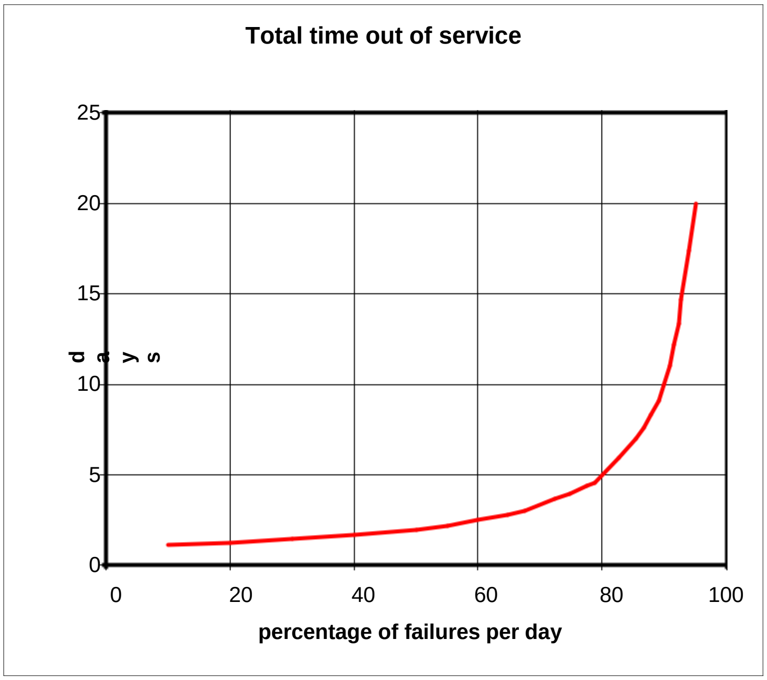
Corrective maintenance

There are two counteracting mechanisms that determine the level of corrective maintenance required. One such mechanism is the reliability of each individual subsystem. With technical improvements the general tendency is that reliability goes up and that the Mean Time Between Failures (MTBF) of subsystems increases. The second mechanism is complexity. Aircraft are becoming more complex and the number of subsystems is on the increase. This has a negative effect on the MTBF of the complete aircraft. This relationship is depicted in the next graph. In this example it is assumed, for simplicity reasons, that all subsystems have the same MTBF. This is of course not the case, but even with a variety of MTBF, the general relationship is the same. The graph shows that, even with very reliable subsystems, the MTBF of a complex aircraft will always be low!



To increase the availability of an aircraft the designer can build in a lot of redundancy. In this way, the failure of a single subsystem will not jeopardize the availability of the aircraft. This is quite common in civil transport aircraft. But redundancy requires extra space and extra weight that is not always available in high performance military aircraft. Consequently high performance aircraft will always have a lower MTBF than transport aircraft.

When an aircraft component fails there are two possibilities. If the effect on the aircraft performance and safety is negligible, it can be decided to postpone repair until the next periodic inspection. If the break-down is more serious an immediate repair is warranted. It is easy to see that expedient repairs require spare parts and technicians, but planning for corrective maintenance is difficult because it is not known when what system will break down. Just having technicians standing by in case of a break-down is expensive as well. A further complicating factor is that there are many types of aircraft technicians who are each only certified to work on a particular set of subsystems. There is a mathematical relationship between the total time-out of service (waiting time plus repair time), the failure rate and the repair capacity. This relationship is depicted in the graph below. The graph shows that when the average amount of repair time needed equals half the repair capacity the average total time out of service is already twice the repair time. If the failure rate increases beyond 50% of the repair capacity, the average time out of service increases rapidly. If the failure rate is 80% of the repair capacity the average time out of service is already 5 times the repair time. This mathematical relationship is not intuitively obvious. Most people expect a much shorter out-of-service time.



In general the following rules apply for corrective maintenance:

- Grouping corrective maintenance is more efficient. The effect of large numbers, for instance of a certain aircraft type, is that the fluctuations dampen out[80]. This enables the more efficient scheduling of the workforce. Secondly it is more efficient in the management of spare parts stocks that do not have to be divided over different locations.
- The more disciplines a technician masters the more efficiently he can be employed. Over-specialization will diminish his effectiveness.
- A lot of corrective maintenance is carried out by just changing the defective part or subsystem. Most avionic systems are easy to remove and even a complete engine change can be completed within a day. But not all components can be removed easily. If, for example, a connectivity problem exists somewhere in one of the many electrical wires, the repair time can be quite long. Changing a part is in itself not always sufficient for completing the repair. A certain level of Follow on Maintenance (FOM) might be required to certify the correct operation. For example, if one of the components of the hydraulic system is changed, the complete hydraulic system must be checked to ensure that no leaks exist and that no air is present in the system. Similar types of FOM are required for other systems.

Spare parts control

Parts that have been removed are sent to a certified repair facility. This can be a depot or one of the back shops on base. The time a part spends in the repair

80 Fluctuation as a percentage of the total number of failures will be less. The absolute fluctuations are of course the same.

process is called the time in the 'pipeline'. It is not uncommon for the pipeline to be several months long, especially when the part has to be shipped abroad for repairs. The consequence is that a lot of 'dead' capital is sitting in this pipeline and that the quantity of required spares will rise.

The number of required spare parts is not always easy to establish. Some goods, for example aircraft tires, have a known and fixed consumption rate, which is expressed in so many takeoffs and landings. As long as the stock is higher than the consumption rate x the time needed to get new ones delivered, the stock is sufficient[81]. The difficult part is when failure rates are not known, or if, after a certain time period, the failure rate suddenly increases. A sudden increase in failure rate is not uncommon in aircraft maintenance. Sometimes ageing problems are the culprit. In this case the stock runs out quickly but the time required to get sufficient new spares can be quite long, often because the manufacturer has to boost his production or, even worse, the production facility has been closed. Setting up a new production line for aeronautical components, with all the required quality assurance checks, cannot be done overnight. Not surprisingly, the user sometimes has to refer to less orthodox methods to get sufficient spares. He will strip usable parts from aircraft out of use due to periodic inspections or long corrective maintenance. This is called 'cannibalization'. Although this eases the spare parts problem, the maintenance hours involved are doubled.[82]

Modifications and configuration management

General

in the present day, the service life of an aeroplane can be much longer than the time served by the average professional airman. A perfect example is the B-52, which entered the service in 1952 and is still flying. It is quite possible that some current B-52 pilots are the grandchildren of the early B-52 pilots. If an aircraft is to stay operational over an extended period and must remain safe and operationally capable, many modifications are required.

It is not uncommon for the first modifications to be designed before the first prototype has been delivered and during the life of the aircraft hundreds of modifications will come into effect. Because not every modification can be carried out immediately, an administrative system is needed to track the modification process. This type of tracking is one part of the configuration control process.

The other part of the configuration control stems from the fact that an aircraft is actually an assembly of many replaceable units. Each individual unit, e.g. an engine or an actuator, has its own time or event-based inspections and applicable modifications that have to be carried out. Of course, not every light bulb needs to be accounted for, but important parts like engines, hydraulic actuators and ejection seats have their own modification standards and their own inspections. These important items are called 'tracked' items because the accumulated lifetime and the modification standard of the item are being tracked.

We will first discuss the functional categorization of modification and then the categorization into systems. We will then go on to explain the modification process and the system that is needed to keep track of all these modifications.

---

[81] Some allowance has to be made for fluctuations in usage and fluctuations in delivery times.

[82] The cannibalized spare part has to be removed as well plus, when the parts are back in stock, the cannibalized aircraft has to be refitted and FOM needs to be performed. This means doing the required job twice, once on the broken down aircraft and once on the cannibalized aircraft.

Functional categorization

The division between functions is to some extent arbitrary and it is quite possible that different services have different categorizations. The chosen approach here is to base the function primarily on what triggers the requirement. That is the reason we discern four categories: flight safety, maintainability, legislation and operational capability.

Flight safety. The most dramatic triggers for a modification are accident investigations that reveal (a) fault(s) in the aircraft or engine design or prove that some parts age much quicker than anticipated. Fortunately, not only accidents but also results from periodic inspections will reveal faster-than-normal degradations of a system.

> *When the F-16 was first fielded it had a small horizontal tail. After several crashes it was proven that the tail area was actually too small to recover from a 'deep stall' and that a pilot could 'depart' the aircraft too easily. From Block 15 onwards all aircraft were fitted with a larger horizontal tail and the older aircraft were retro-fitted with a larger horizontal tail.*

Maintainability. There are several types of modifications that can improve maintainability. The first example is when new hardware has improved reliability:

> *In the 60's most aircraft were equipped with mechanical Inertial Navigational Units. (INU) The mean time between failures (MTBF) of those units was only 60 hours. In the 80's a new type of INU, the ring laser INU was fielded, which had a MTBF of more than 1000 hrs. The introduction of this new INU did improve the operational accuracy only slightly but decreased the amount of overhaul considerably.*

Another maintenance problem is that parts can become obsolete necessitating a new part to be certified and integrated into the configuration, which can trigger other modifications. Especially electronics devices have a short life span and can become obsolete in just a few years.

> *When additional spare engines were purchased for the RNLAF Chinook it turned out that while they did fit into the Chinook, the engine wiring had been changed, which necessitated an additional modification to the airframe wiring.*
>
> *When a new radar computer for the F-16 was introduced during the Mid Life Update in 1998, the –built in chips were already out of production.*

Special attention should be given to parts that are only produced as long as the production line is open. The user and the manufacturer should have a clear understanding of how many spares need to be produced. The economic challenge is that too many produced spares are just dead capital while too few can jeopardize the utilisation of a very expensive aircraft.

> *Examples are aircraft parts like bulkheads, flaps and slats. Normally they are very seldom replaced, but if they are damaged the aircraft is rendered useless.*

Last but not least, the aircraft can be modified to improve maintainability. Some hardware changes can be made to improve the accessibility or some hardware and software changes can be incorporated to track the performance of a module more efficiently.

New legislation. The increase in air traffic has demanded that aircraft be outfitted with better equipment. More accurate altimeters and autopilots enabling smaller altitude separation minima[83], and VHF radios with a smaller bandwidth per channel increasing the number of channels in the allocated spectrum were made mandatory. Other recent improvements were better navigation equipment and a new type of transponder (Mode- S). ATC legislation is not the only reasons for modifications. Civil air transport standards for people and cargo are becoming more and more stringent. Military authorities sometimes wave civil requirements for operational reasons, but often they will try to maintain equivalent safety standards for equivalent types of transport. National environmental regulations are further drivers for modifications. The use of asbestos in aircraft is a good example. Following national regulations these parts have had to be replaced.

New operational capability. The operational envelope of military aircraft normally expands. An example is the F-16, which was initially designed as a light day air-superiority fighter and has since evolved into both an AWX interceptor and a day-and-night bomber. It has also gained some weight over the years. This increase in operational capability has required the integration of new hardware, new software and new weapons.

Operational requirements always exceed available time and money. Therefore most aircraft manufacturers will use a block system. The first models produced will only have a limited capability while each next block will incorporate more of the follow-up capabilities. This block system facilitates the production process of the aircraft manufacturer. At a certain point the design is frozen and production can run smoothly. This is more efficient than retrofitting changes into an ongoing production line. However, there are also drawbacks. If a new block is available the operator will normally want to have his 'old' aircraft equipped with the new capabilities also. If the update is software only, this can be done overnight. But if new hardware has to be incorporated, a special programme is needed. It can be advantageous to carry on with several configurations for some period and add several block changes in one programme. We will now discuss some of the changes in more detail.

System categorization and possible effect on aircraft use

There are many system categorizations, such as aircraft, engine, subsystems, weapons and software. In this book we will group all hardware changes as well as all software modifications together, although in real life this division is somewhat artificial. Many hardware modifications will require software as well, although there are some pure hardware changes, for instance structural changes or rewiring.

New hardware. The last decades have seen a rapid development in sensor and computer technology. The introduction of this new equipment in military aircraft has broadened the operational envelope of the aircraft. Just to mention a few: infrared imaging opened the night window, digital terrain systems made low flying safer, advances in computer power enabled tracking more targets, Head Up Display improved pilot awareness etc. Many systems were retrofitted in legacy aircraft (in the existing fleet), requiring extensive modifications. Because space is a scarce commodity in fighter aircraft and attack helicopters, many systems were built into pods, which could be carried externally.[84]. Hardware modifications also include new systems replacing older systems with lesser capability or lesser reliability (new engines, new radar). To fit new hardware into an aircraft can be a complex task. There are many considerations. The most important are:

---

[83] Initially altitude separation above FL 290 was 2000 ft but with better equipment 1000 ft separation is allowed.

[84] For example targeting pods, but also jamming.

- Does the modification affect the flight envelope of the aircraft? This is always an important consideration when new equipment is mounted externally and seldom if installed internally;
- Is the hardware functional in the complete flight envelope? Can it withstand thermal and pressure fluctuations and vibration levels?
- Does it interfere with other equipment? Every piece of electronic equipment that is not perfectly shielded can become an emitter that could influence the use of other electronic equipment[85].
- Is it incorporated correctly into the software, are all supposed functionalities present and easy to select by the operator?

New hardware is only integrated when the hardware itself already complies with minimum technical standards. This does not guarantee per se that it will function correctly when installed in a particular aircraft. A test and verification programme is normally required for the new hardware. When the hardware is 'form fit and function' equivalent with previously installed hardware there is little need for tests.

Modifying existing hardware. Most of the changes required to existing hardware are necessitated by ageing. New wiring harnesses and re-skinning some parts of the aircraft are common modifications for older aircraft. Even complete bulkheads can be replaced. The certification process for these types of changes is normally straightforward, however the number of work hours involved can be enormous. Certification is easier because nothing 'new' is added, but the airframe is restored to its original strength and capability (or better).

Software. Modern aircraft are software intensive. The software can easily encompass several million lines of code. Most of this code is not a universal programme that can be used in each type of aircraft, but a programme that is only applicable to one type (or even block) of aircraft. The various blocks of F-16's all have different software! If we consider the volume of software and the limited number of users, it is easy to see that the software costs per aircraft are considerable. As a rough indication we can say that 50% of development cost is related to software development. Some modifications are purely software related. For instance, one can change the characteristics of a radar scan or add a new bomb delivery mode through a pure software change. On the other hand, the incorporation of new hardware and weapons will normally also require a software change. Like hardware changes, software modifications must be certified.

Weapons. New weapons are probably among the most complex items that can be incorporated into aircraft. There are hardware and software modifications to start with. But on top of that, the safe employment envelope must be determined (see also chapter 6). The latter can require several test drops or firings, which by nature are very expensive. The process of incorporating a new weapon into the software and certifying it for operational use can easily last for several years.

The process

The aircraft modification process is a structured process, which can be slightly different for each aircraft type. Also, not every type is managed in the same way. The Eurofighter is managed by an international consortium, but the B-1B by a single service. International programmes are of course more difficult to manage. Different countries have different requirements and different budgets, so

[85] The technical expression for this phenomenon is Electro Magnetic Interference or EMI.

compromises must be found, but the following general phases can be distinguished for most programmes:

Requirement definition phase. The operational community is normally leading in defining operational requirements. Initially those requirements are broadly defined. For example incorporate Link-16. In later sessions with operational and engineering experts the requirements become more specific and will stipulate what Link-16 messages have to be incorporated into the software. When the more detailed requirements are known, the number of required changes will normally exceed the finances and time available, which will lead to the need to prioritise requirements.

Engineering phase. In this phase engineers start to develop the hardware, the wiring and the pilot vehicle interface (PVI), plus the test programme. During this phase operators will regularly visit the manufacturer to evaluate the PVI and its progress. Simulators are the most appropriate instruments for evaluating the PVI.

Prototyping. A prototype is not always required for each and every modification but some prototyping is always done. In general, the more challenging the project, the more prototyping is required. If for example the project aims to incorporate a Link-16 into the aircraft, the prototyping may be limited to checking the first new wiring harnesses. Prototyping of the software can be performed in engineering simulators.

System integration lab. The new hardware is tested together with the new software in a software integration laboratory or SIL. In the SIL the actual hardware and software of the aircraft are tested under static conditions.

Development test. After the hardware and software are tested in the SIL, flight tests will be performed to confirm the correct operation of the equipment in the air. During these tests, specific test points will be flown as determined by the contractor and the programme office. During these tests specialized aircraft are used that are modified to capture data for analysis. Typically all data buses of the aircraft and all flight parameters are recorded.

Operational test and evaluation. The last step in the verification and certification process is the operational test and evaluation. Originally the OT&E was purely intended to test the modification in the operational environment and to develop tactics and training procedures. Since most of the changes in an aircraft are software-based, and software errors are hard to find, it has become quite common for new software errors to be detected as late as during the OT&E phase. In fact, even the software released to the service will not be error free. The basic problem is that the scale of the software prohibits the testing of all mode combinations.

Training. New equipment will often require additional training for maintainers and users. Training requirements can range from a small briefing on the new equipment to a full six months' training course. An extensive and long course is typically required after a mid-life update programme in which the majority of all systems are replaced. In fact the latter course is very similar to a conversion course to a new aircraft type.

The process from the requirements phase to fielding for a small modification can be a matter of months. A complete Mid Life Update can take ten years.

Keeping track

Operators presently use computerized system to track the aircraft configuration. Those systems are tailored 'enterprise resource management systems'. A computerized system enables and partly automates tracking functions. However, the quality of the system is primarily determined both by the input quality and the people who use the system intelligently. The Dutch experience with CAMS[86] was that after a few years the system was contaminated with errors that could render it system completely useless. The number of maintenance variables tracked is large and includes part numbers of tracked items, modifications performed and due, aircraft and engine hours and maintenance actions performed, to mention just a few.

Conclusion

A high-quality maintenance and logistics organization is fundamental to the operation of a modern defence organization. But what is true for the defence organization in general is even more valid for air operations. Air operations are performed at the edge of the envelope in a hostile and harsh environment stretching man and machine to the limits of what is possible. Thorough maintenance and a balanced logistics system are essential to make it happen.

[86] Cams = core automated maintenance system

Chapter 8- Navigation

Introduction

Navigation is the science of locating your position, locating your destination and calculating the optimum route between the two. Navigation on land has always been relatively easy. The position was based on landmarks, destinations could be found on a map and a compass gave the direction. Navigation at sea, outside coastal waters, was mainly astronomical navigation. Just after sunset and just before sunrise star elevation angles with the horizon were measured and based on these data the position could be calculated to within a one nm accuracy. With the introduction of aircraft, navigational methods needed to be adapted. Flying above clouds obscures landmarks and astronomical navigation is a slow and tedious process, which is not very practical in fast moving aircraft.[87] Pilots needed a faster way of determining their position irrespective of the weather. Accuracy had to improve as well: when an aircraft is flying to an airport in bad weather it should be guided to a position precisely before the runway and not a few hundred feet to the left or the right.

Position

The positions of the aircraft in longitude, latitude and height (above mean sea level) are just a few of the necessary parameters. For accurate navigation the complete 'state vector' of the aircraft needs to be known. The state vector consists of the following parameters:

- **Latitude and Longitude**. These are normally calculated by the navigation equipment. Instead of longitude and latitude, the position can also be relative to a beacon and be expressed in radial and range. There are several kinds of navigational equipment that can generate this data. These include radio beacons, Inertial Navigation Unit (INU), Global Positioning System (GPS) and Terrain Profile Matching (Terprom). These systems will be discussed in more detail later.
- **Altitude**. Normally altitude is expressed in feet above mean sea level. The most important instrument to measure altitude is the altimeter, which basically just measures the ambient air pressure. To give a correct altitude the altimeter must be corrected for the ambient pressure on the ground, which is changing constantly. With this correction (QNH[88]) the instrument is reasonably accurate at low level. To make the instrument more precise at high level, a correction should be applied that is based on the temperature difference between the actual atmosphere and the standard atmosphere for which the instrument is calibrated. A second method to measure altitude is with the radar altimeter, which just measures the height above the ground. To convert this measurement to an absolute measurement, the elevation of the ground below should be known. The Inertial Navigation Unit (INU)

---

87 Astronomical navigation was used for a while in aircraft, but special sextants were required with an artificial horizon. The calculation of the position took several minutes, so you were only able to determine where you were and not where you are. However, for slow moving aircraft flying over a big ocean, for example maritime patrol aircraft, this was an acceptable navigation method.

88 During the early days of air navigation, a lot of terms were abbreviated in 'Q' code to ease radio traffic which was primarily done in Morse code. Some of those 'Q' code words like QNH are still in use to day.

and the Global Positioning System (GPS) are two navigation systems that are capable of generating an absolute altitude.

- **Bank and Pitch**. The only way to measure angles of the aircraft with the horizon is with a reference gyro. The gyro can be a stand-alone instrument (a dedicated altitude indicator as used by small aircraft), but the data can also be derived from the INU.
- **Heading and Track**. The oldest instrument for measuring heading is a magnetic compass. The magnetic compass has to be corrected for the magnetic field of the aircraft, which is done with small magnets. The remaining error (deviation) is written on a placard and must be applied by the pilot. To change the magnetic heading to a true heading, the local variation that is depicted on aeronautical maps must be applied. An easier way to get the true heading is directly from the INU. Even a true heading will not give the exact direction in which the aircraft is moving. The wind will cause the aircraft to drift. This drift angle can be calculated based on the forecasted wind, but this is not very accurate. The drift angle can also be measured by an RF-Doppler measurement and be calculated by the INU. When the accurate drift is applied, the track of the aircraft over the ground is known.
- **Velocity**. The oldest method for determining aircraft velocity is by measuring the dynamic air pressure. The dynamic air pressure is the difference between the total pressure, which is measured with the Pitot tube[89] and the static pressure measured on the side of the aircraft. The measured or Indicated Air Speed (IAS) must be corrected for position errors[90] and after this correction it is called Calibrated Airspeed (CAS). The calibrated speed is a very important performance indicator, for example to calculate landing velocity, but is not directly usable for navigation. Under standard atmospheric conditions a CAS of 300 knots at sea level is also 300 knots True Airspeed (TAS) but the same 300 knots CAS will be 512 knots TAS at 36,000 ft. If the altitude and temperature are known, every CAS can be recalculated to TAS. This can be done by the pilot or automatically by an air data computer (ADC). If the TAS is corrected for the wind velocity, the true ground speed is known. If the true ground speed is based on forecasted winds it will not be very accurate, but the true ground speed can also be measured directly with an INU or by RF-Doppler measurement.

Reference systems and maps

Each position on earth can be measured relative to a certain fixed point. In many navigational situations this is sufficient. When an aircraft has to land, its relative position to the runway is the only thing that matters. On the other hand, when an aircraft has to fly from one location to another thousands of miles away, it is important that position and destination are measured within the same reference system. As long as we saw the earth as an almost perfect sphere, the reference system was relatively easy. The earth axis and the equator were determined by

[89] The Pitot tube extends from the front of the wing or from the nose of the aircraft and measures the pressure in the forward direction.

[90] The main cause of error is flow field variation around the static ports, which is dependent on Mach number and angle of attack. The aircraft manual has correction charts for IAS to CAS, but it is also possible that the ADC automatically corrects for this error.

measurements. And with a common reference meridian (Greenwich) and radius[91], a global reference system was born. When measurements became more accurate, people discovered that the earth's shape is more like an ellipsoid. Unfortunately a lot of differently shaped ellipsoids have been used for mapping different parts of the world. But the situation got even worse. Local areas on earth were mapped by triangulation each based on its own local reference point (or datum). The result was that there are more than a hundred different types of maps each based on a specific datum and ellipsoid.

> *There is at least one reported accident due to the use of the wrong reference system. When the unfortunate aircraft approached the field, the local radio beacon was not working. The pilot decided to load the position of the beacon into his navigation computer and fly to the airfield based on GPS. Unfortunately the coordinate system that was used for measuring the position of the beacon was different from the coordinate system that the aircraft was using. The result was that the aircraft flew into a mountain killing all on board.*

Presently the most commonly used reference system is WGS-84 and an increasing number of maps are based on this system. There are conversion programmes available to convert a position from one reference system to another.

Accuracy and integrity requirements

The required level of navigation accuracy depends heavily on the flight phase. For en-route navigation a navigational error of 1 nm is quite acceptable; close to an airport the required accuracy is much higher. The highest accuracy and reliability are required for an auto-land system, which requires accuracy to within a few feet (especially in altitude, to make an accurate round out) and a continuous monitoring of the integrity, i.e. the coherence of all onboard data. Part of the integrity monitoring is done by the pilot who has to verify that the correct beacon is selected. Another part is done by automated systems that, for instance, monitor the emitted ILS signal and monitor the correct operation of the onboard systems. Of course, when the required accuracy level of navigation is lower, the integrity of the system does not have to be monitored continuously and periodic checks will suffice.

Military requirements

Military operations require special provisions. The navigation system should be accurate, work over hostile territory, preferably not emit signals (which give away the aircraft's position) and not be affected by electronic jamming. These requirements can not all be met at the same time. GPS is the most accurate system but the GPS signal can be jammed and between mountain ranges the reception can be limited. The INU cannot be jammed and is silent, but its position accuracy decreases over time. Terprom is less accurate than GPS, better than an INU and the system cannot be jammed. However, it emits signals. The best solution is often to have a variety of systems available that supplement each other.

When different systems are integrated into one navigational solution, filtering is required to determine what the most likely position is. The common filtering method used for navigational problems is Kalman filtering[92].

---

[91] If the earth really were a perfect sphere the radius could be measured, but because it is not, the measurement of the radius yields different results depending on where it is measured.
[92] Kalman filtering uses a statistical analysis that takes into account the accuracies of the different sources. Based on these data the most likely position and speed are determined.

The military operator needs more than an accurate own position. He also has to be able to accurately determine the target coordinates. Therefore the target sensors (radar, targeting pods etc.) should give a very precise angle and distance to the target. A targeting pod with laser ranging is one of the most accurate instruments for locating a target.

Equipment and characteristics

This paragraph will discuss the most common air navigation systems. The emphasis will be more on the strong and weak points of the different systems and less on the explanation of their operation.

Inertial measurement Unit (INU)

The INU uses three gyros to measure rotations around the three axis of the aircraft and three accelerometers to measure accelerations in each direction.[93] Before an INU can be used, the start coordinates must be inserted and the platform must be aligned[94]. Starting from a known position and a known orientation of the gyros, the INU integrates over time the rotations to calculate the three angles (bank, pitch and heading) of the aircraft. Integration of the acceleration over time will give velocity and integrating the velocity yields distance travelled.

The system does not emit, gives a continuous output, is not dependent on outside sources and cannot be jammed. Because of these unique characteristics the INU normally constitutes the heart of every navigation system. The only limitation of the INU is that its position accuracy degrades over time, normally by less than one nautical mile per hour. So, to keep it accurate, an INU must be updated at regular intervals.

Terrain profile matching (Terprom)

Terprom uses a digital terrain elevation database (DTED).[95] During flight the contours of the ground are measured with a radar altimeter and compared with stored data. By matching the measured contours with the stored data, the aircraft's position can be determined. To operate correctly, the system also needs to know the heading, speed and altitude of the aircraft. Normally these data are derived from the INU. The accuracy of Terprom is to within 200 ft but will degrade if the aircraft is flying over flat terrain. The system is independent of outside sources, the radar altimeter is hard to jam but the system does emit and this can interfere with stealth tactics.

Visual & radar

The pilot can distinguish significant points on the ground with his eyes, with the radar or the targeting pod. When the reference point coordinates are known in the navigation computer, the pilot can slew the radar, the targeting pod or the Head Up Display (HUD) mark over the correct position to update the navigation computer.

[93] The three axes are the longitudinal axis or x-axis, which is aligned with the fuselage, the lateral axis or y-axis, which is parallel to the wing and the top axis or Z-axis, which is perpendicular to the X and Y-axis.

[94] Alignment means that the platform knows what the level position is (based on the earth's gravity field) and knows where True North is. True North is determined from the earth's rotation, which is sensed by the gyros.

[95] Digital terrain elevation data (DTED) is available in different resolutions, DTED level 1 has a resolution of approx. 100 meter spacing and DTED level 2 of around 30 meter spacing. Not all parts of the world have DTED data in the highest resolution.

If a navigation system is fitted with a good Kalman filter, regular updates will not only correct the present error but also minimize the drift of the system[96].

GPS

Satellite navigation is based on time measurements. The orbits of navigation satellites[97] are known, and each satellite transmits a very accurate time signal. The time difference between actual time and received time is caused by the time the signal needs to travel the distance between the satellite and the receiver. Because the speed of RF signals is known (speed of light) the distance to the satellite is known as well. To calculate its position, the GPS receiver needs to measure the distance to at least three satellites and when the GPS receiver does not have an accurate clock of its own it needs to receive signals from four satellites to calculate its position and time.

The accuracy of GPS is very high. The horizontal position error is around 30 ft and the altitude error is a bit higher, around 75 ft. Initially the accuracy for civil application was lower than military accuracy, but the deliberate error that was added to the civil signal has been removed.

RF Doppler measurement

If a radio signal is transmitted in a forward direction and the ground return of the signal is received, the aircraft's groundspeed can be calculated from the Doppler effect. If the speed is transmitted on the right and left side of the aircraft it is possible to measure the drift angle as well. This RF-Doppler measurement was widely used after World War II. To calculate a position the system needs to know its starting point and the correct heading, which is normally derived from a gyro-compass. The accuracy of the system is normal 1-1.5 knots in speed and better than one degree in drift angle. System accuracy will decrease over time, which necessitates regular updates. The system can not be jammed easily but it does of course transmit.

Beacons

Radio beacons are the oldest air navigation instruments. There are several different types such as VOR, ILS, Tacan. (For a description see chapter 5.) Almost all beacons function based on the principle of phase measurement. Depending on the position relative to the station, the emitted signal phase is different. Some radio beacons will give a distance as well. The distance measurement is based on a transponder function.[98] The accuracy of the systems differs, a VOR radial will be approximately 1° wide, but the glide path accuracy of ILS is better than 0.1°.

Radio beacons are widely used in civil aviation and ILS is the prime landing aid used by civil and military users. During the Blitz in World War 2, the Germans used beacons to fly to targets over England; fortunately the English were aware of the system and jammed it. Presently beacons are no longer used for military applications. Their accuracy over a large distance from the station is insufficient and the signal is easy to jam. But beacons are still the most important landing and navigation aid for civil and military aircraft.

96 A good Kalman filter not only gives the most probable position but can establish the INU drift rate and correct for it.

97 The present GPS constellation consists of 24 satellites, which orbit the earth in approximately 18 hours.

98 The aircraft transponder emits a pulse and when the beacon receives the signal, it will transmit it back to the aircraft. Based on the time delay the distance can be measured.

Conclusion

Accurate navigation has been a big challenge to aviators for many years. In World War II many bombers did not come within 10 nm of the target. With the variety of systems available, navigation has become much easier. The primary challenge for an aviator today is not to know where he is but where the enemy is. This will be further discussed in chapter 11.

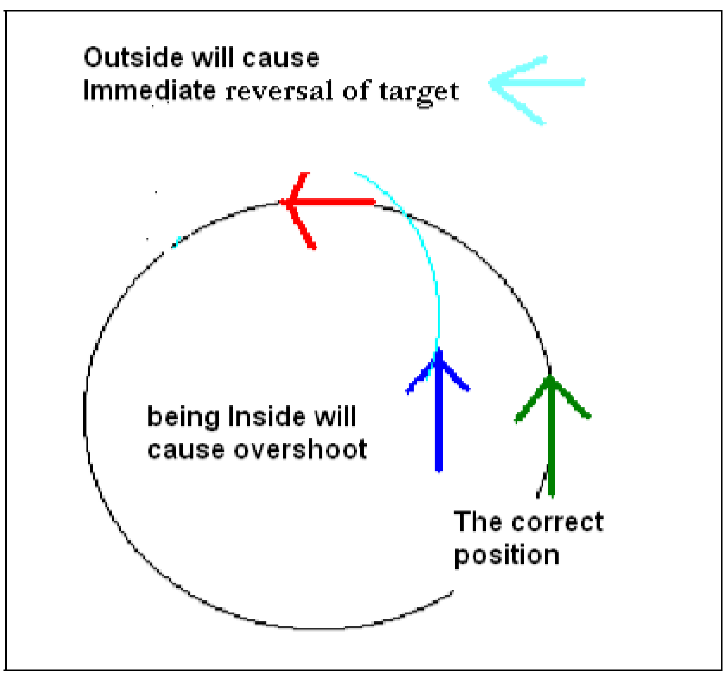
Chapter 9 - Weapon employment

Introduction

There is a large range of weapons that can be employed from aircraft. In the air-to-air arena radar and IR guided missiles are primarily used. But the oldest of all aerial weapons, the gun, is still present in most fighter aircraft. For air-to-ground, guided and unguided bombs are the primary weapons. Against moving targets or when more range is required, guided missiles are the weapons of choice. In this chapter we limit ourselves to describing how an individual pilot will employ his weapons. In reality, the pilot is seldom alone, and the formation leader will brief before the flight on how the formation elements are to support each other. These tactics depend on many parameters, such as threat, experience level and the amount of risk the formation is willing to take. Furthermore, these tactics are under constant review, are adapted over the years and are mostly classified. Therefore formation procedures are outside the scope of this book.

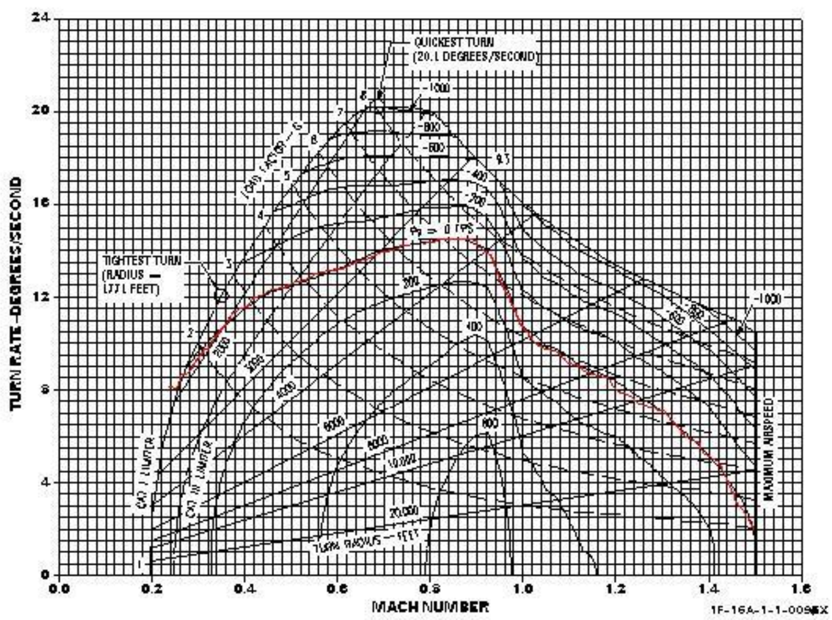
Gunfight

In fighter aircraft, the gun is normally fixed so that the aircraft itself has to be aimed. The most common procedure for a fighter aircraft is to position itself behind the other aircraft at ranges between 1000 and 2000 ft. Because the target will also manoeuvre, it is imperative that the fighter first positions itself on the turn circle of the opponent. The next step is that the fighter closes in to a firing position. This means that while matching the turn circle of the opponent, he must be able to fly faster and, as a consequence must have a higher turn rate and of course a higher g-level.



To be able to do this, the aircraft must have a better manoeuvrability than the opponent. Manoeuvrability however is not a constant but is very dependent on speed. At low speed the aircraft is limited by the maximum angle of attack or stall limit. At high speed the aircraft can not pull more g than the aircraft's structural limit. A further limiting factor is available thrust. The aircraft might at a certain speed be capable of pulling 9 g but if thrust is insufficient, it will decelerate. Finally, if the speed is too high, the turn circle will increase and the fighter will not be able to match the turn circle of his opponent. These limiting factors are all depicted in energy manoeuvrability plots. Comparing these plots for different aircraft can help pilots determine at which speed regime they can outmanoeuvre their opponents[99].

F-16 Clean, 15000Ft AB



A gun fight is all about energy management. The fighter will try to force his opponent into a regime where his manoeuvrability is lower and then close in to the firing position. An opponent is left with several choices: he can try to force the opponent into an overshoot by pulling maximum g or maximum angle of attack, execute a rapid speed reduction, perform rapid flight path changes (jinks) to prevent the opponent from aiming accurately, or execute a 'bug-out' (accelerate away from the fight). Hitting another manoeuvring aircraft requires considerable skill. To help the pilot in the aiming process, most fighter aircraft are equipped with a gun sight that calculates where the pilot should aim, based on radar information.

Theoretically it is possible to avoid the cumbersome manoeuvring needed to position yourself behind your opponent. One could consider making a head-on gun attack. The head-on aiming, however, is very difficult and this engagement is also extremely dangerous to train. In the early eighties an automatic head-on gun mode linking the radar information directly to the autopilot was developed. The results achieved with this system were impressive; however it did not enter the operational arena. Apparently the interest to further improve the gun sight was low. The development of better missile systems has focused the interest away from the gun.

Missile engagement (IR)

The first IR missiles were only able to sense the heat of the engine exhaust and had to be fired from rear aspect, thereby resembling gun attack. The difference with a

[99] This was invented by the famous airpower theorist Col. John Boyd.

gun is that the range at which the missile can be fired is considerably greater and that the aircraft does not have to be aimed as accurately as during a gun fight. Missiles of this type can easily be defeated by flares and a break at maximum g into the missile (turning away the hot exhaust) is effective as well.

The second generation IR missile has a far more sensitive (cooled) sensor. These sensors can track an aircraft from almost every aspect, including head-on. However, if the targeted pilot reduces the thrust in time, the engine and aircraft can be cooled enough to defeat the sensor. The consequence is of course that the aircraft decelerates or descends. This tactic can only be used for a short period during the head-on part of the engagement. When the two aircraft pass each other, a turning fight can develop that resembles the old fashioned attack. This newer type of missile has anti-flare ECCM features but yet flares and breaking into the missile can still be effective.

The third generation missiles have an IR imaging capability. These missiles cannot be defeated by reduced thrust and have an excellent all-aspect capability. The new generation IR missiles are often equipped with a movable nozzle as well. With the moveable nozzle the missile can turn fast at the initial powered phase of its flight. This type of missile is not easily out-manoeuvred and far less prone to be affected by flares.

The seeker head of the missile can be slaved to the radar, the helmet (if the aircraft is equipped with a helmet-mounted sight) or set parallel to the aircraft's axis. When the missile receives IR energy, the pilot will usually hear a tone in his headset. This is also an indication for the pilot to un-cage the missile so it can autonomously track the target. It is advantageous if the seeker has a high off-boresight capability. Instead of turning the aircraft to the threat, now the missile itself can be pointed to the threat. The sensor in the older IR missiles had a limited off-boresight capability (20-40 degrees), so the pilot often needed to point his aircraft to the threat as well. The newest type of IR missile has a much larger off-boresight angle (around 90 degrees). This angle is even larger than the off-boresight capability of the radar. Therefore the pilot needs a helmet-mounted sight to fully exploit this capability. [100]

Missile engagement (BVR)

One of the disadvantages of gun fights and IR missile attacks is that these types of attack can only take place in clear weather. (IR energy does not pass through clouds). Radar missiles do not have these limitations. Also, with radar missiles the enemy can be engaged much further away. This is of course due to the fact that radar has a greater detection range than an IR sensor.

There are two types of radar missiles: passive radar missiles, which sense the radar energy that is reflected from the target and active missiles which, at a certain range, start to use their own radar to track the target. There are several disadvantages to the passive radar missile. The attacking aircraft must illuminate the target with its radar until impact and during that time the attacker will come closer to the target. This illumination will most likely be noted by the target. Also, the attacker can only illuminate one target at a time. The sole advantage of these missiles is that they are relatively cheap. The active radar missile is much more expensive, but does not have these limitations. Its major limitation is that the radar transmitter and receiver in the missile are very small, and limited in power output. Therefore it is only

[100] A pilot should never launch a missile at the maximum off-boresight angle. During the initial acceleration of the missile, when it is not yet turning, the maximum seeker angle might be exceeded. This would lead to a lost missile.

capable of tracking a target at a limited range. That is why most active radar missiles possess a dual mode. During the initial phase of their flight they are guided by an initial guidance unit and only during the final part of their trajectory is the radar switched on. This initial guided mode can be supported by data link updates from the attacker.

A radar missile should preferably be employed head-on at the greatest possible distance from the target. But what is possible depends on many factors. First the attacker can optimise the missile capability by accelerating the aircraft and pointing the nose in the most favourable direction, so that the missile energy will not be depleted for unnecessary turning. After launch the attacker will slow down and/or climb to maintain distance to the target. The crucial point is whether the target is aware or unaware. If the target is aware he can deploy a variety of defensive actions that include EW-countermeasures and manoeuvres. All these actions can result in the loss of a missile and force the attacker to launch a follow-on missile or abort the attack. The capability to outmanoeuvre a radar missile much depends on range. If the radar missile is fired at its maximum range it takes only a small manoeuvre from the target to outperform it. Therefore the attacker has to take into account the expected capability of the target. If he expects a capable target, he has to move in closer before launch in order to increase the kill probability (but thereby also his own vulnerability). One defensive manoeuvre worth mentioning separately is called beaming. When a target moves on a course perpendicular to the attacker, the radar return of the target has exactly the same Doppler as the ground and the radar will normally not be able to track the target against the massive ground clutter.

If the targeted pilot plays it right, he can defeat the radar missile attacks and move in close enough to the attacker to deploy his own IR missiles or even force him into a gun fight. On the other hand, if the attacker plays it right, he will keep his own distance and work together with his formation members to maintain enough distance to the target.

Phases in aerial engagement

Detection is the first phase in every aerial engagement. In most cases, ground radars or airborne early warning radars (AWACS) will be the first to detect a target. Their ranges against small targets can be in excess of 200 nm and well above the range of fighter radars. However, if these assets are not available, the fighter has to depend on his own radar for the detection of the target. There are several factors that can delay detection. Firstly the terrain: in mountainous terrain a target can use the valleys to fly outside the radar coverage of ground and airborne radar. Secondly, the target can be masked by jamming, chaff or possess stealth capabilities. Search radars have an excellent range but a very limited track capability. The basic limitations are a slow update rate and a low altitude discrimination. To get a firing solution, the target must be tracked with greater accuracy, this is done with the radar of the fighter aircraft.

Tracking is the next phase in the engagement. To track a target, six target characteristics need to be resolved: three for the position (latitude, longitude, altitude) and three to determine the aircraft vector (speed, heading and climb angle). The radar has to dwell on a target for some time to achieve accurate tracking. Consequently the radar radiates a lot of energy in the direction of the target at a reasonably short interval. The consequences are that the tracked target will generally become aware that he is being tracked, because the radar warning receiver will detect the tracking. Also, the attacking aircraft's radar has less time available to search for other targets. Consequently the remaining search volume will become smaller if more targets are tracked. Basically the pilot has a choice: tracking few targets with a high accuracy but with the possibility that the adversary notices being tracked or tracking with a lower accuracy but with a smaller detection chance.

The third phase is identification. To identify a friend is relatively easy; to positively identify an enemy is much harder. To identify a friend, there are several options. First, aircraft belonging to one's own coalition should have proper codes in their radar transponders. If the fighter is equipped with an interrogator he can read this code. If he does not have an interrogator, AWACS or ground radar can tell him. Second, own aircraft may be part of the Link-16 network and show up as friendly on the Link-16 data. To positively identify an enemy, there are three possibilities: it is possible that the target has been tracked from the moment it took off from an airbase in enemy territory and therefore is already classified as an enemy. Secondly, an enemy aircraft can be distinguished by its electronic emissions or the Doppler characteristics of the radar return. The radar return may have specific Doppler characteristics that are caused by the rotational speed of the engine. Not all fighters are capable of carrying out this electronic identification. Thirdly, a visual identification may be used.

Proper identification is difficult, witnessed from the fact that there are cases of fratricide during every conflict. Pilots can make errors selecting the right transponder code, equipment may fail, the interrogator can misread the answer and not every aircraft is equipped with Link-16. Added to this, we also have aircraft from neutral countries and civil air traffic to take into account. This is why an airway structure is normally established over the operation area with specific routes for friendly aircraft.

The last phase of the engagement is the actual attack. If a BVR-engagement is authorized, the pilot will preferably proceed with a BVR-engagement such as described above. However, if a visual identification is required, the pilot has to manoeuvre his aircraft in such a way that he has the optimum chance for an unseen entry. This implies that he has to move away from the radar search area of his opponent and move to the stern of the other aircraft. For a proper visual identification, an aircraft has to be within 2 miles of another aircraft. However, if the aircraft is equipped with a targeting pod this can be used to identify targets that are much further away.

Ground attack fixed-wing

A well known saying is that the task of the air force is to put bombs on a target. Consequently there are only two types of persons in the air force, those who put bombs on targets and those who support them. This saying, much disliked by pure fighter pilots, carries a lot of truth.

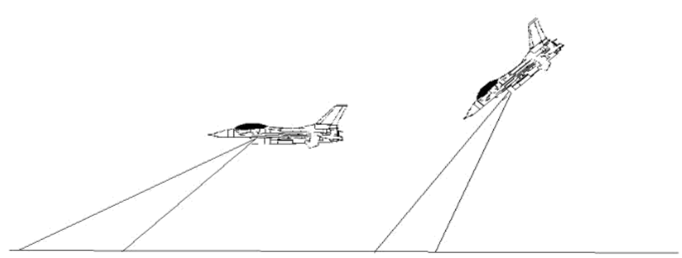
To accurately deliver a weapon on a target from a fast moving aircraft is difficult. To make an accurate delivery, the following parameters have to be known: accurate target position, aircraft velocity, position, altitude and dive angle, bomb characteristics and wind. If all these parameters are known it can be calculated where the bomb will fall in relation to the aircraft.

In the old days, there was no way to calculate all these parameters real time. The only way to make an accurate delivery was to manoeuvre the aircraft to a pre-determined position in relation to the target and deliver the bomb at a fixed altitude, speed and dive angle and aim with the help of a primitive weapon sight. This sight projected a symbol, called pipper, at a fixed depression angle from the longitudinal axis. This depression was calculated before the flight and if the aircraft was indeed at the exact speed, dive angle, altitude and distance from the target, the sight would need to be aimed exactly at the target. Bombing with this type of equipment required a lot of practice and skill.

With the advances in avionics, parameters like aircraft speed, altitude, and distance to the target could be measured continuously and used by fire control computers to calculate where the bomb would fall in relation to the target. The accuracy of bombing with these new avionics therefore improved considerably. Whereas the initial accuracy[101] with a fixed bomb sight was around 40 mils, the accuracy of the newer systems was around 10 mils. There are two basic modes for computerized deliveries. In the first mode (Continuous Calculated Impact Point, CCIP) the computer is continuously calculating where the bomb will hit the ground; this position will be displayed in his gun sight. The pilot will manoeuvre this calculated impact point to the target and release the bomb when they coincide. This mode can only be used when the pilot can see the target. In the second mode the computer has the target coordinates, (from radar data, targeting pod or pre-planned) and based on the aircraft's position, the computer will guide the pilot to the release point. Therefore this mode is called Continuous Calculated Release Point (CCRP). The computer will automatically release the weapon when that position is reached and the pilot will only have to consent to the release.

Because the accuracy is given as an angle, or more precisely a cone around the bomb fall line, it is easy to understand that bombing at low altitude is more accurate than at medium or high level. But the accuracy is also improved when increasing the dive angle. The dispersion is an angle around the bomb fall line: at low dive angles this angle projects a large ellipse on the ground but at steep dive angles the long side of the ellipse will decrease.

There is another advantage to dive bombing: because the aircraft is moving more slowly in relation to the target, the pilot's aim becomes more accurate. The only disadvantage of a steep dive angle is that the pilot may have to deliver from a higher position to prevent him from hitting the ground.



Diving decreases the dispersion area on the ground

With the development of guided bombs and missiles, accuracy has increased even further. Laser guided bombs can be delivered with a CEP of around 10 feet and GPS-guided munitions have an accuracy of around 30 ft. The accuracy of these weapons is not a function of altitude or dive angle. For laser-guided bombs it depends on the tracker accuracy of the laser designator and the quality of the bomb guidance mechanism. For the GPS-guided bomb, accuracy depends on the quality

101 Accuracy is normally expressed as circular error probable (CEP). This means that 50% of all weapons are within the CEP.

of the GPS-signal, the accuracy of the coordinates and the quality of the guidance system.

Delivery at low level

There are several reasons for a low-level delivery. The first is weather: clouds can obscure the target and flying below the cloud deck can solve that particular problem. The threat level is a second consideration: with a lot of small arms or fighters, the preference will be to fly medium level but when there is a considerable radar guided and modern SAM threat, medium level might be a dangerous place.

Flying fast at low level requires skill, practice and the undivided concentration of the pilot. To make an accurate weapon delivery at low level is difficult; the time that the pilot has to acquire the target, aim and deliver the weapon is very limited. To increase his chances of acquiring the target, he will normally perform a pitch-up manoeuvre starting approximately 2-3 nm from the target. During this pitch-up the pilot is constantly changing altitude and heading, making it harder for ground-based AAA to accurately aim at him[102].

Most low-level deliveries are visual deliveries. Very few air forces have practised targeting pod deliveries at low level. This type of delivery is extremely hard because the pilot only has a short period available in the pitch to make all targeting pod slews and selections. A way to circumnavigate this problem is to pitch earlier and higher and stay high until the bomb impacts; the drawback, of course, is that vulnerability increases.

Some aircraft have been especially developed to make low-level bomb deliveries in all weather conditions. Typical examples are the F-111 and the Tornado. These deliveries are normally done in level flight and the target is acquired with the radar or the bombing is on fixed coordinates.

In low-level deliveries, special care should be given to the escape procedure. Not only must the pilot avoid hitting the ground, he must also pull away fast enough to get enough distance from the exploding bomb. Therefore, the pilot will normally perform a three dimensional pull-out that takes him away from the ground but also away from the bomb trajectory. To increase the aircraft bomb separation, the bomb can be equipped with a retardation device. This feature is very useful during level deliveries with a very low cloud deck where the pilot has limited options to pull away from the bomb trajectory.

Toss bombing

There are several reasons for a pilot to keep a certain distance from the target. This can be because the target is very well defended, or to keep well away from the exploding weapon. The latter is of paramount importance when delivering nuclear weapons. A way to increase stand-off is to toss the bomb. At a certain distance from the target (3 to 4 nm) the aircraft is pulled up and the bomb is released at an angle of around 40 degrees. The bomb will follow a parabolic trajectory that will take almost a minute. If the bomb is equipped with retardation devices this could even be longer. This will give the attacker time to turn away and escape before the bomb explodes.

During this delivery the pilot will not be able to acquire the target, so he must navigate to the release point using other navigational landmarks. Naturally, this type of attack can only be used against a target with a known position.

---

102 A typical pitch-up manoeuvre can start with a 30 to 40 degree turn away from the target followed by a 10 to 20 degree climb; at a certain pull-down altitude, the aircraft will be banked 100-120 degrees and pulled towards the target. This latter segment is a typical example of a three dimensional movement.

Toss bombing is also an option for laser-guided bombs that are designated by ground observers. The aircraft is approaching low level, keeps his distance from the target and the forward air controller (FAC) will start to illuminate the target when he receives the call that the weapons are released. The high parabolic trajectory will give the laser-guided bomb more time to guide to the target and will increase its accuracy. However, one should not fire the laser too soon: this can cause the bomb to aim for the target too early, which can lead to undershooting it.

Bomb delivery medium level

Medium-altitude deliveries have been the method of choice during past conflicts. The reduced threat at this level, the increased range, improved quality of medium-altitude deliveries and the advance of precision-guided weapons have been instrumental in this.

There are several methods for delivering bombs medium level. The oldest method is visual delivery. Small targets are hard to recognize from medium level; a compounding problem is that the view directly below the pilot is normally obstructed, so he has to identify the target before it disappears below the aircraft at a range of around 4 nm. Close to the target the pilot will roll in and start a steep dive towards the target (between 30 and 45 degrees); during the descent the aircraft will rapidly accelerate and often the pilot has to reduce power to prevent over-speeding the bombs. This is also the moment at which the pilot has the best view of the target and he should now positively identify his desired impact point. After delivery, the pilot will perform a rapid climb-out to establish himself outside the range of small arms. Because of the large slant range, delivery accuracy is low; CEP is around 150 ft. A positive effect is that the impact velocity is high; this type of delivery is suitable when the bomb has to penetrate concrete structures.

Another way to deliver from medium altitudes is a delivery on fixed coordinates. If the target coordinates and the aircraft's position are known, the fire control computer will give the pilot steering guidance and will release the bomb at the correct moment. This type of delivery will normally be performed in level flight. This delivery is slightly less accurate than dive deliveries because the uncertainty of the aircraft's position adds to the delivery error. When GPS is used for navigation this error is reasonably small.

The best way to operate at medium level is with precision weapons. In clear weather, laser-guided weapons are the weapons of choice. They have the smallest CEP, and with a targeting pod the target can be positively identified. Under adverse weather conditions, a GPS-guided weapon is preferred. To deploy this weapon the target position must be known. If the aircraft is equipped with a SAR, the target can be identified on radar and from the radar picture the exact target coordinates can be calculated.

Targeting pod operation

Modern targeting pods have dual sensors, both in the IR and the visual spectrum. As explained in chapter 4 on sensor technology, the visual sensor has the best resolution but the IR sensor is usable during day and night operations. The targeting pod has several fields of view (FOV). The pilot will normally start with the wide FOV and move to the narrow FOV when he is closer to the target. The FOV is so narrow that a targeting pod is not suitable for searching a target with an unknown position. The pilot must have information from maps, satellite pictures or ground observers about the target. The resolution of the target image is dependent on range and FOV. When the pilot is close enough to the target he can locate the desired point of impact (DMPI); the pilot will than slew the targeting pod to that point and command the pod to track it. Tracking can be based on contrast, or on a hot (or cold) spot. When the pod is not tracking, it will point to an assumed target location and be stabilized by the INU. Targeting pods are able to track moving targets as well, but these targets should not pass under bridges etc., as this would spoil the contrast. Clouds can also spoil tracking and can be even more dangerous if the bomb is already on its way to the target. In that case the laser can not illuminate the target and the bomb will go astray. To counter that problem, a formation can use certain techniques to enable other team members to take over the illumination. During targeting pod operations the bomb is released using the CCRP method. The fire control computer knows the bomb fall time and will start illumination several seconds before impact to give final guidance to the weapon. Delivery of laser bombs has proven to be very accurate with a CEP in the 10 feet range.

Strafing

Gun employment from fixed-wing aircraft against ground targets is called strafing. It is one of the oldest ways of attacking ground targets. In fixed-wing aircraft the gun is aligned with its longitudinal axis. Sometimes it is aligned slightly upwards, which is advantageous in air combat. For ground attack this fixed position is a restraint. To point the gun, the aircraft must be pointed at the target, so a (slight) dive is necessary. The limited weapon range forces the aircraft to come close to the target (and ground). Consequently, the pilot has a very limited time to fire (often less than one second). He also has to break away in time to prevent hitting the ground. Strafing is inherently dangerous; the aircraft is close to the ground and also close to possible target defences. A typical method of performing a strafing attack is from a pitch-up attack. One advantage of a pitch-up attack is that the aircraft is constantly manoeuvring in three dimensions, which is an adequate way to fool AAA. The second advantage is that target acquisition is easier.

Ground attack rotary wing

The main characteristic of rotary-wing combat is that rotary wings, due to their lower speed and hover capability, can either loiter at altitude or fly nap of the earth and use terrain features to take cover. To further minimize exposure time, some attack helicopters are even fitted with sensors on top of the rotor head. The preferred employment starts with a search and identification from a covert hover position. Then the helicopter moves fast into a firing position, shoots and takes cover again. To further minimize exposure time, the guidance time of the helicopter must be short. This is possible with autonomously guided weapons, a turret-mounted gun and a weapon guided with a (rotor head) sensor from a covert position.

A specific advantage of some attack helicopters is the turret-mounted gun. Turning the gun, which is linked to the helmet of the front-seat pilot, is much faster than turning the aircraft. But there is another advantage. The helicopter does not have to fly in the threat direction, which decreases its vulnerability especially against Rocket Propelled Grenades (RPG)[103] and small arms. Another consequence of a turret-mounted gun is that you need an extra crew member. It would be too confusing for a pilot to fly in one direction and aim and shoot in another direction.

The disadvantage of nap-of-the-earth flying and hovering is that the helicopter is vulnerable to small arms fire. If the opposing forces are all mechanized, there is a fair chance that the helicopter crew will detect opposing armour in time and can

[103] It is much easier to aim small arms and RPG at a helicopter that is coming straight at you than at one flying at an angle.

outmanoeuvre and outshoot them. On the other hand, if the opponents are a mixed force of infantry and mechanized armour the risk increases. Individual infantry soldiers equipped with shoulder-launched SAM, RPG and rifles pose a severe threat to low-flying and particularly hovering helicopters.

If the threat consists mainly of small arms, for example in peace support operations, it may also be advantageous for helicopters to move to a higher altitude. This decreases his vulnerability and gives a better view of the ground. Another advantage is that the perceived sound level of the helicopter is much lower, which may enhance its capability to approach undetected.

Chapter 10- Electronic warfare

Introduction

With the invention of radio and radar, the ether has become a new area for warfare. The use of these new techniques increased the speed of command and control. All-weather guidance of aircraft became possible with radio beacons and radar increased the detection capability of enemy aircraft by several orders of magnitude. Not surprisingly, warring parties strive for 'ether supremacy' in the same way as they do for air supremacy. Therefore the objective of Electronic Warfare (EW) can be described as the ability to exploit the ether for own actions whilst preventing the enemy from doing the same.

The first EW event is said to have been the interception by the Japanese of radio signals from the Russian fleet during the Russo-Japanese War (1904). The first EW event in an air campaign was probably the British jamming of the German radio beacons that were used to guide German bombers during the Blitz[104]. It did not take long before EW had grown to major proportions in every air campaign. Chaff-laying aircraft, jammers, and attacks on enemy radars with anti-radiation missiles were essential to the campaigns in Southeast Asia to minimize the impact of North Vietnam's SAM systems on the USAF bomber force. Throughout, the war in the ether has been an ongoing battle, sometimes favouring the ground defender and sometimes the attacking aircraft. During the Cold War in the seventies and eighties, the development of advanced SAM systems was turning the tide in favour of ground-based air defence, but later developments in stealth technology have favoured the attacker.

With the present operating speed and altitude of airpower assets, radar detection and tracking is essential to effective interception. Because radar is the most important sensor, the majority of EW consists of countering radar or ensuring its operation despite any action taken against it. But EW is not all about radar; it also incorporates the interception and jamming of voice data communications, counter-actions against IR systems and direct energy weapons.

In this chapter the most common offensive and defensive principles of EW will be discussed. For sensor principles we refer to chapter 4.

Definitions[105]

Electronic Warfare (EW): Any military action involving the use of electromagnetic and directed energy to control the electromagnetic spectrum or to attack the enemy.

Electronic Attack (EA): That division of EW involving the use of electromagnetic and directed energy to attack personnel, facilities or equipment with the intent of degrading, neutralizing or destroying enemy combat capability.

Electronic protection (EP): That division of EW involving actions taken to protect personnel, facilities and equipment from any effects of friendly or enemy employment of EW that degrades, neutralizes or destroys friendly combat capability.

---

[104] A popular belief is that the British were 'bending the German beams', but in fact they were just using television transmitters to jam the signals.

[105] Based on CJCS Memorandum of Policy (MOP) 6 and CJCS MOP 30.

EW Support (ES): That division of EW involving actions tasked by or under direct control of an operational commander to search for, intercept, identify and locate sources of intentional and unintentional radiated EM energy for the purpose of immediate threat recognition.

The offensive problem

For an offensive force penetrating the enemy air defence system, the problem is surviving the enemy AAA, SAM and fighter threat. Its survival chances can be increased in many ways. The best way to visualize this is to consider the 'kill chain'. In order to destroy a target it has to be detected, identified, tracked, and a weapon must be guided accurately towards it. Both EA and EP measures can be used in every step of the kill chain and preferably airpower assets must have the capability to interfere in all elements of this kill chain. The subsequent steps will be discussed now in more detail.

Delay detection: Several techniques are available to delay detection. The most common are jamming by Stand-off Jammers (SoJ), chaff corridors, and stealth.

Delay identification: Flying at night, jamming identification radars, stealth and camouflage.

Break the tracking: Repeater jammers, lock-braking chaff and Suppression of Enemy Air Defence (SEAD) with anti-radiation missiles.

Create miss distance: Repeater jammers, lock-breaking chaff, (towed) decoys, jamming of the missile-guidance signals, manoeuvring and flares (only applicable against IR missiles).

In the next paragraph these techniques will be further explained.

Techniques for the offensive force

Stealth

Stealth is not just decreasing the radar cross section to a *slightly* lower value but to an *extremely low* value. This can be explained with the radar formula (chapter 4). The detection range decreases with the fourth root of the radar cross section, Therefore, a radar cross section reduction of 50% reduces the detection range by only 16% to 84%. To achieve reasonable stealth, for instance a 90% reduction in detection range, the radar cross section should be reduced to 0.01% of its original value. To achieve this massive reduction in radar cross section three techniques are used. Firstly, the basic shape reflects all EM energy away from the emitting radar. Secondly, the amount of reflected energy is minimized by using EM absorbing materials and thirdly the surface is seamless. Particular attention should also be given to the air intakes and nozzles, which can easily act as radar reflectors.

One problem with stealth is that it works across a limited number of frequency bands. For instance, a stealthy aircraft can be quite effective against modern X-band radar, but visible with an old fashioned VHF radar[106]. It is physically impossible to make an aircraft stealthy in all EM frequency bands. The fact that stealthy aircraft are visible proves that they are not stealthy in the visual band of the EM spectrum. To employ stealth aircraft in the most effective way, use should be made of clever

[106] For low frequency or long wavelength the energy of the airframe as a whole may be seen as an antenna, especially when λ is approximately 25% of the aircraft size.

routing through the radar belt to maintain an optimum altitude and distance from the opponent's radars.

Radar cross section reduction is the most important element in achieving stealth, but attention should also be given to reduction of IR emissions, especially from the hot exhaust. Furthermore, because of their visibility, stealth aircraft are often painted black and preferably used at night[107]. If the aircraft is operating at low level, attention should also be given to noise reduction.

While stealth is an important force multiplier, it is definitely not the answer to all survivability problems but, to quote the JSF programme director: "It sure makes all other problems a lot easier."

Stand-off jamming and escort jamming

Jamming increases the amount of radar clutter and attacking aircraft can hide in this clutter. A stand-off jammer (SoJ) is normally employed outside enemy territory and because radar has a directional antenna it can only jam a certain small sector. Although a SoJ is hiding the attack formation, its employment reveals that some activity is taking place and in addition, from what direction the threat is coming. When the threat approaches enemy radars, the reflected energy can exceed the jamming energy from the SoJ. At this point, which is called the burn-through range, the formation will appear as target on the enemy radar. Another limitation is that only the radars on the line from the jammer to the attacking force are being jammed but radars outside that sector may have a clear radar picture of the attackers. A solution to this problem is to make the jammer part of the formation. This is called escort jamming. The positive effect is that the burn-through range is much smaller and also that escorts can jam all radars looking at the force. A negative point is that you need more aircraft capable of flying over enemy territory.

Chaff corridors

Chaff corridors were deployed for the first time in World War II, when strips of aluminium foil were used to reflect radar energy. Attacking aircraft can fly above or in the corridor. What is true for SoJ also applies to chaff corridors. They may conceal the attacking aircraft but give away the attack direction and the fact that some action is imminent. It did not take long before the allies simultaneously employed fake corridors in order to lure the enemy into launching their night fighters. By the time those fighters had to return for fuel the actual attack force was launched. Chaff has a very high resistance enabling a slow descent, but because of its low speed it can be filtered out by coherent pulse Doppler radars, making them less effective.

Self-protection jamming

All modern fighters and bombers are equipped with self-protection jammers. The most common type of jammer is the repeater jammer. It retransmits the received energy of the tracking radar with some time delay or some frequency change or a combination of both. The tracking radar of a SAM or AAA tracks the target in azimuth, range and velocity. The jamming technique can be optimised against all three tracking functions.

To break the azimuth tracking, the scan frequency of the tracking radar antenna must be known. The jamming energy is then modulated at the radar's scanning frequency which drives the radar antenna away in a new direction, away from the

[107] If a simple fighter has visual contact with a stealth bomber it is relatively easy to shoot it down. The bomber is unable to outmanoeuvre the fighter and even if the radar missile and the IR missile can not track the aircraft the fighter is still able to shoot the bomber with his gun.

target. Trying to jam the azimuth tracking without knowing the scan frequency will be less effective.

To break the range tracking, the retransmitted pulse (which is stronger than the aircraft return) is more and more delayed, which moves the range gate to the wrong position (at a greater range). When the gate is at a sufficient distance from the actual range the jamming can stop and the lock should be broken.

To break the velocity gate tracking, the frequency of the received signal is slowly changed. If jamming is successful the radar will start to track the wrong Doppler filter.

Not all jammers are repeater jammers. Common other types are straight noise jammers and smart noise jammers. Straight noise jammers make it harder for a radar to track the target in range, but the target can be tracked passively in azimuth. Smart noise is modulated in amplitude and frequency, which influences the range finding and the azimuth tracking.

Lock-breaking chaff

Lock-breaking chaff has a larger radar return than the aircraft and should give the tracking radar many false targets that can create a break lock. The chaff is normally dispensed at short intervals to enhance its effectiveness.[108] There are however some limitations to chaff. First the air resistance of chaff is very high, which decelerates the chaff to very slow speeds in a fraction of a second. Most radars, especially coherent pulse Doppler radars, eliminate non- moving targets and in doing so will cancel all chaff returns. A second consideration is the bloom time. It takes a limited amount of time for the chaff to bloom and reach a large radar cross section. Consequently the chaff return is some distance away from the actual aircraft. If the radar is tracking within a very small beam and in a small range gate, the chaff may be too far away from the aircraft to be effective.[109] Finally, the quantity available is always limited and it can therefore only be used for very short periods.

Decoys

Decoys have one big advantage when compared to chaff: they are moving and therefore present the tracking radar with a valid Doppler. If the radar cross section of a decoy is the same as that of the target there is no way for the tracking radar to discriminate between aircraft and decoy. There are two types of decoy. The most advanced type is the free-flying decoy that can either be launched ahead of the main force to attract SAM, or fly with the main force in order to hide the real aircraft. The second type is the towed decoy. The towed decoy should have a larger radar cross section than the target to lure the active radar into locking onto the decoy instead of the target. Towed decoys can be enhanced with some jamming capability as well. The advantage of a towed decoy over a free flying decoy is its price. Disadvantages of a towed decoy are that it limits the aircraft's manoeuvrability and that it is still reasonably close to the towing aircraft. Consequently, depending on the flight profile and the type of proximity fuse of the SAM the aircraft can still be hit.

Flares and DIRCM

A flare is to an IR tracker what chaff is to radar. It presents the IR tracker with a stronger IR source than the actual aircraft and should pull the IR tracker away from the aircraft. Flares are normally employed in a salvo at predetermined intervals

108 These intervals are normally based on field trials and the optimum interval can be different for each type of tracking radar.

109 If chaff is employed with the tracking radar beam, chances for a break lock are highest; the relative range of the target is not changing so chaff and aircraft have the same Doppler.

based on the (presumed) tracker characteristics. There are two types of flares. The standard flare burns with a bright red colour. The advantage is that its IR signature is large but the spectrum differs considerably from the actual aircraft. The newer, so called 'black flare' emits in a spectrum that better resembles the actual aircraft signature. As with chaff, the flare will decelerate quickly, which can result in such a large spacing that it is outside the IR seeker envelope.

If the dispensing of flares is done by the pilot, he must have some idea that the missile is incoming and preferably from what direction. However, with a visual scan the IR missile can easily be missed. An improvement is the Missile Approach Warning System (MAWS). Typically the MAWS will detect the (UV) energy of the missile launch; other MAWS use an active radar to detect incoming missiles. A further automation is possible by letting the MAWS decide when flares must be dispensed.

The number of flares available is always limited, so instead of dispensing flares it is advantageous to emit an IR jamming signal. The system designed to do this is called Directional Infrared Counter Measures (DIRCM). The system senses the approaching missile and blinds it with high-energy IR pulses.

SEAD

The system to suppress enemy air defence with missiles homing in on the EM energy of the SAM radar was developed during the Vietnam War. While the idea is simple, the execution of a SEAD mission is not simple at all. First, the SEAD aircraft needs a very accurate location of the radar emitter. A standard radar warning receiver will only present the pilot with an accurate bearing. So, to find the correct range to the target several bearings from the radar warning receiver must be combined[110]. Secondly, the SEAD aircraft must manoeuvre itself into a position close enough to provoke the radar into tracking, but far enough to prevent being shot at. When the Anti Radiation Missile (ARM) is launched, the intent is that it reaches its target before the SAM hits the SEAD aircraft. Consequently the speed of the ARM should be high which has led to the development of the High Speed Anti Radiation Missile (HARM). A third limitation to SEAD is that (H)ARM can be detected by the radar operators and give them the option to start appropriate counter measures.

Manoeuvring

Without a doubt manoeuvring is the oldest of all counter measures and it is still effective. As explained in chapter 3, a 9-g aircraft can outmanoeuvre a 30-g missile because of its smaller turn radius. Creating a small overshoot may give enough separation to be outside the influence of the warhead. On the other hand, the warhead of a missile is normally not very heavy, permitting enhancement of the missile's manoeuvrability and range.

Manoeuvring is also effective against Anti Aircraft Artillery (AAA). Artillery shells are fired to the position where the aircraft is expected to be after the time of flight of the shell. But manoeuvring will cause a discrepancy with the actual position of the aircraft. This discrepancy may be too small at very short ranges but at larger ranges manoeuvring is usually effective.

110 It is possible to combine directions from different altitudes or from different horizontal positions.

Execution of offensive techniques

SEAD and SoJ aircraft are scarce resources. Therefore these assets will be controlled at the highest level and dedicated to the riskiest and most important of missions. Typically a COMAO in the initial phase of an OCA campaign needs these assets to maximize survivability. The exception are stealth aircraft capable of operating without these resources. In fact, they may operate better without these resources because both SoJ and SEAD will alert the opponent to an incoming attack that might otherwise have been unnoticed.

Most combat aircraft are equipped with chaff, flares and repeater jammers. The individual aircraft can use this equipment in any combination. Tests have revealed that the best results in track-breaking and miss-distance creation are achieved by a combination of techniques. For example: lock-breaking chaff alone does not necessarily create a break lock nor do certain repeater jamming techniques, but if both techniques are used simultaneously and supported by manoeuvring, the effectiveness will increase considerably.

Some techniques are automatic and programmed into the EW management computer. Automatic response is required when a delay in counter actions can be deadly. Typically, an IR missile warning system will automatically release flares when it detects a missile launch. Also self-protection jammers will start jamming when detecting a lock. An advanced system will also select the proper jamming technique against each detected threat.

The defensive problem

The defensive problem is the mirror image of the offensive problem. Defensive techniques must ensure:

- Early detection and identification despite jamming;
- Tracking of a chaff-dispensing, jamming and manoeuvring aircraft;
- Accurate guidance of the weapon to these targets.

Several options are available to enhance the operation including: frequency agility, emission control procedures, integrated sensors, anti-chaff techniques and techniques against repeater jammers. The most common techniques will be discussed below.

Defensive techniques

Emission control

Emissions imply giving away the position and (some) characteristics of the radar. Therefore a good EW game plan starts with a rigid emission control plan. In this plan the frequencies and modes to be used during what period should be stipulated. If, for example, certain frequencies are only used during actual engagements, there is a fair chance that the opponent will not have been able to develop appropriate jamming techniques. The opponent may want to regularly move his radars in order to minimise the success of planned attacks[111]. The bottom line is that a constantly emitting system from a fixed location is prone to attack.

111 Typically a cruise missile is an effective weapon against a radar whose position is known.

Frequency agility

Frequency agility is the capacity to change the frequency of a system fast. If the system can rapidly change the frequency and over a wide band, chances are that the jamming is less effective. For instance, it may only be effective over a relatively small portion of the band. An ever better way is to have the installation work in different frequency bands at the same time.[112] The best frequency agility is achieved when a system can operate in well-separated bands, for instance with a radar and a laser tracker.

Passive detection

Jamming makes it much harder for radar to discriminate the target return from the radar noise. But the antenna can also passively scan and find the azimuth of the emitting aircraft. For accurate tracking and firing a missile, some range information is needed. If two passive tracking radars track the same emitting aircraft, a rough range resolution can be calculated from the cross bearing. Not only a radar but also a missile can use this passive mode, often called 'Home on Jam' to fly exactly to the target.

Electronic features

There is a variety of electronic techniques that can be employed against chaff, flares and jamming. Most of these techniques involve a special type of processing to separate the target from chaff, noise and flares. Normally these techniques are programmed into the software of the radar or missile computer. Its exact characteristics are each system designer‘s well kept secrets. In this paragraph we will only mention some of the used principles as illustrations.

A pulse Doppler radar, which measures angle, range and target closing velocity, will normally cancel out all chaff. Furthermore, range deception (range gate stealer) may not correlate with the measured closing velocity enabling the cancellation of the range gate stealer.

Pulse edge tracking has a very small range cell and only tracks the first part of the return. Because the repeater jammer generated signal is delayed by a small fraction, it will be ignored by the pulse edge tracker.

When an IR tracker, thanks to its spectral characteristic, senses a flare, this can generate a counter action of the sensor, for instance closing its aperture briefly until the flare is outside the field of view of the tracker.

Mono pulse

Mono pulse is a tracking technique that uses a four-quadrant radar antenna with separate receivers to process the signal from four radar quadrants simultaneously. By comparing the signal strength from the four quadrants, the exact angle of the target is measured on a pulse-by-pulse basis. This type of tracking is almost immune to any type of jamming designed to break the angle track of a radar.

Multiple sensors

If sensors are brought into a network the resilience of the air defence network increases dramatically. The network quality will only degrade if multiple sensors in the same area are jammed. So if a single radar is shut down, for instance in the case of a SEAD aircraft attack, this may not affect the capability to launch a weapon. The only requirement is that there are multiple sensors that can accurately track the target.

112 For example the Goal Keeper radar can work in K and X band at the same time.

Bi-static radar

Bi-static radar is a new development that uses a transmitter and receiver that are not co-located. This technique may turn out to be an effective countermeasure against stealth aircraft. The reason for its success is the fact that the stealth aircraft minimizes the radar return in the direction of the transmitter, but reflections in other directions can be substantial. To design and produce effective bi-static radars is however considered complex and costly.

The execution of defensive techniques

The execution of the proper defensive techniques is determined by several parameters. First of all, the whole EW infrastructure must be designed to operate in a hostile environment and be redundant, frequency-agile and integrated into a network. Secondly, there must be a proper EMCON plan and thirdly, the operators must be trained to operate in a hostile EW environment.

The importance of a proper doctrine can be illustrated by comparing operation Desert Storm with operation Allied Force. During Desert Storm, most SEAD was effective against the Iraqi air defence but during Allied Force the Serb Air defence system operators used effective shut-down procedures against SEAD aircraft and the GBAD systems remained operational until the end of the campaign.

Electronic warfare support

EW is only effective if it based on knowledge about enemy systems. This includes locations, frequencies, operating modes, tracking mechanisms and electronic counter measures. Much of this information must be collected in peace time. This will give ample time to design effective counter measures against enemy systems.

Parameters of enemy radar systems can be collected by acquiring those systems, performing some reverse engineering and testing them in an operational environment. Another method is listening to EM emissions and identifying and locating the source. EM emissions may be detected by satellites or by special aircraft such as the KC-135R (Rivet Joint). Because a well-trained operator will move the system on a regular basis (if it is mobile), it is important that this information collection is continued during actual operations.

Most airborne platforms today are equipped with sensors that locate and identify enemy radars. However, most of these sensors lack the ability to accurately locate transmitters and are only used to start proper countermeasures. More modern platforms have sensors are able to locate emitters accurately. If more accurate sensors are available and if the available data is shared on a network, an up-to-date and accurate electronic order of battle may come within reach, significantly enhancing existing operational capabilities.

Chapter 11 – Information superiority and airpower

Introduction

War is about dominance. The pilot wants to dominate the fight, the air force commander wants to dominate the air, the ground commander wants dominance on the ground and the sea commander wants dominance at sea. Whereas fighters, generals and political leadership work within the information environment they all need information dominance, or at least information superiority. Without information superiority it is unlikely that power can be exerted adequately. However, information superiority is not created easily. After all, the quantity of information required is large and large quantities of information are often difficult to disseminate. Furthermore, certain information is hard to collect. Information superiority also requires a continuous effort, starting in peace time and expanding during times of tension and war. Finally, information is not only a prerequisite for other operations but is also a war domain in itself with its own types of operations. Hence the existence of information operations in which information is used as a weapon.

The American Joint Publication 3-13 states the following about information operations[113]:

> *"Information operations are integral to the successful execution of military operations. A key goal of IO is to achieve and maintain information superiority. Information superiority can be described as the operational advantage gained by the ability to collect, process, and disseminate an uninterrupted flow of information while exploiting or denying an adversary's ability to do the same. Information superiority provides a competitive advantage only when it is effectively translated into superior decisions."*

The above description encompasses a wide variety of activities carried out by several organizations. To provide some insight into this process, this chapter deals with the role of information superiority in warfare with a focus on the role of airpower.

Before we start discussing some details about how information superiority is achieved and maintained it is good to determine what information is all about and what general conditions information has to fulfil in order to serve its purpose in warfare. So, starting from the characteristics of (good) information, we have to define the military information requirement, i.e. what information we need "to win the war". Then we have to define, in more detail, what additional information requirements can be listed for air operations, and what role airpower can play in collecting the required information. Thereafter the principles of information operations are described.

113 Since the American military has, quite obviously, developed the most detailed doctrine regarding Information Operations, the part on IO of this chapter is largely based on Joint Publication (JP) 3-13, released 13 February 2006.

The essence of information

Information and data

The literal, etymological definition of information, i.e. the original meaning of the word, derives from the Latin, *informare*, which means "to *give form* to something". So it refers to the act of *informing*, or giving shape to the mind. Informing therefore conveys the sense "to tell (one) of something." Thus, information always needs to be *about* something. That is to say that information is meaningful. Here the difference with data emerges.

Data and information are often used interchangeably, but are actually very different. Data is a set of unrelated information, and as such is of no use until it is properly evaluated. Upon evaluation, once there is some significant relation between data, and they show some relevance for the operation, they are converted into information. Now these same data can be used for different purposes. Thus, unless the data convey some information, they are not useful.

Experience, knowledge and decision-making

In essence, we use information to give shape, understanding and meaning to our experiences and it functions to communicate and recall those experiences. Consequently, one can safely claim that information is ***at the core of our personal and collective perception of reality***. Essential for information is that it can be recognised, interpreted and has a level of novelty. Integrating the novel elements of information creates knowledge. Thus, information is the building block underlying the process of knowledge construction.

Knowledge is a fluid mix of framed experiences, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences. It originates and is applied in the mind of those who know. Thus, knowledge provides the personal capability to act based on new experiences. In other words, it has a direct impact on our decision-making processes. This is why we preserve and retrieve information and also why we send information to each other.

Message

A message is information sent with a particular purpose. In the modern usage of the word, information often even means to give form to a message by moulding it into a shape or pattern that can be communicated. In other words, a message is information materialized. Information is a message that can induce a change in the thinking and acting of the receiver, particularly in clearing his vagueness and his doubt. To that end information needs to be a message received and understood. Information communicated by means of a message from a sender, can only be considered as such when the receiver is capable of understanding it. It assumes the existence of a common language understood by both sender and receiver. Information channels, i.e. channels to convey information in a message, are *sensory channels* (like sight, hearing, sense, smell and taste), *mechanical channels* (like pen and paper or a typewriter), and *electronic channels* (like radar, radio, telephone or computer networks).

The information environment

It has been stated that everyone, from fighter to general, works within the information environment. Consequently, the information environment consists of actors but also of systems or resources that collect, process, disseminate, or act on information. The actors include leaders, decision makers, individuals and organizations. Resources include the materials and systems employed to collect, analyse, apply, or disseminate information. The information environment is where humans and automated systems observe, orient, decide and act upon information, and is therefore the principal environment of decision making. Even though the information environment is considered distinct, it resides within each of the four military domains (air, land, sea, and space). The information environment is made up of three interrelated dimensions: physical, informational, and cognitive.

The physical dimension is composed of the Command and Control (C^2) systems, and supporting infrastructures that enable individuals and organizations to conduct operations across the air, land, sea, and space domains. It is also the dimension where physical platforms and the communications networks that connect them reside. This includes the means of transmission, infrastructure, technologies, groups and populations. The elements of this dimension are the easiest to measure when compared to those of other dimensions, and consequently, combat power has traditionally been measured primarily in this dimension.

The informational dimension is where information is collected, processed, stored, disseminated, displayed, and protected. It is the dimension where the C^2 of modern military forces are communicated, and where the commander's intent is conveyed. It consists of the content and flow of information. Consequently, it is the informational dimension that must be protected.

The cognitive dimension encompasses the mind of the decision maker and the target audience (TA). This is the dimension in which people think, perceive, visualize and decide. It is the most important of the three dimensions. This dimension is also affected by a commander's orders, training and other personal motivations. Battles and campaigns can be lost in the cognitive dimension. Factors such as leadership, morale, unit cohesion, emotion, state of mind, level of training, experience, situational awareness, as well as public opinion, perceptions, media, public information and rumours influence this dimension.

Bottom line

The undisturbed exchange of correct and comprehensible information ensures the dissemination of proper experience and knowledge necessary for adequate decision making. Whereas adequate decision making is at the heart of successful warfare, one of the main objectives of warfare must be to ensure the clearness, correctness and comprehensibility of information exchange, whether in the physical, the informational, or the cognitive dimension.

General information requirements

Quality of information

There are general criteria that define the quality of information relative to its purpose, which are depicted in the box below.

- ACCURACY; Information that conveys the true situation
- RELEVANCE; Information that applies to the mission, task, or situation at hand
- TIMELINESS; Information that is available in time to make decisions
- USABILITY; Information that is in common, easily understood formats and displays
- COMPLETENESS; Information that provides the decision maker with all necessary data
- BREVITY; Information that has only the level of detail required
- SECURITY; Information that has been afforded adequate protection where

Strategic questions

A good appreciation of the information requirement can perhaps best be gained by formulating some strategic questions for the conduct of war, such as:

What are the enemy's capabilities? This could be answered for example with the help of human espionage, with spy planes, by acquiring and testing enemy systems etc. Most of this information is normally gathered continuously during times of peace, tension and war.

Where is the enemy? This information becomes more important during times of tension and is essential during times of war. Reconnaissance and surveillance systems are used to answer this question.

What is the enemy's intent? This is a harder question to answer! Trying to answer this question is the job of spies, diplomats and political analysts.

What is necessary to compel or deter the enemy? This is probably the most difficult question to answer. It requires in-depth knowledge of the enemy, his interest, his weaknesses and his cultural background.

How can I undermine the popular support of the adversary's leadership? This information is vital for the execution of psychological operations (PSYOP). This knowledge can be gathered by studying the values, interests and cultural background of the adversary's population.

What has changed?

When we look at these questions we can argue that basics have not changed since the days of Sun Tzu. This may be correct, but even though the basic questions are the same, two major factors have influenced the position of information in war: airpower and (information) technology.

First of all, airpower has become a new medium for the generation of information, but it is also a consumer of information. In fact, airpower is probably far more dependent on information than other services. A land unit may operate quite independently under certain conditions and even gather its own information. But a pilot always needs a massive quantity of information before he can start his flight and he is seldom able to acquire that independently. Below, these issues are discussed in more detail.

Information technology

Technology and especially the increase in communication and computer capability (i.e. ICT), have made it possible to store, analyse and disseminate a tremendous quantity of data almost real time. These network capabilities have even enabled a new way of warfare: network-enabled capability, which will be covered in more detail in chapter 19 about Command and Control. Technology has also created new vulnerabilities, of which information overload is not the least important.

Advancements in technology have enabled information to be collected, processed, stored, disseminated, displayed, and protected outside the cognitive process in quantities and at speeds that were previously unimaginable. While technology makes great quantities of information available to audiences worldwide, the human perception provides the context in which to translate data into information and knowledge.

The finite amount of time and resources available to obtain information must be considered. Whether decisions are made cognitively or pre-programmed into automated systems, the limited time and resources needed to improve the quality of available information leaves decision making subject to manipulation. Additionally, there are real costs associated with obtaining quality information-that is, information well-suited to its purpose - namely the costs of acquiring, processing, storing, transporting and distributing information.

The position of airpower

Airpower as supplier of information

Airpower has changed the way information is collected. Airborne sensors have the advantage of being able to view a wide area. On top of that, some airpower assets like satellites are almost immune to the opponent's actions. But apart from reconnaissance and surveillance, which is described in more detail in chapter 17 on Support Operations, airpower can also be employed in information warfare. Typical examples are:

- Possible attack on information structure (can be done in counter air, but also in strategic operations, but it might also be done by jamming enemy emissions).
- As platform for radio transmission (PSYOP) or delivering pamphlets.

The special information needs of airpower

Aircraft operate over a wide area that is not under direct observation. Furthermore they have to cooperate with other aircraft and control agencies that are not collocated on the same base. This requires coordination and information. The most important requirements are:

Target locations. Aircraft have the ability to search for their targets, but it is more efficient if an aircraft is directed to a specific target. For air-to-air operations a redundant and extensive surveillance system is used comprising linked ground and airborne surveillance systems. For ground targets the target description and coordinates are needed. For stationary targets information from satellites and reconnaissance aircraft can be used. For information about mobile ground targets surveillance systems like JSTARS, ground observers and Recce aircraft/UAV are required.

Weather info. The operation of aircraft is limited by weather, so the pilot needs forecasted and actual weather information about the target area, his home base, the en-route portion and possible diversion airbases. The level of detail required for air operations is higher than for land or sea operations. Pilots need to know visibility, ceiling, winds at altitude, the altitude where contrails will be persistent, forecasted barometer pressure and all weather-related risks. Since the advances in aviation a large internationally connected network of meteorological services does generate this data.

Aeronautical information. Aircraft need airfields, radio beacons, approach aids etc. Furthermore some parts of the air structure can be restricted. Of course all this information is only available about friendly airspace, but a large portion of the mission is just that. For aviation information an international network has been developed that is employed by civil and military users.

Characteristics, performance and positions of enemy weapon systems. To adequately train against an opponent, the capabilities of enemy weapon systems must be known. Training takes time; therefore it is imperative that this information is acquired well ahead of actual combat. The required information encompasses data on range and lethality but also on sensor performance and characteristics. Knowing sensor characteristics is crucial for the programming of electronic warfare equipment (see also chapter 10). Gathering this type of data is the responsibility of the Intel organization. Real-time information on the actual position of enemy air defence units is also vital for mission planning. This data is normally acquired through pilot reports and (electronic) monitoring systems.

Positions of own assets on the ground. To prevent blue-on-blue fire it is essential to know the positions of own troops. Army positions and expected advances or retreats should therefore be collected and disseminated. Modern warfare, with its high tempo, makes its necessary to regularly update this data. To track own troops, an increasing number of land forces have started to use real-time linked systems. At present it is not possible to link these systems with airborne systems.

Orders, positions and times from cooperating air units. Normally air operations are carried out jointly by aircraft operating from different bases. This requires extensive coordination, for instance on the routing to be followed and the position of tankers. This coordination is done in the Air Task Order (ATO), which is distributed to all units involved.

Required organizations and infrastructure

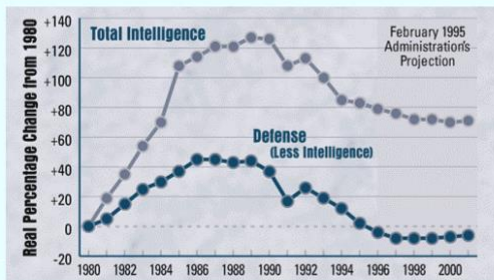
To supply airpower units with the required information requires many organizations. The most important are: the meteorological service, the air traffic control organization, the intelligence organization, the air defence surveillance and reporting system and the command and control organization. These are all large organizations that use far more manpower than needed for the operation of the aircraft. This is however typical for airpower, where most of the work (let us also not forget maintenance) is done on the ground, before an aircraft takes off. To give some understanding as to what is required in terms of personnel and financial resources, we shall give two examples:

Intelligence

One example of the continuous information-gathering effort is the intelligence organization. All larger powers have extensive intelligence information organizations, which collect data about many countries. (friends and foes!) Spy satellites are continuously collecting data, which generate a massive analysis workload for a variety of specialists. Radio and telephone traffic is also monitored.

In the USA, every service has its own intelligence organization apart from the Central Intelligence Organisation (CIA). As an illustration of the size: The annual budget of the NRO, NSA, DIA, NGA and CIA for fiscal year 1997 was estimated at $26.6 Billion.

The upward trend in intelligence spending – up to today's roughly $40 billion total – is clear from this chart produced by the House Permanent Select Committee on Intelligence and published by the Congressional Joint Inquiry into September 11.



Communication and information systems (CIS)

To connect all these organizations also requires an extensive information infrastructure. The infrastructure must be high speed, secure and redundant. During the Cold War most units operated from fixed locations and most of the communications could utilize land lines. The present emphasis on expeditionary operations requires communication systems that can work from remote locations. The only system capable of supporting remote units with high speed secure communications is satellite communication.

The CIS organization is responsible for establishing and maintaining these secure and high-speed communications. To get a feel for the magnitude of the organization and the costs involved some aspects of the ACCS system are described in the box below.

ACCS is the NATO Air Command and Control System, intended to replace the existing air command and control systems in Europe from 2009 onwards. The system was conceived in the 1980's to replace existing air defence systems such as NADGE (NATO Air Defence Ground Environment). The complexity of the system is apparent from the fact that it wasn't until 1992 that the North Atlantic Council agreed to the initial implementation of ACCS to a first level of operational capability in both static and deployable configurations.

When operational, the ACCS will provide a unified air command and control system, enabling NATO's European nations (including new Alliance members) to seamlessly manage all types of air operations over their territory and beyond. NATO members will be able to integrate their air traffic control, surveillance, air mission control, airspace management and force management functions. ACCS will incorporate the most modern technologies, and will make full use of up-to-date data link communications. Through its open architecture, the system is already evolving to meet emerging operational requirements such as those associated with theatre missile defence and it will be able to adapt to a changing operational environment such as network-centric warfare.

At the highest level (planning and tasking) – i.e. the level from which the air battle is run – NADGE comprises the Combined Air Operations Centre (CAOC) supporting the Joint Force Air Component Commander in exercising OPCON over his subordinate forces. At the execution level it provides facilities for aircraft control and production and dissemination of the Joint Environment Picture in control centres: the Air Control Centre (ACC), Recognized Air Picture (RAP) Production Centre (RPC) and Sensor Fusion Post (SFP) – using the generic acronym ARS. The programme comprises both static and deployable elements. Under separate funding, NATO will also procure deployable sensors for the deployable ACCS component (DAC). At present, NATO and the nations have requirements for 5 static and 2 deployable CAOCs, while 18 static and 2 deployable ARS are also planned.

A huge organization is involved in planning and managing the system construction. Overall responsibility for ACCS resides with the NATO C3 Organization, an entity comprising a governing body, the NATO C3 Board; a supporting substructure including the NATO C3 Representatives, and two agencies — NC3A (NATO Consultation, Command and Control Agency) and the NCSA (NATO Communications and Information Systems Services Agency). The programme itself is managed by NACMA (NATO Air Command and Control System Management Agency) with scientific support from NC3A, system and software engineering support form NPC (NATO Programming Centre), logistic support from NAMSA (NATO Maintenance and Supply Agency) and operational support from SHAPE (Supreme Headquarters Allied Powers Europe). The contract to build ACCS was awarded to the Air Command Systems International (ACSI) consortium in November 1999.

The initial NATO contract, worth at the time some US $500 Million, was signed in 1999. Completion of the whole ACCS programme, comprising the provision of hardware, software and communications, will cost NATO and the nations approximately € 1.5 Billion. Not included in this sum is the acquisition of sensors, pushing up the costs ever further.

Information operations

History indicates that the speed and accuracy of information available to military commanders is a significant factor in determining the outcome on the battlefield. It goes without saying that, given this importance of information in warfare, information can be used as a weapon both for offensive and defensive purposes.

Principles of information operations

Success in military operations depends on collecting and integrating essential information while denying it to the adversary. This is the realm of IO, which encompasses planning, coordination, and synchronization of the employment of current capabilities to deliberately affect or defend the information environment. In other words, IO enable the accuracy and timeliness of information required by military commanders by defending our systems from exploitation by adversaries. Furthermore, IO are used to deny adversaries access to their C^2 information and other supporting automated infrastructures.

On the other hand, adversaries are also increasingly exploring and testing IO actions such as asymmetric warfare, that can be used to thwart the military objectives that are heavily reliant on information systems. This requires our military to employ defensive technologies and utilize leading-edge tactics and procedures to prevent our forces and systems from being successfully attacked.

Information challenges

In modern military operations, commanders face a variety of information challenges:

- **Technical challenges** include establishing and maintaining connectivity, particularly in harsh and wide-spread locations.
- **Operational challenges** include the complexities of modern combat against adversaries with growing information capabilities.[114] The global information environment and its associated technologies are potentially available to everyone and as a result, military commanders face another challenge. Our adversaries now have the capability to pass on information, coordinate, exchange ideas, and synchronize their actions instantaneously.
- **Constantly changing content and tempo**. A continuum of long, medium, and short-term factors shapes the information environment for which military operations are planned and in which such operations are executed.
  - Long-term factors that may shape the information environment include the various ways by which humans organize, govern, interact as groups (culture, sociology, religion, etc.), are regionally influenced (stability, alliances, economic relationships, etc.) and are technologically advanced.
  - Medium-term factors may include the rise and fall of leaders, competition between groups over resources or goals, incorporation of specific technologies into information infrastructure; and the employment of resources by organizations in order to take advantage of information technology and infrastructure.

[114] For example, regardless of their size, adversaries, including terrorist groups, can counter our efforts through propaganda campaigns, or develop, purchase, or download from the Internet tools and techniques enabling them to attack our information and information systems, which may result in tangible impacts on our diplomatic, economic, or military efforts.

- Short-term factors can include weather; availability of finite resources to support or employ specific information technologies (IT), and the ability to extend/maintain sensors and portable information infrastructure to the specific location of distant military operations.

- **Span of control.** The pervasiveness of the information environment in human activity, combined with the speed and processing power of modern IT, both enhances and complicates military efforts to organize, train, equip, plan, and operate. Today, technology has opened the way to an ever-increasing span of control.
- **Conflicting messages**. With the free flow of information present in all theatres, such as television, phone, and Internet, conflicting messages can quickly emerge to defeat the intended effects. As a result, continuous synchronization and coordination between IO, public affairs, public diplomacy and our allies is imperative, and will help ensure that information themes employed during operations involving neutral or friendly populations remain consistent.
- **Law of armed conflict.** IO may involve complex legal and policy issues requiring careful review. Beyond strict compliance with legalities, it is policy that our military activities -both in the information environment and in the physical domains- respect societal values and fundamental human rights. Our forces, whether operating physically from bases or locations overseas or from within our boundaries or elsewhere, are required by law and policy to act in accordance with our law and the law of armed conflict (LOAC).

Adversaries in the information domain

Potential information adversaries come in many shapes: traditionally hostile countries who wish to gain information on our military capabilities and intentions; malicious hackers who wish to steal from or harm our government or military; terrorists and economic competitors, just to name a few. Potential adversarial information attack techniques are numerous. Some, particularly attacks by electronic means, can be prevented by the consistent application of encryption, firewalls and other network security techniques. Others are considerably more difficult to counter. Possible threat information techniques include, but are not limited to, deception, electronic attack, computer network attack, propaganda and psychological operations and supporting signals intelligence operations.

Information operations capabilities

IO coordinates and synchronizes the employment of five core capabilities in support of the combatant commander: Psychological Operations (PSYOP), Military Deception (MILDEC), Operations Security (OPSEC), Electronic Warfare (EW), and Computer Network Operations (CNO).

Of the five core IO capabilities, PSYOP, OPSEC, and MILDEC have played a major part in military operations for many centuries. In this modern age, they have been joined first by EW and most recently by CNO. Together these five capabilities, used in conjunction with supporting and related capabilities, provide a commander with the principal means of influencing an adversary and other target audiences by enabling his forces freedom of operation in the information environment.

Psychological Operations (PSYOP) are operations planned to convey selected truthful information and indicators to foreign audiences to influence their emotions, motives, objective reasoning, and ultimately, the behaviour of their governments, organizations, groups, and individuals.

The purpose of PSYOP is to induce or reinforce foreign attitudes and behaviour favourable to the originator's objectives. PSYOP are a vital part of the broad range of activities used to influence foreign audiences and are the only MOD operations authorized to influence foreign target audiences directly through the use of radio, print and other media. PSYOP personnel advise the supported commander on ways to capitalize on the psychological impacts of every aspect of force employment and to create a strategy for developing and planning the dissemination of specific PSYOP programmes, in order to achieve the overall campaign objectives.

Military Deception (MILDEC) is described as those actions executed to deliberately mislead the adversary's decision makers as to friendly military capabilities, intentions, and operations, thereby causing the adversary to take specific actions (or inactions) that will contribute to the accomplishment of the friendly forces' mission.

MILDEC and OPSEC are complementary activities. MILDEC seeks to encourage incorrect analysis, causing the adversary to arrive at specific false deductions, while OPSEC seeks to deny real information to an adversary, and prevent correct deduction of friendly plans. To be effective, a MILDEC operation must be susceptible to adversarial collection systems and "seen" as credible to the enemy commander and staff. A plausible approach to MILDEC planning is to employ a friendly course of action (COA) that can be executed by friendly forces and verified by adversary intelligence.[115] There are always competing priorities for the resources required for deception and those required for the real operation. For this reason, the deception plan should be developed concurrently with the real plan, to ensure proper resourcing of both. To encourage incorrect analysis by the adversary, it is usually more efficient and effective to provide a false purpose for the real activity than to create a false activity. OPSEC of the deception plan is at least as important as OPSEC of the real plan, since compromise of the deception may expose the real plan. This requirement for close-held planning while ensuring detailed coordination is the greatest challenge for MILDEC planners.[116]

Operations Security (OPSEC) is a process of identifying critical information and subsequently analysing friendly actions and other activities.
Its purpose is to identify what friendly information is necessary for the adversary to have sufficiently accurate knowledge of friendly forces and intentions, deny the adversary's decision makers critical information about friendly forces and intentions and cause the adversary's decision makers to misjudge the relevance of known critical friendly information because other information about friendly forces and intentions remain secure.[117]

OPSEC denies the adversary the information needed to correctly assess friendly capabilities and intentions. In particular, OPSEC complements MILDEC by denying the adversary information required to both assess a real plan and to disprove a deception plan. For those IO capabilities that exploit new opportunities

115 However, MILDEC planners must not fall into the trap of ascribing to the adversary particular attitudes, values, and reactions that "mirror image" likely friendly actions in the same situation, i.e., assuming that the adversary will respond or act in a particular manner based on how we ourselves would respond.

116 On joint staffs, MILDEC planning and oversight responsibility is normally organized as a staff deception element in the operations directorate of a joint staff (J-3).

117 On joint staffs, responsibilities for OPSEC are normally delegated to the J-3.

and vulnerabilities, such as EW and CNO, OPSEC is essential to ensure friendly capabilities are not compromised. The process of identifying essential elements of friendly information and taking measures to mask them from disclosure to adversaries is only one part of a defence-in-depth approach to securing friendly information. To be effective, other types of security must complement OPSEC. Examples of other types of security include physical security (PHYSEC), information assurance (IA) programmes, computer network defence (CND), and personnel programmes that screen personnel and limit authorized access.

Electronic Warfare (EW) refers to any military action involving the use of electromagnetic (EM) and directed energy to control the EM spectrum or to attack the adversary. The details of EW have already been covered in chapter 10.

Computer Network Operations (CNO), along with EW, is used to attack, deceive, degrade, disrupt, deny, exploit, and defend electronic information and infrastructure.

CNO stems from the increasing use of networked computers and supporting IT infrastructure systems by military and civilian organizations. CNO is one of the latest capabilities developed in support of military operations. For the purpose of military operations, CNO are divided into three aspects:

- Computer network attack (CNA) consists of actions taken through the use of computer networks to disrupt, deny, degrade, or destroy information resident in computers and computer networks, or the computers and networks themselves.
- Computer network defence (CND) involves actions taken through the use of computer networks to protect, monitor, analyse, detect, and respond to unauthorized activity within MOD information systems and computer networks. CND actions not only protect MOD systems from an external adversary but also from exploitation from within, and are now a necessary function in all military operations.
- Computer network exploitation (CNE) is enabling operations and intelligence collection capabilities conducted through the use of computer networks to gather data from target or adversary automated information systems or networks.[118]

The increasing reliance of unsophisticated military and terrorist groups on computers and computer networks to pass information to C^2 forces reinforces the importance of CNO in IO plans and activities. As computer capabilities and the range of their employment broaden, new vulnerabilities and opportunities will continue to develop. This offers opportunities to attack and exploit an adversary's computer system weaknesses and also requires us to identify and protect our own from similar attacks or exploitation.

Supporting capabilities

Capabilities supporting IO include information assurance (IA), physical security, physical attack, and counterintelligence (CI). These are either directly or indirectly involved in the information environment and contribute to effective IO. They should be integrated and coordinated with the core capabilities, but also serve other wider purposes.

118 Note that due to the continued expansion of wireless networking and the integration of computers and radio frequency communications, there will be operations and capabilities that blur the line between CNO and EW and that may require case-by-case determination when EW and CNO are assigned separate release authorities.

Information Assurance (IA) IA is defined as measures that protect and defend information and information systems by ensuring their availability, integrity, authenticity, confidentiality, and non-repudiation.

This includes providing for the restoration of information systems by incorporating protection, detection, and reaction capabilities. IA is necessary to gain and maintain information superiority. IA requires a defence-in-depth approach that integrates the capabilities of people, operations, and technology in order to establish multi-layer and multi-dimensional protection ensuring survivability and mission accomplishment. IA must assume that access can be gained to information and information systems from inside and outside MOD-controlled networks.[119]

Physical Security (PHYSEC) is that part of security concerned with physical measures designed to safeguard personnel, to prevent unauthorized access to equipment, installations, material, and documents, and to safeguard them against espionage, sabotage, damage and theft.

The PHYSEC process includes determining vulnerabilities to known threats, applying appropriate deterrents, control and denial safeguarding techniques and measures, and responding to changing conditions.

Just as IA protects friendly electronic information and information systems, PHYSEC protects physical facilities containing information and information systems worldwide. PHYSEC often contributes to OPSEC, particularly in the case of MILDEC, when compromise of the MILDEC activity could compromise the real plan.[120]

Physical Attack (PHYSAT) disrupts, damages, or destroys adversary targets through destructive power. PHYSAT can also be used to create or alter adversary perceptions or drive an adversary to use certain exploitable information systems.

The concept of attack is fundamental to military operations. PHYSAT can be employed in support of IO, to attack C^2 nodes so the enemy's ability to exercise C^2 as well as target audiences are influenced. IO capabilities, for example PSYOP, can be employed in support of PHYSAT to maximize the effect of the attack on the morale of an adversary. The integration and synchronization of attacks with IO through the targeting process is fundamental to creating the necessary synergy between IO and more traditional manoeuvre and strike operations. In order to achieve this integration, commanders must be able to define the effects they seek to achieve and staffs will incorporate these capabilities into the commanders' plan.

Physical attack and airpower. Due to the fast-paced conduct of air operations, it is particularly crucial that the planning and execution of both IO and air operations is conducted concurrently to produce the most effective targeting plan.

Counter-intelligence (CI) consists of information gathered and activities conducted to protect against espionage, other intelligence activities, sabotage, or assassinations conducted by or on behalf of foreign governments or elements thereof, foreign organizations, foreign persons, or international terrorist activities.[121]

CI procedures are a critical part of guarding friendly information and information systems. A robust security programme that integrates IA, PHYSEC, CI, and

119 In joint organizations, IA is a responsibility of the J-6.

120 IO plans may require significant PHYSEC resources and this requirement should be made clear to the J-3 as early as possible in the planning process.

121 The CI programmes in joint staffs are a responsibility of the CI and human intelligence staff element of the intelligence directorate (J-2).

OPSEC with risk management procedures offers the best chance to protect friendly information and information systems from adversary actions. CNO provide some of the tools needed to conduct CI operations. For the IO planner, CI analysis offers a view of the adversary's information-gathering methodology. From this, CI can develop the initial intelligence target opportunities that provide access to the adversary for MILDEC information, PSYOP products, and CNA/CNE actions.

Related capabilities

There are three military functions – PA, CIMIC, and DSPD – specified as related capabilities for IO. These capabilities make significant contributions to IO and must always be coordinated and integrated within the core and supporting IO capabilities. However, their primary purpose and rules under which they operate must not be compromised by IO. This requires additional care and consideration in the planning and conduct of IO. For this reason, the PA and CIMIC staffs in particular must work in close coordination with the IO planning staff.

Public Affairs (PA) are those public information, command information, and community relations activities directed towards both external and internal audiences with an interest in MOD.

PA is essential for information superiority, and credible PA operations are necessary to support the commander's mission and maintain essential public liaisons. PA's principal focus is to inform domestic and international audiences of joint operations to support combatant command public information needs.

Principles of public information shall be made fully and readily available, consistent with statutory requirements, unless its release is precluded by national security constraints or valid statutory mandates or exceptions. Rules concerning Freedom of Speech shall be supported in both letter and spirit. A free flow of general and military information shall be made available, without censorship or propaganda, to the men and women of the armed forces and their dependants. Information shall not be classified or otherwise withheld to protect the government from criticism or embarrassment. Information shall be withheld only when disclosure would adversely affect national security, or threaten the safety or privacy of the men and women of the armed forces.

The Ministry of Defence's (MOD's) obligation to provide the public with information on MOD major programmes may require detailed public affairs planning and coordination within the MOD and with other government agencies. The purpose of such activity is to expedite the flow of information to the public.

Civil-Military Cooperation (CIMIC) comprises the activities of a commander that establish, maintain, influence, or exploit relations between military forces, governmental and non-governmental civilian organizations and authorities, and the civilian population.

These actions are conducted across the range of military operations to address root causes of instability and assist in reconstruction after conflict or disaster, or may be conducted independently of other military operations to support our (national) security objectives. CIMIC can occur in friendly, neutral, or hostile operational areas to facilitate military operations and achieve our objectives. CIMIC may include performance by military forces of activities and functions that are normally the responsibility of local, regional, or national government. These activities may occur prior to, during, or subsequent to other military actions. CIMIC may be performed by designated civil affairs (CA), by other military forces, or by a combination of CA and other forces. Certain types of organizations are particularly suited to this mission and form the nucleus of CIMIC. These units are typically CA and PSYOP units. Others, such as, but not limited to, other Special Operations Forces (SOF), engineers, health service support and transportation personnel, military police and security forces, may act as enablers. Personnel skilled in the language and culture of the population is essential to CIMIC.

CIMIC can be particularly effective during peace time and pre-/post-combat operations, when other capabilities and actions may be constrained. Early consideration of the civil-military environment in which operations will take place is important. As with PA, the CIMIC staff also has an important role to play in the development of broader IO plans and objectives. As the accessibility of information to the widest public audiences increases, and as military operations are increasingly conducted in open environments, the importance of CIMIC to the achievement of IO objectives will grow. At the same time the direct involvement of CIMIC with core, supporting and related IO capabilities (for instance PSYOP, CNO, and CI) will also increase. CIMIC, by their nature, usually affect public perceptions in their immediate locale. Distribution of information about CIMIC efforts and results through PA and PSYOP can affect the perceptions of a broader audience and favourably influence key groups or individuals.

Conclusion

Information is at the heart of decision making and decision making is at the heart of warfare. Consequently, one of the main objectives of warfare must be to ensure the clearness, correctness and comprehensibility of information exchange, whether in the physical, the informational, or the cognitive dimension. This is a complex process, involving many actors, huge organizations, and a vast infrastructure. One of the main issues at stake is the actors' awareness of this complexity. Furthermore, the management of information necessary to obtain information superiority, should be at the centre of the commanders' attention. Airpower has an important role to play by offering its unique capability to accumulate information. As a result of the particular dependence of airpower on fast and reliable information, airmen have to place information superiority in the forefront of their activities.

Chapter 12 - Counter air

General

There are two forms of Counter Air (CA): Defensive (DCA) and Offensive (OCA). The aim of DCA is to keep your own airspace free of enemy airpower, while the aim of OCA is to destroy or incapacitate enemy airpower. The preferred result of OCA is the destruction of enemy airpower on the ground before it can be brought into action. The desired end state is a situation where one's own forces can be employed without restriction while the opponent's airpower is incapacitated. The following desired end states can be defined:

Favourable Air Situation is an air situation in which the extent of air effort applied by the enemy air forces is insufficient to prejudice the success of friendly land, sea or air operations. (JWP 0-01.1)

Air Superiority is that degree of dominance in the air battle of one force over another that permits the conduct of operations by the former and its related land, sea and air forces at a given time and place without prohibitive interference by the opposing force. (AAP 6)

Air Supremacy is that degree of air superiority wherein the opposing air force is incapable of effective interference. (AAP 6)

Defensive counter air

The problem

The root problem in DCA is lack of time and assets. The enemy has the initiative and can choose the time and place of the attack. While the attacker can concentrate resources in space and time, the defender has no choice but to spread resources over distance and time.

A key parameter is the distance between the opposing forces. When they are in close proximity to one another, almost every aircraft is within striking distance of the defender's territory. However, when the distance increases, only long-range assets (bombers, ballistic missiles and cruise missiles) are capable of reaching the defender's homeland. In cases where a distant attacker does not have these long-range assets, the need for DCA evaporates.

Depending on the availability of resources, DCA can be set up as a belt defence around the homeland, but if resources do not allow such widespread deployment, concentration around strategically important areas is also an option. The latter is a particularly logical choice for large countries with a limited number of people and resources. A third option constitutes a mix that takes area defence as its basis but with additional resources devoted to the defence of potential high-interest targets (point defence).

Defensive counter air assets

The defender can achieve DCA using Ground Based Air Defence (GBAD) and interceptor aircraft. GBAD can take the form of surface-to-air missile systems (SAM), anti-aircraft artillery (AAA) and shoulder-launched missiles (SLM). Interceptor aircraft can be employed in an airborne Combat Air Patrol (CAP) or they can be placed on Quick Reaction Alert (QRA) on the ground. The capabilities and limitations of these systems are described briefly below.

SAM. Modern SAM systems (e.g. Patriot, SA-10) can engage multiple targets at the same time and are able to engage targets at considerable range (30-40 nm) and altitudes to approximately 60,000 ft. A SAM is normally integrated into an air defence network but can also operate autonomously in the event of network damage. SAM can maintain a state of high readiness over a long period, as it only requires a small number of operators working in shifts. One disadvantage of SAM systems is that they need time (tens of minutes) to reload after all ready-to-fire missiles are depleted. While reloading, they are extremely vulnerable. Another drawback is that it takes considerable time (hours) to move a GBAD system to a new location. At present, GBAD systems are the only assets capable of engaging ballistic missiles.

AAA. AAA was the first GBAD system and is still widely employed. The advantage of AAA is that grenades are cheap and there is an almost unlimited supply. The disadvantage is that grenades are not guided, so accuracy at longer ranges diminishes fast, especially when deployed against manoeuvring targets. AAA is therefore best deployed against low-flying (< 10,000 ft) targets at short range (2-3 nm). When used for point defence, AAA often operates autonomously.

SLM. SLM are widely employed. These missiles are normally IR-guided, which gives them great precision. One limitation is that they are much slower than larger missile systems, restricting their range and limiting their capability against fast-flying targets. Another limitation is that target acquisition relies on visual observation and recognition by the operator, which is a very demanding process. Some systems are equipped with an IFF system to prevent blue-on-blue fire. These systems are best deployed against close-in (2-3 nm) slow-flying targets at low altitude (< 10,000 ft).

QRA. Interceptor aircraft on the ground can maintain a high state of readiness for a long time. When scrambled, they will take between 5 and 10 minutes to get airborne (depending on the alert state). Once scrambled, they will climb in afterburner to the assigned altitude and accelerate to intercept at supersonic speed. The range of a QRA aircraft is dependent on the reaction time available. If alerted well in advance, the scramble can be executed at military power, giving the interceptor a potential range of several hundred nautical miles. However, if the scramble is done in afterburner, fuel will be limited to approximately 20 minutes and the intercept range will be limited to approximately 150 nm. Most interceptor aircraft will carry 4-6 missiles. If the aircraft are scrambled as a two ship, the maximum number of targets that can be engaged is limited to 8-12. A not unimportant feature of interceptors is that they can not only engage but also identify and intervene without attacking other aircraft. These capabilities are extremely important in politically sensitive situations. An example of an F-16 scramble is depicted in the following table.

Event	Time	Total time	nm	Total nm	Fuel used	Fuel left
Alert	0	0	0	0	0	9334
Ground ops	10	10	0	0	200	9134
Take off accel .9	1	11	5	5	800	8334
Climb FL 360	2	13	13	18	1000	7334
Accel 1.5 mach	3	16	40	58	1580	5754
Intercept 1.5	4	20	57	115	2268	3486
Mil Decel	3	23	40	155	150	3336
Shadow	7	30	56	211	336	3000
Recover FL 360	20	50	-160	51	1800	1200
Decent	10	60	-50	1	200	1000
Landing	4	64	-1	0	150	850
Ground ops	10	74	0	0	100	750

CAP. A CAP generates a very high state of readiness over a considerable area and can also operate autonomously, which means it can serve as a fallback option if communications or sensors in the air defence network are damaged. In autonomous mode, the area is limited to the radar range of the aircraft. A CAP operates inside a Fighter Area of Responsibility (FAOR), the typical size of which is approximately 60 by 30 nm, in which the interceptors fly a racetrack pattern. The long leg of the racetrack will be towards the threat, to maximize the time of the radar scan in the threat direction. If an FAOR is manned by four fighters, they may employ a counter-rotating pattern to ensure that one element is always looking in the threat direction. If the CAP is manned by two aircraft, they may fly relatively slowly towards the threat (hot leg) and faster on the cold track, thus maximizing the time the aircraft are facing the threat. The biggest drawback of a CAP is that it consumes a large number of resources. The following example has been calculated by way of illustration.

If an FAOR is manned by two aircraft for 24 hours a day over an extended period, this represents 48 actual CAP hours a day. But how many CAP hours can one aircraft produce, taking into account the limitations imposed by servicing, repairs and scheduled maintenance? In this example we will start by excluding non-scheduled maintenance (repair) and make the following assumptions: time to fly to the CAP is 30 minutes, time on station is three hours and time to recover is also 30 minutes. After each flight, two hours of turnaround time are required. After 50 flying hours, one day (24 hours) of scheduled maintenance is required and after 200 hours, three weeks of scheduled maintenance are required. The result is depicted in the following table.

Cumulative Hours				
Day	Event	Flying	CAP	Maintenance
1	Flying	16	12	8
2	Flying	32	24	16
3	Flying	48	36	24
4	50 hrs maintenance	48	36	48
5	Flying	64	48	56
6	Flying	80	60	64
7	Flying	96	72	72
8	50 hrs maintenance	96	72	96
9	Flying	112	84	104
10	Flying	128	96	112
11	Flying	144	108	120
12	50 hrs maintenance	144	108	144
13	Flying	160	120	152
14	Flying	176	132	160
15	Flying	192	144	168
16	Start 200 hrs inspection	192	144	168
17	End inspection	192	144	672

The result over 37 days is that just over 16% of the time is spent in the CAP. This implies that for every aircraft in a CAP you need at least six available aircraft. If non-scheduled maintenance is also taken into account, the situation becomes worse. It is not unrealistic to assume that the time needed for non-scheduled maintenance is about 50% of that required for scheduled maintenance. In that case only 12% of the time is spent in a CAP and nine aircraft are required for every aircraft in a CAP. A two-ship CAP therefore requires 20 aircraft, more or less a complete squadron! This squadron would produce almost 1400 flying hours with each aircraft, which is about six times the number normally produced in peace time. Consequently, the manpower to produce these hours has to be multiplied by about six as well. It is evident that this cannot be achieved simply by working more overtime! Such a high level of intensity can only be maintained for a limited period of time.

CAP/QRA. It is possible to combine CAP and QRA; moreover the QRA can have different alert states. The lowest level of readiness is 5 to 15 minutes' response time on the ground. This alert state can be maintained over a considerable period. This can be stepped up to a one-minute alert holding at the runway but this can only be maintained for a maximum of a few hours. The third and highest level is to have the QRA holding airborne but not yet assigned to a specific FAOR or threat. This highest level of readiness bears a strong resemblance to a CAP, but it cannot operate autonomously over as wide an area. This option might be efficient when the required response time is high while there are not enough assets available to man each FAOR.

Spreading resources; an example

A good appreciation of the DCA problem can be gained by designing a DCA plan. Let's assume we have to form a 600 nm defensive belt using missile systems. If we want to prevent a situation in which the loss of a single system would blow a hole in our belt, we must ensure that the individual SAM systems are relatively close together. If the range of the SAM is 30 nm, then an interval of 30 nm constitutes a fair distance, as it ensures that every possible target can be engaged by two systems. This means that we need 20 systems to form a belt (or 21 starting at the edge). If each SAM system has 16 ready-to-fire missiles, the maximum engagement capability at any single point is 32 targets. Consequently, the belt is powerless to stop formations consisting of more than 32 aircraft.

If the same area is defended by aircraft on QRA, the situation is as follows. Taking a 150 nm intercept range as the starting point, you need two bases to cover the distance. If redundancy in airbases is required, it may be better to add at least one airbase. However, redundancy can also be achieved by having extra runways and extra aircraft at the same airbase. To achieve the same firepower as the missile belt, you need six aircraft with six missiles at every base. A total of 12 aircraft on QRA can produce the same initial fire power.

Defending the area with autonomous CAPs poses a problem. If it takes four aircraft to man an FAOR and the FAOR is 30 nm wide, 20 FAORs are required. This adds up to 80 aircraft airborne in a CAP, which requires a total of about 800 aircraft!

Analysis of the options

The previous example leads us to conclude that DCA can be done with the least number of assets if QRA is used, while autonomous CAP requires the most assets. But there are several other aspects to consider. To list a few:

Required response time. SAM and CAP have the shortest response and QRA the longest.

Intervention. This can only be carried out by aircraft.

Cost. Even if a complete SAM system is more expensive than an interceptor aircraft, the total system cost might be lower because SAMs do not need airfields, intensive maintenance and do not consume much fuel.

Suitability of terrain. The area you want to defend might not be suitable for SAM systems, due to an intensive ground threat, for example.

Deployment time. Aircraft can deploy quickly to establish a defensive perimeter but deploying a SAM system makes heavy demands on transportation.

How an air defence system is organized in a particular situation depends on all these variables. For a sudden emergency in an expected area, it is easier to deploy air assets. A few interceptor aircraft and some AWACS can quickly establish a reasonable deterrent. If the threat is there to stay, a missile belt could be more cost effective. QRA can then be used as "back stop" (behind the SAM belt) and for intervention. Autonomous CAPs should only be used as a gap filler, for instance when a SAM belt breaks down or during short periods of high tension.

Integrated air defence

The backbone of DCA is the integrated detection, identification and control network. The two most important reasons for an integrated air defence network are that resources can be used more economically, as explained above, and that this network is less vulnerable. The validity of the concept was proven in the Battle of Britain, which can be seen as the earliest example of network-centric warfare. An air defence network requires sensors, communication lines, computers and a centralized command centre. Let's consider the various assets.

Sensors. Early detection and identification increases the engagement options. This requires long-range sensors. The most suitable long-range sensors are airborne early warning radars, which are least hampered by radar horizon limitations. But accurate tracking relies on faster updates than those supplied by airborne early warning. It is therefore important that information from other sensors is also incorporated. This can be data from ground radars, SAM radars and even from fighters equipped with a Link-16 fighter back link. The main sensors used are radar for primary detection, secondary radar for identification and ELINT sensors to detect electronic emissions. Because all transmitting sensors are easy to locate and cannot be housed in protective shelters, they are extremely vulnerable to attack. The AWACS is capable of changing course to avoid the threat but the same cannot be said for land-based systems. A partial solution is to use deployable radars and to relocate them frequently. Another option is to shut down the sensor when under attack from SEAD aircraft, causing the SEAD missile to miss.

Computers are used to integrate the various sensor data into a single Recognized Air Picture (RAP), which is disseminated across the network. It requires a considerable amount of calculation to integrate the data of tens of sensors into one common picture. An RAP may contain several hundred different tracks, some of which represent fast-manoeuvring fighters whose directions are not predictable. Manoeuvring fighters are the hardest targets for a computer to correlate.

Communications are required in order to command the assets and to share the RAP. The most expedient way to exchange data is by means of a digital data link. Several types of data link are used. Link-11 is used between AWACS and land-based Control and Reporting Centres (CRC), while Link-16 is used between fighters, AWACS and CRC. In addition to data-inks, a large amount of voice communication is required to direct aircraft. Much of the data is transmitted through the ether. For the security and stability of the network, it is important that the communication equipment can handle the vast amount of data securely and is able to resist jamming.

The quality of the system

The quality of an air defence system can be expressed in penetration depth and ability to withstand rollback. Penetration depth is the distance an enemy aircraft intrudes into your airspace before it can be engaged and destroyed. Rollback results from the fact that each consecutive aircraft in an attack formation is engaged later and so has the opportunity to come closer to the interceptor or SAM site. Eventually, some attackers will come so close that they can destroy the SAM site or interceptor. Let's investigate this in detail.

Penetration depth. The penetration depth can be restricted by early detection and a quick response by one's own air defence assets. Early detection is hard if enemy airfields are close to the border, if the attacker uses stealth aircraft and if the aircraft fly at high velocity. The defender's reaction time for an interception is dependent on the speed of the interceptor and location of the airfield in relation to the threat. For SAM systems, the reaction time is primarily determined by location and alert state. Considering these parameters, one might conclude that it is advantageous to build SAM sites and air defence airfields close to the border. However, care should be taken that such locations are not open to easy attack by enemy ground forces.

If an AWACS flies a pattern that enables it to look 120 nm over the border, it could detect an aircraft flying at 0.8 Mach 15 minutes before it crosses the border. If we take the example of the F-16 scramble earlier in this chapter as a basis for our calculations, we see that the F-16 can cover approximately 150 nm in 15 minutes. This means that, as long as the airfield is within 150 nm of the point where the enemy crosses the border, the penetration depth will be zero. However, if the same F-16 was on a 15-minute alert, the two aircraft would meet approximately halfway between the border and the base. This gives a penetration depth of around 75 nm.

Rollback. To limit the effect of rollback, it is of primary importance that enemy aircraft are engaged at a great distance and that all engagements are terminated before the enemy can pose a threat to one's own assets. The second consideration is that the chance of success should be high or, in technical terms, the weapon's Pk (probability of kill) should be close to one. A third important parameter is that the SAM or interceptor should have a multi-track and multi-shot capability. The effect of multi-track can be illustrated by the following example:

A SAM system that can only engage one threat at a time with missiles that have an average velocity of 2.0 Mach and a maximum range of 40 nm, comes under attack from several successive aircraft flying at 0.8 Mach and able to launch a weapon at 5 nm. This gives the following result:

Target	Start
1	30.0 nm
2	21.4 nm
3	15.3 nm
4	10.9 nm
5	7.8 nm

A maximum of five targets can be engaged before the SAM site comes under threat. If the same SAM site had a multi-engagement capability, for instance six missiles at a time, the number of targets that could be engaged would be approximately 30.

Threats to the system

There are several threats to the air defence system. Firstly, sensors of all kinds are always exposed and can be detected on satellite pictures, by ELINT or by ground mapping radars. Once detected, sensors are vulnerable to stand-off weapons, cruise missiles and especially anti-radiation missiles. A possible defence against such threats could be AAA, which is optimised to counter incoming missiles.

The second threat relates to the vulnerability of radar. The whole air defence network depends mainly on radar; while stealth aircraft will not always be totally invisible, they could reduce the detection distance to such an extent that engagement is no longer feasible.

Decoys represent a third possible threat. It is not hard to make small cheap drones with a similar cross section to real aircraft with the objective of luring missiles away from the real targets.

Fourthly, as proven repeatedly since World War II, the air defence network is vulnerable to jamming.

Finally, even a modern air defence system with a high Pk and capable of engaging multiple targets simultaneously at great distances is still vulnerable to a mass attack. With the option of choosing the time and the place, the attacker always has a strategic advantage over the defender.

Offensive counter air

General

In a way the Offensive Counter Air (OCA) is the mirror image of DCA. The aim of OCA is to destroy enemy airpower by attacking it from above and in its own territory, preferably before it can be used. An OCA campaign is not only directed against enemy DCA capability but also against its offensive potential, attack aircraft and ballistic missile capability.

There are several historical examples of very effective OCA campaigns. One of them is the Six-Day War (1967), in which Israel destroyed most of Egypt's airpower on the ground during the first hours of the conflict. Desert Storm is another example in which most of the enemy's air assets were destroyed on the ground. The only cumbersome aspect of the offensive was the mission to search and destroy Iraq's ballistic missiles.

Timing OCA

OCA is often the main focus of operations during the first days of a war. The thinking behind this approach is that the sooner air supremacy is achieved the better. This is a logical strategy because air superiority will enhance the effectiveness of one's own airpower. An OCA campaign is offensive by nature and more effective when executed as part of a surprise attack. In the latter case enemy assets are still on the ground and easier to destroy. A negative side of this policy is that it is destabilizing.

There may be a few exceptions where OCA is not the initial priority. Firstly, if the enemy has no offensive assets that can be deployed against you and no effective defensive assets, OCA is a waste of effort. The air war over Afghanistan is a prime example of such a situation.

Secondly, priority should not be given to OCA if there is a serious threat to vital positions on the ground that cannot be held by the army. There is little advantage to winning air supremacy only to discover that your home base has been occupied by enemy forces.

Thirdly, there is no point in embarking on an initial OCA campaign if it is going to jeopardize the mission. For example, the operation to free the hostages during the Tehran embassy siege depended on complete surprise and would have been totally undermined by an initial OCA campaign.

Timing and targets in an OCA campaign

The timing of an OCA campaign is critical, just as timing in the campaign as a whole is essential. Because the effectiveness of the enemy air defence is heavily dependent on the quality of its network, it is logical to target the network first. The command and control facilities appear to be the most lucrative target. But command centres are often located deep underground or housed in an office building that completely disguises its military function. After all, the only assets the defender needs to bring together to form a command centre are space, people, computers and communications.

Sensors are more vulnerable targets and so it could be better to focus on them. Stationary sensors can be attacked using cruise missiles and stand-off missiles. But mobile targets are much harder to hit, especially airborne sensors, which are able to evade any incoming threat. Another drawback to this approach is that sensors are numerous and, when they operate in a network, the quality of the network as a whole will not normally be diminished by the loss of a few sensors.

Another option would be to attack the weapon systems, i.e. SAM systems and aircraft. Given the choice, destroying enemy aircraft should have priority. Aircraft can be used offensively and defensively, so destroying them can have a considerable impact on the outcome of the conflict. Enemy SAM systems only pose a threat to the attacker's airpower, so it is economical to limit the attacks on SAM systems to those that directly hamper attack operations. SAM systems can be combated in several ways:

- **Direct attack.** Using airpower to attack a system that is designed to fight aircraft is extremely high risk. The use of stand-off weapons can enhance the survivability of the attacking aircraft.
- **Suppression.** Firing anti-radiation missiles at a SAM site might destroy it, but a well-trained operator will shut down its sensor in time, so the SAM system will only be suppressed for a short period. This might provide enough time for the attacking force to pass the SAM system but continuing to suppress the enemy SAM systems over a long period requires considerable assets and is therefore not always an economical way to deploy airpower.
- **Stealth.** Stealth could be one of the most effective measures against ground-based air defence. If the number of stealth aircraft is limited, they can be used most effectively to destroy opposing SAM systems. However, if the force consists of 100% stealth aircraft, it might be wise to disregard most SAM systems. It is important to remember that stealth aircraft are still visible and on a clear day they can be spotted and tracked visually. Some

SAM systems are capable of visual tracking and laser range-finding; under these circumstances an attack on SAM systems might still be necessary.

- **Altitude**. Altitude is an effective way to negate much of the SAM threat. Small SAM systems are not effective above approximately 10,000 ft, while larger SAM systems are not always capable of detecting low-flying and fast-flying targets in time. During the Cold War, medium-level SAM systems were considered more dangerous than the small SAM threat and everybody was operating at low altitude. In the 1990s the focus shifted to medium altitude because the larger SAM systems were effectively neutralized by SEAD and stealth assets, while the low threat was still present.

Fighting aircraft. There are various ways to combat aircraft: you can fight them in the air, destroy them on the ground or attack the facilities the enemy needs to sustain them. All these strategies have been employed in the past and each has its advantages and disadvantages. A short summary:

Facilities. Some facilities, such as runways, are easy to find and hit. Another advantage of targeting runways is that they are limited in number. The disadvantage is that runways are easy to repair and have to be attacked repeatedly. Other facilities, such as refuelling points and munitions stores, are harder to repair but are also harder to find.

Aircraft on the ground. Aircraft in the open are easy targets for air attack. One strafing burst can be sufficient to put them permanently out of action. Aircraft protected by a shelter, on the other hand, can only be taken out successfully using deep penetration precision bombs. A further disadvantage of this strategy is that multiple target elements require more missions. It is impossible to tell whether or not a shelter is empty and destroying empty shelters will not affect the outcome of a battle. Mock-up aircraft are quite often mistaken for the real thing: in the past a great many bombs have been wasted on such decoys.

Aircraft in the air. Engaging aircraft in the air is normally the least preferred option. However, some advantages present themselves. First of all, only serviceable aircraft that pose a threat need be engaged. Secondly, if enough fighters can be assigned to the attack, it might be possible to outnumber the defender and change the odds in favour of the attacker. A disadvantage is that the defender will only direct his fighter force against an attacker, so the enemy's ground-attack aircraft remain out of range.

COMAO

OCA missions are normally executed as a complete package or a Composite Air Operation (COMAO). Concentrating the aircraft in such an operation can saturate the enemy's air defence system. This approach also allows all support assets (jammers and SEAD aircraft) to be deployed more economically along the ingress and egress route of the formation, thereby increasing survivability. The drawback of operating in a COMAO is that it requires extensive coordination in terms of planning and execution and so the mission generation will normally be slightly lower compared to the individual tasking of missions. The COMAO can include the following elements:

- Fighter protection with sweep and escort fighters charged with the task of engaging enemy air defence fighters.
- SEAD aircraft will operate in an area (SEAD box) close to the SAM threat and suppress the SAM while the COMAO passes the threat. SEAD aircraft may stay behind after the attack formation has passed the threat (e.g. SAM belt) and

resume suppression when the attack formation returns. Alternatively, some SEAD aircraft can accompany the attack formation to suppress SAM in the target area.

- Jammers are normally deployed outside enemy territory. Some aircraft jam enemy radar frequencies while others jam communication channels.
- Attack aircraft, in the shape of bombers and fighter bombers, will normally attack the OCA targets. The use of fighter bombers has the advantage that less sweep and escort fighters are required; bombers, however, need fighter protection as long as enemy fighters are not neutralized.
- Support aircraft such as tankers, AWACS and reconnaissance aircraft are required for proper execution.

Optimising the OCA campaign

OCA is the most demanding and dangerous part of an air campaign. Its execution requires a great many air assets. Interestingly, the total number of support aircraft often exceeds the number of bombers. Several solutions can be sought to increase effectiveness. Stealth, for example, decreases the need for SEAD aircraft and jammers. Increasing the range of fighter bombers diminishes the need for tanker aircraft. Better munitions could be capable of destroying a runway in such a way that rapid repair is impossible. Improvements to anti-radiation missiles can enable them to remember the location of emission and destroy enemy radars even after they have shut down. Finally, UAV can be used to loiter above enemy airspace and attack air defence radars as soon as they emit[122].

[122] The Israeli HARPY UAV for example.

Chapter 13 – Anti-surface force air operations

Introduction

Historical overview

Not surprisingly, engaging in anti-surface force air operations (ASFAO))[123] was one of the first roles for the aeroplane. Instead of spotting the enemy and guiding the artillery, it made sense to attack the enemy directly from the aeroplane. It made even more sense to attack the enemy over enemy territory before he came within reach of the own ground forces. But ASFAO was not the favourite role for many airpower theorists (see chapter 21). Many regarded this role as a diversion from the main task: gaining air superiority and executing strategic operations. Neither was it popular among aircrew. Fighting other aircrew was often seen as more heroic than moving 'mud' during air-to-ground missions.

If we look at recent conflicts however, it becomes obvious that many of the airpower assets are used for ASFAO today. A few examples:

> *During Operation Desert Storm (Iraq, 1990/1991) 66.9% of attack missions were against enemy ground forces. Most ASFAO sorties were flown before the ground war even started!*
>
> *During Operation Enduring Freedom (Afghanistan, 2001) almost all attack missions were ASFAO.*
>
> *Initially, most of the sorties flown in Operation Iraqi Freedom (Iraq, 2003) were Counter Air and Strategic, but after two weeks 80% of all attack missions were flown against ground troops.* [124]

Several explanations can be given for these high percentages. First, in missions against a less sophisticated enemy, less effort is required to win air superiority (Counter Air), leaving more assets available for ASFAO. A second reason may be that the number of strategic targets, especially against a less developed country, can be small. This also means that more aircraft are available for ASFAO. And lastly, in out-of-area operations and 'empty battle fields'[125] it is often easier and more economical to use aircraft than to ship massive amounts of artillery to the theatre.

ASFAO is never easy, it always requires coordination between army and air force, and targets are often hard to spot and highly mobile. In addition, aircrew have to operate in close proximity to friendly and enemy fire. Modern developments in avionics, in particular precision munitions, targeting pods and data-links have eased some of the problems. Some problems have however remained, in particular the control over air assets. Air assets are expensive and therefore scarce. It is not unlikely that each battalion or even company commander prefers to have a fixed number of air assets assigned to their battalion. However, history has proven that it is more effective to direct these scarce air assets at a higher level. Many Western armies have felt uneasy with a lack of direct control over air assets. This lack of control, added to the availability of helicopter technology, has probably been the most important reason for establishing army aviation. Presently, US army aviation has more airframes than the USAF. Furthermore, modern navies consider aircraft

---

[123] There is a fair chance that ASFAO will be named Counter Land in the future. This will imply that Counter Sea becomes a separate type of operation.

[124] Source: Cordesman, Instant lessons of the Gulf WAR, preliminary report.

[125] The term 'empty battlefield' is used for situations where the enemy is dispersed over a large area, which thinly spreads own resources. A guerrilla type war is a typical case of an 'empty' battle field.

carriers as their core capabilities and often possess other ship and land-based airpower assets in which ASFAO, especially against enemy ships, play a primary role.

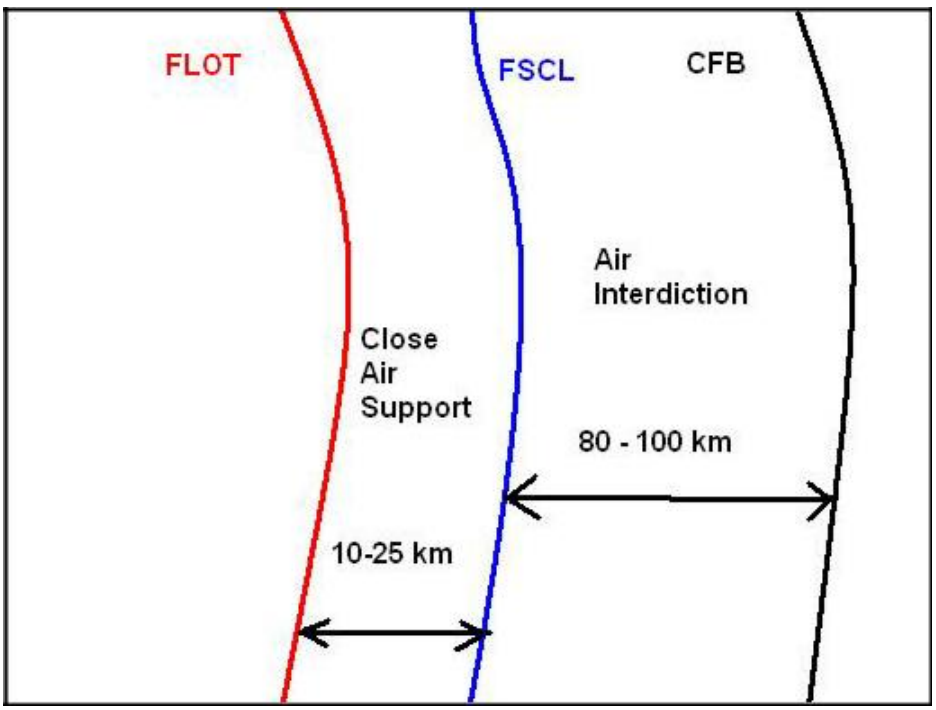
Definitions

A few definitions are required for the understanding of ASFAO.

Forward Line of Own Troops (FLOT). This is the edge of the battlefield. This line is constantly changing and pilots who carry out ASFAO missions should have the latest updated FLOT on their maps.

Fire Support Coordination Line (FSCL). This line is the control horizon of the ground forces in contact with the enemy. They will be able to direct artillery and air assets between FLOT and FSCL. Normally the FSCL is between 10 and 25 km from the FLOT.

Corps Forward Boundary (CFB). This is the planning horizon of the Corps Commander and is 80- 100 km past the FLOT.



Close Air Support (CAS). CAS missions attack targets between the FLOT and the FSCL. CAS is deployed against ground forces. These missions in the vicinity of own troops require a close coordination with ground forces.

Air Interdiction. (AI) Air Interdiction missions have their targets beyond the FSCL. Interdiction missions prevent or delay non-deployed ground forces[126] before they reach the battle area. Interdiction does not require direct coordination with ground forces.

126 Non-deployed forces consist of follow-on forces on their way to the front to reinforce the front troops. But even if there is not yet a front (initial phase Desert Storm), non-deployed troops can be located and attacked.

Forward Air Controller (FAC). The FAC is the ground or airborne controller who guides the aeroplane to the target.

Air Liaison Officer (ALO). An Air Force officer (normally a pilot) attached to a ground unit (brigade or division) to coordinate the air support.

Air interdiction

Aim

The aim of AI is to destroy, damage or hamper enemy forces before they are in contact with own forces. At a minimum AI will create time for own ground forces to deploy, and at best defeat the enemy before they come into contact with own ground forces.

Targets and weapons

Typical AI targets are infrastructure and forces on the move. The characteristic of infrastructure is that the positions of choke points, bridges, railway yards etc. are normally well known, making it relatively easy to plan missions against these targets. In the past AI targets like bridges were hard to destroy, but nowadays, with precision-guided weapons, this has become much easier. Several challenges in the field of infrastructure do remain. When there are many different routes to the front, the number of targets increases and to effectively delay the enemy it is necessary to destroy most of his routes to the front. However, if the same infrastructure is needed for the own advance into enemy territory, it should be safeguarded rather than attacked, for instance by deploying airmobile forces. Not strictly a military consideration but certainly a point to consider is the fact that infrastructure is very expensive and if the burden of the reconstruction effort will be on your shoulders, you should think twice before destroying it.

Military equipment en route normally moves in convoys, which are recognizable from the air, making them an attractive AI target. A further advantage is that the concentration of targets enables one aircraft to attack multiple targets in a single pass. For this type of attack, cluster weapons are ideal, especially with autonomously guided sub munitions. A disadvantage is that when these targets are on the move, the pilot assisted by the weapon control computer must be able to adjust for target movement. A further disadvantage is that the enemy position is not known in advance, so it is hard to pre-plan against these targets. Also, in low-intensity conflicts, with a dispersed enemy, the AI targets may be rather small. Troops may use single civil trucks, which are hard to distinguish from other civil transportation.

Target location

The location of infrastructure targets is relatively easy to determine. Satellite pictures are sufficiently detailed to accurately plan attacks against these targets. To detect enemy convoys is more difficult and is most effective with Air-to-Ground Surveillance radars like JSTARS. Not only are they capable of distinguishing trucks and APC, but they can even discriminate between wheeled and tracked vehicles. Another way of locating these targets is for the pilot to detect the target autonomously. In the latter case the pilot will search over an area or along line features, like roads. Depending on the sensors available in the aircraft, the search is done visually, with a targeting pod or with the onboard radar. If radar is used, this sensor should preferably have a SAR mode.

Attack profiles

The most important parameters determining the attack profile are: weather, threat, type of target and aircraft sensors.

Medium level operations are preferred when air superiority has been achieved and the only threat remaining is small arms fire. The ease of spotting AI targets at medium level depends mainly on their size and the weather. Army columns and infrastructure are normally easy to detect during daytime but during night operations (and against small targets), a targeting pod is required. Large infrastructure can be attacked accurately enough with dumb bombs, but some infrastructure, especially bridges, requires very precise attacks, which is only feasible with an LGB (the most precise weapon). Attacking army columns is preferably done with cluster weapons, which are however not very accurate from a medium level unless they are equipped with a Wind Corrected Munitions Device (WCMD)[127]. It is also preferable to have the cluster weapons radar-fused[128].

When clouds obscure the target, most AI missions will be impossible, the only exception being large area targets (e.g. rail yard staging areas) that are visible on radar or can be attacked based on known coordinates. At present most combat aircraft are not equipped with a SAR capability, but some aircraft like the F-15 Strike Eagle, the B-2, and also the Apache–D with the Longbow radar, have a SAR capability, enabling them to detect and attack small army targets. The preferred weapons are GPS or radar-guided munitions (JDAM and radar-guided Hellfires for example).

The low profile is an option when the air threat at medium level is too high or when clouds interfere with target acquisition, especially when the aircraft is not equipped with accurate air-to-ground radar. AI at low level is not easy, firstly because flying fast at low level generates a high workload for the pilot. Secondly, because the time to spot the target is short. If a target is spotted, the pilot will normally not be able to attack the target right away. There are a few options available: the pilot can execute a manoeuvre to reposition himself into an attack position (during which he will lose sight of the target) or he can order his wingman to carry out a direct attack[129].

Beyond the FSCL, no regular troops are present who can laser-illuminate the targets; consequently delivering precision weapons will require the use of a targeting pod. Operating a targeting pod at low level is extremely hard. The only easy way to operate a targeting pod safely at low level is with a two-men crew or with an automated terrain-following capability in the aircraft.[130]

Operating low level at night is even harder. Specialized equipment like night vision goggles (NVG) and Forward Looking Infra Red (FLIR) sensors are required to present the pilot with an outside reference. An auto-terrain following system can facilitate this operation[131]. In conditions of limited visibility, fast jets at low level can

127 Wind-corrected munitions devices range from INU strap-on kits to conventional bombs that correct for wind-induced trajectory errors. This correction is especially useful for bombs with a long time of flight (high drag).

128 If a cluster weapon is time-fused, the time of fall of the sub munitions will be too long. This will create a wide dispersion and inaccurate delivery.

129 When the lead is searching the road and the wingman flies at a 45° line behind the lead at 1-2 nm he is in an position for direct attack.

130 Low flying requires the pilot to have his eyes outside the cockpit 90% of the time, leaving too little time to operate the targeting pod.

131 The USAF used to carry out night low-level missions with F-16 Block 40 equipped with the Lantirn targeting pod and so did the RAF with the Harrier. Presently only a limited number of assets are still capable of operating a targeting pod at low level.

only operate against fixed pre-planned targets and need terrain-following radar. Under these circumstances it may be more advantageous to use attack helicopters. Because of their reduced speed, attack helicopters can operate in much lower visibility than fast jets.

Phasing

Operating above enemy troops beyond the FSCL is inherently dangerous. The AI operation should preferably only start after the successful execution of the OCA campaign and with air superiority achieved. Historic examples of this are operations Desert Storm and Overlord (Normandy 1944). During Desert Storm almost 50% of Iraqi ground troops were put out of action before the ground offensive even started. During Overlord the German troop movement during daytime was reduced to almost zero.

On the other hand, starting from a defensive position the situation may be completely different. A good example is the Yom Kippur War (Middle East 1973). The attacks on bridges over the Suez Canal without air superiority were a dangerous AI mission that cost many aeroplanes. However, the Israelis had no choice; the country had a very limited depth to defend and the support for the army was bitterly needed.

Conclusion

A well-executed AI campaign can considerably ease the task of the army. Air superiority is a prerequisite for a successful AI campaign. Furthermore, the availability of good sensors (in particular strategic ground surveillance radars and attack aircraft with SAR capability and targeting pods) and precise munitions (LGBs, WCMD and cluster weapons with autonomously guided sub-munitions) are likely to lead to a good result. Without air superiority, AI is still feasible, but high attrition levels should be expected.

CAS

Aim

The aim of a CAS mission is to fight enemy troops in contact. Airpower, with its inherent mobility and firepower, can quickly establish a more favourable position for army units. CAS can be supplied to troops under heavy fire to prevent an enemy breakthrough; it can also strengthen an attack or defend the vulnerable areas (supply lines) of own troops.

A typical example of the latter occurred during World War II when the US Army Air force defended the vulnerable flanks of General Patton's army during his rapid advance through France. Another example is the defence of the road to Baghdad during Iraqi Freedom (2003).

Targets

CAS targets are normally tanks, APC, AFV, artillery positions, and command and control facilities. Most targets will be hard to see due to their camouflage and will be hard to hit because of their movement.

Target location

The location of the target is usually determined by the FAC who will be on the ground with the troops. Airborne FACs cover a larger area and can handle multiple targets while observing the enemy from the same angle as the attackers. Typically, light helicopters and propeller-driven aircraft are used in this role. Their slow speed and relative height give these aircraft a good chance to detect and locate the enemy.

Sometimes even fast jets are used as FAC aircraft (Fast FAC [132]); normally this is well beyond the FSCL and would better be labelled Fast Interdiction. In the past the most common method for ascertaining the enemy's position was to plot his position and then read the coordinates from a map. Presently a FAC is equipped with a GPS and laser rangefinder with which he can accurately measure the target position and even illuminate it.

Coordination and communication

Coordination is the key to a successful CAS mission. At the very least, the attack pilot needs accurate target coordinates, target description and position of friendly troops. In addition, during the attack run the FAC has to guide the pilot to the target with verbal instructions. Guiding a fast jet from the ground is not easy and requires a lot of training. The FAC has to appreciate the speed and turn radius of the attack aircraft but must also visualize how the pilot sees the world from his elevated position. If the FAC is able to give an accurate position and the aircraft is equipped with an accurate navigation system this final guidance task is made a lot easier.

CAS depends on a good communication system. The system should preferably be secure, hard to jam and capable of transmitting data. Secure and jam-resistant radios have been available for two decades, but not all aircraft and ground units have them. The development to use data link instead of voice is more recent, and has an even greater impact. To pass on position data by voice is cumbersome; positions can be misread from the map and understood incorrectly by the pilot. The pilot can make typing errors when entering data into his navigation computer, which is also a time-consuming process. A further development is that the FAC and the pilot can exchange pictures. The FAC can use digital target pictures while the pilot can send his targeting pod video. One of the disadvantages of the standard military UHF radio is that the bandwidth is small, and consequently the data exchange rate is low. This is only a factor when exchanging pictures, not when passing on target position. The 9-line brief:

FAC BRIEFING FORMAT

MISSION C/S ____________ ABORT CODE ______________

NOTE: 1. A,D,G,H IN BOLD ARE MANDATORY READ-BACK (even if "NONE")
2. Heading and bearings Mag unless True is reguested

A. IP ____________________ __________

B. BEARING ________________ __________°

C. DISTANCE ________________ __________NM

D. TARGET LOCATION 1. (UTM) __________

2. (LAT/LONG) __________

E. TARGET ELEVATION __________ __________FT

F. TARGET DESCRIPTION ____________________

G. MANDATORY ATTACK HEADING __________°

H. FRIENDLY FORCES ______________________

I. ATTACK TIME TOT/TTT __________ __________

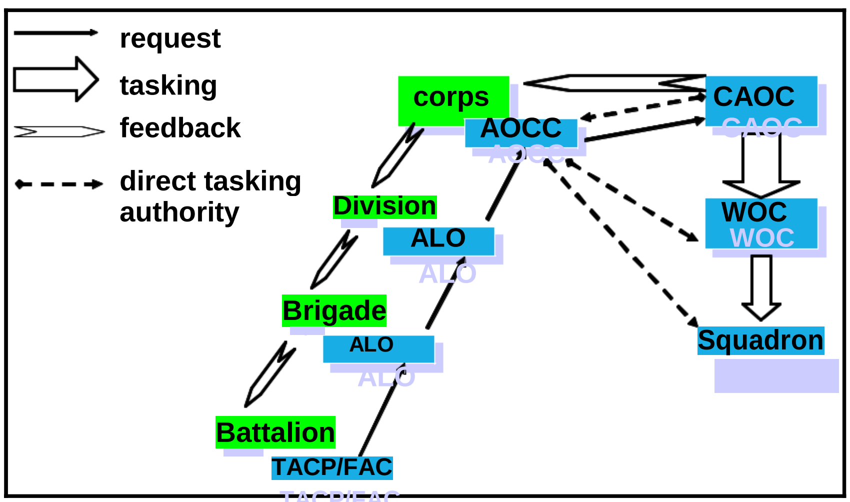
J. ATTACK CLEARANCE FAC C/S TAD

M A N D A T O R Y

132 Fast FAC was used in Vietnam and during Desert Storm

Coordination is also required at higher levels. One of the keys to the success of the German offensive in 1940 was the close cooperation between army and Luftwaffe, with the latter being subordinated to the army. In the Blitzkrieg-concept, German Stuka aircraft operated from locations close to the front, which minimized response time. What was true then still holds today. Two factors are very important in this respect: first determining the right number of aircraft to be designated to CAS. Second, minimizing the response time (sensor-to-shooter time). If too many aircraft are planned to perform CAS, suitable targets may not be available and valuable aircraft sorties are lost. Moreover, if the response time is too long, the target opportunity might disappear before the aircraft has reached the target area.

These two conditions are contradictory, a short sensor-to-shooter time requires that the aircraft is readily available, but scheduling just enough aircraft requires that assets are only assigned when absolutely necessary. During Desert Storm another concept was used: 7 minutes' CAS. Aircraft were tasked to targets behind the FSCL but would hold for 7 minutes before the FLOT to see if some army unit needed assistance. If there was no request during the seven minutes the aircraft would proceed to the alternate target. This is an excellent example of how airpower can be used creatively. The standard NATO CAS request procedure is much slower and is shown in the next picture. It is interesting to note that during most peace support operations, CAS missions are assigned and the actual lines are much shorter. Even individual group and platoon commanders can ask for air support.



Attack profiles and weapons

The attack profile chosen is based on threat level, weather, weapons and sensors available.

The medium level profile is useful when air superiority is achieved and clouds do not obscure the battlefield. The pilot will circle above the target area (the wheel) and will receive all information pertinent to the target from the FAC. Because CAS targets are small in size it is hard to spot them from medium level without a targeting pod. The FAC will use the 'large-to-small' technique[133] to talk the pilot to the target. After the pilot recognizes the target he will roll into a steep dive for the attack. This dive attack brings him closer to the target, enabling better recognition and a more accurate weapon delivery. It also brings him closer to the small arms threat but the exposure time will be short because of the high speed and fast pull-

[133] The FAC 'guides' the pilot via ground features to the target area; starting with large dominant features (radio towers, highway crossings) he tries to talk the pilot's eyes to the target.

out after delivery. If the aircraft is equipped with a targeting pod the dive to the target is not necessary and target recognition is of course much easier.

Technically, LGB can be delivered on CAS missions even if the target is obscured by clouds. In that particular case, the FAC must perform laser illumination. The cloud base should be high enough (>2500 ft) to leave some guidance time for the weapon. Naturally the pilot should have very accurate target coordinates and the Rules of Engagement (ROE) should not prohibit this attack.

Under instrument conditions, CAS from medium level is barely possible. Theoretically a JDAM could be launched on coordinates, but a JDAM has a fixed coordinate and is impractical against a moving target[134]. Another option is a radar-guided weapon (Hellfire).

The low profile may be used when threat or cloud cover prevent operations at medium level. As the front is a busy and dangerous area, coordination with own weapon systems is essential. Not only the own air defence weapons, but also mortars and small arms that are close to the aircraft trajectory should be on weapons hold during the time the CAS aircraft is over the target area [135]. Also, the aircrew will try to limit their exposure time in the front area. Therefore the pilot will receive his briefing while holding in the rear area close to a pre-arranged Check Point (CP). After the briefing is received from the ALO or the rear (area) FAC, the pilot proceeds to an assigned Initial Point, where he will be handed over to the forward FAC for final guidance and proceed from the IP to the target. Before accurate navigation systems came into use, an IP was needed as a reference point in order to proceed to the target. Presently the IP is required less for navigation but will still channel the flow of aircraft to the target.

During low-level CAS attacks it is hard to use a targeting pod. Laser guided bombs can still be used if the FAC is able to designate the target with a laser. Under these conditions it is appropriate to loft the weapon. The loft will give the weapon more altitude and increases the guidance time[136]. A second advantage is that the pilot can keep some distance from the threat. Against moving armour, cluster weapons, Mavericks, Hellfires and large calibre guns (A-10) are most effective.

The relevance of CAS

Many air analysts consider the use of air assets against single targets, which is typical in CAS missions, ineffective. They would prefer to use air assets in an interdiction role where targets can be hit in larger numbers instead of one at a time. Consequently, they like to limit CAS to those periods when the army is not able to hold on its own. This negative opinion of CAS has been contradicted by history many times over. CAS allowed Patton to make his fast advance in France in 1944. As a specialized CAS aircraft, the A-10 was on the list to be phased out time and again, but in every conflict it proved its effectiveness and the phase-out date was delayed.

Based on historical evidence it is safe to say that with air support, the army will be more effective and can maintain a higher tempo. Part of this air support may be organic (army aviation with attack helicopters) but multi-role aircraft will usually be made available as soon air superiority has been established over the battlefield. The mix of AI and CAS with '7 minutes' CAS' is an excellent example of how CAS can be phased in effectively.

134 There are some new developments to correct the JDAM coordinates for target speed.
135 This includes own artillery fire, shoulder-launched AD missiles and own AAA.
136 The FAC should not start the guidance of the LGB too early (before it reaches its apex), the LGB will then be forced into a low trajectory and undershoot the target.

ASFAO against sea power

Using airpower against naval forces constitutes a special type of operation requiring special attention. The first important aspect is that a naval fleet is normally heavily defended, and this has to be neutralized first. A second aspect is that, as it is difficult to hide above and at sea, ships and aircraft will normally be visible to each other over a large distance, which favours the use of stand-off weapons by both sides. Compared to CAS, little attention is needed for own troops. So, one could label ASFAO against a naval fleet as a type of Air Interdiction.

Maritime targets appear in different sizes. The most difficult target is of course a large naval fleet with an aircraft carrier and embedded air defence fighters. On the other end of the scale can be a small landing force, which is not well defended at all.

The mission against a large fleet is initially an Offensive Counter Air operation, with the only caveat that the target positions are constantly changing. This is however far from trivial; a fleet has a very large manoeuvring area. If it performs an adequate emission control, it will be hard to spot without adequate reconnaissance assets.

As has been said before, the sea environment is ideal for stand off weapons. Sea power uses long-range surface-to-air missiles and a fighter force to defend itself. Normally the outer defence ring can be as far away as 200 nm from the fleet and is defended with fighters. The second ring starts at approximately 50-80 nm from the fleet and is defended by long-range missiles. The ships may be equipped with short-range defence systems, which can fire on aircraft but also on incoming missiles. But stand-off weapons are not only the prerogative of the navy. Airpower has a variety of stand-off weapons that can be used against sea power. During the Falklands War (1982) the French Exocet became known for its destructive capability against British ships. The Russians also developed a variety of stand-off missiles for use against American naval fleets.

Against smaller naval threats other weapons like the Norwegian Penguin missile and the Maverick guided air-ground missile are very effective.

The balance

Since the Falklands War no major air-against-sea power encounter has taken place. During that war the vulnerability of ships was dramatically exposed but also the effectiveness of the Sea Harrier in defending the ships was proven. This type of war is typically a high-tech war, which favours the contender with the most modern equipment.

Chapter 14 - Strategic operations

Introduction

Not long after aircraft were introduced in the First World War their strategic potential was realized. It seemed to make sense not to engage in large battles but to pass over the front in order to attack the enemy's heartland. Ultimately this approach was expected to shorten the war and to prevent the high casualty rates encountered in World War I. A more complete description of the development of this strategic doctrine is given in chapter 21. This chapter focuses on the execution of a strategic campaign, with particular attention given to the problem of target selection.

Definition

Strategic operations are directed against the will and the ability of the enemy to continue fighting. It targets the vital elements of the enemy's systems.

The problem

The problem in executing an efficient strategic campaign lies in identifying those targets that will yield the best result in relation to the effort. This means that the problem of a strategic campaign is threefold: an intelligence problem, a psychological problem and an allocation problem. We will discuss these elements in more detail.

Intelligence

Good intelligence is needed to identify strategic targets. In World War II the USA lacked strategic information about the German industry. The planners of the bomber campaign had to revert to information from banks (which had given loans to the German industry) to compose a target list. So, not having a loan from an American bank was initially sufficient to prevent that particular industry from appearing on the target list. Building a good intelligence picture takes time; satellite pictures have to be processed; human intelligence sources need to be verified and all this work needs a lot of old-fashioned manpower. Strategic information can also become outdated rapidly, especially the whereabouts of the leadership and positions of mobile strategic assets (for instance positions of Scud missiles and mobile laboratories). The lack of correct intelligence is the most frequently mentioned reason for the failure of a strategic campaign.

Psychological

The psychological problem is that the reaction of the enemy to a strategic attack is hard to predict. In World War II a lot of effort was put into bombing cities by the Germans and British alike. Contrary to the assumptions of most airpower theorists, the effect on morale was less than expected, so the results of the outcome of the conflict were hard to substantiate. This does not prove that bombing, or the threat of bombing the population is never effective. The very limited attacks on cities during World War I caused a great unrest. In a similar vein, nuclear deterrence during the Cold War proved to be effective. What is true for the population is also true for other strategic targets. If a country loses most of its industrial power it might be ruined, but its destruction will not necessarily force its leadership to comply with diplomatic demands. Especially dictatorial leaderships tend to accept a lot of suffering of their own people and will not readily give in, particularly if the position of the head of state is at stake.

Timing is also an important factor in the psychological effect. If the strategic campaign is a long lasting, low-intensity effort, the enemy might adapt to the situation and to the suffering involved. A comparable phenomenon is the public's reaction to the yearly death toll from road traffic accidents when compared to a few hundred casualties caused by a natural catastrophe. Another example of habituation is the World War II bombing campaign against the German oil supply system. Although in itself effective, it was slow enough to allow the Germans to make synthetic oil from coal.

The choice of strategic platform

The traditional platform used for strategic bombing since the 1930-ies is the dedicated bomber aircraft. However, since the Vietnam War the B-52 has been used in ASFAO, while a variety of tactical aircraft and missiles has been proven capable of attacking strategic targets.

Cruise missiles and ballistic missiles can be used against fixed targets at large ranges for which the coordinates are known. These weapons do not need air superiority for their deployment, which makes them ideal for the initial phase of the campaign.

Fighter bombers with precision weapons and using air refuelling have proven to be capable of strategic attack. If air superiority is not achieved, the fighter bombers can be deployed in a COMAO to enhance their survival. Moreover, the use of stealth enables deployment without the support of other assets (SEAD, jammers etc).

The strategic bomber in most cases is the most effective platform to be used against targets at long range. When deployed over enemy territory without air superiority, fighter protection may have to be supplied unless the bomber is stealthy. Then it can be deployed without fighter support under low visibility conditions and at night. During clear weather it is still vulnerable to enemy fighters, which could detect the bomber visually. Advantages are that the bomber is capable of attacking multiple targets and can be re-tasked during flight. Another option for the strategic bomber is to launch cruise missiles outside the range of the enemy air defence system.

Airpower advocates should not forget that other means can be used effectively in strategic attacks. Special Forces and resistance movements are worth mentioning in this respect.

The nuclear dimension

The invention of the atomic bomb changed the concept of strategic warfare completely. Its instantaneous destructive power was orders of magnitude larger than society had ever experienced before. Suddenly it seemed possible to subdue a country with a limited air attack in a way early airpower theorist had predicted (see also chapter 21). It was of course logical that the atomic bomb became a typical airpower weapon. Its destructive power was so great that the weapon had to be delivered from a safe distance or altitude to allow for a safe escape. This was initially only possible with air assets. During the Cold War therefore, strategic air operations became synonymous with nuclear operations and Strategic Air Command was the prime nuclear asset of the USA. But with nuclear weapons three problems arose:

firstly the issue of morality, secondly- after the acquisition of a nuclear capability by the USSR- the issue of vulnerability and thirdly the threat of proliferation.

The moral aspects

The nuclear weapon was initially not very accurate but had a large destructive potential. It was impossible to direct it only against industries or armies so the only ´logical´ targets were cities. Contrary to the International Law of Armed Conflict (see also chapter 20) it was aimed directly against civilians. Some believe that because of this, atomic weapons should be banned. Others believe that weapons could never be banned. They exist, so we have to live with them, whether we like it or not. To have the threat of the ´ultimate weapon´ available can furthermore discourage any opponent. That of course also depends on the mutual vulnerability, which will be discussed next. Considering the fact that the Cold War remained cold one could conclude that this latter, more practical, approach was justified.

Vulnerability

During just a short period after the Second World War the USA was the only country capable of nuclear attack. Not vulnerable itself, and with the capability to strike everywhere with devastating power it had unparalleled strategic superiority. In this situation the only problem was when the nuclear card should be played, or in other words how low the nuclear threshold was. It was obvious that the nuclear option should not be used lightly. So even in a situation with nuclear dominance a considerable non-nuclear force remained necessary to deal with 'minor' conflicts.

When the USSR developed its own nuclear weapons the situation changed. With both the USA and USSR now vulnerable, there were basically two options. One: to strike first in such a way that the opponent's nuclear potential was neutralized. With both nations having the inherent capability to gain nuclear superiority by striking first, the situation was technically unstable and resembled the strategic situation before the First World War. The country that mobilized last was expected to lose, which favoured the attacker. The second option was to ensure that the nuclear capability of both could not be destroyed. With the development of submarine-launched missiles and missiles with multiple warheads the situation changed to a situation where both opponents had a second strike capability. This second strike capability was so strong that the situation was labelled: Mutual Assured Destruction (MAD). This situation lasted till the end of the Cold War and proved stable.

New technological developments made it not only possible to develop very powerful nuclear weapons like the hydrogen bomb, but also small nuclear weapons such as nuclear grenades and depth bombs. The problem with smaller nuclear weapons is that they lower the nuclear threshold and have the inherent risk of changing a stable situation into a situation where gradually more nuclear weapons are used.

Proliferation

The third problem with nuclear weapons is proliferation. The most severe proliferation risk is when non-territorial groups acquire nuclear weapons. These groups are almost immune to retaliation, whereas the fear of (nuclear) retaliation was probably the most important reason why the nuclear arms race never evolved into a nuclear exchange. One cannot control the spread of nuclear knowledge in the present open society. The only way to control proliferation is to control fission material and nuclear technology. This work is primarily done by the UN Nuclear Control Agency.

Nuclear outlook

In today's world it seems that nuclear weapons are less and less important. It can also be argued that the mere existence of these weapons is enough to stabilize the situation between the larger countries. This leaves us to fight the 'smaller' but far more frequent wars and wars against non-state actors.

Vulnerability and phasing

Vulnerability is also an important consideration in conventional wars. One's own vulnerability dictates the perception of a strategic offensive[137]. If two countries can both launch a strategic attack against one another, it may be advantageous to strike first and direct the attack against the enemy's strategic potential. If one country is vulnerable to strategic attack but incapable of attacking the opponent's heartland, his only option is to execute a defensive counter air campaign or, if a front exists, he can attack at the front. Given this situation the opponent may want to delay his strategic air campaign until he has achieved air superiority. History has shown however that the situation is sometimes more complex. During Desert Storm, Saddam Hussein did not have the option of attacking the USA directly, but his Scud missile attacks against Israel were definitely strategic attacks directed against the cohesion of the alliance between the USA and the Arabic partners. Given the indirect strategic capability of the opponent, the allied forces probably decided to start their strategic campaign as early as possible. This would require a combination of cruise and/or ballistic missiles and stealth aircraft.

Targeting

In World War I limited experience was gained with strategic attack, but soon after the war, airpower theorists like Douhet already envisaged the future. He foresaw that in the next war airpower would be used for strategic attack, which was synonymous with attacks against the civilian population. In World War II a large part of the strategic effort was directed against the (war) industry and against oil production and supply. But a systematic analysis of what targets should be taken into consideration for strategic attack was missing. Colonel John Warden was the first scholar to make a systematic classification of strategic targets[138]. In the next paragraph Wardens system will be described briefly and some considerations will be given concerning the different target groups.

Warden

Col John Warden was the architect of the air campaign plan during Desert Storm. He was convinced that new technologies like precision bombing and stealth would change the way air warfare was conducted. Now airpower had the capability to reach and strike targets as never before.

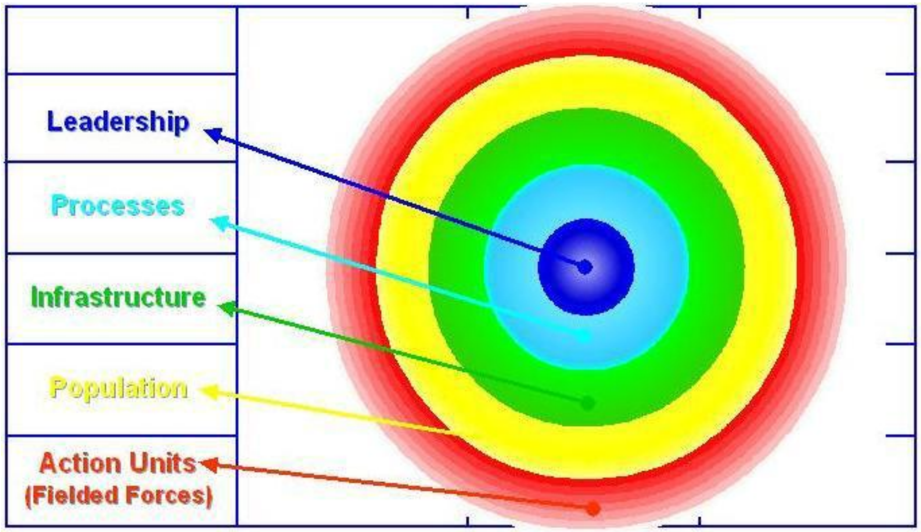
Warden compared the enemy with a (living) organism, with the brain as the central command and control unit. The living system has essential organs like lungs and a heart, and the blood vessels and nervous system can be seen as its infrastructure. The individual cells are its population. Finally, it has a defence mechanism: the white blood cells. In a similar way Warden classified the enemy, who has a brain: the leadership. It has vital organs: essential industries and power plants. It has an

137 In his book "The Air Campaign', John Warden describes four types of vulnerability and distinguishes between two parties being able to reach the heartland and/or the front.

138 This idea was initially presented by Warden in PowerPoint briefings and used during Desert Storm. In 1995 an article on this subject entitled "The enemy as a system" was published in the Airpower Journal. Airpower Journal 9, no. 1 (spring 1995): 40-55

infrastructure and a population and finally it has a defence system as well: its armed forces.

Warden visualized his concept with concentric rings.



Warden argues that as with a living system, each ring can be the focus of attack and the elimination of each individual ring will cause the breakdown of the complete system. Furthermore, he argues that the number of targets increases from the inside to the outside. So the inner rings require less effort and represent better strategic targets than the outer rings.

These target categories in itself are not new, but the visualization and the comparison of an enemy with a living body is new. Apart from focusing on each individual ring it is possible to make parallel attacks on all rings, maximizing the synergetic effect. This is called parallel warfare. In the past attacks on multiple rings would be considered a waste of effort. The resulting division of resources would give the opponent the capability to recover and it was considered better to focus on war industry or transport. But with the present precision weapons and the high operational tempo of warfare these attacks on multiple rings can strengthen the total effect and lead to a strategic paralysis of the enemy. Or, to use the analogy of the human body once again: if in a very short period a person loses an arm, a leg and a lung and has a severe headache he will feel paralysed.

The Warden model has been widely used, widely discussed and also widely criticized. But first and foremost we have to realize that it is just a model, a simplification of reality to shape our thinking. A model is never perfect, neither is Warden's, but his model helps in designing ways to attack the enemy.

Some limitations of the model are worth discussing in more detail. One limitation is that it is not always clear how the destruction of a certain part of the enemy's system will contribute to the overall campaign objective. The second limitation is that in operations other than war it is sometimes hard to define the system of the enemy. We will discuss these limitations in more detail.

The campaign goal

Campaign goals can vary considerably. Unconditional surrender as was demanded by the Allied Powers of Japan and Germany during World War II overshadowed the strategic air campaigns against those countries. In 1982, the Falklands campaign goal was that Argentina abandon the islands. In 1999 during Allied Force the campaign goal was that President Milosevic of Serbia agree with the Rambouillet peace agreement. The latter example was no threat to the autonomy of Serbia but was only to limit the influence of Serbia in Kosovo.

The target categories given by Warden may tell us what is necessary to paralyse an opponent. If an unconditional surrender is wanted, it is likely that the incapacitation of the enemy is a prerequisite for the subsequent occupation of his country. Under these conditions the leadership will not easily surrender and will defend its position as long as possible. So, in this case the model may help us to detect ways of achieving our goal with a minimum level of effort.

But how much damage is enough for a campaign with a limited objective? When a war has a more limited objective, we often use the term coercion. Coercion is not new but the use of the term is. Unfortunately there is not a generally accepted definition of coercion. For this book we will define coercion as the use of force or the threat of force against an enemy in order to change his behaviour. In contrast, during a total war the enemy is eliminated, which makes his behaviour irrelevant. In coercion the destruction of the enemy is not required, it could even be counterproductive. In the Kosovo example it was unwise to attack the leadership because the objective was to have the same leadership sign the Rambouillet peace agreement. The aim here was to use enough pressure to enforce this change of behaviour and not to paralyse the leadership.

What is the system?

In classical state-against-state warfare, it is easy to see the enemy as a system, but in non-classical warfare, especially in counter-insurgence and counter-terrorist warfare, it is harder to clearly identify the elements of this system. However, it is not impossible, as Warden has pointed out. He uses an example identifying the five rings for a drug cartel. An interesting problem is that its leadership may be more inspirational than actually controlling. Inspirational leadership is hard to tackle; even if the spiritual leader is captured or killed, his ideas may flourish. Another peculiarity is that leadership is often hard to distinguish within the system. Do the terrorists form one system, does one have to cope with several systems, or is every terrorist a system in itself. A practical approach is needed here. If a specific group can be identified as an enemy, it should be possible to treat it as a system with its specific vulnerabilities. Not all its rings may be easily identifiable, but it is not impossible. Even if some parts of the system are hard to tackle, like inspirational leadership, other elements of the system may be more visible and vulnerable to attack. For example: terrorists still need to recruit, which is a vital part of their vulnerable infrastructure, forcing them to expose themselves. So, a practical use of the system for non-conventional warfare is required.

Target categories revisited

This paragraph will discuss the pros and cons of each target category.

Leadership. The leadership, especially a dictator, is an obvious target but hard to attack. It is normally well protected and highly mobile. In a democratic society, leadership is more balanced and the country's position will generally not change when its leadership comes under attack.

Essential industries. Attacking vital power resources will definitely paralyse a modern society, but might be less effective in an underdeveloped country. The more developed a country is, the more important its vital industries are to civilians and the more likely it is that their destruction jeopardizes the population's livelihood. The effect of attacks on vital industries may not always be directly felt. If an oil refinery is attacked it may take some time before supplies run out. Attacks on power plants are normally felt sooner and are more harmful to civilians than to military forces who have ample emergency power units.

Infrastructure. A modern society has extensive infrastructure, which is vital to its functioning. It encompasses roads, railroads, waterways, the sewage system, the water system, gas and oil pipelines, the communication grid (telephone, cable and internet) and last but not least the electricity grid. Destroying the communication grid is effective in undermining the government's capability to control the country. But then again, attacks on infrastructure can also cause civilian casualties.

Population. Bombing a population into submission was deemed quite appealing in World War II. Today this type of warfare is considered inhumane by all official standards. On the other hand, the population is the favourite target for terrorist groups and attacks can be quite effective. After the terrorist attack on a railway station in Madrid (11 March 2004) the newly-elected Spanish government withdrew from Iraq.

Fielded forces. Strategic operations will normally try to circumnavigate the fielded forces in order to reach the head of the snake. But, based on the above considerations it will be clear that hitting a lot of these target categories has severe drawbacks, especially if we want to prevent collateral damage. Therefore the enemy's military forces may be a more acceptable target. The obvious drawback of course is that these forces are trained to counter attacks and to use camouflage and sheltering. However if the enemy is in no position to attack your own heartland, the bombing of fielded forces by airpower may be more efficient than moving in a large army over the ground. Some examples of strategic attacks against fielded forces:

In the later phases of Allied Force many missions were tasked against Serbian fielded forces; after the war the results were much less than initially expected.

During Desert Storm the air attacks against the fielded forces of Iraq before the ground war started were instrumental in the fast ground victory.

Application of airpower during Operation Enduring Freedom in Afghanistan can be labelled as ASFAO (with the assistance of Special Forces) or as strategic attacks on fielded forces. It showed that the application of airpower could tip the balance in the battle between the Northern Alliance and the Taliban.

Execution

Strategic operations can be carried out by individual airpower assets if air superiority is achieved, if the stealth characteristics of the platform give enough protection or if the platform is practically immune to attack (cruise missiles). In all other cases the strategic campaign needs the protection of fighters, jammers and SEAD-assets, so the classical method is using a COMAO.

The most effective initial strategic target is the leadership, because the number of leadership targets is quite often limited. It is relatively easy to incorporate these limited numbers of targets in the first waves of the air campaign and even concomitantly with the OCA campaign.

Conclusion

The strategic potential of airpower was conceptually realized early in the 20th century, well before airpower was able to fulfil the expectations. The wars in the second half of that century have seen the fulfilment of many expectations. On the other hand, this has confronted us with dilemmas. Not all strategic targets are acceptable on humanitarian or practical grounds, not all destruction leads to the desired political result. We may state that the correct application of strategic power is in essence still an unsolved problem.

Chapter 15 - Helicopter operations

Introduction

It is arguable whether in the national context helicopter operations should be treated as a species of airpower. One might reason that they are synonymous to air manoeuvre. Air manoeuvre complements ground manoeuvre and therefore land doctrine should govern air manoeuvre operations. Indeed, this seems to be the case in the Netherlands, where the *Leidraad Air Manoeuvre (LD 6)* is part and parcel of the land forces doctrine series. However, this particular volume has been approved not only by the CinC of the Army, but also by the CinC of the RNLAF. This is a consequence of the fact that in this country the vehicles of air manoeuvre are part of the air force organization. Thus, the RNLAF is deeply committed to the doctrine of air manoeuvre.

Apart from this joint commitment, which in the national setting comprises joint operations with the Marine Corps, helicopters can be employed in air force and independent operations. During crisis operations, attack helicopters have cooperated with fighter aircraft and transport helicopters have complemented missions of fixed-wing air transports. Combat Search and Rescue (CSAR) is a typically independent helicopter mission, designed for the rescue/pick-up of fighter crews or Special Forces in enemy-held terrain. Furthermore, attack helicopters have come to assist ground patrols involved in shoot-outs with irregular forces. The aim of this chapter is to provide a short overview of air manoeuvre, following the LD 6.

Historical overview

Air manoeuvre operations stem directly from airborne operations as they developed during WW II. Airborne troops at the time could be divided into air landing units and parachute units. Air landing units were light infantry put on the ground by fixed-wing aircraft and/or gliders, while parachuted infantry left their vehicles in the air. In both cases the units had to re-assemble before being able to fight. Additionally, the establishment of an airhead in enemy territory implied the need for a speedy link-up with (mostly mechanised) ground units that had broken through enemy lines and reached the airhead. German airborne operations were carried out in the vicinity of The Hague in May 1940 and on Crete in 1941. Allied airborne troops landed in June 1944 in Normandy and played an important role during operation Market Garden in September 1944. During the colonial crises Western powers used airborne troops often on a smaller scale. A national example is the raid on Yogyakarta in December 1948 when Dutch Special Troops raided the headquarters of Sukarno, the leader of the provisional Indonesian government.

From the war in Vietnam on, airborne operations were gradually replaced by airmobile operations. Transport helicopters became capable enough to transport sizeable numbers of troops into the combat zone and pick them up on completion of the mission. The last step would have been impossible to execute by transport aircraft or gliders and added an essential quality to this type of operation. The armament of helicopters became the next logical step, first as an addition to the transport helicopter being capable of suppressing enemy fire in the landing zone, later as armed helicopters evolved into attack helicopters. This new category became dedicated to offensive actions in concert with airmobile troops on the ground or to attacking the enemy independently, often in cooperation with fixed-wing aircraft. When special units were formed to exploit the combination of airmobile troops on the ground and the firepower of attack helicopters, a new form of manoeuvre called air assault developed.

Definitions

Air manoeuvre comprises the following combinations of the employment of infantry and helicopter components:

Air assault: integrated operations of helicopters and specialized infantry

Air mechanised: mainly attack helicopters

Air mobile: specialized or specially trained infantry transported and dropped by helicopters

Airborne: parachuted specialized infantry

The most obvious characteristics of air manoeuvre are identical to those of airpower, but air manoeuvre is an offensive operation aimed at eliminating adversaries or *occupying terrain*. Because of the independence of terrain during the manoeuvre, this type of operation will always try to use surprise as a cornerstone of the mission plan. Like airpower, air manoeuvre will be employed over relatively large distances, possibly deep into enemy terrain. The 101 (US) Airborne Division (Air Assault) is able to operate in an area of 300 x 1300 km, as employment in Iraq in 1991 has shown.

Types of combat

The air manoeuvre types of combat that are leading for the doctrine of the Dutch 11 Air Manoeuvre Brigade can be categorized as follows.

Offensive operations

Under symmetrical conditions *deliberate attacks* can be planned with an optimal balance between manoeuvre and fire power. The same goes for a *deep attack* executed in enemy terrain. This type op operation can bring decisive results but implies considerable risk, necessitating careful planning. A *hasty attack* is carried out when a sudden and favourable window of opportunity opens. Often attack helicopters will bear the brunt of such an operation. An air manoeuvre unit may furthermore be used for a *counter attack* or a *pre-emptive attack*. In *both* cases the upper echelon will try to (re-)gain the initiative while the enemy is in a vulnerable position. Lastly, the *raid* is used for a swift penetration of hostile territory in order to secure information, confuse the enemy or destroy its installations. After the raid the unit will withdraw.

Defensive operations

Air manoeuvre units are not very suitable for defensive operations in sectors, more so if a lack of air superiority occurs, a precondition for helicopter operations. On the ground, airmobile troops show limited mobility and fire power. Only when attack helicopters can be used can the unit be of vital importance.

Delaying operations

The air manoeuvre unit is able to deploy in almost any tactical situation faster than any other manoeuvre element. This characteristic paves the way for employment in delaying operations where territory can be exchanged for time. The ground elements will normally be used for continuous observation while the attack helicopters deliver the heavy blows. Transport helicopters can only be employed during some degree of air superiority.

Helicopter planning data

While planning helicopter operations, the following conditions have to be taken into account:

Helicopter performance is influenced by weight, altitude and temperature. Engines and rotor systems respond to air density. A decrease in air density through height and/or temperature (*hot and high*) may severely limit maximum lift capacity and manoeuvrability. During operations in Iraq in 2004, temperatures sometimes exceeded 50° C with even higher cabin temperatures, putting a halt to Cougar flights. Chinooks can be outfitted with internal fuel tanks, which can limit the payload for other reasons. External loads limit the manoeuvrability of the helicopter, resulting in lower cruising speed and higher altitude above terrain and obstacles.

Weather can provide tactical advantages to helicopter operations, when limited visibility allows unobserved approaches. On the other hand sensor performance of attack helicopters, especially forward-looking infrared, is restricted in moist conditions, resulting in a non-ability to use weapon systems at maximum range. Night-vision goggles help transport pilots at night, giving minimum moonlight or so-called rest light, the absence of which will bring operations to a halt. Though most helicopters can fly under instrument conditions, the approach to the landing site is dependent on air and ground instruments that are sensitive to weather as well. But even with instruments, the approach will still have weather limitations. Finally, wind gusts, mountain downdrafts and turbulence can lead to the cancellation of operations.

Safety and security. Air safety measures aim to limit or, even better, to exclude the negative impact of losses in personnel and materiel. They are part and parcel of every air and/or air manoeuvre operation. Under extreme circumstances air safety can have a dominant effect on planning, for instance when rotor whirlwinds in an environment of sand or snow make landings challenging. Crew rest directly affects air safety. Depending on operational urgency and earlier commitments, continuous flying is limited to 10 hours in daylight or 5 hours during darkness. Security is a different matter altogether, being a normal military consideration. Helicopters have to be secured against the enemy air defences or against heavy infantry weapons. Therefore transport crews include door gunners. During operations airmobile troops will provide close protection, while the brigade will take care of the air defence. Further security measures imply the use of ingress and egress routes and, if practical, camouflage. Dispersing assets will always impede maintenance and encumber security by ground personnel.

Logistics are the backbone of every helicopter operation. While air mobile units operate as lightly as possible, a helicopter unit is of necessity heavy. In the theatre of operations most logistics support will travel by road, but it can be necessary to lift some essential personnel and/or equipment by air. Moreover, forward operating bases (FOB) and forward arming and refuelling points (FARP) can be situated in adverse terrain and be supplied by the scarce transport helicopter capacity. During operations, planning for a downed aircraft recovery team (DART) has to be done. Maintenance planning offers limited flexibility, upward or downward, but will always have some effect on the overall capacity.

Complexity. The number of helicopters and the diversity of types furthermore affect the operational planning. Multi-ship formations have to adapt to the speed of the slowest aircraft. It takes time to build up a multi-ship formation, while the pick-up of personnel and loads needs careful planning. Operations at FARPs get more complex given the number and types involved.

NBC circumstances significantly impede helicopter operations. Airframes are sensitive to NBC detergents.

Command and control

The success of an operation depends heavily on a well-balanced planning among air and land staff officers who take the strengths and weaknesses of both components into account. It is most difficult to anticipate the circumstances upon landing, when dispersed units have to be assembled in order to mass slow moving infantry. Central planning and decentralized execution are the key words to describe command and control arrangements for air manoeuvre. Operational integration of both elements will be reached through common procedures, organizational integration and unity of command. Crucial functions conditional to the employment of transport helicopters are organized in the ground component, like pathfinders, riggers/marshallers, landing point commanders and helicopter handling instructors. All personnel have gone through basic training aimed at working with helicopters. Loadmasters and gunners form part of the crew of transport helicopters.

The commander who will direct the combat action on the ground is responsible for the ground tactical plan, in which the order of the higher level has been worked out. The ground tactical plan, usually achieved via backward planning, dictates other planning, like the landing plan, the air transport plan, the loading plan and the pick-up plan. For fire support, air defence and logistics, separate plans will be made.

Operational framework

As air manoeuvre assets basically have the same characteristics as other airpower assets, they are extremely suitable for deep operations. In this type of operation the adversary will have to be found, engaged and defeated by way of surprise. His freedom of action will have to be limited at an early stage, preferably by hitting his most vulnerable elements, like C2-installations, radars, fire support or critical logistics. A deep operation always fits into a broader operation plan, for which the higher echelon will be responsible. This echelon will provide the intelligence pertinent to the operation and will initiate the employment of other assets, e.g. C2 platforms, communications, fighter escort, SEAD and long-range artillery. Often, a deep operation will not be limited to a raid (strike and return), but involve the employment of ground mechanised units as a follow-up to the air manoeuvre. By the same token, air manoeuvre assets can be used against adversaries more closely in enemy terrain unsuitable for ground mechanised units.

Alternatively, to counter surprise attacks, using air manoeuvre assets is often the only way to thwart an adversary who has succeeded in penetrating the own troops' rear area. Obviously, attack helicopters will play a major role in such an operation, when thorough preparations are out of the question and the involvement of higher echelon assets would be time consuming. More so than in a planned attack, the security of the staging area, the forward operating base or the forward arming and refuelling points are at stake.

Air manoeuvre can play an essential part in the creation of an *airhead*, a kind of bridgehead in enemy terrain. An airhead is defined as a 'designated area in a hostile or threatened territory which, when seized and held, ensures the continuous air landing of troops and materiel and provides the manoeuvring space for projected operations.' An airhead may serve as a stepping stone for further operations, e.g. deep operations. After link-up with a ground mechanised unit the forward momentum of the mission can be safeguarded.

Helicopter assets

Attack helicopters

The heavy combat power of every air manoeuvre unit lies mainly with the attack helicopter, the infantry being lightly armed. However, the long reach of the helicopter weapon system makes it the weapon of choice in open terrain, more than in or above woods. As mechanised units also prefer to operate in open terrain, the qualities of the attack helicopter can be brought to bear here. By virtue of its long range weapons the attack helicopter reaches farther than most vehicle-borne weapons. The second advantage lies in the choice of weapons: attack helicopters can use precision-guided missiles, area weapons and guns depending on the nature of the targets. Moreover, they possess two or three sensor systems (infrared, radar and electro-optical) that complement each other by day and night. The resulting flexibility is of utmost importance when switching from one task to another: attack helicopters can be used in show-of-force, reconnaissance, observation and security operations. It has even occurred in crisis situations that a helicopter was used as a C2 platform, providing a platoon commander in a tight spot with valuable directives. More and more, attack helicopters are being used in urban warfare. Exposure over unknown terrain makes them vulnerable, but there is a lot of compensation in their stand-off capabilities.

Attack helicopter capabilities make them at least as suitable for Close Air Support missions (CAS) as fixed-wing aircraft. While the same procedures apply in that type of employment, it has to be ensured that the joint staff (usually at the brigade level) have the proper communications assets available. The quality of the network system may be even more instrumental when it comes to intelligence. Attack helicopter crews are dependent on timely field intelligence and target information. Although the helicopter is generally able to find and locate a target, it is preferable to decrease its vulnerability by having the most recent data provide by so-called eyes on target. The less information a crew gets, the more time for preparation is needed, resulting in a slower reaction time. Most helicopter units work with minimum essential information lists or the like when it comes to employment in situations where they have to compromise. Usually, attack helicopters prefer to work at a minimum in pairs when the threat level is high.

Transport helicopters

Military transport helicopters have been designed for operations under adverse circumstances. The design usually comprises elements of protection, security and redundancy. Protection implies EW-suites, e.g. radar warning receivers, chaff and flares, on-board machine guns and Kevlar blankets. Security in systems has many faces, notably in communications: secure voice, GPS and other passive systems. Redundancy often means for example twin engines, double hydraulics systems separate from each another, multiple radios etc. All these qualities put together make helicopters true military systems that can operate independently or with minimum added escort.

Transport helicopters are the backbone of air mobility. The Chinook can lift 12 tonnes at sea level, move 25 troops with backpacks or three vehicles as external loads. The Cougar lifts 3.5 tonnes, 15 troops with backpacks or one external load.

In an optimum choice situation Chinooks will lift materiel and Cougars personnel. It is evident that the nature of the threat may require escort from fighters or attack helicopters.

Transport helicopters usually have a two-pilot crew, with the senior pilot acting as aircraft commander. The crew further consists of a load master/technician/gunner and often an assistant load master/gunner. Riggers/marshallers are qualified members of the ground team who prepare loads to pick up as under-slung loads and direct the crew to fly the helicopter to a position overhead the load, enabling the loadmaster to connect the load to the helicopter. They are part of the landing point team. Several landing points form one landing zone. Pick-up points allow the pick up of personnel and, again, several points form a pick-up zone.

Cooperation with fighter aircraft

ASFAO

Fixed-wing fighter aircraft in ASFAO operate under conditions of (limited) air superiority, which is prerequisite to the success of any land operation. In the land environment operations aimed at neutralizing, delaying or destroying enemy forces can be divided into CAS and Air Interdiction (AI). The latter does not need integration with fire and movements of own troops during its execution. However, the planning of air and land operations does need careful integration in order to be complementary. CAS takes place in the immediate vicinity of the own troops and therefore requires precise coordination with their fire and manoeuvre. It should be noted that CAS will only be called in if the fighting power of the own troops appears to be insufficient. Through its concentration and the surprise element, CAS can thus make the difference. Planning from the brigade level up is done by an Air Liaison Officer (ALO) who coordinates with the higher air echelon and internally with artillery and anti-aircraft artillery. The actual control in the field will be carried out by a Forward Air Controller (FAC), who will indicate the targets to the fixed-wing pilot.

JAAT

A specialized form of CAS are Joint Air Attack Team (JAAT) operations. Joint Air Attack Team (JAAT) implies joint and simultaneous employment of fighter aircraft, ground fire support and attack helicopters in order to attack a deployed mechanized unit. In most cases attack helicopters will be able to mark or suppress targets, e.g. command vehicles (antennas) or air defence vehicles (radars) from their positions. This will leave the tanks to the surprise and firepower of fixed-wing aircraft and artillery. Artillery observer tasks can be carried out by helicopter crews in order to minimize frictions. Some jet aircraft have been especially developed for JAAT types of missions, notably the A-10.

JSEAD

Joint Suppression of the Enemy Air Defences (JSEAD) must be considered an integral part of ASFAO operations and has as its objective to temporarily suppress, neutralize or destroy the enemy air defence systems or parts thereof. The latter can be achieved together with air support, artillery and EW-assets or a combination. Land commanders must be involved in the planning and carrying out of JSEAD, this is particularly important when considering CAS operations. The more JSEAD effort ground forces can provide, the better attack aircraft can concentrate on their primary ASFAO task. By the harassment and destruction of “soft” and mobile air defence targets, artillery in particular can significantly increase the effectiveness of CAS missions. Furthermore, attack helicopters in complementary operations can concentrate on the destruction of less conspicuous elements of the air defence chain. This 'preparation of the battlefield' will increase the chances of success for fixed-wing formations penetrating enemy airspace. It should be noted that forms of SEAD have to be applied consistently in order to limit the effectiveness of remaining elements in the air defences.

CSAR

Combat Search and Rescue is a combined mission where transport helicopters, attack helicopters and fixed-wing aircraft act together to rescue downed airmen from behind enemy lines. The mission may require action deep into enemy territory, which may exceed the standard range of the helicopter. Therefore CSAR helicopter assets are often equipped with air refuelling capability. To exploit surprise and increase chances of survival, missions are often carried out by night. CSAR helicopter assets are normally outfitted with additional equipment and armament enabling them to fly low at night and under adverse weather conditions. The role of fixed-wing air assets in this mission is to provide air cover and local air superiority. The CSAR mission being a rather complicated one, it is carried out in most armed forces by a dedicated unit.

Organizational build-up

Like every other expeditionary formation, air manoeuvre elements, after disembarkation, rally in an *assembly area* (AA). The stationary time in the AA will enable the unit to put together all the constituting elements. After this, the air manoeuvre unit will move to a *staging area* (SA), from where the *area of operations* (AO) can be reached. The SA serves as a base for the formation of task forces charged with the execution of the mission in the AO. Centrally located in the SA are the logistic elements of the squadrons. They have the least inherent protection for which the tactical units will compensate in an outer ring. In the SA all normal logistic tasks can be fulfilled, for which facilities will be built up, like FARPs and maintenance areas.

Conclusion

Air manoeuvre operations, although in terms of C2 and doctrine very much an army affair, are in the national context unthinkable without the involvement of the helicopters of the air force. Apart from the aim of occupying terrain, air manoeuvre units have identical characteristics to other airpower assets, of which flexibility, mobility and reach are the most prominent ones. More so than most other military operations, air manoeuvre operations are seldom characterized by simplicity. They are mostly high-risk missions and therefore demand a very careful planning of every single aspect. This is partly due to the complex logistics, partly to the fact that air manoeuvre is often carried out over and in enemy terrain. Helicopter operations are severely affected by considerations of weight, altitude and temperature. Attack and transport helicopters may be considered as complementary: attack helicopters escort transports and transports carry the ammunition (and in some case the fuel) that attack helicopters need. In many instances fire support by artillery and anti-surface force air operations by fighter aircraft assist air manoeuvre in the successful execution of their mission.

Chapter 16 – Operations other than war

Introduction

Military power and military equipment are not used solely to wage war or to prevent it, but can also perform many other tasks, such as disaster relief, emergency evacuation, fighting drug criminals, counter-terrorism, riot control, peace enforcement and peace support operations. Some of these tasks merely require the use of military equipment without military force. Disaster relief is a perfect example of such a task. It requires military transport, medical assistance and military engineering capability but not the use of force. On the other hand, counter-drug and counter-terrorism operations, let alone peace support and peace enforcement operations can require the use of (limited) force. These operations in the low spectrum of violence have increased considerably since the end of the Cold War. Low-intensity conflicts confront us with special problems, making the tactics and strategies used for war not practicable. In this chapter we will evaluate the possibilities of airpower in these types of conflicts.

Humanitarian operations

Natural disasters like the tsunami (Indian Ocean, Christmas 2004) and the hurricane Katrina (New Orleans, 2005) are good examples of the role military equipment can play in these situations. During humanitarian disasters the standard infrastructure is often unusable, which makes air transport the sole method for reaching the affected area. Especially large transport helicopters, like the Chinook, are capable of evacuating many people and bringing large quantities of relief goods to inaccessible areas. Tactical and strategic fixed-wing airlift can be used to bring relief goods, communication equipment, engineering equipment and even mobile hospitals from supporting countries to 'bare base' airports close to the disaster area. But it is not only airlift that matters, military reconnaissance aircraft are capable of reconnoitring the area.

Another factor is that humanitarian disasters are normally unexpected and require an immediate response. Military organizations, more than others, are trained for such immediate response operations.

A final consideration is that armed forces use their own support organization, so they can deploy and work without any third-party support.

Given the unique capabilities of airpower, it speak for itself that civilian authorities will depend on the military to perform these tasks. There is however a drawback. Military equipment is presently widely used in many operations around the globe, which limits the number of assets available. In an emergency situation it is quite possible to stop training and recall people from leave, but it is much harder to divert assets that are already employed in the support of troops involved in peace operations. So a common mistake is to expect too much. In real emergencies only a limited number of assets are readily deployable. If fighting natural disasters were to become a prime responsibility of the military, the total number of assets required should be re-evaluated.

Operations other than war requiring the use of force

Counter-terrorism and counter-insurgency warfare differ considerably from conventional warfare in terms of objectives, strategies, targets and weapons. They do however present a modern air force with similar problems. The enemy is very difficult to locate and even when he is visible it may only be for a short period of time.

Opponent capability

The non-conventional opponent has some characteristics that are worth mentioning. First of all, the knowledge of non-conventional adversaries is limited. Some observations about their equipment can be made, for instance that they may have old-fashioned weapons. This may be so, but slow-flying aircraft can be downed by simple rocket-propelled grenades and car bombs can be constructed from fertilizer and detonators. The weapons used in the 9/11 attacks were normal aeroplanes, but they were seized by using simple box cutters[139]. This is not the complete story. Thanks to the large advances in the field of communication these unconventional opponents, like everybody else, can communicate with each other from almost anywhere in the world, using internet and cellular phones. They can use these same phones to trigger explosive devices from a distance. By rapidly changing devices and using prepaid telephone cards they are able to hide their identity. Through the internet they have access to a lot of information and if they wish they can also communicate with each other in fairly secure ways. Encoding programmes can be downloaded freely. On top of that, the internet may also be used as an advertising and recruiting medium. And with the present cheap means of international transport, people can move quickly and freely as long as they are not linked to terrorist groups. In short, our open, highly technological society can be used to their advantage.

So it appears that our adversaries are presently using low-technology weapons in a high-technology environment. We should be even more careful when we make assumptions about 'the man'. There is not one homogeneous group and researchers have not come up with a generally accepted theory on why someone becomes a terrorist.[140] Therefore, it would be better to focus on how terrorists operate in order to expose their possible weaknesses[141]. We may assume that terrorists have a problem organizing and recruiting[142]. Recruiting is a long process that slowly draws people into a system. To recruit the right mix of personnel takes time. To train and test (in secrecy) is much harder than with a conventional army. This might be a reason why until now most terrorist have used relatively simple methods of attack. Even highly skilled personnel[143] require much training. Several specialists are needed to make more advanced weapons and the use of these requires even more training. History has shown many examples of terrorists who failed to produce adequate chemical and biological weapons. Making weapons of mass destruction is more complicated than is often suggested. It may well be that, because simple methods are still effective, the need for more complex approaches is not urgent.

139 David Clark Technology & Terrorism 2004
140 Friedland Becoming a Terrorist, p82.
141 John Horgan The Psychology of Terrorism 2005 p31
142 John Horgan The Psychology of Terrorism 2005 p96
143 Within the ranks of terrorist one can find many highly educated people, It's more common that highly educated people (who are deprived from political power) and relatively wealthy people revert to terrorism than the lower class 'have-nots' who are more concerned with there daily survival.

Air defence against unconventional opponents

The air defence network has always been the most sophisticated part of airpower. It was the first to use a network-centric approach with coupled radar stations, centralized command and control agencies that direct SAM batteries, and highly sophisticated fighter aircraft equipped with beyond-visual range missiles. However, the whole system is designed to fight an enemy from the outside, in which case the radars will give ample warning time to launch fighters. But the system is far less capable of countering a threat from within, as was proven on 9/11. The readiness state on 9/11 was not high, but even with today's higher alert states, defending against a 9/11 type attack remains difficult. Let us consider the physics first.

We will assume an aircraft in the air is hijacked and directed to a target 30 minutes' flying away. Assuming a speed of 0.8 Mach (standard), the aircraft will fly for approximately 240 nautical miles. Air defence has two options to counter this hijacked plane. The best result is achieved by having aircraft fly Combat Air Patrols and directing them to the threat if required. However, to have two fighter aircraft (the minimum) airborne 24 hours a day and 7 days a week is extremely expensive and would take around 20,000 flying hours a year[144]. Furthermore, a single CAP can only defend a limited airspace of around 200x200 nm.

A more economical way is to have aircraft on quick reaction alert (QRA) and scramble them only when needed. But scrambling aircraft takes precious minutes and calculations show that it is hard to intercept an aircraft that is past the scrambling base and flying away. On the other hand, if the hijacked aircraft flies straight towards the threat the situation is more favourable. The distance covered by the hijacked aeroplane and our interceptor can be added up, meaning that an interceptor can easily be scrambled against a target that is still 300 nm away and approaching. It will complete the intercept in 20 minutes at half distance, which will leave 10 minutes to decide what to do and get permission to fire if conditions so dictate. The latter example always applied in a classical war situation, where the enemy would be detected when flying towards our country and would be met head-on by our scrambled fighters. But in a terrorist scenario this will not work. First of all, it is not always clear what the hijacked aeroplane's target is and while we may make assumptions about the most likely target(s), we cannot reposition our airfields overnight. Theoretically, fighters should be located close to possible target(s), enabling them to meet the threat head-on. The bottom line is that fighters on alert can intercept some aircraft, but certainly not all aircraft and the shorter the time is from the hijack moment to the target impact the harder it will be.

A third way to defend the air space is to use SAM-batteries. This may be a tricky choice. There is no way to ascertain whether a suspect aircraft is experiencing navigational or communication problems. The only options are to shoot or hold fire. It is possible to have a no-fly zone for certain special areas and for certain time periods, but experience has shown that most real scrambles are against planes that are not hijacked at all but have some other kind of problem.

Air defence was never easy. In all classical wars, hostile aircraft have been able to penetrate air defence system. Air defence against a terrorist threat is not easy either. It is dangerous (when using SAMs), has limited effectiveness (when using QRA aircraft) or is cost prohibitive (when using CAPs). From this analysis it is clear that prevention is the only feasible option when trying to defend against a 9/11- type of scenario. We should also be aware that 9/11 was just one. Terrorists might invent

144 24 x 2 x 365 = 17,520; adding 15 % for moving to and from the CAP makes 20,148. The RNLAF produces around 25, 000 F-16 flying hours per year.

other ways to get control over an airliner or to terrorize airspace. Building a simple cruise missile is theoretically within reach. The detection of and the defence against such a threat pose a different kind of problem. It is difficult to detect and identify small and slow-flying objects. On the positive side, these objects will probably not carry a heavy payload.

Reconnaissance

Reconnaissance is the oldest use of airpower. The high ground has always been the favourite position from which to watch the battle and the use of balloons and aircraft greatly improved the quality of reconnaissance. During the Cold War, the most common reconnaissance platforms were fast combat aircraft equipped with cameras. Apart from a limited number of specialized aircraft such as the SR-71 and the U-2 (TR-2), which could fly unchallenged above (some) enemy territory, all other aircraft were vulnerable to enemy fighters and SAM and needed speed and agility. So, combat aircraft were used, but they had little endurance and could not loiter over the target area. Neither could reconnaissance satellites stay above the target area. The result was that all information gathered by reconnaissance assets was outdated by the time the aircraft returned. Low-intensity conflicts, where one party has air superiority, have however made it possible to employ a different kind of airframe: the reconnaissance Unmanned Air Vehicle. The combination of large wings, low speed and low weight in a theatre without a considerable threat, has made it possible to loiter over the target area sometimes for more than a day using a relatively simple and cheap UAV. The tactical consequence has been that operations have been able to upscale from reconnaissance to surveillance. No wonder Israel was one of the first users of UAV. Having air supremacy above its own country and part of southern Lebanon, the Israeli Defence Forces used simple UAV to spot incoming rocket attacks on its settlements in the north of Galilee. When a rocket firing was spotted, operators were able to track the shooters and follow them to their hideouts. Within half an hour of the attack, a combat aircraft equipped with laser-guided weapons would attack. This perfect example of networking and the power of surveillance has been emulated by American armed forces with the Predator UAV. To shorten the sensor-to-shooter cycle even further, some Predator-Bs have been equipped with Hellfires so that they can immediately react to a threat. This is an interesting development: whereas the UAV was initially a cheap and expendable aircraft, after integration of weapons into the system it has become more complex. The more expensive it is, the more the owner will seek to prevent it from crashing (which is what current UAV tend to do at a much higher rate than combat aircraft). To prevent it from crashing, the designer adds redundancy into the system, which will make it even more expensive and more complex.

Some limitations apply to UAV. At 15,000 feet you may be able to oversee an area of more than 20 square nautical miles, but the sensor will only look at a small spot at the time, typically an area of 500 by 500 feet[145]. The picture also needs to be interpreted. To carefully scan the complete area (a few seconds in each direction) will take 4 hours, so airborne surveillance will only work if the operators have additional information and thus will know where to look. If an explosion triggers the operators, they can immediately direct their cameras to the trouble spot. It will be hard to find any trace of the suspects of a car bomb that was positioned hours earlier with an explosion triggered from a distance. It will also be difficult to prevent

[145] That is slightly more than the size of a soccer field. It should be clear that this is probably the minimum detail an observer would like to see to scan an area. In many situations more detail is warranted but that will directly affect the total scanning time.

the taking of hostages, let alone stop suicide attacks. Only if the enemy has a limited area of operation and exposes himself, like for example during the insurgence in Fallujah, is the UAV able to make the difference. It is difficult to perceive a role for an UAV in the prevention of terrorist attacks like those that took place in London (2005) and Madrid (2004).

Ground attack

The final airpower role to discuss is ground attack. Airpower today is able to strike at any ground target with precision once it has been located and identified. During unconventional battles, this timely location is at the heart of the problem. In addition, ROE of the mission in many instances demand a plausible identification. Target lists such as exist in an all-out war are normally not available. An enemy's truck is hard to spot by air-to-ground surveillance if it does not stand out in the normal traffic flow. Also, instead of proceeding to a target from the list, attack helicopters and attack aircraft in low-threat scenarios are orbiting or on ground alert until they are called up to intervene. What is true for air defence combat air patrols also holds true for ground attack air patrols. They consume a large number of flying hours and are expensive to maintain. When on ground alert, the attack helicopter has an advantage over an attack aircraft. It is independent from an airfield, so it can be on ground alert much closer to the (expected) target area and can even hold with engines running (but with disengaged rotor) without consuming much fuel. Still, even in this waiting position it uses up engines hours. Given the empty battlefield, airpower is still the logical choice because of its ability to protect a large area with relatively few assets. Good and fast cooperation between spotters and shooters is a prerequisite for effective employment.

Because of the efforts involved in having expensive fighter bombers or attack helicopters loiter, it is logical to look into the option of arming UAV as well. If the UAV is already above the target area engaged in surveillance and also attacks, the sensor-to-shooter cycle is reduced considerably.

Considering the three airpower options to attack ground targets, each weapon system has its pros and cons:

- *Attack helicopter*: have a short response time because of their closeness to the target area. They can operate under conditions of low ceiling and reduced visibility. They are integrated into army procedures and can stay on alert for a reasonable period. However, they are vulnerable to shoulder-launched SAM, RPG and machine guns. They are also expensive and speed is low, limiting their use over a larger area.
- *UAV* have a very short response time and are relatively cheap. However, they can only operate safely at altitude so need a high ceiling and good visibility. Speed is low, which limits their use over a large area. The weight increase for armament and fire control will negatively influence their endurance. Experience with armed UAV is at this moment still limited and it will take development and experience before the full capabilities and restraints will be known.
- *Combat aircraft* have a high speed allowing them to cover a larger area, they can operate high and low and are relatively immune to ground-launched weapons. However, when not in orbit the response time can be long because they have to use well equipped airfields that are not always close to the target area. They are expensive, and when they operate at low level it is much harder for them to keep the target under surveillance. They need good visibility to operate safely at high speed.

Presently some research is done into the development of a dedicated aircraft for counter insurgency tasks. The aim here is to combine a relatively cheap turbo prop driven airframe with small weapons. It would not require a large runway, it would not be expensive to exploit and would be capable of delivering small precisions weapons in a low-threat environment. These plans are still only on the drawing board, but if the present emphasis on counter insurgency remains, they may well be implemented.

Special attention should also be given to the weapons used. It is possible that most weapons are still oversized for the role. If the price of stopping the insurgence in an urban environment is ruining the city, that price may be excessive and the resulting situation may backfire and trigger more unrest. There is a definite requirement for smaller and if possible non-lethal munitions or temporarily incapacitating weapons, so that collateral damage is much smaller. Precision alone is definitely not good enough. It should further be noted that although the employment of airpower in an urban environment in principle is not much different from its use in other environments like mountains, the essential difference is the presence of civilians, the hearts and minds of whom may be the centre of gravity of the whole mission.

Conclusion

Airpower can be a vital instrument for the government in humanitarian relief. In situations of national crises, air transportation and especially large transport helicopters will always be in high demand.

In operations against a non-conventional opponent the biggest problem for airpower (and ground power) is to detect the opponent in time. This forces airpower into a reactive position that is expensive to maintain. On the other hand, in an asymmetric war the conventional force normally owns the sky. This gives him freedom of movement and observation. Furthermore, airpower is still the asset with the best response time in an 'empty' battle field. So, although expensive, it is still the most effective means of fire support for the man on the ground.

Chapter 17 - Support operations

Introduction

Airpower cannot achieve its objectives without the support of transport, tankers, surveillance and reconnaissance aircraft. Since the end of the Cold War, out-of-area operations became the common modus operandi and the demand on tankers and transport increased tremendously.

> *To give just a few numbers, air transport carried out 18,466 sorties during the deployment and redeployment of Desert Storm, transporting some 482,000 passengers and 513,000 tons of cargo. During Desert Storm some 110,000 combat sorties where flown, 15,434 sorties were tanker sorties. During those sorties 45,955 aircraft were refuelled. Of the 110,000 combat sorties less than half were dedicated to attack missions (42,600).*

During the Cold War most of tanker assets were used solely by Strategic Air Command to give global reach to its strategic bomber force. Most fighters and fighter bombers were located close to the border and could easily operate without tanker support. The only transportation needed was for bringing reinforcements from the USA to the theatre. Not surprisingly the USA maintained a sizeable transport force while the European NATO partners only possessed a limited transport capacity mostly for tactical use. The table below depicts the slightly improved situation in 2001.

Aircraft type	EU (total)	USAF
Strategic transport	55	338
Tactical transport	352	519
Tankers	72	606

We call air transport a support operation, and most transport operations are, but it is not completely accurate. Sometimes air transport is the operation itself. The classic example is probably the Berlin Air Bridge, which saved the city of Berlin from starvation during the Russian blockade in 1948, showing that transportation can be a mission in its own right.

Another support operation is surveillance and reconnaissance, which were the first roles of airpower in World War I and still applicable in the 21st century. The emphasis however has shifted from reconnaissance to surveillance, which requires long-endurance aircraft and a more permissive environment (air superiority).

There are some other support operations that will not be covered in this chapter. First of all, the EW support (SEAD and jammers) was already covered in chapter 10. Secondly the specialized work of Combat Search and Rescue aircraft has been covered in chapter 15.

Air-transport

Definitions

Air transport is divided into two categories: strategic and tactical. Strategic transport brings passengers and cargo to the theatre while tactical transport aircraft moves passengers and cargo inside the theatre. Strategic transport aircraft are larger, have intercontinental capability (>2500 nm), use normal runways and seldom have self-protection equipment. Tactical aircraft on the other hand are smaller, can use non-hardened runways, are equipped with self-protection equipment and often have a more limited range. Typical strategic transport aircraft are the C-141 and the C-5; the most common tactical transport aircraft is the C-130. The C-17 is a modern and capable hybrid.

Why strategic transport?

Compared to the volume of traditional civil air transportation, the military transport effort is almost negligible. Considering the fact that both strategic airlift and civilian airlift use the same type of airport, the same size aeroplane and operate over the same distance, a reasonable question is: why not use this massive potential instead of maintaining a military strategic air transport capability? First of all, the military does use civil transport. The USA has its Commercial Reserve Air Force (CRAF); this service was activated during operation Desert Shield and commercial aircraft transported about 60% of troops and 27% of cargo airlifted to the Middle East. However, most civilian carriers cannot transport outsized cargo or heavy equipment like tanks because they use standard wide-body aircraft like the Boeing 747 for cargo transport, which are not outfitted with ramps. Furthermore, heavy and outsized equipment is normally transported by ships. Another problem is that civilian operators may be reluctant to divert a large portion of their transport capacity to serve the military instead of serving their regular customers. There is a fair chance that the regular customers will permanently divert to another operator, so serving the military can negatively affect their competitiveness. A third reason is a simple economic one: to have at least some strategic airlift capability can be used as price leverage.

What is special about tactical air transport?

Tactical air transport is a typically military operation requiring specialized aircraft capable of performing tasks seldom performed by civil operators. They include:

Operating from unprepared and short runways. To enable this type of operation requires an aircraft with a low landing speed, a firm undercarriage with many (low pressure) tires, high positioned engines to prevent dirt ingestion and consequently also a high wing[146].

Operating over hostile territory. To protect the aircraft over hostile territory first of all requires the aircraft to be equipped with self-protection equipment against GBAD. Secondly the aircraft should be able to load and unload very rapidly to shorten the time on the ground when it is extremely vulnerable. To accomplish this, an aircraft with a high wing and a roll-on-roll-off capability is ideal. Thirdly, the aircraft can be equipped with Kevlar blankets to protect its occupants against small arms fire. Fourthly, the aircraft should be able to operate low level, preferably during both day and night. This latter option would require the use of Night Vision Goggles or Forward Looking Infrared equipment.

[146] See also chapter 2-platforms.

Paradrop. To enable an aircraft to perform Para drop requires several modifications. First of all, the pilot needs a system to calculate the exact drop point based on drop height, type of parachute and the prevailing wind. Secondly the aircraft should be equipped with special Para doors, which can be opened safely in flight and have a deflection structure that routes the air in a favourable way past the opening. The rear ramp is also an excellent opening to be used for Para drop. Furthermore, the aircraft should be equipped with a warning system for the paratroopers. For Para drop at altitude a high-quality oxygen system is required. Dropping paratroopers is not possible in all weather conditions, with especially the wind as a limiting factor. Normally the speed limit for airdrop is 18 knots and even in worse wind conditions the expected number of wounded paratroopers is significant. The normal expected casualty rate is 2% by day and 5% by night. In adverse wind conditions these figures will be higher. A special requirement for the dropping zone is that it should be safe from enemy actions for at least 20-30 minutes, which is the time needed for the paratroopers to assemble by day. By night it may require as much as 40-60 minutes.

Cargo drop. There are reasons for dropping air cargo and several ways of doing it. The reasons for dropping cargo instead of landing are straightforward. There may not be an adequate runway or the situation on the ground may be too dangerous for landing. There are several ways to airdrop cargo. It can be delivered by parachute from medium altitude. The advantage of this method for the aircraft is that it is well outside the reach of small arms fire, but the delivery accuracy is rather low. A new development is to equip the cargo with steerable parachutes and a GPS system[147]. Another method is that the aircraft approaches the drop zone at very low altitude (10-20 ft) at low speed (110 knots) and that the cargo is extracted from the aircraft by parachute. This delivery method is called LAPES, which stands for Low Altitude Parachute Extraction System. The drop zone should be reasonably flat to allow the aircraft to fly low safely and the absence of ground structures will permit the cargo to decelerate without colliding. The aircraft will perform this delivery with gear down to prevent an unintended ground contact causing damage to the aircraft. This method is not without risk: the extraction of the cargo causes a rapid centre-of- gravity shift bringing on a pitch input for which the pilot must correct immediately. A variation to this technique is that the aircraft approaches at a small dive angle, very close to the ground, the aircraft is then pulled up and this climb attitude causes the cargo to glide outside.

Transport requirement

To deploy a combat unit requires a large transport capacity. The following table is an example of transport requirements for a number of Dutch units. The underlying presumption for these numbers is that a reasonable host nation support is available. If host nation support is lacking, some of these numbers may even double. Transport capacity required is expressed in passengers (pax) and lane metres for a roll-on-roll-off ship. Generally speaking a mid sized roll-on-roll-off ship has a capacity of 2000 lane metres. For complete transportation by air very large aircraft like the C-5 or the An-124 are needed to carry all trucks and tanks. The An-124 has a cargo floor length of 36.5 metres and a width of 6.4 metres. So, considering this width of 6.4 metres and a two-lane capability an AN-124 has a capacity of approximately 70 lane metres.

147 The SPADES system for example, developed by Fokker Space industries in the Netherlands.

	Air transport		Sea transport		***If sea transport is done by air***
Unit	Pax	Flights KDC-10	Lane metres	RO/RO ships	***Flights AN-124***
Marine battalion	520	4	1040	1	***15***
Patriot squadron	245	2	2700	2	***39***
Tactical Helicopter group	2171	14	5500	3	***79***
Armoured Infantry battalion	682	5	4150	3	***59***
F-16 squadron	352	3	2500	2	***36***
Humanitarian battalion	1200	8	2600	2	***37***

The choice of transport

The choice between sea transport, land transport and air transport is not always easy. We will examine the different modes of transport.

Sea transport is considered the cheapest mode of transport, because personnel required is limited[148], fuel consumption is low and ships are able to carry a substantial cargo. However, this is not necessarily so. The classic example is passenger transport. Flying passengers by aircraft only requires a cabin crew to serve some meals. For transport by ship, passengers need full hotel accommodation for several days, which may be costly even if the fuel costs per passenger are lower on a ship.

Cargo transport ships require the smallest amount of fuel per transported ton and carry the largest load. But when comparing various modes of transport, it should be realised that transport capacity is load times speed. As is shown in the following table, the capacity of a C-17 approaches that of a mid-sized RO/RO ship in theory. In reality, the transport capacity may differ considerably. Ships can normally not proceed directly from port to port in a straight line but have to follow coastlines. Aircraft might also not be granted flyover rights, forcing them to follow a less than optimal route.

Transport	**Load** (tons)	**Speed** (knots)	**Capacity** (ton nautical miles/hr)
C-17	76	450	34200
Truck	10-20	40	400-800
Train	800	40	32000
RO/RO ship (2500LM)	3000	16	54000
Containership	120,000	18	2,700,000

148 Some large containerships have a crew of less than 20 men.

Strategic transport is just one part in the transport chain. To make a correct comparison, both strategic and tactical transport have to be taken into account, including possible trans-shipment delays. This makes each transport situation unique, requiring its own analysis in order to determine the optimal capacity.

Apart from fuel efficiency and transport capacity, there are other important factors. Not all destinations can be reached by ship. Operations in landlocked countries like Afghanistan show the limitations of sea transport. In situations of time pressure, air transport is also the required mode of transport. Finally, there are also safety factors involved. In hostile territory, transport by truck and train is particularly dangerous, as it poses an attractive target for terrorist who might use mines, car bombs and improvised explosives to disrupt the flow of goods. Transport aircraft, and especially tactical transport aircraft, can bring goods close to the final destination speedily and relatively safely. This combination often gives air transport a significant advantage over all other modes of transportation. It is therefore no wonder that the demand for transport aircraft has grown considerably over the past years.

Tankers

Why tankers

The use of tankers has become so commonplace in current military operations that the question as to why we use them is seldom addressed. Two decades ago Burt Rutan proved that an aircraft could fly around the world without refuelling. The Global Hawk can even remain airborne for several days, so why are aircraft not carrying their own fuel? The problem is that flying, and especially flying at high speed, requires a lot of fuel. Carrying extra fuel for more range increases weight, thereby increasing induced drag, which will in turn increase fuel consumption over the whole mission. So each aircraft design has a certain point where adding extra fuel will not increase but decrease range! If an aircraft design must allow for an extremely long range, this can only be achieved by using a low-speed aircraft with large wings. These aircraft will not be able to withstand high g-forces and are not useable for every type of military operation, with the exception of long-endurance surveillance.

Ways of refuelling

There are two systems for air refuelling: the probe and drogue system, and the boom. The probe and drogue system is the oldest system. It consists of a hose with a drogue at the end to stabilize it behind the tanker and to give an aft force on the hose. The receiver is equipped with a probe, which fits into the receiving end. When the receiver pushes the drogue forward, the pressure will open a valve and fuel will start to flow. The advantage of a drogue system is that a tanker can refuel multiple receivers at the same time, but the transfer speed is limited to approximately 2500 lbs/min. The second system is the boom system, developed by Boeing in 1950. The advantage of the boom is that the maximum transfer speed is much higher, approximately 9000lbs per minute. A second advantage is that the pilot of the receiver has more manoeuvring room. He does not have to make the connection, the 'boomer' in the tanker does the hook-up. The pilot in the receiver aircraft is directed to the correct position by receiver lights, which are installed under the belly of the receiver. When the hook-up is made the directing lights are controlled by the position of the boom. During air refuelling the receiver intercom system is connected to the boomer. There are two different boom systems. The first is the Boeing KC-135 boom system, which is hydraulically controlled and the second is the boom developed by McDonnell Douglas for the KC-10 which is controlled by airfoils.

There are three ways for an aircraft to be refuelled. The first is flying in formation with the tanker (fighter drag). This is applicable for most ferry flights. The maximum number of fighters that can pair up with a tanker is ± 6. A ferry flight however always has to take into account the possibility of a tanker contingency, such as a refuelling system breakdown. Especially over the ocean this can cause a critical situation for a ferry flight. Should this happen, all fighters in the fighter drag must be able to divert to a suitable alternate airfield. Therefore, fighters in a fighter drag will never wait to refuel until they are almost empty, but will refuel early to maintain an adequate reserve. A ferry flight can also meet the tanker at a predetermined refuelling point. The tanker flies in a pre-arranged air refuelling orbit and after refuelling the receivers proceed to their destination or to the next refuelling point. The disadvantage of this method is that the trip will last longer for the receivers, as they have to spend some time in the orbit. The advantage is that, as the tanker only has to fly to and from the orbit and not cover the whole route, he has more fuel available for the receivers. On war missions, the tanker orbit is the common procedure, bombers will receive extra fuel before they enter a hostile area and may also have to refuel on the way back.

Joining up with a tanker in orbit is not always easy. For a fighter-bomber with a large manoeuvring envelope this may be relatively easy, but tankers also refuel bombers and other large aircraft with a small manoeuvring envelope. It is also possible that this refuelling has to be carried out when there is low flight visibility. Because the tanker is equipped with a TACAN the receiver has an indication of the position of the tanker even without radar.

Important to note is that air refuelling is costly. As a rule of thumb: it takes a pound of fuel to bring a pound of fuel into orbit.

Reconnaissance and surveillance

Reconnaissance was the earliest application of airpower. It was first used on the 28th of June 1794 during the battle of Fleuris, when the French Captain Coutelle used a balloon to spot the Prussian positions and relay this information to General Jourdan. Historians claim that this information was instrumental to the victory of the French over the Prussians. For the first time in history man could use the high ground to reconnoitre at will. Aeroplanes and satellites have now made it possible to spot almost anything anywhere. That does not mean that there are no problems in the field of reconnaissance and observation, for intelligence always require more than can technically be delivered. The heart of the problem is the nature of the enemy. If the enemy is a country, it is relatively easy to gather data about its military installations, its infrastructure and to assess its military strength, given enough time.[149] But when fighting insurgence and terrorism, gathering data becomes an order of magnitude more difficult. The exposure of the enemy may be extremely short, he is not bound by geographical limits, he may be disguised as civilian and may use weapon systems that are hard to identify.

The amount of data that has to be gathered is therefore enormous and some pre-selection has to be done. Also one has to realise that the vast majority of all

149 Even under those circumstances intelligence can be insufficient. Recently the CIA was wrong about weapons of mass destruction in Iraq. The largest and best funded intelligence organization in the world also failed to foresee the fall of the Berlin Wall in 1989, the Russian atomic bomb in 1949 and was wrong about the effect of the invasion in Cuba in 1962.

gathered intelligence data is useless and finding the few important facts will require a lot of work. Finally, all data needs to be interpreted. Interpretation is mostly done by trained humans and consequently takes up a considerable amount of time. Some help may be expected from intelligent computer programmes that are able to recognize certain objects or activities, but the task remains enormous.

One of the central questions in reconnaissance is: what area has to be covered and at what resolution. The resolution is dependent on the purpose, for example if an enemy has a new radar and technical intelligence wants a close-up of the wave tube[150]. In order to estimate in which frequency band the radar is working, a picture with centimetre accuracy is needed. On the other hand, if the requirement is just to spot a column of armoured trucks, a resolution of 30 cm could suffice. The lower the resolution the larger the area that can be surveyed at the same time, so it is advantageous to move to the lowest acceptable resolution in order to watch a larger area.

A large part of time-sensitive data gathered will be disseminated via satellites. Here another limitation surfaces: the available bandwidth. It is not possible to transmit to a satellite with an infrared laser, you have to use UHF or microwave data link. The data capacity of a communication satellite may seem large (comparable to a hundred or more TV channels), but there are reconnaissance sensors that are able to gather more data than the input of several TV channels, so a few high-capacity reconnaissance assets can easily consume all available bandwidth. However, in the near future some help may come from onboard analysis and compressions techniques.

Types of intelligence

One can distinguish between tactical and strategic intelligence. Strategic intelligence is concerned with the capabilities and intent of the opponent. This data is collected mostly by satellites, espionage aircraft like the TR-1 and by human intelligence, and mostly during peace time.

Tactical reconnaissance is carried out to support the ongoing battle and is gathered during hostilities. Although this type of reconnaissance is mostly carried out by tactical reconnaissance aircraft that have a good chance of survival over the battle area, when air superiority is quickly gained or not challenged at all, UAVs are a better choice. The main advantage of UAV over combat aircraft is their long endurance. As explained in chapter 3, being unmanned is essential for long endurance. A tactical benefit of UAVs is that they can be used for surveillance as well.

Surveillance can be carried out in two ways. If the threat is low, the platforms can operate over the battle area. Surveillance can also be carried out from a safe distance (and in own airspace) if the sensor range is large enough. AWACS and J-STARSs (both radar sensors!) fall in the last category; visual and IR sensors should be closer to the target for optimal performance.

Sensors

All sensors described in chapter 5 are used for surveillance and reconnaissance. They include air surveillance radar, ground surveillance radar (SAR), infra red (IR) sensors and electro optical (EO) sensors. Furthermore, electronic intelligence can be gathered by intercepting radar, radio, television and mobile telephone transmissions. As explained in chapters 4 and 5 the resolution of the sensors

150 The wave tube guides the radar energy from transmitter to antenna. The size of the wave tube depends on the radar frequency.

depends on their aperture, the kind of sensor (primarily the wave length) and distance. The highest resolution is achieved with EO sensors and IR comes second. SAR resolution is normally the weakest and around 1 ft. On the other hand, SAR resolution is not dependent on distance while IR and EO are.

Platforms

Most of the sensor platforms are described in detail in chapter 4. In this paragraph we will repeat the most common platforms with their most important parameters.

Satellites employ IR and EO sensors and operate in low orbits (200-600km). Their time over target is normally limited to several minutes per orbit. Some satellites have an orbit that will revisit the same spot every 86 minutes. Other orbits may have revisiting times of up to two weeks.

The **AWACS** is the standard USA and NATO airborne surveillance platform. It is based on the old civilian Boeing 707 airframe. Its sensor range exceeds 200 nm. The AWACS can be air refuelled to increase orbit time.

JSTARS. Is the best known radar ground surveillance platform used by the USAF and was first operational in Desert Storm. The JSTAR uses a SAR, enabling it to distinguish ground targets with a resolution of less than 1 ft in all weather conditions. European variants for ground surveillance radar include the French Horizon system (helicopter), the Italian CRESO system and the English ASTOR system. Presently many European NATO partners are working together in the CEASAR project to get a common ground surveillance radar capability

The **TR-1** is a further development of the U-2. It is a strategic reconnaissance aircraft that can fly at ±80,000 ft. It is equipped with IR, EO and SAR sensors. Until ground-based air defence missile systems became widespread the system was immune to enemy attack.

The **Global Hawk** is the most recent unmanned strategic reconnaissance platform. It has EO, IR and SAR sensor, flies at ±65,000ft and can stay airborne for 42 hours.

The **Predator** is presently the most used tactical reconnaissance sensor platform. It has EO, IR and SAR sensors. The Predator-B can be equipped with Hellfire missiles as well. The Predator is designed to loiter for 40+ hours and to operate at 15,000 ft.

Tactical reconnaissance aircraft have been the backbone of the tactical reconnaissance effort during the Cold War. Presently there are still many nations that equip their tactical aircraft with a reconnaissance pod to perform tactical reconnaissance. The advantage of a tactical reconnaissance aircraft is that it can cover a vast area in a short time and is able to survive in a hostile environment. The drawback is that, without a data link, it takes time before the aircraft is back at its home base and the data can be processed. With the present high speed of operations the data may have lost its relevance by the time it is processed.

The **RC135-R Rivet Joint** is an ESM reconnaissance aircraft, which is based on the civilian Boeing 707. Its mission is to identify all types of transmitters.

Methods

There are several ways to employ a reconnaissance platform. Fast assets can carry out an area reconnaissance. The faster the aircraft, the larger the area that can be mapped. The satellite is probably capable of mapping the largest area. A different method is to send the platform to different points of interest that have to be surveyed (For instance TBM sites, airfields etc.). It is also possible to reconnoitre along certain line features: roads, railroads or waterways, to spot any military activity. One important form of reconnaissance is post-strike reconnaissance, which is necessary to perform battle damage assessment (BDA). Planners like to know how well the strike package has done its job and whether some targets need to be re-attacked. The reconnaissance assets may be part of the strike package and follow the bombers at a short distance. This distance should be large enough to allow debris and clouds caused by the explosions to settle, but also short enough to stay close to the protecting escort aircraft. The slower reconnaissance assets (e.g. the Predator) are mostly employed in a surveillance role and will more or less stay in the same area.

Future outlook

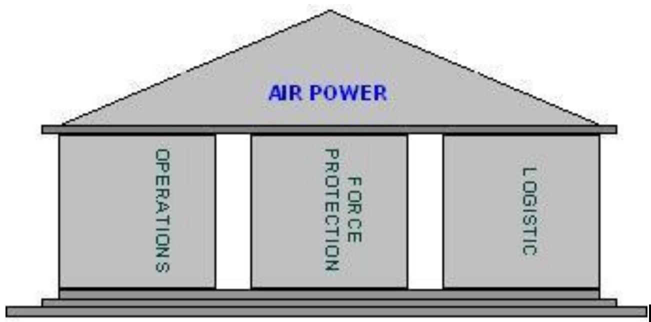
Over the past decades the reconnaissance situation has changed dramatically. One of the most striking changes is that sensors are no longer scarce. Every modern combat aircraft is equipped with radar and quite often with a targeting pod as well, so in the near future every flying platform is a potential reconnaissance source. The 'only' thing to do is link them in a network. In some fields this is already achieved to a certain extent. With Link-16 the fighter back link provides AWACS with the (more accurate) position of targets as determined by the fighter radar and this is integrated into the recognized air picture. In principle it would even be possible to turn the AWACS radar off and just combine the picture from the fighters. A further but not yet exploited advance in technology could emerge if every sensor were automatically scanning that part of the sky most beneficial to the overall picture. In the ground reconnaissance arena there is not one common data network for all ground sensors. A big problem is the immense amount of data collected by all ground observation sensors. To share that data in one common network is a big challenge. Techniques for on-line compression will improve over time. Furthermore, onboard processing should sift information from data; a possible method could be to compare ground pictures with stored data and only transmit changes. It should also be noted that every individual soldier on the ground could be considered a sensor and these ground sensors (especially in operations other than war) should also be incorporated. Even a massive number of integrated ground sensors would not lead to sufficient vital information, especially in operations other than war. This was covered in more detail in chapter 16.

Chapter 18 - Force protection

Introduction

"You have to be lucky all the time, the opposing forces have to be lucky only once"

Airpower is more than the employment of jets, helicopters and guided missiles. Operations and logistics cannot perform without force protection. Therefore, airpower comprises three interlinked building blocks. Force protection, together with logistics, must be seen as enablers of operations and thus prerequisites for the use of airpower. This chapter deals with the most common aspects of Force Protection (FP).



Every opponent must be expected to target our forces and our security measures must prevent this from happening. For good reasons security is one of the principles of war, protection being a key component of security. A fundamental principle is that military units at any level must be able to defend and protect themselves. FP therefore is a responsibility of all commanding officers, whereby the staff assignment is usually carried out by a FP specialist, despite the basic duty of all military personnel in this field. NATO lists FP as an Essential Operating Capability (EOC).[151]

"Survivability and Force Protection to minimize the effects of any adversarial action, to include the effects of WMD, whilst ensuring Allied Forces freedom of action and force effectiveness"[152]

It is self-evident that the protection of fielded forces is a key factor in maintaining public and political support for any mission.

What is force protection

Firstly we have to understand what is meant by FP. Most of the time nations operate in a mission that is led by NATO, the European Union (EU) or in a coalition of the able and the willing. Usually these coalitions use NATO procedures and doctrine. Therefore in this chapter the Allied Joint Publication (AJP) 3-14, dealing with NATO FP doctrine is used to cover the theoretical part of FP.

151 Allied Joint Publication 3-14 Force Protection, page 1-1
152 Military Committee Document 400/2

NATO defines FP as:

> *"Measures and means to minimize the vulnerability of personnel, facilities, material, operations and activities from threats and hazards in order to preserve freedom of action and operational effectiveness thereby contributing to mission success."*[153]

Whilst FP aims to protect both operation and personnel, the level of risk the latter have to accept has to be balanced against FP. This is always a dilemma; a 100% protection will frustrate operations; insufficient FP endangers personnel and weapon systems and therefore, in the long run, operations as well. To strike the right balance between the two is easier said than done. FP is applicable from deployment, through operational employment, and into redeployment.

FP can be seen as a system. The first step is to deprive the opponent of his will to attack (influence the mental component); the second step is to limit his opportunities for attack (irregular use of routes, departure times, OPSEC), the third step is to counteract his capabilities (SHORAD, Electronic Warfare) and the final step is to minimize the consequences of an enemy attack (protective clothing, bunkers, revetments).[154] It is important to realize that in a conflict not only the physical component but also the mental component will be attacked. Therefore FP must also protect this component of our forces.[155] By doing this we enter into the domain of Information Operations (IO); this overlap has not yet been fully developed.

FP can be divided into two areas, notably active and passive measures. Active measures contribute to FP with traditional military tasks: deter, prevent, nullify, or reduce the effectiveness of an adversary's attack. Passive measures involve the steps taken for the physical defence and protection of personnel, essential installations and equipment in order to minimize the effectiveness of hostile action. Total FP is unachievable and unaffordable, even in the most benign environment. NATO FP must therefore be based upon effective risk management - that of minimizing risk to forces, not eliminating it. It is thus not possible to protect every asset against all kinds of threat, all of the time.

Threat and risk analysis

The necessary level of FP depends on the threat a unit faces. A commander can estimate the threat to his unit based on the intelligence he receives from various sources: national intelligence, security organizations of the host nation, police organizations, etc. Counter Intelligence (CI) organizations are in charge of a general threat analysis specific to a mission or location.

To make it easier to work with threat analysis, generic threat levels are used for the purpose of definition. Based on the escalation from small civil disorder to major military conflicts (war), three threat levels are defined: low, medium and high.[156]

Low

In peacetime and during Peace Support and Crisis Response Operations (PSO/CRO), this is the lowest threat level defined. A general threat exists and there is a risk of peacetime incidents (such as accidents and fires) through civil disorder up to sabotage and attacks by terrorist organizations, normally small units. Air

153 Allied Joint Publication 3-14 Force Protection page 1-1 (ratification draft)
154 UK Joint Doctrine Pamphlet 1/99, Force Protection in Joint Operations.
155 Landmacht Doctrine Publicatie 1, Militaire Doctrine, 1996, Koninklijke Landmacht, hoofdstuk 7.
156 Allied Joint Publication 3-14 Force Protection, page 3-1.

attacks and the use of nuclear weapons are not expected. However, terrorists may use biological and chemical weapons or large quantities of explosives.

Medium

The medium threat level recognizes the existence of a more localised threat without the specific nature, target or timing being defined. It envisages a short-duration engagement. Attacks can include the simultaneous interdiction of several NATO facilities by tactical air, land or maritime forces. Forward NATO formations and areas may be attacked by conventional weapons against vital facilities, and may include the use of cruise and/or ballistic missiles. The threat ranges from terrorism, through attacks by small but proficient and well-armed aggressors, up to and including the use of Special Operations Forces (SOF). The use of nuclear weapons remains extremely unlikely, although chemical and biological weapons, radiological or fission-grade nuclear material and high explosives could be employed.

High

The high-level threat environment assumes that a specific threat exists or that an incident has occurred. The threat may encompass the full range of attack options from widespread terrorism through to large-scale assaults. Though it is highly unlikely that nuclear weapons are employed, the risks posed by environmental hazards, including ROTA (Release Other Than Attack); B/C contaminations caused by industrial disasters and NBC warfare remain.

Threat analysis/operational risk management

Threat analysis calculates the threat against own assets, enemy target analysis identifies and prioritises own vital assets, vulnerability analysis assesses their weaknesses. Commanders determine the vulnerability of their facilities and operations through a vulnerability assessment (VA) process. The question here is what has to be protected, against what kind of threat; to what extent FP will impede the continuation of operations and how much money is involved. This process is called Operational Risk Management (ORM) This is a reiterative process of identifying and assessing risks, and implementing reasonable controls to reduce the risk of threat and hazards. Commanders should furthermore continuously evaluate their FP measures in order to ascertain whether they are still in line with the current threat and operations.

Prioritisation

FP must embrace the whole force, but it is unlikely that there is enough capability to protect all force elements to the same level. Priority must therefore be given to C2 facilities, main weapon platforms and concentrations of personnel and CIS equipment.

Host nation

When operations are outside the own country, FP measures should be coordinated with the Host Nation (HN). Preferably, already existing FP structures should be used and only complemented when required. It is often cheaper to procure heavy and voluminous materiel (concrete, barbed wire, steel plates, etc.) in the HN if available. However, FP remains a national responsibility.

Command and control

Because missions are mostly combined, FP measures should be coordinated between participating countries. Sometimes there is a difference of opinion about threat levels and required FP measures. Not all nations interpret the threat in the same way and therefore implement different FP measures. Ideally FP measures should be the same at one location.

In Crisis Response Operations (CRO) nations sometimes show a different perception of the threat level and accompanying FP measures; e.g. against the same threat of small arms fire, some nations wear full protective clothing (helmet, flak jacket) and use armoured vehicles for movements; other nations use open vehicles and wear their protective clothing only if urgently required. Both manners however give an inconsistent perception to the population. This must not be tolerated by the Force Commander.

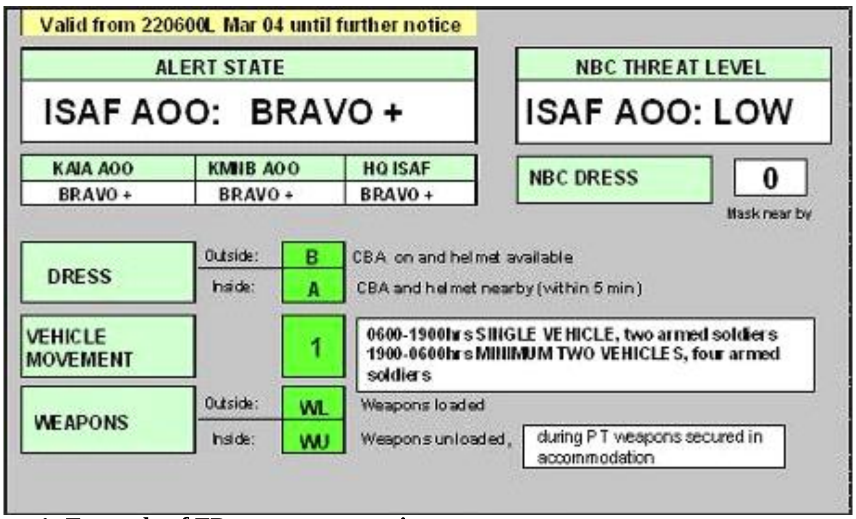


Figure 1: Example of FP measures overview.

Force protection components

While not all-inclusive, the essential component pillars of FP are illustrated in the figure below.

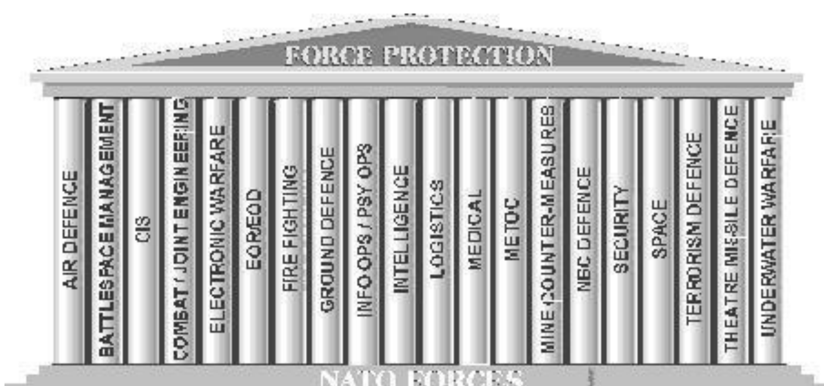


Figure 2: FP Component Pillars [157]

157 BiSC Force Protection Directive 80-25

FP covers a wide range of activities. In peacetime, FP measures will essentially be those of Protective Security and those services provided by the emergency services (i.e. police, fire, rescue and medical). However, as tensions rise and the forces move through Peace Support/Enforcement Operations to war fighting, the needs of FP will evolve to include the more traditional war fighting capabilities. FP, however, may be viewed throughout the entire spectrum in four capability areas, that of Protective Security, Active Defence, Passive Defence and Recuperation. FP measures are normally carried out by personnel in a primary or secondary function.

Protective security

Protective Security is defined as the organized system of defensive measures with the aim of achieving and maintaining security. It consists of physical security, information security (e.g. information on a need-to-know basis) and personnel security. It is proactive and preventative by nature. Such programmes are also related to programmes for fire protection, law enforcement, environmental health & safety, medical and road/traffic safety. As a unit moves through crises to conflicts, war fighting elements of FP will increasingly apply; however, the basic elements of security remain an integral part of FP. Examples of Protective Security are:

- (Mobile) fences, gates, concertinas ;
- Entry control, illumination;
- Splinter protection (HESCO's, sandbags, earth walls, Kevlar blankets, etc);
- Fire extinguishers, fire plans, etc.;
- Counter Intelligence (CI) is a key measure in this area.



Figure 3: Observation denial Bullet protection Anti rocket screen

The protection is organized as a layered defence, inside and outside a compound or airbase. The inner defence is carried out by policing (patrolling and access measures) and component defence activities mostly by dedicated unit personnel or as a secondary task. The outer defence is performed by manoeuvre-related activities (infantry tasks) and is considered to be an active defence measure (see the next paragraph). The aim is to prevent unauthorized personnel from accessing vital or key areas. A system must be in place to detect and delay the enemy in order to give ample time to defend the location when the need arises.

Active Defence

Active Defence involves measures to deter, prevent, nullify or reduce the effectiveness of an enemy attack, including defence against surface air and missile attack. It is about taking the battle to the aggressor and either denying his hostile intent or neutralizing his ability to attack or pose a viable threat. Active Defence measures include Surface-to-Air Missiles and TMD (as described in chapter 12), Short Range Air Defence (SHORAD), dedicated ground combat assets and control of access/entry. Ground defence should also be organized as a system. By bringing depth into the defence, the effects of opposing forces' fire are degraded. Entry to vital areas must be controlled, intruders have to be detected. Once detected, there must be a response by a reaction team. Vital areas and equipment are surrounded by time-delaying infrastructure (locks, armoured glass, walls, solid doors, etc.), the implication being that an intruder will take a certain time to break through the delaying infrastructure. The delay time should not exceed the response time of a reaction team. Examples of Active Defence are:

- Foot/motorized patrols, (dog) guards, Quick Reaction Team/Force (QRT/F);
- Securing convoys;
- VIP protection;
- Local air defence by STINGER.

A 100% protection of a location against intrusion by opposing forces, especially during darkness is a very difficult job. The size of an airbase is considerable. The runway often has a length of approximately 3000 metres. The fence around an airbase is often more than 10 km long. An observation post can normally cover approximately 100 metres as a field of fire, assuming the availability of night vision equipment. Covering the fence with observation posts would take a lot of military personnel. This problem can be solved by using automated electronic detection devices. An alternative or combination is found in guarding and defending only the vital areas, weapon systems, mission-essential equipment or personnel.

Fig. 4: Dog handler

STINGER

An example of a SHORAD system is the STINGER. This weapon can be fired from the shoulder and has a short range (max 10 km). It is a shoot-and-forget weapon, which means that once launched, the missile seeks the target autonomously.

Figure 5: Active Ground Defence

Passive Defence

Passive Defence involves measures to protect own personnel and essential equipment against conventional and NBC weapons. Passive Defence implies the physical protection of personnel and materiel from all forms of attack. It also includes camouflage, concealment, dispersal and Individual Common Core Skills (ICCS). Examples of Passive Defence measures are:

- Establishing NBC recce and decontamination teams;
- Establishing toxin-free areas;
- Providing NBC protective clothing;
- Defence Against Mortar Attack (DAMA);
- Decoys;
- Counter IED measures;
- Civil and Military Cooperation (CIMIC);
- Psychological Operations (PSYOPS).

During Operation Allied Force allied air forces targeted Serbian dispersed air defence units. Camouflaged air defence units deployed in the field are very hard to find. Additionally, Serbian forces placed identical looking decoy air defence systems made of cardboard and gunny. After the war it became clear that allied bombers had 'destroyed' decoy air defence systems, thus wasting many sorties and bombs on fake targets.

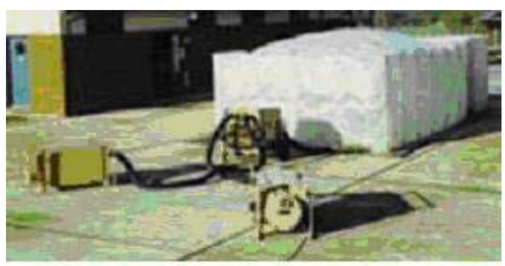


Figure 6: Collective Protection

Recuperation

Recuperation covers all measures necessary to recover from the effects of an attack, restore essential services and enable operations to continue with a minimum of disruption. Recuperation functions may include, but are not limited to, Damage Control (DAMCON), Post-Attack Reconnaissance (PAR) and Assessment, Explosive Ordnance Reconnaissance and Disposal (EOR/EOD), fire fighting, rescue, Restoration of Aircraft Operating Surfaces (RAOS), casualty handling and NBC decontamination. Examples of Recuperation are:

- Establishing NBC decontamination teams;
- Medical treatment facilities (casualty collection points, hospitals, etc.);
- Repair of Essential Services and Facilities (RESF), restoring services and facilities that are essential for the continuation of operations;
- Fire fighting;
- Explosive Ordnance Disposal (EOD) and taking care of Improvised Explosive Devices (IED).

Practical application

What are the problems in applying this knowledge in a real-time scenario? Suppose you were a FP officer. How all this information is applied depends on the situation your unit is in. WE will now discuss the application of FP.

Static location (airbase/compound)

Contrary to land units who can avoid being attacked by manoeuvring away from the threat, air force units regularly use a static location to operate from. Much effort and money goes into building an airbase. Redeployment to another base is often impossible due to geographical and financial restrictions. Lack of freedom in this respect has both advantages and disadvantages. It is disadvantageous for the location to be known so that an opposing force can determine the time and place to attack. On the other hand, being static can be advantageous, in the sense that more time and money can be spent on permanent measures to upgrade FP. At all times protective security measures have to be executed by the emergency services (i.e. police, fire, rescue and medical). Information Security (INFOSEC) focuses on what to tell whom, plus what information to hide when. Deception management regarding information on units, strengths, plans, movements, intents, rules for the use of force, etc. provides the wrong information for an enemy threat analysis and leaves the opponent in uncertainty. Ideally it will be harder for him to attack if he doesn't know our response. Active Defence measures will comprise entry control and patrol of vital areas. If the area to protect is too large, vital assets, scarce equipment and personnel will be secured separately. This can be done using electronic devices like mobile field radar. These kinds of measures are normally taken on a permanent basis. SHORAD against incoming aircraft will be executed during an actual air threat. SHORAD is limited in range and height.

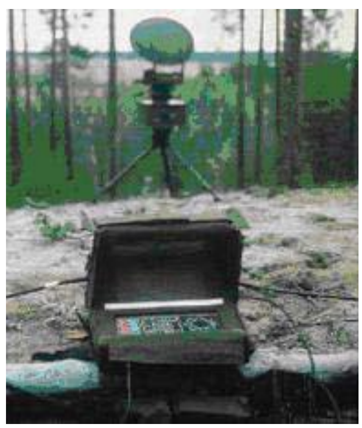


Figure 7: Field radar

As previously stated, Passive Defence includes the physical protection of personnel and materiel from all forms of attack. Time and money can be spent on camouflage and concealment. The use of camouflage nets is complemented by hiding vital assets in the surrounding terrain so that they are harder to find. During the Second World War for example, the Luftwaffe built bunkers with red roofs that resembled neighbouring farmhouses. When a pilot of an attacking aircraft, already under the stress of being illuminated by an air defence radar, has less time to find his target, he may aim and deliver the weapon inaccurately. If alerted in advance to an upcoming attack, sheltered personnel in protective structures will suffer minimal casualties. An alerting system using (NATO-) standardized signals must be installed in such a manner that all personnel can be reached. Dispersing mission-essential equipment and personnel will be done on a (semi-) permanent basis. Protection against explosives and B/C agents will be upgraded by building semi-permanent concrete structures. To give an indication of how much material has to be used to create a protecting wall against small arms fire: rock and stone requires a width of 0.5 metre and when using sandbags and stamped earth 1 metre.

To protect against B/C agents and industrial hazards (ROTA), mission-essential equipment and personnel can operate from buildings that provide coverage against direct attack with liquid agents. Even in a more permanent environment COLPRO (collective protection) facilities can be built so that under B/C conditions essential operations can be performed from a filtered location without the need for protective clothing. Working in a contaminated area is physically and psychologically very distressing and NBC conditions severely hamper operations in terms of time and quality, meaning that the job has to be done by more people and will take more time.

It is evident that preventative measures never stop a base from being attacked. In that case the operational facility has to recover from the resulting damage as soon as possible in order to resume operations, whilst preventing unnecessary casualties. Priority has to be given to restoring essential services and infrastructure enabling the continuation of operations (for a flying unit this is the runway, taxi tracks, fuel facilities, operation rooms, mission-essential personnel, etc). NATO criteria for Restoration Air Operation Services (RAOS) time frames are applied.

Fig. 8: Airfield Damage Repair

Fig. 9 Explosive Ordnance Reconnaissance

Post Attack Reconnaissance (PAR) can be prepared in advance by assigning specific areas to specialized personnel, augmented by regular unit personnel. It may be expeditious to have pre-planned recovery plans available. It is essential to subsequent operations to create an overall picture of the damage. It may not be possible to remove all UXOs (unexploded ordnance) on time. UXO's that do not hamper the continuation of operations may be removed later, though on a static base the final removal of all unexploded ordnance has to be planned. A robust fire-fighting system on a modern base comprises an automated fire-alert system and sprinklers in crowded and mission-essential areas. After an attack the inflicted damage to essential operational surfaces (runways, taxi tracks, platforms, etc) has to be repaired. Casualties have to be stabilized and prepared for transport to a hospital at casualty collection points. In case of B/C contamination, mission-essential equipment and surroundings have to be decontaminated. Of special concern are Improvised Explosive Devices (IEDs). IEDs can be produced at low cost, are hard to detect and can be very effective. The defence against IEDs is organized as follows: firstly with defensive measures (detection, neutralization and protection against the effects) and secondly with offensive measures focusing on the system of recruiting, financing, transporting and manufacturing IEDs. NATO is currently working on an IED Defeat Strategy.[158]

158 Countering IED, Opleiding en Training, nr. 38 December 2006, p. 19-21

Protecting a temporary (forward) base

Protecting a forward base, e.g. a Forward Area Refuelling Point (FARP), poses several problems. On the one hand, a relatively short time use gives the advantage of surprise because the opponent has little time to prepare for an attack. INFOSEC and PSYOPS may then be the most important FP measures. On the other hand, it is not possible to provide a forward base with the same level of protection as a static base, which uses concrete structures, barbed wire, COLPRO, intelligence, etc. At a temporary location, very vulnerable weapon systems (aircraft, GBAD, etc), large quantities of ammunition and fuel are more or less out in the open. Additional FP measures such as improvised protection constructions (earth walls, containers, use of natural cover in the terrain) and the availability of a fire fighting system may then become important.

Aircraft

Protecting an aircraft requires specialized procedures and personnel. For aircraft using allied airbases in peacetime, FP is part of the regular FP measures and structures of the host unit. During a peacetime visit the host nation security unit will take care of visiting aircraft. Few additional measures, such as preventing unauthorized personnel from approaching the aircraft, are required. If a host nation is not able to provide sufficient FP, own forces have to do so.

Aircraft security team (AST)

The aim of the Aircraft Security Team is to secure the aircraft and its crew if it is expected to land on a base where the host nation does not guarantee security. The AST will then ensure that the crew (un) load the aircraft and depart in a secure manner. Normally an AST is a small team and therefore only capable of countering small threats.

Downed aircraft security team (DART)

Helicopter operations in the field can always involve aborted flights due to a technical malfunction. In that case the helicopter must land as soon as possible. If a helicopter crew is downed in a hostile environment, it can be decided to launch a Combat Search and Rescue (CSAR) mission. This is outside the scope of FP and will therefore not be discussed here. If a helicopter is downed in a friendly environment, a quick recovery of crew and aircraft is required. Securing the crew, the helicopter and the surrounding area is an FP activity. Entrance to the area will only be allowed to authorized personnel. A DART can be moved by road or by air.

Convoys

The protection of convoys is mostly executed by ground forces who follow land operations procedures.

Finally

Air forces can only 'provide' airpower if operations are supported by solid logistics and customized FP measures. Force Protection should be based on the existing threat and should be in balance with the mission. Achieving a 100% protection is impossible and undesirable because this will hamper the execution of the mission too much. A variety of FP measures commensurate to operations can be applied. It is the duty of every commanding officer to select appropriate measures, based on the threat, risk analysis and vulnerability assessment.

Chapter 19 - Command & control

Introduction

Commanding a modern defence force, with assets placed around the globe, combining several nationalities and all services, is complex and dangerous. Mistakes can easily lead to unnecessary casualties and the loss of a battle. A well executed campaign on the other hand can paralyse an opponent and lead to a quick victory with a limited number of casualties.

There are many factors involved in the successful execution of an air campaign. To mention the most obvious: correct and timely intelligence that is disseminated speedily; correct coordination between the services; the choice of the right targets; optimisation of force composition and routing; and last but not least morale.

We can see an air campaign as a process. There is a required output or result and there are 'knobs' to turn to make the process move in the right direction. Because of the complexity of command and control we will start this chapter with a short side-step to the simpler mechanical process or control theory. Evaluation of control in simple mechanical systems shows that even those systems have control problems. We can therefore safely assume that these problems will also exist in the more complex processes like an air campaign. That is why we will start with the 'easy' problems and then move on to more complex processes. Another reason for this approach is that it is a good stepping stone to John Boyd's well-known OODA loop, which is at the heart of the command and control process in an air campaign.

Definitions

Command and control are typically military expressions, whereas in civilian life the terms leading and managing are more commonly used. The shortest definitions of the civilian expressions probably are:

Leading is choosing the right thing to do;

Managing is doing it the right way.

Leading however is more; it also requires a moral authority and a charismatic personality. When you are a leader it is not enough to know what the right way is, you must also be able to convince and inspire your men. Managing is doing it the right way, which requires correct decision-making and insight into the process.

It is obvious that more managers than leaders are needed, but it is very likely that more people aspire to be a leader than 'just' a manager. Without efficient management however, no leader can reach his objectives.

Command and control is a special form of leading and managing, applicable only to combat operations. Running a defence organization in peace time is not commanding, but just leading and managing. The term commander means that someone has the authority to lead and manage his unit in combat situations. He has the authority and obligation to lead, within the boundaries set by his superior. To lead in combat requires moral authority and inspirational leadership. Because a commander is making the final decision he must also posses managing and decision-making skills. We always talk about "command and control" because without control no effective command is possible. Command is giving the orders and control is taking all necessary action to disseminate and execute those orders and report back to the superior. As we will discuss later, this feedback loop is essential to controlling a process. It is also an essential part of all mechanical systems, as will be discussed in the next paragraph. Feedback is also an essential part of the OODA-loop which will also be explained later in this chapter.

It is obvious that military control of a complex air campaign is only possible with a fast, redundant, high-capacity and secure communication and information network. In literature you often find the term C4I, which stands for Command, Control, Communications, Computers and Information. C4I is in a way a strange combination, because command and control are processes, computers and communication are the required hardware, and information is the required 'software'. In this chapter we will concentrate on command and control (the processes) and assume that the required hardware and software are available.

The theory of command and control

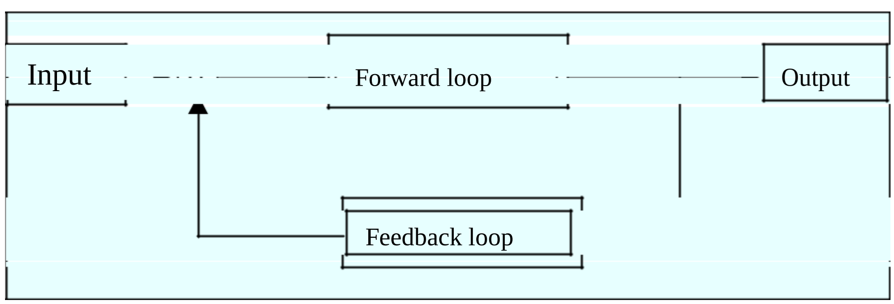
The 'simple' mechanical approach

The most simple mechanical control process only has a 'forward' loop. An example is a system with one knob or valve that initiates an action, the result of which is not checked but just presumed. A good example is an automated lawn sprinkler system, which is turned on every evening for an hour whether it has rained that day or not. A forward loop system is normally cheap, fast and useful when the result is:

Not critical

Predictable

A slightly less predictable process such as the temperature regulation of a house, already requires a more complex system. Simply turning the heating on for a certain period of the day may not give the required result. The way engineers solve this is with a feedback loop. The actual temperature is fed back into the thermostat and the required temperature can be maintained, at least in theory.

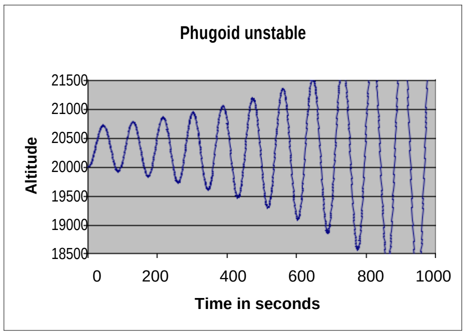


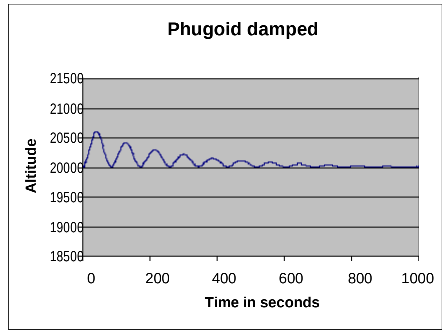
One of the results of this feedback loop is that the response is now **time-**based and the control theory proves that several other scenarios are also possible:

The capacity of the heating system may be so great that when the desired temperature is reached the residual heat in the radiator will cause the temperature to overshoot the required temperature by a fair margin.

When the temperature drops it may take the heating system a long time before the radiators are at the correct temperature, which may lead to a temperature much lower than the set value

As depicted in a graph, the response can be steadily oscillating, damped or even unstable. For a illustration we turn to aircraft mechanics. If an aircraft that is trimmed for level flight, is disturbed from its flight path, it may enter an oscillatory mode[159], which is more or less damped depending on the friction of the aircraft. This aircraft mode is also initiated by feedback in a similar way as the temperature fluctuations are initiated by the feedback loop to the thermostat.





Engineers will try to prevent these fluctuations by using some clever design methods. For instance, a thermostat may not just turn the heating on or off if the temperature is below or above the desired value, but it may set the heating system to burn for short intervals to allow the temperature of the room to catch up.

The first thing we can learn from control theory is that every feedback system has a certain **inherent frequency** based on '**mass**' and '**resilience'**. In aircraft mechanics the 'mass' can be the inertia moment around an axis and the resilience can be the change of pitching moment caused by an angle of attack change. Mass can also be the number of people working in an organization with the resilience being their desire not to change the way they perform their duties. One of the important lessons we can learn from control theory is:

It is impossible to change a system faster than its inherent frequency[160].

We have to respect the own frequency of the system or organization in order to make command and control effective. The classical example of what happens if one does not respect the frequency of the system is the pork price example.

> *Henry Ludwell Moore(1914) investigated cycles in the prices of agricultural produce. He found a cycle in the price for pigs. It turned out that when the price of pigs was high more farmers would turn to pig breeding. However, their pigs would all be ready for slaughter at the same time which would lead to a fall in prices. Thus it is hard to make a good price if you don't respect the frequency.*

The second important variable is **friction**. Friction is caused by the rate of change of a system. It has the tendency to slow down the rate of change but it also enhances stability. Friction is a term that was introduced into military strategy by the famous military philosopher Carl Von Clausewitz. Von Clausewitz was right, friction is a universal characteristic that always appears when something has to change and is not only limited to mechanical processes. Friction will always cause the outcome of a process to be slower than 'theoretically' possible.

The third important thing to learn from control theory is that one has to know the 'knobs'. Even a mechanical system can have a fair number of control functions, each with their own time response and interacting with each other. These processes may enhance but also counteract one another. Therefore a 'simple' mechanical

159 The technical term for this slow oscillatory motion is phugoid.
160 Often called 'own value'.

system can become quite complex. Below are two classic examples of counteracting modes:

> *The Airbus (A-330) that crashed in Toulouse in 1994 was in the altitude capture mode when one of the engines failed. The aircraft was no longer able to climb at the initial rate but instead of lowering the nose to maintain speed the altitude capture routine forced the nose to an even higher position, which lead to the crash of the aircraft. In the engine failure had occurred in any other autopilot mode the accident would not have happened, but in this particular case the altitude capture mode conflicted with the one engine out mode.*
>
> *Almost everybody has experienced computer 'hang ups'. The fact of the matter is that a lot of different processes are running at the same time. However, not all possible combinations have been tested. Some combinations of operations force the computer into an unsolvable loop and a hard restart is necessary, making even a 'simple PC' a process that can become unstable.*

With the knowledge of control theory we will now look at human organizations and see if we can identify processes, 'knobs' and response times.

Human processes

General We have established that mechanical processes can become complicated. It is easy to understand that the situation is even worse in complex human organizations. First of all, not all the processes (with their inherent frequencies) are always known, nor are all the 'knobs'. This is probably one of the reasons why so many management text books have been written. Researchers discover new processes and new 'knobs' (culture of the organization, chain management, management by objectives, inspirational management etc.) and write about it.

Processes in combat In war time the situation is even more critical. The pure survival of the contenders is at stake and this will force them to move along paved and unpaved routes. Or, we can make the following analogy:

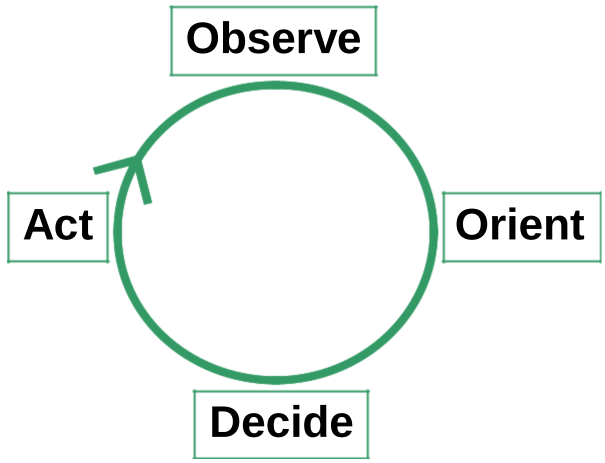
> *War is a kind of chess game where unknown pieces may enter the board, which is changing in size all the time and halfway through the game the rules are changed, of course without telling you.*

It is self-evident that when the outcome is uncertain you need a system with ***feedback****. This is not enough when the game is changing in character, you need a system which is* ***adaptive*** *and will change* ***itself*** *when necessary. A special extra step is needed for a system to become adaptive. John Boyd developed such a system called the OODA Loop.*

> ***John Boyd*** was a fighter pilot who served in Korea and Vietnam. His dog-fighting skills were excellent[1] and from his experience with dog fights he developed the OODA loop principle. The OODA loop principle has far wider applications than just for aerial dog fights. Not only can it be used for a complete campaign, it is also used outside the military, for example in business science. John Boyd however was more than the inventor of the OODA loop. He developed the Energy Manoeuvrability concept for the development and comparison of fighters[1]. His ideas led to the development of the F-15 Eagle and the F-16. After his retirement he continued researching warfare. His findings were never published but only put on slides[1]. With his bag full of slides he gave many presentations throughout the country and influenced many military professionals.

The OODA Loop

In the OODA loop quickly shows the feedback principle. The loop needs to be repeated until the objectives are met. Feedback is of course essential when uncertainty is high. We also note that Boyd distinguishes four steps: Observe, Orient, Decide and Act. Let us examine all four more closely.



Observation is more than good intelligence. Observation is a state of mind, and particularly an open mind. One of man's greatest shortcomings is that he only tends to see what he assumes to be there. Magicians make a living from exploiting our weakness in this field. Particularly in armed conflict we should not limit out observation by our preconception of what is reasonable, or what seems logical.

Orientation is the most crucial part in Boyd's loop and essential to making an adaptive system. The orientation phase should be more than just working out some alternatives and deciding which one yields the best results. Orientation should be an in-depth analysis to discover trends, opportunities and threats. It should increase the Situational Awareness (SA). Vital to a good SA is an understanding of the commanders' intent. We have already explained that the orientation phase can only be successful if we realize the time delays involved in the different processes and know what knob(s) to turn.

Decision What knob should be turned and which knobs are available at what level? It can be frustrating to a commander when there is only one knob. If the commander has less freedom to decide, the system is less adaptive and will react more slowly to changed circumstances. Therefore his command will be less effective. There may be several reasons why his freedom to decide is limited. Political restraints are often in place, but lack of sufficient assets and lack of correct intelligence will also limit the decision options of the commanding officer.

> *If your only tool is a hammer, all problems will look like nails.*

Decision making is often considered a logical knowledge-based[161] process of choosing the best alternative. However, psychological research shows that our decisions are not always the outcome of a rational process. Some of our decisions are made intuitively and rationalized afterwards to defend our initial, intuitive

161 Rasmussen distinguishes three human processes: skill-based (like walking), rule-based (like working according to a checklist) and knowledge-based.

choice. The danger of this is that we look favourably upon each argument that strengthens our initial choice and tend to downplay each argument against it. On the other hand, the intuitive decision does not come out of the blue. Our intuition is formed by previous experiences and experience is in itself a valuable resource. Furthermore, even a purely logical process will not always give a correct answer. First of all, which logical process do we use? Are we minimizing risk or maximizing possible gain, or are we seeking a compromise? A second disadvantage of the logical process is that not all relevant information is known and we may have to make a decision based on many uncertainties. In general it can be argued that taking the correct military decision is more an art than a science.

This principle, which leaves as much authority as is workable at the lower levels, was first developed in the German army during the First World War and was called 'Auftragtaktik'. It is good to mention here that an army has many more levels at which command and control take places. In an air campaign the number of assets seldom exceeds a thousand, and all assets are heavily dependent on each other. If fact, the air campaign can be considered as just one orchestrated action. Therefore command and control are only executed at the highest level. At the lower level (mission level) there is the freedom to execute the mission based on skills and experience. So for airpower operations the adage is 'centralized *command* and decentralized *execution'*.

Act To act one needs assets and trained war fighters. The act is normally the best known part of a military operation. That is what most of the training is focused on. The processes involved with the act are normally clearly understood.

Another insight that Boyd gained was the fact that if one completes the OODA loop **faster than the adversary** he will be outmanoeuvred and you - rather than the opponent - will have the initiative. There is some analogy with the Evolution Theory. If a species can adapt faster, it may be able to survive under changing ecological conditions. A species that adapts more slowly may however become extinct.

Command and Control practice

Loops, processes and knobs in aerial combat operations

In the previous paragraphs we discussed the theory of command and control. In this part we will discuss how command and control is carried out in aerial warfare practice. If we don't limit ourselves to the actual combat operations but also look at the preparation of armed forces we will discover that there are many loops, each with their own time delays and knobs. The most obvious ones are:

Research and development. This is definitely one of the longest loops. It may take several decades before research will lead to a superior weapon or weapon system. The knob for research is not only the amount of money that the government is willing to spend but also the ability to select the right projects.

Procurement. This is another long process. Presently the time from the initial requirement phase to initial operational capability is between 10 and 20 years. During the procurement and development phase many knobs need to be turned.

Money is the obvious one, but many decisions on equipment, performance, capabilities and force composition have to be made.

Intelligence is also a long cycle. It may take several years before adequate intelligence about an opponent is gathered. Even year-long efforts have not always led to adequate intelligence[162]. The hardest question to answer of course is: how much intelligence is enough. The knobs here are personnel, money, assets; the decisions as to what intelligence to gather and on which countries or organizations needs special attention.

Training. It takes several years to become an experienced operator. This not only applies to aircrew, but also to technicians, fighter controllers, logisticians etc. On the other hand, if a well-trained organization already exists it may only take from a few weeks to several months to train for a new environment or to adopt new tactics. A well-trained human operator is flexible and will easily incorporate new tactics and procedures. The knobs for training are: available flying hours, adequate training areas, realistic threat scenarios, simulators, weapon ranges etc. , all of which are extremely expensive.

Campaign planning. It can take several months before a campaign is fully planned. This long planning cycle may jeopardize the effective use of airpower. That is the reason why a lot of planning against possible adversaries is carried out during peace time. These are called contingency plans. The most important knobs for campaign planning are good intelligence and brain power. Campaign planning will be discussed later in more detail.

Force generation and deployment. It may take several months to generate enough forces, especially in large combined operations, but there is always a need for at least a minimum force to be available at very short notice. This is the reason for NATO countries to earmark units for the NATO Response Force or NRF[163]. During the time these units are earmarked for a specific NRF they should be ready to deploy with minimum notice. After the forces are moved into the theatre the Transfer of Operational Command (TOA) of the units to the Joint Task Force Commander will take place. The fact that forces are earmarked may however not prevent a delay of actual deployment through political decision-making. Another limiting factor is the availability of strategic transport capability. The knobs for force generation and deployment are: readiness state, strategic transport and political approval.

Campaign execution. The actual combat can last from a few days to many years. During command and control of the battle the most important knobs are allotment, apportionment and targeting. We will discuss the campaign cycle in more detail later.

Tactical cycles. There are many tactical cycles. A dogfight can last less than a minute, a ballistic missile engagement just a few seconds. It is paramount that in this type of combat the whole OODA loop only takes a couple of seconds. Orientation in that case will be more based on doctrine drills and skills than on in-depth analysis. The knobs in tactical cycles are normally weapons and tactics.

Bottom line: Know the processes, know the knobs, know the time delays and orient. The enemy is adapting also.

162 It is interesting to note that the best paid intelligence service in the World the CIA was wrong on at least three important issues: They did not expect the Russians to have an atomic bomb that early, they did not predict the Fall of the Berlin Wall and they were wrong about the Iraqi weapons of mass destruction.

163 And the European Union has its Battle Groups.

The NATO[164] campaign planning

General

The NATO campaign planning is in itself a cycle. Planning does not stop when the campaign plans are made, but plans are adjusted when circumstances so dictate. A planning process is an iterative process between the commander and his staff. In this process we can also recognize the steps from the OODA loop.

Planning considerations

When a political decision is made for military action, the campaign planning will start. The end results are (politically approved) Operation Plans. The most important 'knobs' in the campaign planning process are:

- Concept of operations
- Force composition

Translating the political objectives into a military strategy is a process with many 'known unknowns'; but the more dangerous problems arise, of course, from the 'unknown unknowns'. Typical questions to answer are: what are the centres of gravity (CoG) of the enemy, what are his Critical Capabilities (CC), Critical Requirements (CR) and Critical Vulnerabilities (CV)? What are the opponent's possible courses of action? Will a strategic air operation be enough to compel the opponent or is an occupation necessary? History gives many examples where initial military assumptions were proven wrong. (See also chapter 1- airpower history.) It is good to realise here that a thorough cultural and psychological knowledge of the opponent is as important as good military insight. A further limiting factor is that political goals are not always easy to translate into simple military objectives. For example: "Stabilise the situation in IRAQ" is difficult to define. How many violent incidents a day define a stable country? If the goal is vague, it will also be difficult to establish what processes are needed and what time lines are involved. Another example of the complexity: is it only necessary to find the terrorist groups and destroy them, or is it essential to achieve a change of mind that discourages each potential terrorist in order to achieve lasting stability? In the latter case, the time required can be decades.

Joint

Campaign planning and execution must be Joint[165]. This lesson seems obvious, after all the process in itself is already complicated enough. Two men at the wheel, or two men steering different wheels is just not smart. But inter-service rivalry has been a problem for many years and has caused problems during the Vietnam War, operation Eagle Claw[166] and in operation 'Urgent Fury' in Grenada (1983). The last operation was the trigger for the Goldwater-Nichols Act of 1986. This act streamlined the chain of command and ordered the services to develop a joint doctrine for the integrated employment of joint military operations.

> *The doctrine-oriented Army was the first service to embrace joint doctrine. The result was a land-centric approach to warfare. For the Air Force, this initially meant that airpower was viewed as a support mechanism to the land commander on the battlefield. Despite the land-centric focus, the Air Force did make some inroads into the establishment of joint doctrine. Most notable was the emergence of the concept of the Joint Force Air Component*

---

164 A campaign might be led by NATO, the EU or a coalition of the willing; but in almost all cases the planning will done according to NATO procedures.
165 In a joint campaign all military services are working together.
166 Operation Eagle Claw was intended to free the hostages in the USA embassy in Iran (1980)

Commander or JFACC. The JFACC concept provided a framework for integrating and employing airpower from all service components under a unified organizational structure. This concept also had a significant impact on the development of air strategy and the conduct of joint air operations. The JFACC concept is a subject for another lesson. The Goldwater-Nichols Act had a far-reaching effect on the Air Force. This legislation was the impetus for moving the Air Force away from the era of doctrinal excursions toward employment as a decisive contributor in joint operations[167].

COA= Courses of action

CONOPS= Concept of operation

OPPLAN= Operation plan

Each stage depicts staff responsibilities (left) and Commanders' responsibilities (right). Also briefings to Commanders are represented: Mission Analysis Brief, Course of Action Decision Brief, the OPPLAN Brief and Revised Decision Brief.

The operational planning process (OPP) is the responsibility of the commander and the planning team[168]. It is a deductive reasoning process moving from general points of fact to a specific option. The OPP establishes procedures for analysing a mission, developing Courses of Action (COAs) against the threat, comparing friendly COAs against the commanders' criteria and each other. War gaming can be used as an analysis tool for the selection of the optimum COA and the preparation of an operation plan for execution.

The design of the OPP provides the planning staff with maximum freedom to develop ideas and concepts.

From planning to execution

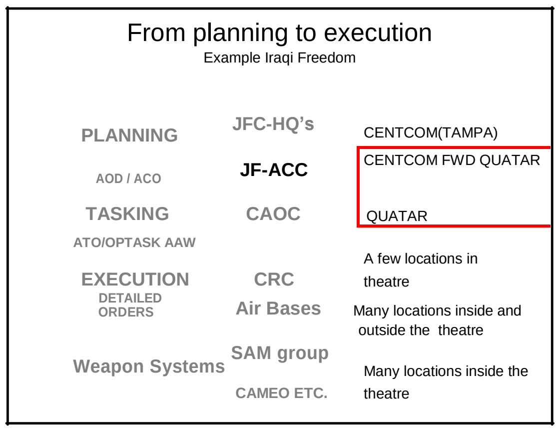
Plans at different levels can theoretically be valid for the complete operation, however if conditions warrant, they can be updated. The OPLAN must be put into action. This is the task of the Commander of the Combined Joined Task Force

167 http://www.iwar.org.uk/military/resources/aspc/text/excur/goldnich.htm d.d. 4 December 2006

168 The OPP can be done in a Permanent Joined Head Quarters e.g. CENTCOM or SHAPE, or in a deployed HQ or in a combination. Eventually the OPPLAN needs political approval.

(Com CJTF), also called the Joined Forces Commander J-FC, and his component commanders. The following components can be part of the task force: Land Component, Maritime Component, Air Component, Psychological Operations and Special Operations Component. The J-FC will coordinate with his components commanders and will task them. For the Joined Forces Air Component Commander this tasking is the Air Operation Directive (AOD) and the Air Coordination Order (ACO).

The vehicle for daily command and control is the Air Task Order (ATO), which is made at the Combined Air Operations Centre (CAOC) and is used to task individual units. The CAOC is normally inside the theatre while the Combined Headquarters can be located inside and outside the theatre. Modern communication networks facilitate working with a large part of the staff outside the theatre, this is called reach-back. The following graph depicts the route from OPPLAN to ATO to the individual units.



The knobs to steer the air campaign

The most important knobs to steer an air campaign are:

Targeting priorities. Thanks to intelligence efforts the task force will have a targeting list (TL). In the targeting list some targets will be classified as restricted or prohibited. Permission at the highest political level (North Atlantic Council NAC) is required before restricted or prohibited targets are selected, so this targeting knob is turned at the highest level. Targets are selected from the list by the components and a Joined Integrated Prioritised Targeting List (JIPTL) is drawn up and approved by the Joined Coordination Board, so this knob is turned at the J-FC level.

Allotment is the temporary allocation of forces to a subordinate commander.

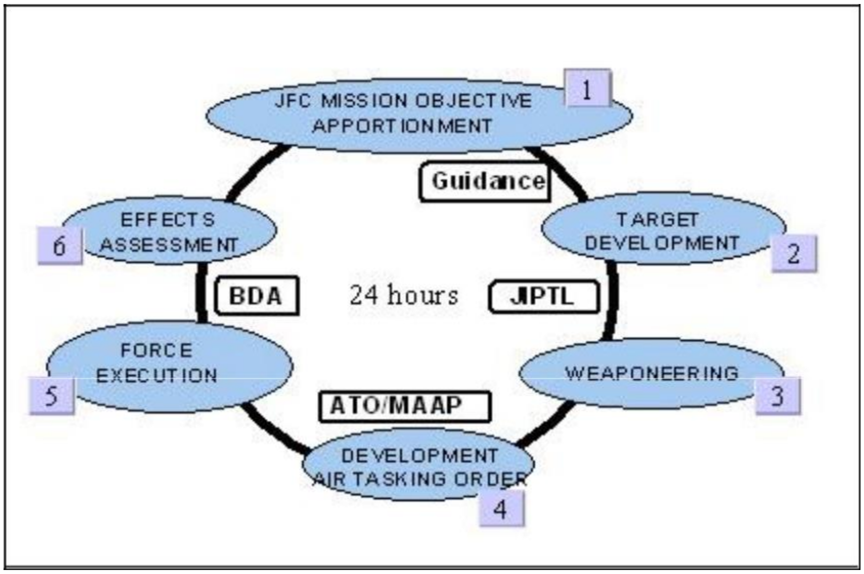
Apportionment gives the percentage of available assets that will be allocated to a certain mission. (An apportionment can be 40% OCA, 20% DCA, 20% Strategic and 20% ASFAO).

Allocation is translating the apportionment of assets into the number of sorties per aircraft type and per task.

Tasking is translating the allocation into detailed orders for each unit.

The ATO cycle

The ATO cycle is repeated every 24 hours, meaning that while number one is executed, number two is being finalized and the concept of number three is drafted. The system is in itself quite rigid. However, as has been stated before, an air campaign is an orchestrated event needing careful planning. This does not preclude an airborne mission from being rerouted to an alternate target or even re-rolled to another mission. The picture below shows the ATO process.



The different phases:

Phase 1: In this phase the JFC and his Component Commanders (CCs) evaluate the progress made and design future plans. The CCs will voice their requirements. The result may a rearrangement of priorities and objectives and the apportionment of assets.

Phase 2: In this phase the target development takes place. The Guidance Apportionment Tasking cell (GAT) nominates targets from the Joint Target List (JTL) based on the objectives, apportionment and priorities set by the JFC, resulting in a Joint Integrated Prioritised Target List (JIPTL) and an Air Operations Directive (AOD).

Phase 3: Weaponeering. In this phase detailed target information is used to determine the Desired Mean Point of Impact (DMPI), the preferred number and types of aircraft and the preferred weapon types. Furthermore attention will be given to the probability of collateral damage and success. All findings will be presented in the Master Air Attack Plan (MAAP).

Phase 4: ATO generation. After the JFACC has approved the MAAP, ATO generation can take place. The ATO encompasses not only the different targets, times and routings for each mission but will stipulate the Special Instructions (SPINs) that are in force. To regulate all air movements an Air Coordination Order is issued. This serves two purposes. It will de-conflict all movements, and it will facilitate the identification of friend and foe.

Phase 5: Execution. This is self-explanatory.

Phase 6: Effects assessment. The effects assessment is based on the after-action reports, reconnaissance missions and Battle Damage Assessment (BDA) from pilots.

The target cycle may give the impression of a rigid mechanism that does not allow for fast response or fast targeting. This is not the case. Time Sensitive Targeting (TST) and flexible targeting are key issues in modern aerial warfare. There are several ways to realize TST within an ATO. Missions can be tasked on ground alert (QRA) or can be assigned to do a CAP and be ready to act when called upon. Another possibility is to reroute targets on a mission to a higher priority target.

New strategies and command and control

Two new concepts of operation have emerged in the last decades. The first is Network Centric Warfare (NCW), which is presently more often labelled Network Enabled Capabilities (NEC) and the second is Effect Based Approach to Operations (EBAO). Both developments have their repercussions on command and control and we will examine what these are.

Network centric warfare

The physical components of Network Centric Warfare (NCW) are an integrated sensor grid, an integrated shooter grid and a command grid. This physical grid is of course highly dependent on high capacity and secure communication. Furthermore the computers in the weapon system and in the command centres should be able to communicate with each other over a secure network. But the physical part is just the entry fee to NCW. The improved networks give the war fighter better information and consequently a better situational awareness (SA). Improved SA leads to better decisions and better actions will yield a better effect, and this is what counts. This will be discussed in more detail in the paragraphs on EBAO. It is evident that an improved SA throughout the organization will facilitate other ways of command and control. This will be discussed later but we will first consider some of the historical developments.

The Royal Air Force is probably the first service that started with NCW, (without naming it as such, and also without computers and data links). From a conceptual point of view there is a good reason to consider the British Air Defence system during the Battle of Britain as the first NCW battle. All radar information was gathered in the control rooms and a centralized command via radio was able to effectively steer the fighters to their targets. In Boyd's terms, the British OODA loop was faster than the OODA loop of the German attackers, nullifying their advantage of surprise. The nice thing to note here is the combination of NCW and centralized command.

The air defence arena is probably most advanced in NCW. The vehicle for NCW in air defence is the Link-16 system, which connects shooters (aircraft and guided missile systems), sensors and decision makers[169]. In a modern Air Defence scenario[170] NCW has changed a lot. With Link-16, fighters have access to the Recognized Air Picture (RAP), which gives an almost instant picture of all friends and foes in their vicinity. On top of that they may be able to feed target data from their fire control radar back into the system, which further enhances the accuracy of the RAP. This transfer of data is much faster than relaying target data by voice over the radio, it is more accurate and gives a better situational awareness to the pilot. In essence, the higher situational awareness makes observation and orientation quicker, which in turn helps make your OODA loop faster than your opponent's.

169 For specifications of Link-16 see chapter 3.

170 Most modern air forces are presently in the process of equipping their fighters with Link-16 , but it may last until beyond 2010 before all units are Link-16 equipped.

A significant other benefit is that the probability of a blue-on-blue engagement is reduced significantly.

We must realize that a similar system for air-to-ground missions does not yet exist. Conceptually it would of course strengthen situational awareness and increase the speed of the OODA loop. The problem in the air-to-ground arena is that there are many potential ground targets. In fact, a few orders of magnitude more than there are aerial targets. A lot of those targets may also be small and camouflaged and only visible for a short period of time[171]. We simply lack sensor capability to cover everything.[172] Even if in the future we were able to have a complete ground picture in the cockpit, we could easily become overwhelmed by the volume of data. Another challenge is that only a few sensors are air force-owned, most ground observers are from the army and the navy. At this moment many countries are equipping their armies with a NCW capability, but different countries are developing different systems. There is still a long road to travel before a complete NCW infrastructure is present. On the other hand, on a smaller scale air-to-ground data links are becoming more and more common for CAS, and Predator pictures that are sent right away to the CAOC may initiate an immediate response.

The question is not if but how NCW will change command and control. If fact two routes are possible. If the situational awareness is higher, every participant can 'self-synchronize' his actions as long as he acts in accordance with the commander's intent. This will lead to a loose control of the battle and more freedom for the individual. The second route is that the higher commander will become more directive because he has almost the same situational awareness as the pilot. Theoretically the best way is for the higher level commander to do what he has to do and the lower level commander to do as much as he can. In reality however, the route that will be taken is probably more dependent on cultural background than on technical and theoretical considerations.

With the incorporation of more NCW capabilities some risk factors can be identified:

Micro-management at the higher levels, which may lead to bureaucracy;

Information overflow. The more information is available the less time is left to study it. This was illustrated fairly well by what happened with e-mail after somebody found out that you can CC each mail to everybody.

Effect-based approach to operations

The Effect Based Approach to Operations (EBAO) works from the principle that all actions should be optimised to reach a certain effect. The idea was conceived during the Gulf War. Faced with the enormous increase in accuracy and penetration capability, planners became aware that many more targets than ever before could be attacked with the same assets. Instead of sequentially attacking targets, parallel targeting became possible. The increased tempo resulting from parallel targeting had a paralysing effect on the enemy. In his essay, General Deptula [173] gives as an example that power plant operators were so afraid of an air attack that they shut down their installation to prevent destruction. This meant that the required effect was reached without a single asset dedicated to these targets.

171 For example a SCUD launcher will be out of a camouflaged position for only a short time to fire the missile and will most likely hide right after launch.

172 This is the reconnaissance law of physic, you can either scan a large area at a low resolution (which will cause you to miss a lot of smaller targets) or scan a smaller area at a high resolution (which will cause you to miss all targets outside the scanned area).

173 See ALSO Effect Based Operation – Change in the nature of warfare- by Brigadier General David A. Deptula.

EBAO however has different implications at different levels. We will briefly discuss the operational tactical levels and the strategic level.

As General Deptula's article points out, the result of effect-based operations at the operational level was proven during operation Desert Storm. In a way it can be considered as simply a better way to make a MAAP and more options for the JFC to achieve his objectives. It this sense EBAO is a better way to make an OPPLAN and a better way to make a MAAP. The only caveat is that this principle has not yet been proven against non-state actors like insurgence and terrorist.

On the strategic level EBOA means that all national or coalition efforts should be optimised to achieve the desired effect. This however does not look very different from the way things were done in the past. The military option has always been only one of a nation's options

Conclusion

Good command and control are essential ingredients for success. The Airpower command principle of centralized planning and decentralized execution is a proven concept. The real difficulty however in Command and Control is that political desires have to be converted into proper military actions against a constantly changing adversary. Adaptation by proper orientation is the only possible way to keep the upper hand.

Chapter 20 – Airpower in legal perspective

Introduction

As history has shown, military aviation has developed from an interesting, but in a military sense insignificant, method of warfare into a devastating power capable of projecting an enormous force with high precision at selected targets around the globe. However, power projection is not without bounds. All modern societies are based on the rule of law and their administrations must operate within domestic and international legal frameworks. Like political, operational, logistical and financial concerns, legal aspects govern the possibilities to launch and sustain a military operation abroad.

Commitment of airpower in a conflict today is hardly confined to national territory and airspace, and can be stretched far beyond a state's own boundaries (e.g. expeditionary operation). Unmanned aerial vehicles can be guided from a command centre thousands of miles from the area of operation and collected information can be sent straight from the battlefield to the headquarters back home. However, the air force units involved can only be partly controlled and supported from the homeland. To a large extent, units will have to be staged in forward areas abroad. To this end, host states must grant permission and mutual arrangements on the position of the visiting forces have to be agreed upon. The applicable international rules and regulations are this chapter's first point of consideration.

Advanced systems enable the attack of targets with a minimum of collateral damage while at the same time limiting one's own risks. This befits a modern, Western approach to military action in which an ever decreasing level of acceptance of loss of human life seems to exist. Both civilian casualties and own losses have to be kept at a minimum and preferably avoided completely. This can only be achieved by thoroughly controlling the commitment of forces. Bearing this in mind, all forms of military use of force are regulated from different angles. International law plays an important role from the moment it is first decided to commit military means and continues to play this role during the actual operation. International legal norms regarding the use of force are the second and most important point of consideration of this chapter.

This chapter is organized to deal firstly with the law governing visiting forces within the framework of expeditionary action. Among others, Status of Forces Agreements (SOFAs) and Memoranda of Understanding (MOUs) will be considered. It then proceeds to discuss the rules regulating the use of force. This part starts with the right of a state to use armed force against another state (ius ad bellum), which provides the legal basis for the actual commitment of military means. The use of force *during* an operation is further regulated by the Rules of Engagements (ROE), which have to be drafted in accordance with the law, including international humanitarian law, better known as the law of armed conflict (ius in bello), the final part of this chapter.

Law of visiting forces

The tasks of modern armed forces are many-sided and complex. Therefore, most states will cooperate with each other in the preparation and execution of their respective missions. For purposes of education, training and exercise they will regularly admit personnel and materiel to each other's territory and during the actual deployment as well as in crisis management operations they will work together from foreign soil. The presence of foreign forces in a country could imply an infringement of the territorial sovereignty of a state. Subsequently, questions may arise regarding the legal position of visiting military forces: for example, who is responsible for damage caused by the visiting forces and which country can prosecute a visiting soldier who violates the local laws of the host state? Finally, some thought should be given to the question whose local cooperation needs to be sought in order to accomplish the mission.

Sovereign states are equal and independent of each other so that they have full authority over their own territory without the obligation to follow directions of other states. This principle of sovereignty is one of the fundamental principles in international law and implies that a state cannot enter another state's territory or airspace without proper justification: a legal basis is required (**ius ad praesentiam**). Besides the possibility of a country invading another country with military force and bringing it under its authority (occupation), that legal basis must be founded on the host state's consent with the foreign presence on its territory. In general, a state will grant permission explicitly, depending on the relationships between both states and the purpose of the foreign visit. For example, the 1944 Chicago Convention on International Civil Aviation explicitly states that no military aircraft '*shall fly over the territory of another State or land thereon without authorization by special agreement or otherwise, and in accordance with the terms*' (article 3).

Permission for the long-term stationing of forces, as was fairly common during the Cold War period, is laid down in comprehensive base rights agreements between the states involved. For activities taking place within a shorter timeframe, such as conducting training or exercises, less formal arrangements by way of non-treaty agreements such as a Memorandum of Understanding (MOU) are generally sufficient. Furthermore, over flight and landing of military aircraft abroad is generally agreed upon on a case-by-case basis through the issuing of diplomatic (over flight or landing) clearances by the ministry of foreign affairs. Standing, long-term procedures based on treaties can be agreed upon as well and generally require only a timely notification of a particular over flight or landing.

These procedures cover the permission for entry and stay in a foreign state, but so far nothing has been said about the rules applicable to the visiting units while present in that country (**ius in praesentia**). In the past it was commonly accepted that if a country consented to the visit of foreign forces, it also accepted that these forces would be subject to the visiting state's criminal jurisdiction. This basically meant that if personnel of the visiting force would commit a criminal offence in the host state, the courts of the visiting state would have the right to initiate criminal proceedings against the suspect, thereby excluding the host state's courts. Today that assumption is not so evident anymore and the host state's court can in cases also have jurisdiction. Thus the status of visiting forces is not very different from that of any other foreign visitor. To avoid all confusion, states draw up agreements on the legal position of visiting forces while present on host state territory: Status of Forces Agreements (SOFAs). These agreements can have any form: formal treaties, exchanges of letters or notes, and occasionally non-treaty agreements like MOUs et cetera. Sometimes they are incorporated in other arrangements like base rights agreements or transit agreements so that the *ius ad praesentiam* and *ius in praesentia* are covered in the same document.

Typical of all SOFAs is that they stipulate criminal jurisdiction over the visiting forces. Additionally, a number of other provisions are formulated to accelerate and facilitate the entry and stay of visiting units. For example, SOFAs can contain provisions to the effect that personnel will not need visas to enter the host state, that military supplies and personal belongings are exempted from customs duties, that designated air fields and other infrastructure can be used and SOFAs may deal with the manner in which claims will be processed. A SOFA with only the most essential host state restrictions is critical for a commander to be able to operate effectively from foreign territory.

Especially in cases of short-notice deployments, the lengthy process of concluding SOFAs can be an impediment. States that send their military forces abroad generally want to retain as much authority over their troops as possible. Therefore, the host state will be requested to grant the visiting unit criminal immunities so that only the visiting state's judge has the authority to decide on criminal cases involving visiting state's personnel. However, this request goes against the territorial sovereignty of the host state who in turn will generally try to establish some authority over visiting forces. It must be accepted that negotiations may take time, especially taking into account the additional time needed in many states for consultation of parliament on this sensitive issue.

Besides agreements regarding the entry and stay of the visiting state's forces abroad, further arrangements may be necessary, e.g. on host state support, co-operation with coalition partners and mutual logistic support. In most cases MOUs are used. They are relatively easy to conclude as parliamentary approval is generally not required. The process can even be simplified and speeded up if the MOU is based on existing, special arrangements. For example, the US has concluded a so-called Acquisition and Cross-Servicing Arrangement (ACSA) on mutual logistic support with a number of allies. The procedures that have been agreed upon in the ACSA can easily be 'activated' for a particular operation or other activity through the conclusion of an MOU. As a result, support in the form of services or provision of goods can be exchanged simply by filing the necessary forms mentioned in the ACSA.

Ius ad bellum

With the conclusion of the Charter of the United Nations (UN) in 1945, the right of a state to use military force against another state was regulated within a collective security system. Since then, all states are obliged to settle their international disputes by peaceful means and are to refrain from threat or use of force against another state. However, a state always retains the inherent right of self-defence if an armed attack occurs. Furthermore, the UN Security Council can decide to take measures, including the use of armed force, when it has determined the existence of any threat to the peace, breach of the peace, or act of aggression.

A Security Council resolution that authorizes military action contains the mandate for the mission and forms the basis for the deployment of armed forces. On the international level, it creates the authority to use armed force. However, in reality the Security Council is not always able to come to a timely decision. This may result in action being taken by a state or a group of states that falls outside the collective security system of the UN as occurred in 1999 when NATO launched a series of air strikes against Serbia. In the opinion of the NATO-member states the humanitarian situation in Kosovo justified military action, as within the Security Council no agreement could be reached on the situation. Instead of a UN-mandate, the NATO decision then provided the basis for military action. This chapter, however, is restricted to the more common situation where a UN Security Council resolution forms the basis for a crisis management operation.

The Security Council mandate defines the purpose of the mission and stipulates the basic conditions for the use of military means. In the past the mandate could be rather restrictive: the use of force by peace-keeping forces was extremely limited, and allowed the troops to use restrained force in self-defence only. In the 1990s the nature of operations started to change and robust action was needed to fulfil the mission. In the first years of that decade the mandates did not take into account the new tasks at hand. As a consequence, this left the peace-keeping forces without the necessary elbow room to accomplish their mission. As a result, UN operations in for example Somalia, Bosnia and Rwanda were not as satisfactory as was hoped for. Nowadays it is common practice to include a paragraph in Security Council Resolutions authorizing the member states of the UN to take '*all necessary measures*' or to use '*all necessary means*'.

Rules of engagement (ROE)

The choice of wording such as '*all necessary means*' in the UN-mandate offers a broad range of possible courses of military action. Based on political, legal and operational considerations the possibilities are further determined by the Rules of Engagement (ROE): the rules or guidelines for the application of force, or actions that might be construed as provocative. In order to implement this rather short list of guidelines effectively, other documents are of importance as well.

During the drafting of the ROE, political interests already play an important role. The Security Council resolution that authorizes an operation is a political compromise that is reached after intensive negotiations between the 15 members of the Council. Furthermore, states must consider if they are prepared, based on the UN-mandate, to participate in the operation. They will only approve of participation if it fits in with their national political framework. Within NATO the importance of political considerations is underlined by the inclusion of a so-called Political Policy Statement (PPS) in the ROE. This statement explains in broad terms the overall policy aim, thus enabling the military commanders to frame the Rules of Engagement in the proper context of the mission. The PPS is often complemented by a Political Policy Indicator (PPI). The PPI clarifies the trends regarding the foreseeable or expected developments of the mission. It may indicate NATO's wish to minimize its involvement, to maintain the status quo or NATO's acceptance of the risk of escalation.

The legal considerations for drafting the ROE are a rather constant factor. All military action must be in accordance with national and international law. However, even in the legal field time does not stand still and new developments and insights influence an operation. For example, in the past the applicability of the rules of the Law of Armed Conflict during crisis management operations has been discussed extensively. Nowadays it is beyond doubt that if (military) force is used during these operations the Law of Armed Conflict must be respected and complied with.

Eventually it must be operationally possible to accomplish the mission with the available means. The ROE are an essential instrument for a commander to reach that goal. They must support an effective commitment of forces which, however, does not preclude restrictions on the ROE. Present crisis management operations take place in complex environments. A force can consist of units from different states while at the same time other organizations might be active in the area of responsibility. In Afghanistan for example two missions have been in progress for a longer period of time: Operation Enduring Freedom (OEF) and the International Security Assistance Force (ISAF). To prevent units from obstructing or even causing damage to each other or to the units of the other mission, it is necessary to coordinate the actions of own units within the ROE.

In the end, the combined political, legal and operational considerations will determine the rules or guidelines for the use of force: the ROE. Strictly speaking, this is not much more than a list of rules that military action must comply with. The list counts only a few pages of text and forms a rather summary and clearly-structured narrative. At first sight, the decision to use force in certain circumstances during an operation does not require a complex consideration: does a situation fall within a defined rule or not? In practice, however, insight into a succession of other documents is required to reach a balanced decision.

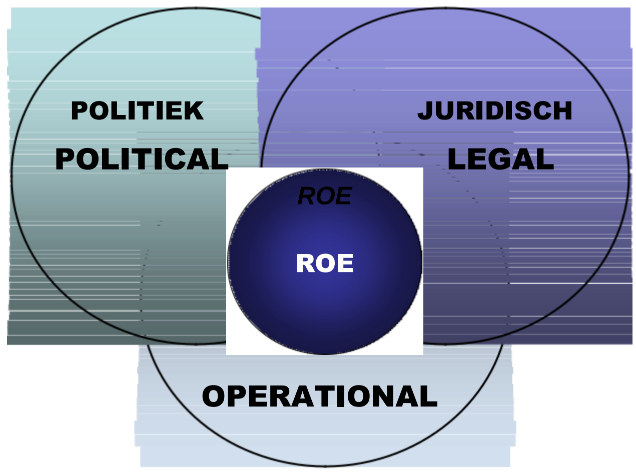


Fig. 1. Factors that influence the ROE

It makes a difference if a country conducts the operation on its own, as part of a coalition or under the auspices of an international organization like the UN, EU or NATO. Every coalition and organization has its own procedures for drafting and developing ROE. Details may differ but in outline the purpose will be the same. Let us take a major NATO operation as an example.

Within NATO the Supreme Allied Commander Europe (SACEUR) lays the foundation for the conduct of military action in the Concept of Operations (CONOPS), which in turn forms the basis for the Operation Plan (OPPLAN). In stock annexes, certain subjects are amplified whereby Annex E, '*Use of Force*', describes the ground rules for armed action and the use of force in accordance with the ROE. The ROE themselves are not part of the SACEUR OPPLAN, but are established in a separate procedure based on the NATO-document MC 362/1, NATO Rules of Engagement.

The MC 362/1 includes, besides a number of broad sections on the use of force, the procedure for drafting the ROE. Furthermore, an extensive catalogue ('*the compendium*') of possible rules on the use of force is included in an annex. The compendium is further subdivided into categories of activities ('*Series*') that cover the full spectrum of possible causes of action and range from the geographical positioning of the own forces to attacks. All series are summarily referred to with a two-digit number. Within each series different rules have been formulated in a succession of possible courses of action: from an explicit prohibition up to (conditional) permission. Each specific rule is summarily referred to with a three-digit code that consists of the category code followed by a one-digit number.

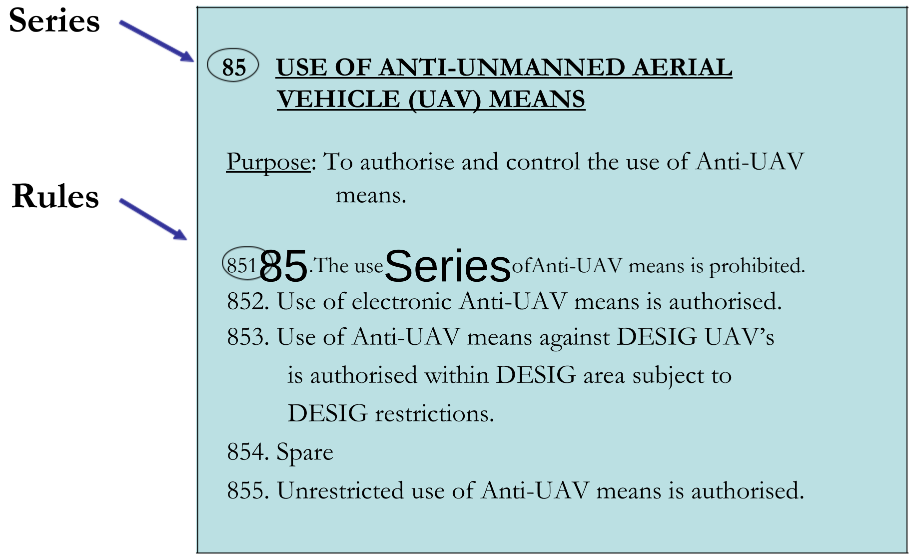


Fig. 2. ROE example (fictitious)

For each and every operation SACEUR will have to examine which series can be applied and consequently, based on the purpose of the mission and the SACEUR OPLAN, he has to decide which specific rule(s) are needed for the operation. Rules allowing more elaborate actions with few restricting conditions are in line with a robust mandate. If the circumstances so require, additional rules can be drawn up and selected for the operation at hand. These selected rules, together with some general provisions, for example on self-defence, constitute the ROE for an operation.

Through the implementation of the ROE, the mission commander receives the necessary authority to accomplish the mission. He will not, however, retain all power, but he will in part pass on permissions granted to him to his subordinate commanders. This procedure will be reflected in a matrix, which notes for every specific rule to what level authority has been passed on. The highest level will be the Regional NATO commander. The lowest level is marked as '*unrestricted*', meaning that each individual soldier is authorized to apply a specific rule.

The matrix will be copied into the plans and orders of the mission commander, who is directly responsible for the execution of the operation and therefore for the implementation of the ROE. His operation order will contain more detailed instructions regarding the application of the ROE. It is possible that amplifying rules are inserted regarding the use of e.g. bomber aircraft or attack helicopters. This can be achieved by drafting a targeting directive in which, for example, the division of authority to use certain forms of force is laid down in more detail, or a procedure to estimate or accept collateral damage is described.

Besides these international plans and orders, national directives must also be taken into account. Based on domestic policy or law a country may want to deviate from the ROE or issue amplifying instructions to their own, national units. In practice, these restrictions on the ROE are referred to as '*caveats*' and must be passed on to the international commander, for example at the Transfer of Authority (TOA) of the unit to the international commander.

Law of armed conflict

In the second half of the nineteenth century states developed the notion that the devastating effects of war had to be reduced by protecting the victims of armed combat and limiting the means and methods of warfare. The earliest treaties in the area of the Law of Armed Conflict were drafted and were further developed in the period thereafter. An important example, which is still relevant today, is the 1907 Hague Convention IV Respecting the Laws and Customs of War on Land. The military use of airspace was then in its infancy and was until the First World War primarily confined to the use of balloons. Therefore, Law of Armed Conflict treaties focused on the war on land or sea. In these treaties reference was made to aerial warfare only sporadically.

Ever since World War I it became more and more evident what airpower was capable of. It was, however, not feasible to draft a comprehensive treaty for aerial warfare. An initiative in 1923 did lead to the Rules of Air Warfare but the rules were never adopted in a legally binding form. Since 2004, Harvard University has been hosting a project '*to make an informal contemporary restatement of customary international law governing air and missile warfare*'. The focus of the project is not the drafting of a formal, new treaty but framing a manual for the benefit of the military and policy makers.

The legal appreciation of the use of airpower in armed conflict must be based on the general treaties in the field of the Law of Armed Conflict and on international customary law. Most of the treaties have been drafted or amended as a result of conflicts in which airpower played an important role. For instance, the four 1949 Geneva Conventions were revised (Convention I - III, on the wounded, shipwrecked and prisoners of war) or adopted (Convention IV on civilians) in the aftermath of World War II. The period that followed saw a growing number of conflicts, e.g. in Vietnam, that gave rise to the need for further protection of the victims of war. These events led to the adoption of two Additional Protocols to the Geneva Conventions in 1977.

Where treaties cannot provide the answer in a certain situation, international customary law and general principles continue to play a role. This is embodied in the so-called Martens Clause, which was inserted in the 1907 Hague Convention IV for the first time and was repeated in treaties of a later date. This clause states that in cases not included in the provisions '… *the inhabitants and the belligerents remain under the protection and the rule of the principles of the law of nations, as they result from the usages established among civilized peoples, from the laws of humanity, and the dictates of the public conscience.*'. The Law of Armed Conflict is thus for an important part technically independent and can be applied without difficulties in modern times.

The principles of the Law of Armed Conflict (military necessity, distinction, proportionality and humanity) are enshrined in the general thought that forms the foundation of the Law of Armed Conflict: to reduce as much as possible the horrors of an armed conflict. Both these principles and the treaty provisions that elaborate on the principles are directly applicable to air operations.

The principle of **military necessity** is closely connected with the ultimate objective of a military operation. Traditionally, that was the submission, in whole or in part, of the adversary. In modern times the objective can be less all-encompassing, depending on the international mandate of an operation. Military necessity allows the use of force insofar as it is necessary to secure the object of the operation and is not prohibited in the Law of Armed Conflict. The latter provision precludes a successful appeal to the principle of necessity in case of a breach of the Law of Armed Conflict. The nature and level of the use of force must be restricted to what is strictly necessary for reaching the objective.

The principle of **distinction** basically provides for operations to only be executed against military targets. This implies that the parties involved must always make a distinction between the civilian population and combatants; similarly between civilian objects and military targets (article 48 Additional Protocol I (AP I) to the Geneva Conventions, generally referred to as the '*basic rule*'). Therefore, it is important to know what precisely military targets are. Article 52 AP I describes them as objects that by their nature, location, purpose or use make an effective contribution to military action and whose total or partial destruction, capture or neutralization, in the circumstances ruling at the time, offer a definite military advantage.

This definition leaves room for interpretation and can, for example, give rise to problems regarding objects with both civilian and military use like power plants and oil tanks: the so called '*dual-use facilities*'.

> In 1999 NATO conducted an air campaign against the Federal Republic of Yugoslavia, codenamed Operation Allied Force. On 23 April 1999, at 02.20 hours, NATO intentionally launched an aerial attack on the headquarters of the state-owned Serbian TV and Radio Station (*Radio Televizija Srbije* -- RTS) in the centre of Belgrade. At the time of the attack an estimated 120 civilians were at work in the building. One missile hit the building causing the deaths of between 10 and 17 people. A news broadcast was disrupted but RTS resumed broadcasting 3 hours later.

NATO defended the attack against the RST inter alia because of the role the service played in the Serbian propaganda efforts. However, the argument that a broadcasting station is an instrument of propaganda and suppression is not sufficient to conclude that the attack offers a definite military advantage. A stronger argument that NATO advanced was the fact that the station was believed to be integrated into the command, control and communications system of the Yugoslav armed forces.

An amplification of the principle of distinction can be found in article 51, AP I, that prohibits indiscriminate attacks. This type of attack is not aimed at a specific, military objective, nor does it include the use of means or methods the effects of which cannot be limited as required by the Protocol. The former case prohibits the indiscriminate use of weapons, the latter the use of indiscriminate weapons. An example of the indiscriminate use of weapons was the series of Hezbollah attacks with Katusha rockets in the summer of 2006 on Israel. The attacks were generally not directly aimed at specific objects and were intended to land in populated areas resulting in a great deal of suffering by the civilian population. The rockets as such are not indiscriminate per se. For example, rockets can be launched against a formation of armoured vehicles in open fields. The same applies to cluster bombs. The sub-munitions of these weapons cover a large area while a part of the sub-munitions will not detonate (the '*dud rate*'). This makes the use of cluster bombs in populated areas questionable, but when aimed at a large military airbase the weapon is not necessarily considered indiscriminate.

Distinction as a principle places high demands upon the planning and execution of aerial attacks and puts the attacker in a dilemma. With the present state of technology it is possible to identify and attack targets with high precision. But does that imply that the use of '*dumb*' munitions is now prohibited and only precision weapons can be used? Not really, but many factors must be taken into account. E.g. the supply of precision weapons can be limited so that the assumption that a conflict may last for a lengthy period of time and precision weapons must be available throughout that period must be taken into consideration when deciding whether or not to launch precision weapons. This consideration may lead to the decision to attack an object with a less accurate weapon and save the precision weapon for later use. On the other hand, precision weapons are not that expensive anymore. It is possible to equip unguided, free-fall bombs at rather low costs with a GPS-guidance system. In short, whatever means are available, a decision is required every time whether the military advantage balances the damage to civilian property in those particular circumstances.

Targets that in the past would not qualify for attack by aircraft, for example because they were situated in urban surroundings, can nowadays be attacked with a diminished chance of collateral damage. Weapon systems and the people that operate them are not infallible, implying that a target can be missed and collateral damage still occurs. Therefore, tactics and weapons are being developed that diminish the physical effects of an attack. For example, the so-called '*focused-lethality munitions*' concentrate the explosive powers of a weapon on a smaller area, thus reducing the chance of collateral damage. Other weapons are designed to eliminate specific targets. For example, munitions that disseminate large quantities of carbon fibre filaments are intended to cause short-circuits in power plants and distribution stations, thus reducing the chance of additional damage to the surroundings.

The principle of **proportionality** is closely connected with the previous principles. It prohibits the launch of attacks that would be excessive in relation to the concrete and direct anticipated military advantage (articles 51 and 57, AP I). Proportionality is not an absolute rule: the extent of the damage is not as such conclusive but what matters is how the damage relates to the expected military advantage. In practice, emphasis is placed on the loss of human life. Proportionality has, however, a broader meaning than that. Indirect damage such as economical harm must be taken into account as well. For example, an attack on a bridge that serves a limited military goal but at the same time cuts off the population in the rear area from the rest of a country can thus be disproportional. In practice, it is not always easy to determine what is proportional or not. Various factors must be weighed against each other: what is the importance of the target, is enough target information available, what weapons are available, is the attack influenced by weather, terrain, visibility et cetera, are civilians present in the vicinity of the target and finally: what risks do the own forces run.

At the same time, an attacker must take precautionary measures to spare the civilian population, for example through the spreading of leaflets. There is an exception to this obligation: the population must be informed unless the circumstances do not permit such action. An attack for which the element of surprise is essential, relieves an attacker from his duty to warn the population. If the civilian population does not respond to the warning they will not necessarily lose their protection. Attackers must still take the civilian presence into account when preparing and executing the attack on a target. It may very well be possible that not everybody is capable of escaping the area, for example because of physical defects or because all escape routes are blocked.

Fig. 3, leaflet dropped by Israeli Air Force July 2006 in Lebanon

<u>To all citizens south of the Litani River</u>

Due to the terror activities being carried out against the State of Israel from within your villages and homes, the IDF is forced to respond immediately against these activities, even within your villages.

For your safety!

We call upon you to evacuate your villages and move north of the Litani River.
The State of Israel

As a counter balance to the principle of military necessity, the principle of **humanity** emphasizes the human factor. Chivalry is considered by some to be the origin of this principle. In mediaeval times feudal knights observed certain customary codes of honour, parts of which have evolved into the modern rules of the Law of Armed Conflict. For instance, it is not lawful to use the UN-emblem or internationally recognized protective signs like the Red Cross, on aircraft that are involved in combat missions.

In a broader sense, this principle limits nations in their choice of means and methods of warfare. The prohibition of weapons that may cause unnecessary suffering, like incendiary weapons, can be viewed along this line. Over the years a number of treaties have been concluded that prohibit the use of weapons that are considered to be excessively injurious. Recently, research has been started into munitions reinforced with depleted uranium. Although the use of these types of munitions is not forbidden, radiation may have longer-term effects on people and the environment that are difficult to identify. Planners must bear in mind that the use of depleted uranium weapons may be limited or perhaps prohibited entirely in the near future.

Restrictions in order to protect the natural environment reflect a more recent amplification of this principle. The NATO air attacks in 1999 included strikes against chemical plants and oil installations. As reports mentioned the release of pollutants, the Office of the Prosecutor (OTP) of the Yugoslav Tribunal in the Hague reviewed the possible breach of articles 35 and 55 AP I. The OTP concluded that the air strikes did not cause widespread, long-term and severe damage as mentioned in AP I. The OTP's review shows that during the planning of aerial attacks on industrial targets the probable long-term effects of these strikes should be taken into account as well.

Besides the general principles of the Law of Armed Conflict, specific treaty provisions are relevant to the use of airpower. The most important provisions have already been mentioned in the paragraphs above that dealt with the four principles. Other provisions concern the protection of hospitals, medical facilities, transport and personnel (Geneva Convention I) to be recognized by the 'Distinctive Emblem':

The protection of property that is of great importance to the cultural heritage of every people, such as monuments of architecture, art or history, manuscripts, archives, museums et cetera is covered by the Cultural Property Convention of 1954. During the 1991 Gulf War, the coalition air forces were, of course, prohibited to attack any of these objects. A problem arose as Iraq started to use historical sites for stocking military materials. Although these objects may lose the protected status under the Cultural Property Convention, the coalition policy did not change. For one part, the continuation of the protective regime was motivated by the fear of negative publicity in case of accidental damage to cultural property during an attack. For another part, the military necessity for the destruction of military objects simply vanished. It is clear that MiG fighter aircraft are military objectives. During the conflict Iraq put two of these aircraft on low-bed transporters and moved them to an important archaeological site some kilometres from the nearest airstrip. At that time the coalition had already established absolute air superiority. By moving the aircraft to a location that made it close to impossible to launch any of them within a reasonable timeframe, they lost their threat to the coalition forces. Therefore, it was decided to monitor the position of the aircraft but not to undertake any aerial attack against the now inert objects. Furthermore, AP I prohibits the attack or destruction of objects indispensable for the survival of the civilian population (article 54). Also, works and installations containing dangerous forces, like dykes and nuclear installations, may not be attacked if there is a risk of the release of the dangerous forces that could lead to severe losses among the civilian population (article 56). To facilitate identification, these objects may be marked with a special sign consisting of three orange circles:

Literature

Fleck, D. (ed.), Handbook of the law of visiting forces, Oxford 2001. Lazareff, S., Status of military forces under current international law, Leiden, 1971.
Rogers, A.P.V., Law on the battlefield, Manchester 2004.
Ronzitti, Natalino and Venturini, Gabriella (ed.), The law of air warfare, contemporary issues, Utrecht 2006.
Schmitt, Nichael N., Precision attack and international humanitarian law, International Review of the Red Cross, Vol. 87, Number 859, September 2005.
Shaw, M.N., International law, Cambridge 2003.
Essays on the future of the Air Force, Clingendael Centre for Strategic Studies (CCSS), The Hague 2006.

Chapter 21 – The history of airpower strategy

The origins of airpower

Technology enables military use

Strategy for the employment of air assets in principle does not differ from other fields of strategy and should be seen as a species of the genus 'general strategy'. The notion of an air strategy came into existence around the turn of the previous century when technological developments led to conquest of the air by the use of dirigibles and sail planes. Gas-filled and hot-air balloons had already extended *Feldherrnhügel* and reconnaissance positions for that matter, to higher elevations since Montgolfier in 1783, but the Wright brothers in 1903 achieved a major breakthrough by using a combustion engine to drive a propeller on a winged contraption. Some theorists soon grasped the implications of this industrial age innovation and started developing a new line of thought: it will soon be possible to use powered flight for military purposes and attack the enemy from the air.

Independent or supporting air operations

In essence this is the reason why a separate field of strategy had to be developed: the art of using air assets in order to achieve political goals in a conflict at acceptable costs. As this definition of airpower was simple, much debate arose around the question whether air assets could achieve these goals independently or always jointly. A straightforward answer to this question cannot be given, mainly because crises, conflicts and wars are one-time experiments that can be learned from but never be repeated. By studying history and keeping in step with technological developments, strategists may at least succeed in avoiding the failures of the past. It is noteworthy that over time air philosophers have agreed that airpower can be employed at all levels of conflict: in peace operations, times of crisis and all-out war. This does not mean that air planners can fall back on standard methods; as in other fields of strategy every new situation requires adaptive application. Independent or joint use should be seen as options in a strategic menu for winning the military part of the operation. Winning the peace is an entirely different matter, in which airpower can only play a supporting role.

Douhet

The best known air philosopher of the first hour and supporter of the independent line of thought was an Italian officer by the name of Giulio Douhet. As is often the case with new technologies, airpower was overestimated for many decades. Douhet assumed that modern warfare would not distinguish between combatants and non-combatants, that successful offensives by land forces would become impossible and that it would be fruitless to take defensive measures against an offensive aerial strategy. Therefore, a nation had to be prepared to launch massive bombing attacks against the enemy centres of population, government and industry, compelling the enemy government to sue for peace. Consequently, a nation had to build up an independent air force armed with long-range bombing aircraft as a primary military requirement. Other types of aircraft were nice to have, but not absolutely necessary. Douhet was an airpower strategist in extremis: he advocated the existence of only one service, thus army and navy were no longer relevant.

Word War I

First military use

When World War I dawned, field commanders had not yet taken the aircraft seriously as a military tool. The only probable use for aircraft was initially deemed to be as an extension of the eyes of the ground commanders, just as balloons had been used throughout the 19th century. The great mobility and range of powered aircraft rapidly widened this narrow scope, with pilots reporting enemy positions much more efficiently than the cavalry. Reconnaissance – it was called observation in those days – became the stepping stone to the more elaborate role of artillery spotting. As spotter aircraft represented an obvious threat, they were from then on challenged by enemy aircraft who attempted to shoot them down with handguns and rifles, later with machine guns mounted on own aircraft. Thus reconnaissance and pursuit roles were the first to clearly emerge.

Tactical roles

One innovation was the tactical support of forces engaged in combat, where aircraft guns and bombs would be directed against troop positions on the ground, with the aim of either assisting the advance of one's own troops or thwarting the advances of the enemy. Used in this manner, usually referred to as the attack role, the aircraft operated either close to the troops or at short distances in the enemy's rear against rallying points, such as supply dumps, key intersections, military headquarters and railheads. Today we would call these tactical roles close air support and interdiction. By the end of the war, another vision emerged, that of aircraft operating independently of armies and navies.

The first air force

This vision that implied entry into the strategic domain, was spurred by the German raids with Zeppelins and Gotha bombers over England. Neither army nor navy had been able to counter these attacks effectively, so the British government had decided to pull all air assets together into one service, the Royal Air Force, established on April 1st, 1918. The task of this new organization was to attack targets far removed from the battle lines, with the primary aim of destroying essential enemy war capabilities. At the back of the air commanders' minds were aerial operations against industrial and population centres on a vast scale. Implicitly, the tactical roles developed in WW I became secondary and subordinate. When the war ended, airpower had not brought about such supremacy. As instrument of war it was still in its infancy, having played an occasionally spectacular, but nonetheless unessential part in the outcome. The fight for the prominence of airpower however was to begin from here.

The interwar years

Mitchell

Together with Douhet (1869-1930), whose *Il dominio dell'aria* (1921) after translation gradually found its way to the military staff libraries of all modern states, several other airpower theorists made their cases for the ubiquitous employment of aircraft because of their speed and elevation. The American brigadier general and army pilot William Mitchell (1879-1936) was not so much an original thinker – he borrowed his ideas from an international community of airmen, which he joined in WW I – as a forceful advocate for centralized coordination of all air assets under one autonomous air command, freed from its dependence on the army. Contrary to Douhet, Mitchell stressed the importance of the tactical roles of airpower. He became notorious in navy circles after experiments with the bombardment of ships in the early twenties and was court-martialled. These tests became a bone of contention in the debate between airpower protagonists and conventional sea power supporters for many years to come. Unwittingly, he provoked the development of aircraft carriers, initially in scouting for the main battle fleet, later in power projection roles. Japan, with its brief naval tradition less committed to traditional ships of the line and more concerned with power projection, was the first to grasp the advantages of tactical airpower, of which the surprise attack at Pearl Harbor in December 1941 was clear testimony.

Airpower in the Netherlands

The Netherlands had been neutral during WW I and thus had missed important air developments. During a large part of the interwar years the aviation organizations (*Marine Luchtvaartdienst, Luchtvaartafdeling KL* and *Luchtvaartafdeling KNIL*) followed the progress of the most important nations. The idea of one independent air force was investigated in the early thirties but rejected by the Navy as counterproductive to the concept of 'harmonious' naval forces with a global reach. The Dutch-Indian Army denounced the idea, fearing that it would snatch her aviation trump card in the competition with the Navy for the primary defence role in the archipelago. The bomber force concept in the Indies fared better. Both ideas were strongly voiced by the Dutch army officer and air observer J.G.W. Zegers (1891-1953) whose influence as a publisher and member of two state committees on defence was felt among senior officers and politicians. Before the outbreak of the war with Japan, the Netherlands Indies possessed a fleet of nearly 120 bomber aircraft and supplementary fighters. J.J. van Santen (1893-1937), a prominent aviation commander at the time, conceived what he called *De Wet der Quadraten*: a geographically small nation like the Netherlands would need a bomber fleet x times the size of that of a much bigger neighbour in order to compensate for the number of extra targets, leading to the conclusion that a Dutch bomber force would be a waste of effort. By the same token, a colony like the Netherlands Indies could benefit immensely from a fleet of bombers in wartime, not against the enemy heartland but in countering an invasion of a troop transport fleet.

British developments

The man with the most extensive practical experience after WW I was Air Chief Marshal Hugh Trenchard, the first British aviation wartime field commander. His experiments with the Independent Air Force in the course of 1918 had been inconclusive and as Chief of the Air Staff he suggested the use of aircraft for policing functions in the colonies. This peacetime air control emphasized air presence, coercion and minimum application of force, leading to the substitution of air for ground forces notably in the Middle East and saving the Treasury considerable sums. In the mid-thirties a thorough doctrine for this employment had been worked out and was taught at the senior officers' courses. In the operational arena the strategic thinker Basil H. Liddell Hart (1895-1970) furthermore established the theoretical framework for the air-land team in armoured warfare, from which the German *Blitzkrieg* concept was at least partly derived. Ironically, this concept put aircraft under centralized command on a par with tanks and motorized infantry, an arrangement Douhet would have thought inefficient. J.F.C. Fuller (1878-1966) also contributed to army-air cooperation and developed what he called the Constant Tactical Factor: the idea that there never had been or would be an absolute weapon. The pendulum would always swing from offensive to defensive; the defence would always catch up with the offence.

This notion was not widespread in British government circles during the interwar years. Air operations against an enemy's material and moral resources justified an independent air force in order to avoid a repetition of the slaughter in World War I at all costs. Trenchard especially advocated the attack aimed at the enemy's morale, resulting in the major RAF commitments to Bomber Command. Cooler heads like Air Chief Marshal Hugh Dowding (1882-1970) switched to air defence and strengthened Fighter Command with two types of new fighter aircraft, and a comprehensive command and control system including the introduction of radar. This emphasis on defence came just in time and did not reflect the views of the majority of air staff at the time. The Battle of Britain – a narrow escape from a *Luftwaffe* victory – could hardly have been won without this switch in strategic thinking and the application of new technology.

World War II

American developments

In the United States, the Army's Air Corps Tactical School developed the doctrine of employment for operations against the enemy's industrial web. The role of bombing aircraft in this doctrine became dominant from 1932 onwards. The American planners could not envisage fleets of the size implied by Douhet and examined methods that would identify critical targets whose destruction would bring an entire industry to a halt. Identifying targets was one thing, hitting them from the air was something else. With the advent of the B-17, airpower theorists seemed to be able to bridge the gap between theory and practice. This bomber had the range, speed, altitude and bomb-carrying capacity deemed necessary. The new Norden bomb sight reinforced this optimism: the day might not be far off when a fleet of perhaps 100 B-17s could take off, fly at high altitudes above the effective height of the enemy's anti-aircraft guns and reach the target several hundred miles away, all the time being able to defend themselves and the formation as a whole. Over the target area, they would carefully sight in the target with the new bombsights, release their bomb loads and return to base. They would leave a badly crippled, if not devastated, factory behind and, because of its strategically perfect selection, implicitly an entirely crippled industry. It was decided that such attacks should take place in broad daylight, when accuracy was expected to be optimal. The question whether the bomber fleet should be escorted to the target by fighter aircraft was answered in the negative: the bombers could defend themselves and, in any case, fighters with equal range did not exist.

Strategic bombing

In the course of World War II this theory appeared to have many flaws. First, the unspoken assumption that precise intelligence regarding enemy targets would be available. Second, a tendency to overestimate existing capabilities, while at the same time minimizing the probable effects of limiting factors such as weather. Third, a gross optimism regarding the self-defence capability of bomber aircraft against a daring air force with a redundant command and control system. The Army Air Corps learned quickly from these experiences however and soon introduced the P-51 Mustang as escort fighter. Meanwhile, RAF Bomber Command kept trying to attack enemy morale through city bombing by night. Most historians today agree that, apart from the ethical aspects of attacks on civil population, the strategic bombardments of World War II suffered from an imbalance between input and output. An exception are the two nuclear bombs that brought the war against Japan to a close much faster than would have been possible with conventional means. The strategic use of nuclear airpower brought about a victory that had not been within reach with the use of conventional airpower and set the strategic scene for the Cold War.

Tactical roles

Tactical air employment in World War II on the other hand, showed spectacular successes. *Blitzkrieg* operations, close air support in the western desert of North Africa, Patton's use of airpower during his march through France: all were splendid examples of the decisive employment of aircraft in support of land forces. The same holds true for the endeavours of Bomber Command in the interdiction of German targets in Northwest France on the eve of the allied invasion on the Normandy beaches. The temporary use of bomber aircraft for interdiction instead of city bombing proved that it is possible to successfully downscale military means. On the other hand the *Luftwaffe* had experienced that it is nearly impossible to upscale tactical aircraft for strategic purposes, as the Battle of Britain had shown.

Airpower at sea

Japan's army and navy each had an air organization, but only the naval air arm developed a long-range striking force of some importance. Nevertheless, four months after Pearl Harbor the Doolittle raid, launched form the carrier Hornet against Tokyo, made Japan's strategic vulnerability clear to its planners, something they had not taken into account earlier. Sea power in World War II no longer depended on battleships, but on carrier-based airpower. In the Battle of the Coral Sea naval forces engaged for the first time without ever seeing each others' fleets.

Subordinate or co-equal

Much debate during World War II centred on the question of centralized command and control of air assets. The notion that land power and airpower are co-equal and independent slowly received acceptance during the North African campaign in 1943. This usually meant that air forces would achieve air superiority primarily in the theatre of operations, before concentrating on isolating the battlefield by attacks on enemy forces and supplies out of artillery reach. This role, called battlefield interdiction, would have priority over close air support, directly against enemy troops on the battlefield. The latter was seen increasingly as a less efficient use of air assets, but indispensable in unfavourable tactical situations. It should be noted here that the USA and Britain were both maritime nations, allowing for positions of equality between services. Germany and the Soviet-Union, being continental nations, subjugated the other services to the army. For them a strategic bomber force seemed to make less sense. Their air forces both designed frontline aircraft specialized for close air support.

Cold War

Domination of nuclear operations

After World War II the United States Air Force and the RAF perceived strategic roles as their primary mission, which was understandable in the nuclear age. The international political situation during the Cold War demanded a high state of readiness of the strategic bomber forces, later supplemented by intercontinental ballistic missiles and submarines. This implied a certain neglect of tactical roles, as became clear during the Korea conflict. General O.P. Weyland (1902-79), Patton's air commander in France and highest air officer in Korea, was famous for his remark that "*what was remembered from World War II was not written down, or if written down was not disseminated, or if disseminated was not read or understood*". Meanwhile, air superiority and (battlefield) interdiction were the roles that prevented defeat in Korea. Airpower could not prevent defeat in Vietnam however. Nuclear strategy had so much priority that even the tactical air forces had concentrated on implementing tactical nuclear roles. Conventional tactical roles had subsequently suffered. Interdiction campaigns and Offensive Counter Air (OCA) over North Vietnam provided limited results. It has to be said that political restraints in the application of force (e.g. rules of engagements), driven by fear of antagonising China and the Soviet-Union, compelled the USAF to half-hearted efforts. No precedents existed for using airpower to attain limited goals. Having said this, two strategic air offensives in 1972, *Linebacker I and II*, forced North Vietnam to the negotiation table. The secret for this success lay in the political consent for an all-out endeavour to turn the tables. The entry of helicopters in Vietnam on a large scale provided the US Army with its own airpower and considerably increased its tactical mobility, but could not become a decisive factor in winning the war. Yet, the Vietnam episode matured the role of the helicopter in airmobile and air assault operations, out of which a completely new doctrine emerged. Eventually, the US Army transferred its responsibilities, including its air assets, to the army of South Vietnam, but this political decision only marked the beginning of the end.

Decolonisation

In many respects the war in Vietnam resembled a post-colonial conflict, in which notably Britain, France and the Netherlands used airpower in an army-supporting mode. Its most effective use was in non-firepower roles, like reconnaissance, transport and liaison. Transport aircraft were sometimes employed to drop parachutists. The air raid on Yogyakarta in 1947 whereby Dutch commandos took rebel leader Sukarno prisoner, stands out as a major feat in the history of the intervention. However, never in the post-colonial era could airpower gain the status the RAF had achieved in policing the colonies from the air during the interwar years.

Airpower in crisis and war

The importance of surprise

Of the principles of war, surprise and concentration of force stand out as particularly applicable to airpower. The OCA attack on the Egyptian air force in the Six-day war of 1967 resembled the Japanese attack on Pearl Harbour and made the Israeli air planners instantly famous. This success could not be repeated in 1973 due to the Egyptian strategic surprise this time and the build-up of surface-based anti-aircraft defence, both SAMs and radar-directed AAA. It took an army unit to cross the Suez Canal and suppress those air defences from the ground, leading to the establishment of Israeli air superiority. The geographical and political situation of Israel gradually led to the innovative use of airpower, a development encouraged by the US. After a fruitless raid in 1980 by Iranian F-4 Phantoms on the Osiris nuclear reactor near Baghdad, Israeli F-15s and F-16s successfully destroyed the reactor a year later. A similar long-distance air raid on the PLO Headquarters in Tunisia in 1985 also had the desired strategic effect. Equally spectacular was the liberation of Israeli hostages from dictator Idi Amin in Uganda by Israeli C-130 Hercules transport aircraft. This raid on the airfield of Entebbe, carried out in 1976, appealed to the popular imagination because of the surprise that could be achieved by air assets. Surprise was also the decisive factor a year later, when Dutch F-104 jets in afterburner roared over a hijacked passenger train in Drenthe in 1975, paving the way for a marines' ground attack.

Transport

Transport aircraft are indispensable to air forces for many logistical reasons, but have equal importance for army and navy support. The 1948/49 Berlin Airlift, an independent operation, was the heyday of military air transportation. In its strategic setting this operation enabled the Western allies to bring Soviet expansionism in Eastern Europe to a halt. This peaceful exploitation of airpower for humanitarian purposes serves as an example of what the military can do for the civic powers. Crisis operations in the heart of Africa and Central Asia nowadays depend heavily on the employment of military transport assets. Air-to-air refuelling is rightly seen as a strategic capability, making power projection over large distances possible.

Helicopters

Helicopters are often seen as aircraft that do not need runways. This does not do justice to the true characteristics of rotary-wing aircraft: being able to move around freely in the terrain at a self-chosen speed or at an altitude overhead the battlefield, beyond the reach of light and medium-calibre arms. The multi-sensor AH-64 Apache is able to hover at 10 km distance and launch its Hellfire missiles with precision on the target. This attack helicopter set the scene for *Operation Desert Storm* in 1991, when it took out Iraqi radar sites as a kind of Suppression of the Enemy Air Defence (SEAD) mission, enabling fighter-bombers to proceed through the gap to targets around Baghdad. Armed Chinook helicopters have been the backbone of nearly every military operation since Vietnam. Jets badly need Combat Search and Rescue (CSAR) helicopters when pilots or crews have to be picked up in enemy-controlled zones, as USAF Captain Scott O'Grady experienced in 1995 during *Operation Deny Flight* over Bosnia.

NATO airpower

NATO airpower, or rather aerospace power, is defined as *the military employment of any system that operates in or passes through air or space*. During the Cold War employment of airpower focused on wartime operations. Since the early nineties war operations have been overshadowed by peacetime and crisis operations in the coalition, leading to a change of assets in most NATO countries: less fighter aircraft, more transport aircraft and helicopters. Operations *Desert Storm* and *Iraqi Freedom* stand out as exceptions to the rule: both relied heavily on the use of fighter-bomber based airpower, but eventually ground forces had to occupy the terrain. An interesting phenomenon was *Operation Allied Force* in 1999. Relying solely on airpower, NATO compelled the Serbian authorities to give up their occupation of Kosovo. Although the desired end-state was finally reached, NATO had limited its own freedom of choice considerably by showing it had no appetite for a ground offensive.

Modern technological developments

Airpower has become a weapon of choice for politicians in Western democracies. Because of its main characteristics (altitude, speed and reach) the perceived political effects of air employment have an immediate and global reach, with a lower risk of casualties than that associated with the use of land power. The above characteristics can be subdivided into: flexibility, mobility, speed of reaction, concentration precision, penetrating capacity and omnipresence. It goes without saying that modern technological developments like air-to-air refuelling (AAR), precision-guided munitions (PGM), global positioning system (GPS), airborne warning and control system (AWACS) and joint surveillance, targeting and reconnaissance system (JSTARS), satellites, stealth and (near) real-time information have immensely contributed to the enhancement of those sub-characteristics. Network-enabled capabilities (NEC) originate from the airpower domain. Limitations of airpower still lie in the duration of its employment and the effective loads airborne platforms can carry. Of course, the occupation of terrain will always be an impossibility. The development of unmanned aerial vehicles (UAV) for power projection is still in its infancy, although surveillance missions complement other types of airpower extensively. It can be compared to the position of fixed-wing aircraft on the eve of World War II and rotary wing on the eve of Korea.

Modern theorists

Theories on the use of airpower have improved over time. The most prominent airpower strategist is USAF Colonel John A. Warden III (1945-), whose *The Air Campaign* (1989) and the model of the five strategic rings have already withstood the test of time. The five rings somewhat resemble the targets identified by Giulio Douhet and are at the same time a refinement of the *Ziel* in Clausewitz' (1780-1831) *Vom Kriege*. Warden's *Centre of gravity* (COG) is a new dimension to Clausewitz' *Schwerpunkt*. A COG is a capability within the enemy system whose deactivation or destruction will have fatal consequences for the enemy's ability to continue the war. Warden recommends analysing the most important enemy target systems and selecting the COG from those. The five (concentric) rings are: leadership (military and civic command & control), organic essentials (energy production and vital industries), infrastructure (notably transport-oriented), population and fielded forces. The essence of his theory is that the more one selects the outer rings, the more limited the effects in time and place will be.

John R. Boyd (1927-97) is often called an airpower theorist. He was a USAF colonel as well as a competent fighter pilot, but his influence on war philosophy goes far beyond airpower. His OODA (observation, orientation, decision, action)-loop has become a corner stone of decision-making theories. The essence is to make the right selection of choices in the orientation phase and run through the cycles faster than your adversary, all the while increasing friction (a Clausewitzian notion) for the adversary and decreasing it for oneself.

Coercion by airpower

The American professor of political science Robert A. Pape (1960-) has researched strategic air offensives. He regards these as a form of *coercion,* meant to change the enemy's behaviour. Pape argues that the strategic air offensive has had little success, implying that attempts to have the enemy population suffer in order to provoke a regime change have never been fruitful. On the other hand, it is likely that the Iraqi army deserted massively in 1991 due to the B-52 attacks, executed after the inconclusive air offensive against the leadership. Pape concludes that attacks on enemy vital systems (*punishment*) do not produce the desired results, whereas attacks on enemy fighting power (*denial*) do. Coercion implies both *deterrence* and *compellence*. Deterrence is applied to induce an adversary to refrain from taking any intolerable actions he intends to take, while compellence is meant to induce an adversary to reverse an intolerable action he has taken earlier. Jan F.W. van Angeren (1947-), RNLAF colonel and scholar, has developed a paradigm for the opportunities and limitations of compellence strategies. He views the employment of airpower as part of a compellence strategy aimed at curbing the adversary's decision-making process as a more usable option than the intended defeat of an enemy through the use of air attacks.

Irregular warfare

Success for airpower in irregular or asymmetrical warfare may be the most difficult to accomplish. Robert A. Pape points at positive American results during Operation *Enduring Freedom*, when air and land power were jointly employed in a 'hammer and anvil' mode. When the Taliban amassed in order to withstand ground attacks, they became an air target; when they scattered in order to avoid air attacks, they could be mopped up by land forces. Thus, air superiority must be seen as a major advantage for Western powers seeking to combat irregular forces. On the downside, airpower operating independently may be confronted with sensor-to-shooter time lags. When UAVs spot irregular forces launching ground-to-ground missiles, airpower must react immediately. If it does not, the irregulars will disappear into the crowd after launch. Experiments with Predator UAVs outfitted with Hellfire missiles have shown promising results. It is self-evident that longer-range missiles need more complicated launching systems and are therefore easier to spot. In urban environments and mountainous or wooded terrain the advantage of air superiority will be harder to attain. Continuous air surveillance will undoubtedly enhance the intelligence picture in all cases. When avoidance of collateral damage is a political imperative, PGMs or non-lethal bombs may provide an answer. Much heavier ordnance will have to be used against enemy forces hiding in bunkers or caves. Much study has yet to be done on airpower roles in counter-terrorist operations.

Because of its inherent mobility and flexibility, airpower will be asked to play a role in almost every tactical and strategic situation. Therefore it will continuously focus on the application of modern technology. To prevent air force officers from overemphasizing new and promising innovations, it is mandatory that they study the employment of airpower in the past and develop a doctrine that is based equally on technology and earlier experience.

List of Abbreviations

AAA	Anti Aircraft Artillery
AAR	Air-to-air refuelling
ACSA	Acquisition and Cross Servicing Agreement
ADC	Air Data Computer
AI	Air Interdiction
ALCM	Air Launched Cruise Missile
ALO	Air Liaison Officer
AO	Area of Operation
ARM	Anti Radiation Missile
ASFAO	Anti Surface Forces Air Operation
AST	Aircraft security Team
ATC	Air Traffic Control
ATO	Air Task Order
AWACS	Airborne Warning and Control System
BDA	Battle Damage Assessment
BVR	Beyond Visual Range
CAP	Combat Air Patrol
CAS	Calibrated Airspeed
CAS	Close Air Support
CCIP	Constant Computed Impact Point
CCRP	Constant Computed Release Point
CEP	Circular Error Probable
CFB	Corps Forward Boundary
CI	Counter Intelligence
CIMIC	Civil Military Cooperation
CIS	Communication and Information System
CNA	Computer Network Attack
CND	Computer Network Defence
CNE	Computer Network Exploitation
CNO	Computer Network Operations
COA	Courses of Action
COAC	Combined Air Operations Centre
CoG	Centres of Gravity
COMAO	Composite Air Operation
CONOPS	Concept of Operations
CP	Control Point
CRAF	Commercial Reserve Air Force
CRC	Control and Reporting Centre
CSAR	Combat Search and Rescue
DAMA	Damage Assessment
DAMCON	Damage Control
DART	Downed Aircraft Recovery Team
DCA	Defensive Counter Air
DIRCM	Directed Infra Red Counter Measures
DMPI	Desired Mean Point of Impact
DTED	Digital Terrain Elevation Data
EA	Electronic Attack

EBAO	Effect Based Approach to Operations
EM	Electro Magnetic
EOC	Essential Operating Capabilities
EOR/EOD	Explosive Ordnance Reconnaissance/ Explosive Ordnance Disposal
EP	Electronic Protection
ES	Electronic Support
EU	European Union
EW	Electronic Warfare
FAC	Forward Air Controller
FALSTAFF	Fighter Aircraft Loading Standard For Fatigue
FAOR	Fighter Area of Responsibility
FARP	Forward Air Refuelling Point
FFT	Fast Fourier Transform
FLIR	Forward Looking Infra Red
FLOT	Forward Line Own Troops
FOB	Forward Operating Base
FOM	Follow on Maintenance
FP	Force Protection
FSCL	Fire Support Coordination Line
GBAD	Ground Based Air Defence
GHz	Giga Hertz
GPS	Global Positioning System
HARM	High Speed Anti Radiation Missile
HESCO	Splinter protection wall
HF	High Frequency
HN	Host Nation
HUD	Head-up Display
Hz	Hertz
IA	Information Assurance
IAS	Indicated Airspeed
ICCS	Individual Common Core Skills
ICT	Information and Communication Technology
IED	Improvised Explosive Device
INU	Inertial Navigation Unit
IR	Infra Red
ISAF	International Security Assistance Force
IT	Information Technology
JAAT	Joint Air Attack Team
JDAM	Joint Direct Attack Munitions
JMEM	Joint Munitions Effectiveness Manual
JSEAD	Joint Suppression of Enemy Air Defence
JSTARS	Joint Surveillance and Target Attack Radar System
KEK	Kampfeinsitzerkommando's
LF	Low Frequency
LGB	Laser Guided Bomb
LOAC	Law of Armed Conflict
LOS	Line of Sight
MAAP	Master Air Attack Plan
MAD	Mutual Assured Destruction
MAWS	Missile Approach Warning System
MF	Medium Frequency
MFD	Multi Functional Display
MHz	Mega Hertz

MILDEC	Military Deception
MOD	Ministry of Defence
MOU	Memorandum Of Understanding
MTBF	Mean Time Between Failure
NATO	North Atlantic Treaty Organisation
NBC	Nuclear Biological and Chemical
NCTR	Non Cooperative Target Recognition
NCW	Network Centric Warfare
NEC	Network Enabled Capabilities
NVG	Night Vision Goggle
OCA	Offensive Counter Air
OEF	Operation Enduring Freedom
OIF	Operation Iraqi Freedom
OPPLAN	Operation Plan
OPSEC	Operation Security
ORM	Operational Risk Management
OTP	Office of the Prosecutor
PA	Public Affairs
PAR	Post Attack Reconnaissance
PGM	Precision Guided Munitions
PHYSAT	Physical Attack
PHYSEC	Physical Security
Pk	Probability of Kill
PPI	Political Policy Indicator
PPS	Political Policy Statement
PRF	Pulse Repetition Frequency
PSO/RSO	Peace Support Operations/ Crises Response Operations
PSYOPS	Psychological Operations
PVI	Pilot Vehicle Interface
QNH	Local sea level barometric pressure
QRA	Quick Reaction Alert
RAP	Recognized Air Picture
ROTA	Releases Other Than Attack
RPG	Rocket Propelled Grenade
SA	Situational Awareness
SA	Staging Area
SAC	Strategic Air Command
SACEUR	Supreme Allied Commander Europe
SAM	Surface to Air Missile
SAR	Synthetic Aperture Radar
SEAD	Suppression of Enemy Air Defence
SHF	Super High Frequency
SHORAD	Short Range Air Defence
SLM	Shoulder Launched Missile
SOFA	Status of Force Agreement
SOJ	Stand of Jamming
TACAN	Tactical Air Navigation
TAS	True Airspeed
TERPROM	Terrain Profile Matching
TMD	Theatre Missile Defence
TOA	Transfer of Authority
UAV	Unmanned Aerial Vehicle
UHF	Ultra High Frequency

UN	United Nations
VA	Vulnerability Assessment
VC	Viet Cong
VHF	Very High Frequency
VOR	VHF Omni Range
WCMD	Wind Corrected Munitions Dispenser