Part I

Engineering issues specific to entry probes, landers or penetrators

This part of the book is intended to act as a guide to the basic technological principles that are specific to landers, penetrators and atmospheric-entry probes, and to act as a pointer towards more detailed technical works. The chapters of this part aim to give the reader an overview of the problems and solutions associated with each sub-system/flight phase, without going into the minutiae.

1

Mission goals and system engineering

Before journeying through the various specific engineering aspects, it is worth examining two important subjects that have a bearing on many more specific activities later on. First we consider systems engineering as the means to integrate the diverse constraints on a project into a functioning whole. We then look at the choice of landing site for a mission, a decision often based on a combination of scientific and technical criteria, and one that usually has a bearing on the design of several sub-systems including thermal, power and communications.

1.1 Systems engineering

Engineering has been frivolously but not inaptly defined as 'the art of building for one dollar that which any damn fool can build for two'. Most technical problems have solutions, if adequate resources are available. Invariably, they are not, and thus skill and ingenuity are required to meet the goals of a project within the imposed constraints, or to achieve some optimum in performance.

Systems engineering may be defined as

> the art and science of developing an operable system capable of meeting mission requirements within imposed constraints including (but not limited to) mass, cost and schedule.

The modern discipline of systems engineering owes itself to the development of large projects, primarily in the USA, in the 1950s and 1960s when projects of growing scale and complexity were undertaken. Many of the tools and approaches derive from operational research, the quantitative analysis of performance developed in the UK during World War II.

Engineering up to that epoch had been confined to projects of sufficiently limited complexity that a single individual or a team of engineers in a dominant discipline could develop and implement the vision of a project. As systems

became more sophisticated, involving hundreds of subcontractors, the more abstract art of managing the interfaces of many components became crucial in itself.

A general feature is that of satisfying some set of requirements, usually in some optimal manner. To attain this optimal solution, or at least to satisfy as many as possible of the imposed requirements, usually requires tradeoffs between individual elements or systems. To mediate these tradeoffs requires an engineering familiarity and literacy, if not outright talent, with all of the systems and engineering disciplines involved. Spacecraft represent particularly broad challenges, in that a wide range of disciplines is involved – communications, power, thermal control, propulsion and so on. Arguably, planetary probes are even more broad, in that all the usual spacecraft disciplines are involved, plus several aspects related to delivery to and operation in planetary environments, such as aerothermodynamics, soil mechanics and so on.

While engineers usually like to plough into technical detail as soon as their task is defined, it is important to examine a broad range of options to meet the goals. As a simple example, a requirement might be to destroy a certain type of missile silo. This in turn requires the delivery of a certain overpressure onto the target. This could be achieved, for example, by the use of a massive nuclear warhead on a big, dumb missile. Or one might attain the same result with a much smaller warhead, but delivered with precision, requiring a much more sophisticated guidance system. Clearly, these are two very different, but equally valid, solutions.

It is crucial that the requirements be articulated in a manner that adequately captures the intent of the 'customer'. To this end, it is usual that early design studies are performed to scope out what is feasible. These usually take the form of an assessment study, followed by a Phase A study and, if selected, the mission proceeds to Phases B and C/D for development, launch and operation. During the early study phases a mission-analysis approach is used prior to the more detailed systems engineering activity. Mission analysis examines quantitatively the top-level parameters of launch options, transfer trajectories and overall mass budget (propellant, platform and payload), without regard to the details of subsystems.

In the case of a planetary probe, the usual mission is to deliver and service an instrument payload for some particular length of time, where the services may include the provision of power, a benign thermal environment, pointing and communications back to Earth.

The details of the payload itself are likely to be simply assumed at the earliest stages, by similarity with previous missions. Such broad resource requirements as data rate/volume, power and mass will be defined for the payload as a whole. These allow the design of the engineering system to proceed, from selecting

among a broad choice of architectures (e.g. multiple small probes, or a single mobile one) through the basic specification of the various subsystems.

The design and construction of the system then proceeds, usually in parallel with the scientific payload (which is often, but not always, developed in institutions other than that which leads the system development), perhaps requiring adaptation in response to revised mission objectives, cost constraints, etc. Changes to a design become progressively more difficult and expensive to implement.

1.1.1 The project team

The development team will include a number of specialists dedicated to various aspects of the project, throughout its development. In many organizations, additional expertise will additionally be co-opted on particular occasions (e.g. for design reviews, or particularly tight schedules).

The project will be led by a project manager, who must maintain the vision of the project throughout. The project manager is the single individual whose efforts are identified with the success or otherwise of the project. The job entails wide (rather than deep) technical expertise, in order to gauge the weight or validity of the opinions or reports of various subsystem engineers or others and to make interdisciplinary tradeoffs. The job requires management skills, in that it is the efforts of the team and contractors that ultimately make things happen – areas where members of the team may variously need to be motivated, supported with additional manpower, or fired. Meetings may need to be held, or prevented from digressing too far. And this demanding job requires political skill, to tread the compromise path between constraints imposed on the project, and the capabilities required or desired of it.

A broadly similar array of abilities, weighted somewhat towards the technical expertise, is required of the systems engineer, usually a nominal deputy to the project manager. A major job for the systems engineer is the resolution of technical tradeoffs as the project progresses. Mass growth, for example, is a typical feature of a project development – mass can often be saved by using lighter materials (e.g. beryllium rather than aluminium), but at the cost of a longer construction schedule or higher development cost.

A team of engineers devoted to various aspects of the project, from a handful to hundreds, will perform the detailed design, construction and testing. The latter task may involve individuals dedicated to arranging the test facilities and the proper verification of system performance. Where industrial teams are involved, various staff may be needed to administer the contractual aspects. Usually the amount of documentation generated is such as to require staff dedicated to the maintenance of

records, especially once the project proceeds to a level termed 'configuration control', wherein interfaces between various parts of the project are frozen and should not be changed without an intensive, formal review process.

In addition to the hardware and software engineers involved in the probe system itself, several other technical areas may be represented to a greater or lesser extent. Operations engineers may be involved in the specification, design, build and operation of ground equipment needed to monitor or command the spacecraft, and handle the data it transmits. There may be specialists in astrodynamics or navigation. Finally, usually held somewhat independently from the rest of the team, are quality-assurance experts to verify that appropriate levels of reliability and safety are built into the project, and that standards are being followed.

In scientific projects there will be a project scientist, a position not applicable for applications such as communications satellites. This individual is the liaison between the scientific community and the project. In addition to mediating the interface between providers of the scientific payload and the engineering side of the project, the project scientist will also coordinate, for example, the generation or revision of environmental models that may drive the spacecraft design.

The scientific community usually provides the instruments to a probe. The lead scientist behind an instrument, the principal investigator (PI), will be the individual who is responsible for the success of the investigation. Usually this means procuring adequate equipment and support to analyse and interpret the data, as well as providing the actual hardware and software. An instrument essentially acts as a mini project-within-a-project, with its own engineering team, project manager, etc.

For the last decade or so, NASA has embraced so-called PI-led missions, under the Discovery programme. Here a scientist is the originator and authority (in theory) for the whole mission, guiding a team including agency and industry partners, not just one experiment. This PI-led approach has led to some highly efficient missions (Discovery missions have typically cost around $300M, comparable with the ESA's 'Medium' missions) although there have also been some notable failures, as with any programme. The PI-led mission concept has been extended to more expensive missions in the New Frontiers line, and for Discovery-class missions in the Mars programme, called 'Mars Scout'.

A further class of mission deserves mention, namely the technology-development or technology-validation mission. These are intended primarily to demonstrate and gain experience with a new technology, and as such may involve a higher level of technical risk than one might tolerate on a science-driven mission. Some missions (such as those under NASA's New Millenium programme, notably the DS-2 penetrators) are exclusively driven by technology

goals, with a minimal science payload (although often substantial science can be accomplished even with only engineering sensors). In some other cases, the science/technology borderline is very blurred – one example is the Japanese Hayabusa asteroid sample return: this mission offers a formidable scientific return, yet was originally termed MUSES-C (Mu-launched space-engineering satellite).

Whatever the political definitions and the origin of the mission requirements, it must be recognized that there is both engineering challenge and science value in any spacecraft measurement performed in a planetary environment.

A dynamic tension usually exists in a project, somewhat mediated by the project scientist. Principal investigators generally care only about their instrument, and realizing its maximum scientific return, regardless of the cost of the system needed to support it. The project manager is usually confronted with an already overconstrained problem – a budget or schedule that may be inadequate and contractors who would prefer to deliver hardware as late as possible while extorting as much money out of the project as possible. One tempting way out is to descope the mission, to reduce the requirements on, or expectations of, the scientific return. Taken to the extreme, however, there is no point in building the system at all. Or a project that runs too late may miss a launch window and therefore never happen; a project that threatens to overrun its cost target by too far may be cancelled. So the project must steer a middle path, aided by judgement and experience as well as purely technical analysis – hence the definition of systems engineering as an art.

1.2 Choice of landing site

Technical constraints are likely to exist on both the delivery of the probe or lander, and on its long-term operation. First we consider the more usual case where the probe is delivered from a hyperbolic approach trajectory, rather than a closed orbit around the target.

The astrodynamic aspects of arrival usually specify an arrival direction, which cannot be changed without involving a large delivered-mass penalty. The arrival speed, and the latitude of the incoming velocity vector (the 'asymptote', or V_∞, unperturbed by the target's gravity, is usually considered) are hence fixed. Usually the arrival time can be adjusted somewhat, which may allow the longitude of the asymptote to be selected for sites of particular interest, or to ensure the landing site is visible from a specific ground station. Occasionally this is fixed too, as in the case of Luna 9 where the descent systems would not permit any horizontal velocity component – the arrival asymptote would only be vertical at near-equatorial landing sites around 64°W.

The target body is often viewed in the planning process from this incoming V_∞: the plane going through the centre of the target body orthogonal to that vector is often called the 'B-plane'. The target point may be specified by two parameters. The most important is often called the 'impact parameter', the distance in the B-plane between the centre and the target point. For a given target body radius (either the surface radius, or sometimes an arbitrary 'entry interface' above which aerodynamic effects can be ignored) a given impact parameter will correspond to a flight-path angle, the angle between the spacecraft trajectory and the local horizon at that altitude. This may often be termed an entry angle.

The entry angle is usually limited to a narrow range because of the aerothermodynamics of entry. Too high an angle (too steep) – corresponding to a small impact parameter, an entry point close to the centre of the target body – and the peak heating rate, or the peak deceleration loads, may be too high. Too

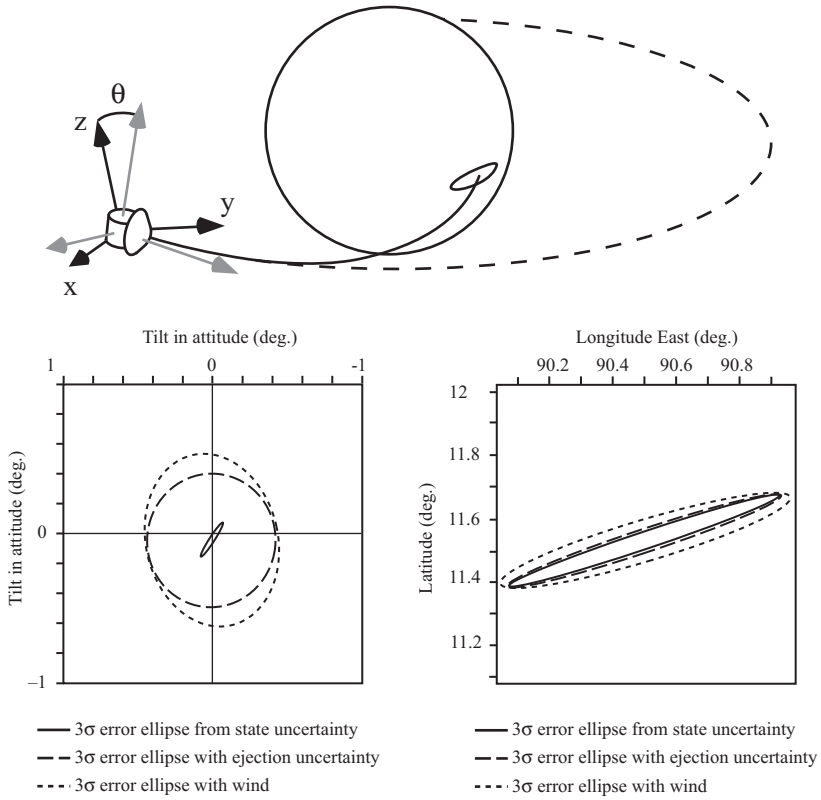

Figure 1.1. Top: cartoon illustrating lander-delivery uncertainty arising from uncertainties in the state vector at deployment. Bottom: attitude and landing-error ellipses for Beagle 2 (adapted from Bauske, 2004).

shallow an angle may result in a large total heat load; in the limiting case of a large impact parameter, the vehicle may not be adequately decelerated or may miss the target altogether.

The entry protection performance may also introduce constraints other than simple entry angle. For the extremely challenging case of entry into Jupiter's atmosphere, the ~12.5 km s^{-1} equatorial rotation speed is a significant increment on the entry speed of ~50 km s^{-1}. Heat loads vary as the cube of speed, and thus by aiming at the receding edge of Jupiter (i.e. the evening terminator, if coming from the Sun) the entry loads are reduced by a factor $(50+12.5)^3/(50-12.5)^3 = 4.6$, a most significant amelioration.

The second parameter is the angle relative to the target body equator (specifically where the equatorial plane crosses the B-plane) of the impact parameter. A B-plane angle of zero is on the equator; 90° means the entry point falls on the central meridian as seen from the incoming vector.

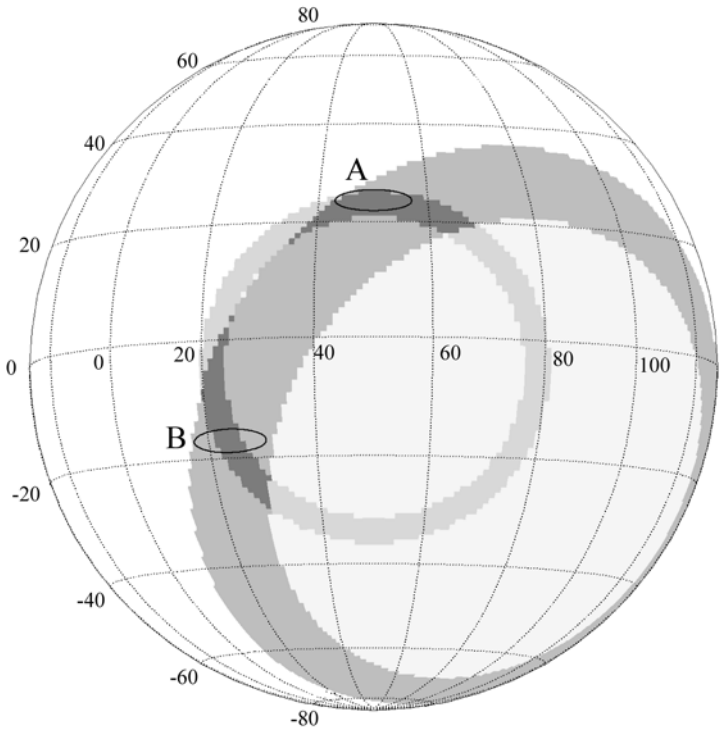

Figure 1.2. View of Titan from the arrival asymptote of the Huygens probe, with overlapping annuli reflecting the constraints on entry angle (light grey) and solar elevation (darker grey). Of the choice of two target locations where the regions overlap (A and B), only A accommodates the probe's delivery ellipse.

Other constraints include the communication geometry – if a delivery vehicle is being used as a relay spacecraft, it may be that there are external constraints on the relay's subsequent trajectory (such as a tour around the Saturnian system) which specify its target point in the B-plane. Targeting the flyby spacecraft on the opposite side of the body from the entry probe may limit the duration of the communication window. Current NASA missions after the Mars mission failures in 1999 now require mission-critical events to occur while in communication with the Earth: thus entry and landing must occur on the Earth-facing hemisphere of their target body.

Another constraint is solar. The entry may need solar illumination for attitude determination by a Sun sensor (or no illumination to allow determination by star sensor!), or a certain amount of illumination at the landing site for the hours following landing to recharge batteries. These aspects may influence the arrival time and/or the B-plane angle.

So far, the considerations invoked have been purely technical. Scientific considerations may also apply. Optical sensing, either of atmospheric properties, or surface imaging, may place constraints on the Sun angle during entry and descent. Altitude goals for science measurements may also drive the entry angle (since this determines the altitude at which the incoming vehicle has been decelerated to parachute-deployment altitude where entry protection – which usually interferes with scientific measurements – can be jettisoned).

The entry location (and therefore 'landing site') for the Huygens probe was largely determined by the considerations described so far (at the time the mission was designed there was no information on the surface anyway). The combination of the incoming asymptote direction and the entry angle defined an annulus of locations admissible to the entry system. The Sun angles required for scientific measurements of light scattering in the atmosphere, and desired shadowing of surface features defined another annulus. These two annuli intersected in two regions, with the choice between them being made partly on communications grounds.

There may be scientific desires and technical constraints on latitude. Latitude may be directly associated with communication geometry and/or (e.g. in the case of Jupiter), entry speed. For Mars landers in particular, the insolation as a function of latitude and season is a crucial consideration, both for temperature control and for solar power. Many Mars lander missions are restricted to 'tropical' landing sites in order to secure enough power.

So far, the planet has been considered only as a featureless geometric sphere. There may be scientific grounds for selecting a particular landing point, on the basis of geological features of interest (or sites with particular geochemistry such as polar ice or hydrated minerals), and, depending on the project specification,

these may be the overriding factor (driving even the interplanetary delivery trajectory).

A subtle geographical effect applies on Mars, where there is extreme topographical variation – of order one atmospheric scale height. Thus selecting a high-altitude landing site would require either a larger parachute (to limit descent rate in the thinner atmosphere), or require that the landing system tolerate higher impact speed.

The landing sites for the Mars Exploration Rover missions (MER-A and -B) were discussed extensively (Kerr, 2002). The not-infrequent tradeoff between scientific interest and technical risk came to the fore. As with the Viking landing site selection, the most scientifically interesting regions are not the featureless plains preferred by spacecraft engineers. The situation is complicated by the incomplete and imperfect knowledge of the landing environment.

One constraint was that the area must have 20% coverage or less by rocks 0.5 m across or larger that could tear the airbags at landing. Rock distributions can be estimated from radar techniques, together with geological context from Mars orbit (while rocks cannot be seen directly, geological structures can – rocks are unlikely to be present on sand dunes, for example), and thermal inertia data.

Although there are no direct wind measurements near the surface in these areas, models of Martian winds are reaching reasonable levels of fidelity, and these models are being used to predict the windspeeds at the candidate landing sites. Winds of course vary with season (e.g. the Martian dust-storm season, peaking soon after solar longitude $L_s = 220°$, is best avoided!) and time of day.

The Pathfinder lander, for example, landed before dawn, at 3 a.m. local solar time, when the atmosphere was at its most stable. The MER had an imperative (following on in turn from the Mars Polar Lander failure) that it must be in communication with the Earth during its descent and landing. This requires that it land in the afternoon instead – when winds are strongest! Here, perversely, a politically driven engineering uncertainty introduces a deterministic (i.e. certain!) increase in risk.

On Earth, a handful of landing sites are used. The US manned missions in the 1960s and 1970s relied upon water landing; the mechanical properties of the ocean are well understood and uniform over some 60% of the globe, with the only variable being uncertain winds and sea state. Other missions (unmanned capsules and Russian manned missions) have landed on large flat areas, notably the Kazakhstan steppe, and Utah was used for the Genesis solar wind and Stardust comet-sample-return missions. A significant factor in the choice of landing is the accuracy with which the capsule can be targeted (oceans may be less desirable landing sites, but they are hard to miss) and whether a particularly rapid retrieval (e.g. for frozen comet samples) is required.