Proc. Indian Acad. Sci. (Engg. Sci.) Vol. 5, Pt. 3, September 1982, pp. 197–258. © Printed in India.

A review of satellite lifetime and orbit decay prediction†

J DE LAFONTAINE and S C GARG*
Institute for Aerospace Studies, University of Toronto, 4925 Dufferin Street, Downsview, Ontario, Canada M3H 5T6
*Present Address: Ford Aerospace and Communications Corporation, Palo Alto, California, U.S.A.

MS received 1 October 1981; revised 17 July 1982

Abstract. This review considers the modelling and the analytic methods necessary to predict satellite lifetimes. After a brief discussion of orbital perturbations, models for Earth's potential field, luni-solar attraction and atmospheric drag are reviewed. Detailed static and dynamic density models are presented followed by a discussion of satellite properties (projected area and drag coefficient). Closed-form solutions to the prediction problem using these models are examined and compared. In conclusion, an assessment of the state of the art and the identification of areas requiring more research are given.

Keywords. Satellite life time; satellite trajectory; orbital perturbations; orbit decay; earth's potential field; luni-solar attraction; atmospheric models; dynamic density model; static density model; projected area; drag coefficient; perturbation modelling; decay stage; hydrostatic equation

1. Introduction

In the past two decades, as the number of satellites in low Earth orbit has increased, scientists have become concerned with the particular problems associated with such orbits. First, low orbits involve drag disturbances which are large compared to the other orbital perturbations. Hence, missions that require accurate trajectory prediction need a realistic model of the atmosphere and of the forces it exerts on the satellite. Secondly, as the satellite loses its energy through atmospheric friction, its orbit contracts to the point that re-entry eventually occurs. Break-up, ablation and volatilization follow and there remains a usually small but non-negligible probability of ground impact. Some examples of actual re-entry are given by King-Hele (1976), King-Hele & Walker (1958) and King-Hele *et al* (1969). Eberst (1978) shows that, on the average, about 570 decays per year occurred during the period 1974–1977. This represents about 57% of the number of satellites launched annually. Falling satellites may represent hazards to people, aircraft and other objects. *Cosmos* 954 and *Skylab* immediately come to mind, but these two cases are not exceptions; many other examples are given by Drago & Edgecombe (1974).

In view of this challenge, many research projects were devoted to satellite motion in an atmosphere, more specifically in trajectory prediction (dispersion analysis) and lifetime estimation.

†This work is the revised version of UTIAS Review No. 43, December 1979.
A list of symbols appears at the end of the paper.

Research on trajectory dispersion analysis is carried out in order to:

(i) predict with better accuracy the position of a satellite at any future time,
(ii) provide the analytical tools necessary to derive the properties of the atmosphere (chiefly density and its variation with altitude),
(iii) give some insight into the particular and sometimes unexpected behaviour of satellites subject to atmospheric drag, like early re-entry, lateral dispersion, decreasing inclination, etc.

Research in the lifetime estimation of a satellite under the influence of air drag is directed towards:

(i) computing in advance the expected lifetime of satellites, permitting adequate mission planning,
(ii) in cases where the satellite is uncontrolled, predicting the re-entry time so as to avoid aircraft hazards, body injuries, etc.,
(iii) in cases where the satellite is still under attitude control, obtaining re-entry predictions as a function of the attitude, allowing the control with reasonable accuracy of the re-entry point and the impact region and thus avoid populated areas.

Lifetime prediction is a repetitive process; as the design of a satellite evolves, changing physical or orbital parameters require new estimations of the lifetime for mission planners. During flight, as tracking stations provide increasingly accurate information on the state of a satellite, lifetime predictions are updated accordingly. Hence, by the very nature of the problem, computer-efficient methods of prediction have a distinct advantage in terms of time and cost. Moreover, most predictions involve lifetimes of a few months to a few years and consequently, numerical integration of the original equations of motion is not desirable. These factors have contributed to the rapid development of approximate but rapid analytic (closed-form) solutions to the lifetime problem. It is the purpose of this paper to review some of these theories.

After a discussion of orbital perturbations, this review gathers the various means of mathematically modelling these disturbances. Then, a survey of the available literature on analytic predictions of orbit decay and reentry of near-Earth satellites (perigee altitude below about 500 km), assuming an uncontrolled satellite, is given.

2. Orbital perturbations

A number of factors play a dominant role in perturbation modelling for near-Earth satellites. We define the " dispersion parameters " as " any parameter, variable or constant, which affects the dynamic behaviour of a satellite, its trajectory and lifetime ". A list of the major dispersion parameters will be given later (table 3).

In the unperturbed motion, the Earth and the satellite are considered point masses interacting *in vacuo* without the presence of any other disturbing fields. The resulting trajectory is a conic section, lying in a fixed plane with respect to an inertial frame. Since we are concerned with bounded orbits only, elliptic paths

are dealt with. The special case of a circular orbit is included. The classical orbital elements (a, e, i, ω, Ω, θ or M) are often used to completely describe the orbit although there exist any number of alternatives.

In a more realistic situation, the masses in question are finite and non-symmetric and other perturbations are present (atmosphere, Sun, Moon, charged particles, magnetic field, etc.). These cause perturbations in the path which are superimposed on the basic elliptic motion. For small perturbation, the resulting trajectory approximates an ellipse but for large ones, it may be quite different (for instance, a re-entry trajectory).

Alterations to the nominal ellipse are classified here into two groups: (i) environmental perturbations and (ii) specific perturbations.

2.1 *Environmental perturbations*

Environmental perturbations are those caused by the physical bodies exterior to the satellite and surrounding it. Their magnitude usually depends on the spatial position. Table 1 gives the main environmental perturbations affecting Earth satellites.

Solar radiation pressure becomes important for high altitude orbits (Zee 1973) but for near-Earth satellites, drag perturbations are many times larger, except for balloon-type satellites (large area-to-mass ratio) (King-Hele 1964, p. 7; Rowell *et al* 1962). Relative to the drag, solar pressure is therefore neglected for the majority of the satellites. Similarly, the influence of the Earth's magnetic field and that of the ionospheric charged particles is negligible. Meteoroid impacts may have negligible effect on the lifetime of a satellite but larger meteoroid effects have a very low probability of occurrence. Similar to the previous perturbations, meteoroid impacts are not usually considered in the literature on lifetime prediction.

Due to the dissipation of energy through friction, atmospheric drag leads to decreasing values of eccentricity e and semi-major axis a. Clearly, this represents the main agent causing orbital decay. Consequently, most papers on lifetime prediction concentrate on the air drag influence and neglect other environmental perturbations mentioned above.

Apart from these effects on a and e, the perturbation in the trajectory is also seen in other orbital elements. The inclination of the orbit i and its right ascension Ω are affected by the rotation of the atmosphere. The argument of perigee ω is affected by both the atmospheric rotation and flattening. But even in an ideal stationary and spherical atmosphere, ω would still be perturbed.

Table 1. Environmental perturbations

(a)	Solar radiation pressure,
(b)	Earth's magnetic field,
(c)	Presence of charged particles,
(d)	Meteoroid impacts,
(e)	Presence of an atmosphere,
(f)	Luni-solar attractions,
(g)	Non-symmetric mass distribution of the Earth.

As will be shown later, many variables are needed to accurately describe the behaviour of the atmosphere. For example, it has been shown by Ladner & Ragsdale (1964) that the diurnal bulge may alter the lifetime of a satellite by 4% (circular, 200 km-high orbit) and up to 20% (at 500 km). The unexpected early re-entry of *Skylab* was due to unpredicted solar activity which caused the atmospheric density to increase (Smith 1978).

Obviously, luni-solar attraction and the non-symmetric mass distribution of the Earth cannot by themselves lead to orbital decay as these perturbations are conservative. However, they can induce large oscillations in the shape and the orientation of the orbit. Since the drag experienced by a satellite is proportional to the density encountered, which in turn depends on the position of the satellite, these gravitational perturbations, coupled with the atmospheric drag, alter considerably lifetime predictions obtained when drag only is considered. In particular, the perigee height and its position (given by ω) are important for the following reasons. At perigee, the velocity is maximum and the density is, in general, the highest throughout the entire orbit. Since

$$a_D \sim \rho v^2,$$

where a_D = drag deceleration, ρ = density and v = velocity of the satellite, then, the drag perturbation is the highest at the perigee. Consequently, any changes in the perigee radius r_p or in its argument ω are bound to affect the orbit substantially.

The perigee height is a function of a and e. Any oscillations induced in these two elements can produce such a low r_p that the satellite could unexpectedly enter the lower atmosphere. Similarly because of the non-spherical symmetry of the atmosphere (the diurnal bulge, for instance), oscillations in ω could locate the perigee in a denser region of the atmosphere, for a given perigee height r_p. These examples show how oscillatory variations are still important when they are large enough.

The non-symmetric mass distribution of the Earth brings large secular variations in Ω and ω, small periodic perturbations in i, Ω, ω and moderate oscillations in e (King-Hele 1964). Coupling between Earth's oblateness and air drag is discussed by Lubowe (1970), Zee (1971b), Brouwer & Hori (1961), Pimm (1971) and others.

Luni-solar attraction causes small oscillations in all the orbital elements. In particular, highly eccentric orbits are likely to experience large oscillations in r_p (King-Hele 1964). Furthermore, the Moon and to a lesser extent the Sun induce a precessional motion of the perigee. This precession, proportional to a, e and i may be up to 0·18° per year (Vanguard I) and thus is important for a long-life satellite (Blitzer 1959). Luni-solar perturbations are less important than those due to the geopotential harmonics, especially for low-eccentricity, near-Earth orbits where they can be neglected with little loss in accuracy.

For short-term lifetime predictions, it is customary and still accurate to consider drag perturbations only and to neglect the gravitational influences. These influences do not have the time to acquire a dominant importance. For long-term predictions, the analyst must be cautious in interpreting the results, especially if critical situations (*e.g.* perigee located in the diurnal bulge due to Earth's oblateness) are likely to occur. For trajectory predictions, the three dominant environmental perturbations must

be included in the formulation and for low-eccentricity orbits, drag and geopotential perturbations at least are required.

2.2 *Specific perturbations*

Specific perturbations are those which depend on the specific physical properties of the spacecraft in question (shape, dimensions, mass, material, etc.). They alter the effects that the environmental perturbations may have on the satellite attitude motion and through coupling, the orbital motion is influenced. Table 2 gives the most important parameters at the source of these specific perturbations.

The finite dimension of a satellite brings about the problem of the attitude determination of an uncontrolled satellite and the non-negligible torques due to gravity gradients, aerodynamic forces and solar pressure forces. These torques may be amplified by non-symmetric mass distribution and shape. In all cases, variable cross-sectional area and variable drag coefficient (due to the varying angle of attack) lead to non-negligible trajectory dispersion as shown by Hunziker (1970). A non-symmetric shape also generates aerodynamic lift which may induce large deviations from the expected trajectory.

Large area-to-mass ratio satellites may experience solar radiation-induced perturbations of the order of the drag deceleration itself (Rowell *et al* 1962).

The variations in the mass of a spacecraft may be controlled (firing of thrusters, stage separation) or uncontrolled (boiling-off of propellant, rupture of appendage under air drag or meteoroid impact, ablation and breakup upon re-entry). Any loss of mass increases the magnitude of the drag deceleration since:

$$a_D \sim 1/m$$

and consequently lifetime estimates are affected as shown by Ladner & Ragsdale (1964).

In particular, ablation causes the shape and the chemical properties of the surface of the body to change. It is well known that temperature, Mach number and shape variations greatly affect C_D and A (Bailey & Hiatt 1972; Henderson 1976). Errors in prediction are directly related to the uncertainty in C_D and A. Consequently, the re-entry phase of a satellite, during which ablation and breakup are likely to occur due to the severe temperature and pressure conditions involved, represents a period where large trajectory dispersions originate. However, this phase has a very short duration compared to that of the slow decay of the orbit.

Although some discussion on satellite properties variations is included in this survey, most of the papers in the literature consider the satellite as a point mass, with

Table 2. Specific perturbations.

(a) Finite satellite dimensions, non-symmetric shape and mass distribution,
(b) Area-to-mass ratio,
(c) Variable mass,
(d) Ablation and break-up of the satellite.

constant physical and chemical properties but being large enough to experience atmospheric drag. Actually, under reasonable assumptions concerning the attitude motion of the satellite and the range of applicability of the theory (see §§ 4 and 5), the " point mass " approximation can be shown to be physically relevant. In this case, the specific perturbations described here are neglected and some knowledge of the inaccuracy introduced can be obtained. This is convenient in developing a more general theory. Furthermore, considering specific perturbations would require the analyst to include the attitude motion equations in the formulation of the problem. This increase in the complexity of the theory is unnecessary because this degree of sophistication would largely exceed that currently available for atmospheric density models.

To summarize the previous remarks, under realistic assumptions and restricted range of validity for the theory (to be given later), this literature survey will exclude the effects due to: (a) attitude motion, (b) lift generation, (c) variable mass (ablation, breakup), (d) variable shape and dimensions, and (e) large satellite extent (large area-to-mass ratio).

The remainder of this review is organized into three parts:

(i) Survey of the different mathematical models expressing the three largest environmental perturbations, along with the modelling accuracy (§§ 3, 4 and 5).
(ii) Survey of the various analytical methods for predicting orbit decay and lifetime, along with a discussion on the merits of these various closed-form solutions (§ 6).
(iii) A general assessment of the state-of-the-art will close this review (§ 7).

3. Perturbation modelling

As stated in the previous section, we will be concerned with the three dominant environmental perturbations. For a more detailed analysis, these disturbances will be dealt with according to their associated dispersion parameters, given in table 3. The conditions under which the specific perturbations can be neglected will also be included in the sections dealing with the atmospheric drag and the satellite properties models.

3.1 *The geopotential*

It is known that due to the non-symmetric distribution of the Earth's mass, its potential field is not strictly of a $1/r$ form. The potential is written in terms of the Legendre polynomials $P_n^{(m)}$ (Croopnick 1972):

$$\Phi = \frac{\mu}{r}\left[1 + \sum_{n=2}^{\infty} (r_{Ee}/r)^n \sum_{m=0}^{n} (C_{nm} \cos m\lambda + S_{nm} \sin m\lambda) P_n^{(m)} (\sin \delta)\right], \quad (1)$$

where the dependence on longitude λ and geocentric latitude δ is shown. C_{nm} and S_{nm} are the cosine and sine gravitational coefficients. Other forms are sometimes used, for instance (Guttman 1965):

Table 3. Dispersion parameters.

(a)	Relative to the atmospheric drag:
	—altitude dependence of the density (scale height and its variations),
	—atmosphere flattening,
	—diurnal bulge,
	—semi-annual, seasonal and solar activity-induced fluctuations in density,
	—sporadic density variations,
	—atmospheric rotation.
(b)	Relative to the potential field:
	—asymmetry of the Earth's mass,
	—luni-solar attraction.
(c)	Relative to the satellite:
	—variations in the projected area A,
	—uncertainty in the proper value for C_D,
	—generation of lift,
	—gravity-gradient and other torques,
	—solar pressure on large area-to-mass ratio satellites,
	—mass loss, breakup.

$$\Phi = \frac{\mu}{r}\left[1 - \sum_{n=2}^{\infty} J_n \,(r_{Ee}/r)^n \, P_n^{(0)}(\sin\delta) + \sum_{n=2}^{\infty}\sum_{m=1}^{n} J_n^{(m)} \,(r_{Ee}/r)^n \, P_n^{(m)}(\sin\delta)\cos m\,(\lambda - \lambda_n^{(m)})\right], \tag{2}$$

where J_n, $J_n^{(m)}$ and $\lambda_n^{(m)}$ are constants, specific to the Earth's potential.

The evaluation of the zonal (J_n) and tesseral harmonics ($J_n^{(m)}$) has been done by many (Newton *et al* 1961; Kosai 1961b; King-Hele *et al* 1963).

The dependence of the Earth's potential on longitude is usually neglected since the order of magnitude of this influence is smaller than that depending on latitude. Then, the geopotential takes the following form:

$$\Phi(r) = \frac{\mu}{r}\left[1 - \sum_{n=2}^{\infty} J_n \,(r_{Ee}/r)^n \, P_n(\sin\delta)\right]. \tag{3}$$

These harmonics represent the variation of the Earth's mass from spherical symmetry. The first term in the infinite series J_2 is the one that takes most of the oblate mass distribution of the Earth into account; J_3 stands for its " pear-shape ", etc. (King-Hele 1964).

From the values obtained by King-Hele *et al* (1963), it is clear that

$$|J_n/J_2| < 0{\cdot}003, \text{ for } n \geqslant 3, \tag{4}$$

and consequently spherical harmonics of order greater than 2 are usually neglected and the most common geopotential model encountered in this survey considers the largest oblateness coefficient of the Earth only:

$$\Phi(r) = \frac{\mu}{r}\left[1 + \frac{3}{2} J_2 \frac{r_{Ee}^2}{r^2}(\tfrac{1}{3} - \sin^2 \delta)\right]. \tag{5}$$

The effect of Earth's oblateness on the motion of a satellite is a very well understood problem (Sarychev 1962; Zee 1969; Lubowe 1970; Zee 1971a, b; Santora 1976). For a more detailed analysis of geopotential perturbations of orbit, Kosai (1961a), Croopnick (1972) and Guttman (1965) present good analysis. Closed-form solutions for the motion of drag-free satellites under the influence of a non-spherical Earth mass are given by Brouwer (1959), Kosai (1959) and Liu (1974) among others.

3.2 *Luni-solar attraction*

The tidal forces cause oscillations in all orbital elements as well as precession of the orbital plane. In general, the greater the distance of the trajectory from the Earth's centre, the greater are these forces. Thus, the resulting perturbations are significant only for eccentricity larger than approximately 0·5, if near-Earth satellites are considered.

A simple approach to this problem considers the Sun and the Moon as rings of matter around the Earth (Blitzer 1959), having the same total mass. Then, average torques on the orbit may be calculated by the gyroscopic theory and precession rates are obtained. Here, the precession rate must be small compared to the rotation of the disturbing bodies around the Earth.

A more complicated but more accurate technique consists of including in the equations of motion the various attracting forces exerted by the Sun and the Moon (Zee 1972, 1973).

Since most near-Earth satellites are launched on low eccentricity orbits, luni-solar perturbations are usually neglected in lifetime studies.

3.3 *Atmospheric drag*

Actually the drag force is only a part of the total aerodynamic force applied on a body moving in a fluid. In fact, aerodynamic effects can be resolved into (see figure 1):

(i) a force component parallel to the flow (DRAG),
(ii) a force component perpendicular to the flow (LIFT),
(iii) a torque about the centre of mass.

For controlled spacecraft, the aerodynamic torque can be balanced and a stable flight configuration results. However, the available theories on lifetime predictions have their range of validity restricted to uncontrolled satellites only. Under this condition (which will now be tacitly understood throughout this paper), the aerodynamic torque will normally cause the satellite to tumble, as pointed out by King-Hele (1964). The axis of spin is aligned with the axis of maximum moment of inertia of the satellite, since this is the only mode that is stable when energy dissipation

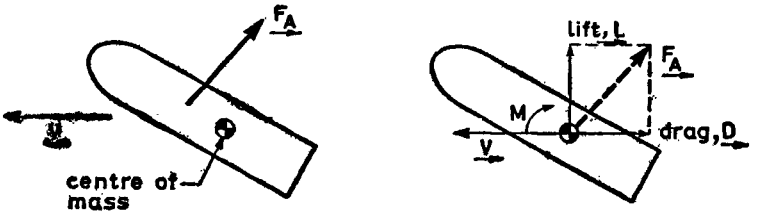

Figure 1. The aerodynamic force on a satellite.

and small external torques (aerodynamic, geomagnetic and gravitational) are present. The spin axis is approximately fixed in space apart from a slow precession induced by the small torques just mentioned. The spin rate slowly decays due to magnetic and aerodynamic dampings. According to King-Hele, under this condition of uncontrolled tumbling, near-spherical and cylindrical satellites (with $l/d > 1$) have almost no resultant lift acting on them. For disk-shaped satellites (with $l/d \ll 1$), the lift becomes more significant but, because the lift-to-drag ratio (L/D) is usually small for satellites (L/D $< 0{\cdot}1$), even for disk-like satellites, the lift is negligible. Thus, lift generation does not need to be considered under the above-mentioned condition.

In the present problem, atmospheric drag is a difficult perturbation to model because it depends on a large number of parameters (see table 3), many of which are known to a poor accuracy. Furthermore, the effects of atmospheric drag are strongly dependent on some parameters related to the specific perturbations (C_D and A). This will be evident in the mathematical expression for the air drag.

Although various mathematical, expressions represent the drag deceleration (Jastrow & Pearse 1957; Billik 1962b), the more widely used is the following:

$$\underset{\rightarrow}{a}_D = -\frac{C_D A}{2m} \rho v_R^2 \underset{\rightarrow}{1}_{v_R} \tag{6}$$

$\underset{\rightarrow}{a}_D$ = drag deceleration vector

$\underset{\rightarrow}{1}_{v_R}$ = unit vector along v_R.

It is simpler to write:

$$\underset{\rightarrow}{a}_D = -\frac{C_D A}{2m} \rho v_R \underset{\rightarrow}{v}_R. \tag{7}$$

The drag coefficient is taken so that the following equation holds true;

$$C_D = \frac{2m|\underset{\rightarrow}{a}_D|}{A \rho v_R^2}. \tag{8}$$

The drag coefficient thus defined varies greatly with flow conditions, angle of attack,

velocity range, type of fluid and mean free path of the molecules. Its determination is not easy especially due to the irregular shape of most satellites.

The cross-section area A of the satellite is directly related to the shape, the dimensions and attitude motion of the satellite. It represents the projected area of the satellite on a plane perpendicular to the fluid flow.

Finally since the atmosphere is not stationary, we have

$$\underset{\rightarrow}{v} \neq \underset{\rightarrow}{v}_R \tag{9}$$

and discrepancies must be considered.

In view of the above remarks, the following dispersion parameters have to be discussed in more detail:

(i) the atmospheric density ρ and its variations in space and time,
(ii) the atmospheric rotation, ω_A,
(iii) the drag coefficient C_D and the cross-section A.

4. Atmospheric models

This section deals with analytic (closed-form) models of atmospheric density and atmospheric rotation. Before getting into the details of these models, it is wise to first point out the main techniques used to obtain these models. These are classified here into empirical and theoretical techniques.

The empirical methods essentially use the techniques of curve fitting. The basic steps are:

(i) Obtain numerical values of the density as a function of the spatial variable (and the time if required) from satellite observations and/or rocket explorations.
(ii) Using a mathematical expression with undetermined coefficients, obtain numerical values for these coefficients to get the best fit to the data.

The main drawback of these methods is the fact that, as the solar heating and geomagnetic activity change within the diurnal, seasonal, semi-annual and sunspots cycles of density variations, the fitting coefficients also vary and tables of these coefficients can become as cumbersome as tabular density models. (The latter models use tables of observed density data with interpolation algorithms.)

Theoretical methods rely on the mathematical formulation in terms of the equations of state. The steps are:

(i) Obtain the equations of state interrelating the various properties of the atmosphere.
(ii) Solve these equations.
(iii) Obtain the density as a function of the other variables.

So far, the best theoretical models cannot describe all the density variations and are, for some range of the independent variables, very far from the actual measured

values. As the accuracy of the basic equations increases, it becomes useless to try to obtain closed-form solutions and numerical analysis is required. This explains why theoretical results are very often combined with the empirical techniques. This leads to the semi-empirical techniques. They usually follow the pattern:

(i) Obtain an expression of the form $\rho = \rho(r, t, a_1, a_2, \ldots, a_n)$ from theoretical techniques where the a_i are coefficients resulting from the derivation.
(ii) Use this mathematical expression to fit data obtained from observations and experimentation. This is the empirical part: the n coefficients a_i are used as parameters of the curve fitting process.

An example is the exponential density model which, in theory, is only valid for a small altitude range. Other examples are the Jacchia static density models (Jacchia 1965, 1970, 1971b, 1977). In these cases, the temperature and mean molecular mass profiles are obtained empirically from observations and the state equations are then numerically integrated. The resulting models are tabular in nature.

The advantage of the semi-empirical models is the fact that they represent observational data accurately and provide simple analytic expressions relating the various parameters of the model.

Sometimes, the expressions like $\rho\,(r, t, a_1, a_2, \ldots, a_n)$ are used to fit standard atmosphere models (CIRA, ARDC, Jacchia, USSA, etc.) which are themselves obtained through numerical solution of theoretical methods as well as observational analysis.

Having discussed the techniques of getting the models, we now differentiate between the two types of model: static and dynamic models.

4.1 *Static density models*

In a static model, the only independent variable is the altitude h and it represents the mean atmosphere averaged over diurnal and seasonal density variations under moderate solar activity.

4.1a *Altitude dependence.* For small changes in altitude, some papers consider the density as being constant (Karrenberg *et al* 1962). In order to improve this approach, approximate results can be obtained by applying the principle of variation of parameters to the constant-ρ solutions (Perkins 1958), so that ρ is allowed to vary with h.

However, due to the gravitational attraction of the Earth, we may expect that the density will vary substantially with altitude so that a deterministic relationship between ρ and h must be considered for greater accuracy.

In the simplest case, we may suppose the atmospheric gas to be in hydrostatic equilibrium under the Earth's gravity field. We further assume that the atmosphere:

(a) is a homogeneous mixture of all its constituents so that:

$$P = \Sigma_i P_i,$$

$$M = \Sigma_i M_i x_i, \tag{10}$$

where P_i = partial pressure of constituent i, M_i = molecular weight of constituent i, x_i = volumetric fraction of constituent i, P = atmospheric pressure, M = molecular weight of the atmospheric gas (28·964 kg/kmole).

(b) each component obeys the perfect gas law.

Then the hydrostatic equilibrium equation is obtained (see Appendix A) as:

$$\frac{d\rho}{\rho} = -\frac{g}{k}\,dh - \frac{dk}{k}, \tag{11}$$

where g = local acceleration due to gravity and $k = P/\rho$. From Appendix A,

$$k = RT/M = nRT/m, \tag{12}$$

n = number of moles in mass m, m = mass of the gas under consideration. It is interesting to note that the proportionality parameter between ρ and P, k, is directly related to the square of the velocity of sound in the atmosphere. In fact, assuming the rarefactions and the compressions in a sound wave to occur isothermally, then the bulk modulus is equal to the pressure (Halliday & Resnick 1962):

$$B_{\text{isothermal}} = P. \tag{13}$$

Since the velocity of sound is given by:

$$v_s = (B/\rho)^{\frac{1}{2}}, \tag{14}$$

then,
$$(v_s^2)_{\text{isothermal}} = P/\rho = k. \tag{15}$$

The parameter k is thus the square of the velocity of sound in the air under isothermal conditions. However, it is recognized that the sound wave involves adiabatic processes rather than isothermal. In this case,

$$B_{\text{adiab.}} = \gamma P, \tag{16}$$

where γ is the ratio of specific heats for the atmospheric gas ($\simeq 1{\cdot}4$). Then,

$$k = \frac{v_s^2}{\gamma} \approx \frac{v_s^2}{1{\cdot}4}. \tag{17}$$

Before proceeding further, it must be noted that the assumption of a homogeneous mixture of the various species is not valid for altitudes above 85 km (according to USSA 1976) to 100 km (Jacchia 1977). Photo-dissociation and diffusive separation make M variable and the hydrostatic equilibrium must undergo some alterations.

For an easy solution to (11) we assume

(c) constant temperature T,

(d) constant molecular weight M,

(e) $g = g_E = \mu/r_E^2$, (r_E = Earth's radius).

Integrating (11) yields:

$$\rho = \rho_0 \exp\left[\frac{-g_E}{k}(h - h_0)\right],$$

where subscript 0 denotes a reference point. Following King-Hele & Cook (1963), we define the density scale-height H as:

$$H = -\rho\,(d\rho/dh)^{-1}, \tag{18}$$

and the solution becomes

$$\rho = \rho_0 \exp\left(\frac{h_0 - h}{H}\right), \tag{19}$$

where $$H = \frac{k}{g_E} = \frac{RT}{Mg_E}. \tag{20}$$

Equation (19) represents the exponential density model. It is widely used because of its mathematical simplicity (Parkyn 1958; Brouwer & Hori 1961; Billik 1962a; King-Hele 1962; Zee 1963; King-Hele & Cook 1963; King-Hele 1964; Cook & King-Hele 1965; Brofman 1967; Lubowe 1970; Zee 1971b; Santora 1975, 1976).

Now (19) can be used in a semi-empirical technique by taking H and ρ_0 as matching coefficients. When ρ_0 is evaluated at h_0 (which is usually the case, if one needs a good fit at h_0), then (ρ_0, h_0) is a fitting point. However, the assumptions of constant temperature and constant molecular weight imply large discrepancies with the actual density profiles and the exponential model with constant matching coefficients is thus valid only for a restricted range of altitude.

To circumvent these difficulties many authors added some refinements to this model:

— the " layer-technique " (or multi-fit),

— scale-height variation.

The layer-technique consists in breaking the atmosphere into n layers with respect to altitude. Within each layer, the matching coefficients are constant so that in the ith layer we have:

$$\rho_{0i},\ H_i,$$

$$\rho_{0i} \neq \rho_{0j},\ H_i \neq H_j \text{ for } i \neq j.$$

As long as the satellite is in the ith layer, the ith set of matching coefficients is used. The altitude range of each layer is, of course, related to the range of validity of the exponential model in that region. Although this technique improves accuracy, it is numerically inefficient, especially for high eccentricity orbits where many layers are traversed within a single revolution or when a nearly circular orbit is on the boundary of two layers and oscillates on each side of it. Of course, the mathematical simplicity of a pure exponential model is lost.

King-Hele (1964), Santora (1975), Willey & Pisacane (1974) and others avoid the layer-technique by assuming variations in the scale-height. In the COSPAR

International Reference Atmosphere (Kallmann-Bijl *et al* 1961) the following approximate value is obtained:

$$dH/dr \approx 0{\cdot}1. \tag{21}$$

From 40 km at 200 km altitude, H increases to about 60 km at $h = 400$ km.

Willey & Pisacane (1974) assume a quadratic form for the scale-height. Santora (1975) considers its variations due to the diurnal bulge. Since H appears in the exponential term, its variation with altitude greatly complicates the analytical expression. Instead, King-Hele (1964) gives the form:

$$\rho = \rho_p \left[1 + b\,(r - r_p)^2\right] \exp\left[\frac{r_p - r}{H_p}\right], \tag{22}$$

where p represents a variable reference altitude (perigee height). The values of b and H_p are considered constant over one orbital revolution but between revolutions H_p takes the form:

$$H_p = H_{p_0} + 0{\cdot}1\,(r_p - r_{p_0}).$$

This form gives a better approximation to the real density variation at the expense of a relatively more complex expression.

As a last comment on the exponential model, table 4 gives the characteristics of this static model.

A more realistic model, based on the hydrostatic equilibrium equation, is obtained when temperature and molecular-weight variations are considered, taking

$$g = \mu/r^2. \tag{23}$$

The following derivation is taken from Escobal (1968). The molecular weight is assumed to vary according to:

$$M = M_0 \frac{T}{T_m}, \tag{24}$$

Table 4. The exponential model (characteristics)

(i) Mathematical simplicity.
(ii) Upon integration of the Lagrange planetary equations, a closed-form solution can be obtained as a series of Bessel functions of imaginary argument, see King-Hele (1964).
(iii) Theoretically valid for small altitude range.
(iv) The scale-height must vary with altitude for wider range of accuracy, since temperature and molecular weight vary.
(v) Used as an empirical tool, the layer-technique or a varying scale-height are usually introduced at the expense of complications in the analytical manipulations and in the computations. The fit, however, is acceptable for wider altitude range.

where T_m is called the molecular-scale temperature (the temperature where $M = M_0$, a reference value) and T_m is a function of the altitude. This dependence is approximated by the layer-technique in which the relation $T_m = T_m(h)$ is linear within each layer:

$$T_m = T_{moj} + S_j (h - h_{oj}) \quad j = 1, 2, \ldots, q, \tag{25}$$

where q = number of layers, T_{moj} = molecular-scale temperature at h_{oj}, h_{oj} = reference altitude for the jth layer, S_j = slope of the approximate relation between T_m and h in the jth layer.

Values of T_{moj} and S_j are usually given in tables as a function of the altitude. Substituting equations (23), (24) and (25) into the hydrostatic equation and integrating yields:

$$\rho = \frac{T_{mo}}{T_m} \rho_0 \lambda^a \exp(-\beta), \tag{26}$$

where $T_{mo} = T_m$ at $\rho = \rho_0$,

$$\lambda = \left[\frac{T_{moj}}{r_E + h_{oj}}\right]\left[\frac{r_E + h}{T_{moj} + S_j (h + h_{oj})}\right],$$

$$a = KS_j/[T_{moj} - S_j (r_E + h_{oj})]^2,$$

$$\beta = K(h - h_{oj}) [T_{moj} - S_j (r_E + h_{oj})]^{-1} [(r_E + h_{oj}) (r_E + h)]^{-1},$$

$$K = \frac{M_0 \mu}{R}.$$

It can be shown that when

$S_j = 0$ (T_m = constant),

$h_{oj} = h_o$; $T_{moj} = T_{mo}$,

$h \ll r_E$ so that $r \approx r_E$ and $r_0 \approx r_E$,

one obtains the exponential model

$$\rho = \rho_0 \exp[-(h - h_0)/H],$$

where $1/H = g_E/k_0$,

and $k_0 = RT_{mo}/M_o$,

and $$M = M_0 \frac{T}{T_m} = M_0 \frac{T}{T_{m0}}.$$

From these last three equations, it is clear that we get

$$H = RT/g_E M$$

similar to (20).

For obvious reasons, this model is theoretically more valid; it gives better fits when used as an empirical tool. However, mathematical simplicity is lost.

As stated above, the assumption of homogeneous mixing is erroneous above approximately 85 km altitude. As the United State Standard Atmosphere (USSA 1976, p.1) puts it, diffusive separation and relative change in the composition bring about the need for more complex theoretical considerations of the vertical component of the flux for the individual species. Although this more complex approach is interesting from a physical and theoretical point of view, it usually fails to provide closed-form solutions for the equations of state. Therefore, it is generally unnecessary in an analytical approach to our orbital problem where closed-form solutions are sought. The discrepancies involved in the simple models so far described are best taken into account by an empirical curve fitting.

A purely empirical expression is given by the power-law model of the form:

$$\rho = \rho_0 (h_0/h)^\eta, \tag{27}$$

where η and ρ_0 are matching coefficients, h_0 is a reference altitude and usually, h_0 is chosen so that $\rho_0 = \rho$ at $h = h_0$. It must be noted that the unusual case where ρ_0 is different from ρ (actual) at h_0 may occur. For instance, when one defines h_0 to be 150 km but needs ρ to be exact at 250 km as well as at 300 km, then the points (ρ_1, 250 km) and (ρ_2, 300 km) are taken as fitting points and consequently, the density at h_0 given by (27) may be slightly off the actual value.

According to Billik (1962b), comparisons between (19) and (27) favour this last model. Similar to the exponential model where H had to be varied for a better fit, η varies with altitude. A table of η versus altitude is given in the above reference.

With η constant, transforming (27) into

$$\rho h^\eta = \rho_0 h_0^\eta = \text{constant},$$

we can see that the power-law model is approximated by the following model, used by Egorov (1971):

$$\rho r^a = \rho_0 r_0^a = \text{constant}, \tag{28}$$

where $$a = \frac{gr}{(dP/d\rho)} \text{ (assumed constant).}$$

As a particular case of the power-law model, the form

$$\rho = k\,(\lambda + \sigma/r), \tag{29}$$

(where σ and λ are matching coefficients) has been used by Newton (1962). While being very handy in mathematical manipulations, (29) is accurate over a small altitude range so that the layer technique should be used for long-lived satellites or eccentric orbits.

Concluding this section on static models, it is interesting to note that any theory giving the lifetime as a function of an integral of the form

$$\int_{h_0}^{h} \rho f(h)dh \quad \text{or} \quad \int_{h_0}^{h} \frac{f(h)}{\rho}\,dh,$$

(where f is any function of the altitude h) is bound to be very useful since no detailed expression for a static density model is needed. These integrals can be tabulated using observational data and consequently modelling inaccuracies introduced by (19), (27) and (29) are avoided (see Perkins 1958).

The authors do not claim to have included all static density models in the literature. Many other models, some of them very accurate, are not discussed here. However, it is felt that the most relevant models which are attractive to any orbital theory are given above.

4.1b *Flattening corrections.* So far, static models have been derived as a function of altitude referred to the Earth's surface. However, the equations of motion for a satellite are given with respect to an inertial frame centred at the Earth's centre. For the static models to be useful in orbital theory, they must be given as functions of the radius vector from the origin of the inertial frame. A change of variable from h to r is thus in order.

Most theorists consider the simplest change of variable:

$$h = r - r_E, \tag{30}$$

in which the Earth's radius, r_E is here assumed constant. Sometimes, r_E is taken as the mean Earth's radius but more often it is taken as the equatorial radius, r_{Ee} (assumed constant here):

$$h = r - r_{Ee}. \tag{31}$$

Using (30) with r_E constant or (31) is equivalent to the assumption of a spherical Earth. Since the relation $\rho = \rho(h)$ of the static model is tacitly assumed to be identical at all latitudes and longitudes (discrepancies are usually included as dynamic effects), it is also equivalent to the assumption of a spherical atmosphere. Equation (31) thus yields atmospheric density models referred to variously as spherical, sym-

metrical or equatorial density models. This terminology is applied whenever the Earth's flattening (and other asymmetries) is not considered. For instance,

$$\rho = \rho_0 \exp\left(\frac{r_0 - r}{H}\right), \quad r_0 = r_{Ee} + h_0,$$

is a spherical exponential model.

In reality, the Earth is not spherical nor is the atmosphere. The change of variable (30) still holds if r_E undergoes the proper variation. The Earth's figure is approximated for our purposes by an oblate spheroid, symmetric about the equatorial plane. The flattening at the poles is

$$f_E = \frac{r_{Ee} - r_{\text{pole}}}{r_{Ee}} \approx \frac{1}{298{\cdot}25}. \tag{32}$$

Still assuming the relation $\rho = \rho(h)$ to be the same at all latitudes, the atmosphere is thus flattened by an amount

$$f_A \cong f_E = f. \tag{33}$$

It follows that on a spherical surface around the Earth, the density is greater at the equator than at the poles. The resulting static models are termed 'oblate' or 'flattened' static density models.

The Earth's radius can be approximated by the relation

$$r_E = r_{Ee}(1 - f \sin^2 \delta), \tag{34}$$

where δ is the geocentric latitude. Substituting (34) into (30) yields the proper expression for h to be inserted into spherical models. A constant-altitude surface has a variable radius, denoted σ (King-Hele 1964), which is now approximately given by:

$$\sigma = \sigma_e (1 - f \sin^2 \delta), \tag{35}$$

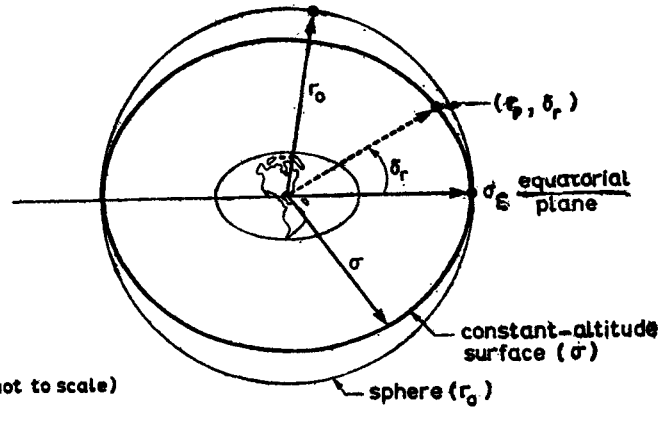

Figure 2. Atmosphere flattening

where σ_e = radius of the reference spheroid at the equatorial plane, and δ = geocentric latitude of the point in space considered. The error in (35) is of the order of $f(\sigma_e - r_{Ee})$. In particular, (35) reduces to (34) at the Earth's surface. In general, the spheroid of reference may be chosen tc pass through any point (r_r, δ_r) so that: (see figure 2)

$$r_r = \sigma_{er}(1 - f\sin^2\delta_r), \tag{36}$$

where subscript r in σ_{er} is a reminder that the spheroid contains the point (r_r, δ_r). Consequently, from (36),

$$\sigma_{er} = r_r/(1 - f\sin^2\delta_r).$$

A constant-altitude surface passing through the point (r_r, δ_r) is seen to obey:

$$\sigma_r = \frac{r_r}{(1 - f\sin^2\delta_r)}(1 - f\sin^2\delta). \tag{37}$$

The oblate exponential density model is thus given by:

$$\rho = \rho_r \exp\left[-(r - \sigma_r)/H\right]$$

where $\rho_r = \rho(\sigma_r) = \rho(r_r, \delta_r)$.

Another means of getting around the flattening correction problem is to introduce a factor which takes into account the difference between the average effective density profile in a spherical and an oblate atmosphere. This method is applicable to circular orbits (Lee 1962). The factor is defined as:

$$F_{Af} = \frac{1}{2\pi}\int_0^{2\pi}\frac{\rho_f}{\rho_s}\,d\theta, \tag{38}$$

where ρ_f and ρ_s refer to flattened and spherical density models. Integrating (38) with an exponential density model yields:

$$F_{Af} = 1 - K_f/2 + 3K_f^2/16 - 15K^3/288 \pm \ldots$$

where $$K_f = \frac{r_E f\sin^2 i}{H}.$$

This expression for F_{Af} used with lifetime predictions based on a spherical model should give more realistic results.

The constant-density surface introduced by King-Hele seems the most adequate procedure for flattening correction. It has also been used by Santora (1975, 1976).

4.2 *Dynamic density models*

In these models, the independent variables are the space variables and the time. They are an extension to the static models and attempt to predict, at any time, the

structure of the atmosphere. These models, also called time-varying models, are accurate up to a factor of about 2, according to Escobal (1968).

Dynamic density variations are mainly induced by the heating of the atmosphere due to the Sun's electromagnetic radiation. It is generally recognized that the ultraviolet light (UV), absorbed in the upper atmosphere, is responsible for these temperature variations. Obviously, this energy flux is hardly measurable from the Earth's surface. However the energy of the radio waves (at wavelengths 10 to 20 cm) is closely correlated to that of the UV rays. Since the former are not absorbed by the atmosphere, their recording provides a useful index for the amount of the Sun's energy incident on the atmosphere. The 10·7 cm radio wavelength is commonly used and the index, denoted by $S_{10\cdot7}$ (or $F_{10\cdot7}$) is called the "10·7 cm solar flux index". It usually varies between 70 and 220 (in units of 10^{-22} watt/m^2-Hz) depending on the activity of the Sun (Jacchia 1977). An increase in $S_{10\cdot7}$ involves an increase in air temperature at a given altitude.

Two sources for this UV radiation are defined (Jacchia 1977): (a) the active regions of the solar disk, and (b) the solar disk as a whole.

The radiation emitted by the active regions varies greatly in occurrence and intensity from day to day. The axial rotation of the Sun (having a period of about 27 days relative to the Earth) causes these active regions to reappear periodically and, despite the very random nature of this phenomenon, periodic variations in density are discernible, with a 27-day period (King-Hele 1959; King-Hele & Walker 1960a; King-Hele 1964; Cook 1970). The intensity of active sources is related to the number of sunspots. However, these intensity variations are basically random in nature. Plots of the solar flux index $S_{10\cdot7}$ versus time show large random fluctuations (see for example Cook 1970, 1971; King-Hele & Walker 1971b and Walker 1972b).

The measurement of the contribution of the entire solar disk is simply given by the average value of $S_{10\cdot7}$, $\bar{S}_{10\cdot7}$, over a few solar rotations (Jacchia 1977). This is the component of the UV radiations that slowly varies with the 11-year sunspot cycle.

In addition to the UV radiations, the solar wind also interacts with the Earth's magnetic field, producing an increase in temperature dependent on the magnetic latitude. Variations in temperature are directly related to density fluctuations. Since the interaction perturbs the Earth's magnetosphere, a readily available index of this flux of energy incident on the atmosphere is the geomagnetic planetary amplitude A_p (or sometimes its quasi-logarithmic value, the geomagnetic planetary index, K_p) which provides magnetic activity measures around the world (King-Hele 1964). In periods of quiet geomagnetic conditions, $A_p = 0$. During strong geomagnetic storms, A_p may reach 120 or 130 (King-Hele & Walker 1971a). All the above-mentioned references have graphs showing correlations between A_p and the density ρ.

One can expect that the influence of these two sources of heat will be largely modulated by the position of the Earth relative to the Sun. Furthermore, these effects vary with altitude. This results in space-dependent and time-dependent asymmetries in density. These are negligible below about 180 km altitude, except for the semi-annual variations which extend down to 150 km or even lower.

The dynamic variations, which are added to the static models of § 4.1, can be

grouped into five classes: (a) the day-to-night variations (diurnal bulge), (b) the semi-annual variations, (c) the sunspot cycle variations, (d) the unpredicted day-to-day variations, (e) the latitudinal-seasonal variations. Each of these is now analyzed separately.

4.2a *The diurnal bulge.* Constantly heated by the Sun, the atmosphere on the sunlit side of the Earth expands in a big hump. This bulge may be as high as 100 km at 500 km altitude (King-Hele 1964) which represents a ratio

$$f_B = \frac{\rho_{\max}}{\rho_{\min}} = \frac{\rho_{\text{day}}}{\rho_{\text{night}}} \tag{39}$$

of about 5 ($f_B = 8$ at 600 km for moderate $S_{10.7}$). This can affect estimates of lifetime by a factor of up to 3. This factor decreases to about 1·5 for perigee height near 200–250 km (King-Hele 1978). This effect is usually negligible below 150–200 km.

Thus, on a constant altitude surface, the density increases as one moves towards the Sun. However, due to atmospheric rotation, the bulge centre is not exactly situated at the sub-solar point. It lags by an angle of about 25 to 30° (Ladner & Ragsdale 1964). Thus, the maximum air density at a given altitude occurs around 14.00 hr, local time. The declination of the bulge centre is, in many papers, assumed to be that of the Sun.

One of the first analytic expressions of the diurnal bulge is due to Jacchia (1960). He used a semi-empirical equation giving the variation of the density with altitude and bulge angle:

$$\rho = \rho_s\,(h)\,[1 + 0{\cdot}19\,(e^{0{\cdot}0055h} - 1{\cdot}9)\cos^6\phi/2], \tag{40}$$

where $\rho_s\,(h)$ represents the night-time static density model and ϕ is the bulge angle, the angular distance from the bulge centre to the point where density is evaluated. Altitudes (h) are expressed here in kilometers. Equation (40) assumes axial symmetry about the Earth-bulge centreline. Comparisons of (40) with actual models (Ladner & Ragsdale 1964) show that the agreement is very acceptable for altitudes in the range

$$117 \text{ km} < h < 700 \text{ km}.$$

For $h \simeq 117$ km, the factor

$$(e^{0{\cdot}0055} - 1{\cdot}9),$$

vanishes and below this value for h, the bulge effect is absent, as noted before. Another feature of this equation is the fact that the bulge factor is an increasing function of altitude (the bulge factor is the coefficient of the periodic function in ϕ).

A simple expression, devised by Cook & King-Hele (1965), assumes a sinusoidal variation of the density with angular displacement from the bulge centre:

$$\rho = \rho_s\,(h)\,(1 + F\cos\phi), \tag{41}$$

where $\rho_s(h)$ is the average density and F is constant (matching coefficient) in the first approximation. In order to obtain an exact fit at $\phi = 0°$ (where $\rho = \rho_{max}$) and at $\phi = 180°$ ($\rho = \rho_{min}$), F was taken as:

$$F = \frac{f_B - 1}{f_B + 1} = \text{bulge factor.} \tag{42}$$

From (39), we get:

$$F = \frac{\rho_{max} - \rho_{min}}{\rho_{max} + \rho_{min}}.$$

The variation of this bulge factor, $(f_B - 1)/(f_B + 1)$, has been obtained by Harris & Priester (1962) for $S_{10.7}$ equal to 70, 160 and 250. For moderate solar activity (150), F varies from 0·05 at 200 km to 0·8 at 600 km. Another assessment of the importance of the diurnal bulge as a function of altitude is provided by King-Hele (1978). Thus assuming F constant in (41) is theoretically incorrect. Therefore, while both (41) and (40) assume a symmetrical bulge, King-Hele and Cook considered a constant bulge factor for simplicity in their derivation. Furthermore, it has been shown that such a symmetry about the bulge-Earth centreline does not actually exist. Newton (1970) showed that a strong positive latitudinal density gradient exists at latitudes above 40°. He supposed this was caused by atmospheric heating induced by the higher geomagnetic activity present at these northern latitudes. Jacchia & Slowey (1967) show that, at least above 500 km altitude, the bulge is not symmetric about the centreline. In fact, from satellite observations, the bulge seems to be elongated in a north-south direction and its centre does not move much from the equator.

Despite the relative inaccuracy of (41), its use in the orbital theory is attractive because of its mathematical simplicity. That is why (41) is widely used (Cook & King-Hele 1965; King-Hele & Walker 1970; Santora 1975 and 1976).

More details on the day-to-night effects are given by King-Hele & Quinn (1966), King-Hele & Walker (1960b), Walker (1972a, 1978), King-Hele (1959), Martin & Priester (1960), Groves (1958) and Jacchia (1977).

As a final note, Santora (1975) describes the effect of the diurnal bulge on the scale-height, H, in the form;

$$H = H_0 + S\cos\phi; \quad H_0 = \frac{H_{max} + H_{min}}{2} = H_{av}; \; S = \frac{H_{max} - H_{min}}{H_{max} + H_{min}}. \tag{43}$$

Other small variations in density (Jacchia 1977) and the generation of small winds (Bedinger 1970) are associated with the day-to-night effects but they are usually neglected.

4.2b *Semi-annual variations.* When the day-to-day random and diurnal fluctuations in the density are ignored, the remaining trend of the density variation is semi-

annual (King-Hele & Walker 1971a). The amplitude of this effect is smaller than that of the 11-year solar cycle or the diurnal bulge but it is more important for satellite lifetimes (King-Hele 1978) because it extends as low as 150 km (probably even as low as 90 km (King-Hele & Walker 1971a)) with no substantial loss in magnitude and its period is of the order of the lifetime predictions most commonly made. Despite the fact that its intensity and phase vary from year to year (which renders its modelling quite inaccurate), peak densities are usually observed in March-April and October-November, while minimum densities occur in January and July. The July minima are deeper than those in January. The following ratios, for $h = 185$ km, were obtained in 1968-1969 by King-Hele & Walker (1971a):

$$f_{sa} = 1{\cdot}43 \text{ (April/July)}; f_{sa} = 1{\cdot}32 \text{ (October/July)};$$

$$f_{sa} = 1{\cdot}22 \text{ (April/January)}; f_{sa} = 1{\cdot}22 \text{ (October/January)};$$

where the 'semi-annual' ratio of density f_{sa} is:

$$f_{sa} = \rho_{max}/\rho_{min}.$$

The difference between the mean density and the largest variation varies by a factor of 1·6 at 200 km to 3 at 500 km.

Early research projects tried to relate the semi-annual variation to the temperature differences induced by the solar activity. Two superimposed effects were thus identified (King-Hele 1964). First, the variable distance from the Earth to the Sun (maximum in July) could cause lower densities in July. Second, the inclination of the Earth's orbit (7°) to the plane of the Sun's equator could induce increasing density every 6 months, when the Earth crosses this plane. However, Cook (1970) and Jacchia (1971a) later rejected this dependence.

According to Jacchia (1971), the most appealing explanation is the following. Due to the change in the inclination of the Earth's magnetic pole with respect to the direction of the solar wind during the year, fluctuations in the shape of the magnetosphere are induced. Semi-annual variations in the geomagnetic index have already been recorded which confirms this supposition. Variable interactions between the magnetic field and the charged particles involve variable heat generation in the atmosphere leading to fluctuations in density. This seems to contradict the assumption of independence of the semi-annual variation on solar activity (which is related to the solar wind). Actually, it is probable that the influence of the solar cycle upon changes in magnetosphere is weak. More details can be obtained from Cook (1967, 1970 and 1971) and Walker (1972a).

Jacchia (1977) gives empirical expressions for the semi-annual density variations in terms of

$$\Delta_{sa} \log \rho = \log (\rho/\rho_{av}),$$

where ρ_{av} is the average density over the semi-annual cycle. He gives:

$$\Delta_{sa} \log \rho = f(h)\, g(t), \tag{44}$$

where $$f(h) = \left[0{\cdot}04\left(\frac{h}{100}\right)^2 + 0{\cdot}05\right] \exp\left(-0{\cdot}25\frac{h}{100}\right) (h \text{ in km}), \tag{45}$$

$$g(t) = 0{\cdot}0284 + 0{\cdot}382\,[1 + 0{\cdot}467 \sin(2\pi\tau + 4{\cdot}14)] \sin(4\pi\tau + 4{\cdot}26). \tag{46}$$

In the time-dependent equation, τ is a periodic function of the fraction Φ of the tropical year T corresponding to the time t:

$$\tau = \Phi + 0{\cdot}0954\left\{[\tfrac{1}{2} + \tfrac{1}{2}\sin(2\pi\,\Phi + 6{\cdot}04)]^{1{\cdot}65} - \tfrac{1}{2}\right\}, \tag{47}$$

where $$\Phi = \frac{t - t_0}{T}\ (t_0 = \text{January } 1{\cdot}0). \tag{48}$$

While $f(h)$ varies from 0·070 at 100 km to 0·332 at 1000 km, $g(t)$ takes the following minima and maxima:

minimum in January: $g(t) = -0{\cdot}188$,
maximum in April: $g(t) = +0{\cdot}361$,
minimum in July: $g(t) = -0{\cdot}522$,
maximum in October: $g(t) = +0{\cdot}478$.

From these data, it is seen that the October maximum is greater than that in April.

These equations are obtained from 12 years of satellite-drag data covering a wide range of altitudes. They are part of the 1971 atmosphere model (Jacchia 1971b). Sometimes, the semi-annual effect is decomposed into an annual and a semi-annual term (Volland *et al* 1972). The annual term is seen to be nearly independent of altitude. Jacchia (1977) gives:

$$\Delta_{sa} \log \rho = f_1(h)\, g_1(t) + f_2(h)\, g_2(t), \tag{49}$$

where subscripts 1 and 2 refer to the annual and semi-annual components respectively and

$$f_1(h) = 0{\cdot}03 \tanh\left(0{\cdot}6\frac{h}{100}\right), \tag{50a}$$

$$f_2(h) = \left[0{\cdot}017\left(\frac{h}{100}\right)^2 + 0{\cdot}015\right] \exp\left(-0{\cdot}25\frac{h}{100}\right), \tag{50b}$$

and $$g_1(t) = \cos\,[2\pi(\Phi - 0{\cdot}047)], \tag{51a}$$

$$g_2(t) = \cos\,[4\pi(\Phi - 0{\cdot}296)]. \tag{51b}$$

A comparison between (49) and (44) is shown in Jacchia (1977).

4.2c *Sunspot cycle variations.* We have previously stressed the importance of the solar flux, $S_{10.7}$, as well as its average value, $\bar{S}_{10.7}$. The average flux seems to be directly proportional to the number of sunspots which varies in a roughly predictable

cycle of 10 to 11 years duration. The intensity or the date of a maximum cannot be predicted accurately and scientists must rely on past cycles (Smith 1978) to have an approximate measure of the future cycles. The intensity of the solar wind also has some relation to the sunspot number (Jacchia 1971a).

Near 600 km, sunspot variations may cause the density to vary by a factor of up to 20, decreasing to between 2 and 5 near 200–300 km altitude. This is the dominant dynamic effect. The error in lifetime predictions (over several solar cycles) varies between factors of 0·2 and 2·0 at perigee height of 300 km and decrease to 0·5, 1·5 for 200 km perigee altitude (King-Hele 1978). Ladner & Ragsdale (1964) obtained variations in the predicted lifetime of $-$ 60% and $+$ 150% from the nominal value.

Short-term predictions are usually not much affected. For long-term predictions, average values of density over many cycles are commonly used. In all cases, it must be kept in mind that extrapolation of the solar flux obtained from past cycles gives only crude results and sometimes rapid and unexpected orbit contraction may occur, *e.g.*, *Skylab* (Smith 1978).

4.2d *Unpredicted day-to-day variations.* Superimposed on all the previous density variations are random density fluctuations induced by ephemeral solar disturbances (King-Hele 1964) and solar flares (Nonweiler 1958). King-Hele & Walker (1971a) show that a very close correlation exists between the density and the geomagnetic index, which reflects the response of the atmosphere to the solar wind-geomagnetic field interaction. Two strong geomagnetic storms (1 November 1968: $A_p = 122$; 15 May 1969: $A_p = 131$) caused increase in density of 30% and 70% respectively. The correlation between the density and the solar flux index $S_{10{\cdot}7}$ is less obvious at low altitudes. However, above 200 km, the random fluctuations in $S_{10{\cdot}7}$ can strongly influence the density. Below this altitude, an increase in $S_{10{\cdot}7}$ from 150 to 200 caused an increase in density smaller than 10%.

More recently, King-Hele (1978) noted that near 150 km, the random variations in ρ are of the order of $\pm$ 10%, increasing to a factor of 2 and 6 at 200 km and 600 km respectively.

Despite the random nature of these fluctuations, the 27-day recurrence of maxima and minima, as previously noted, can be observed. Predicted lifetimes of about 4 days were shown to be too large by a factor of 2 if a major magnetic storm were to occur (King-Hele 1978).

These sporadic variations are the least predictable fluctuations in air density. Unfortunately, the trajectory dispersions they provoke cannot always be neglected. Obviously, a deterministic approach to this problem is in vain and one must then rely on a stochastic approach.

4.2e *Latitudinal-seasonal variations.* Density profiles are not symmetric with respect to the equatorial plane. Small variations are present as shown by King-Hele & Walker (1960a, 1960b), Martin & Priester (1960) and Newton (1970).

From observational data, Martin & Priester (1960) identified two super-imposed effects: (i) a seasonal effect, related to the declination of the Sun, (ii) a terrestrial effect. The first of these two produces maximum density at the latitude corresponding to the Sun's declination. The terrestrial effect involves a higher density above the equator.

The interaction of the solar wind with the magnetosphere (King-Hele & Walker

1971a) produces higher density depending on the magnetic latitude. This phenomenon may explain the terrestrial effect obtained by Martin and Priester. Newton (1970) considers a probable correlation between the thermospheric heating and the auroral oval. He observed strong positive latitudinal density gradient at latitudes above 40° which, he supposed, may be influenced by the geomagnetic field.

These fluctuations in density are usually small and are generally neglected. To the knowledge of the authors, no lifetime prediction theories have considered them.

4.3 *Rotation of the atmosphere*

The upper atmosphere is not stationary but rotates at a slightly higher angular rate than that of the Earth in a west-to-east direction (King-Hele & Scott 1966, 1967; King-Hele *et al* 1969, 1970; Bedinger 1970; King-Hele 1971; Ching 1971; Schuchardt & Blum 1976). This rotation is usually expressed in terms of the ratio Λ:

$$\Lambda = \omega_A/\omega_E, \tag{52}$$

where ω_A and ω_E are the angular velocity of the atmosphere and the Earth respectively.

This rotational ratio is determined from the secular variations it induces in the inclination of the orbital plane of a satellite (i). King-Hele (1964) derived an expression relating Δi to Λ. The early measurements of Λ (King-Hele & Scott 1966) gave:

$$\Lambda = 1{\cdot}27 \pm 0{\cdot}18 \text{ (rms)},$$

for the altitude range (200–300 km). In this paper, no clear dependence of Λ on latitude or time was noticed. Later, King-Hele *et al* (1970) compiled the data obtained from 32 satellites for the range (200–400 km). A quadratic fit to these data revealed the following altitude variation for Λ:

altitude (km):	200	300	400,
Λ:	1·10	1·35	1·45.

There was no evidence that Λ varies from year to year with solar activity. A table of values of Λ is given in King-Hele *et al* 1970 for measurements made between 1964 and 1970. In a subsequent paper, King-Hele (1971) confirmed the previous values of Λ for low altitude range and obtained a linear decrease for Λ from 350 to 500 km altitude. The results were:

Low altitude: linear increase from $\Lambda = 1{\cdot}1$ ($h = 200$ km)
to $\Lambda = 1{\cdot}4$ ($h = 350$ km)

High altitude: linear decrease from $\Lambda = 1{\cdot}0$ ($h = 420$ km)
to $\Lambda = 0{\cdot}7$ ($h = 500$ km).

Only 4 data points were available for the high altitude range and it is suspected that large uncertainties in Λ are present in this interval. It was also shown that Λ undergoes wide variations over short-time intervals for the range 120 km $< h <$ 230 km.

A recent evaluation of Λ for altitude above 375 km did not confirm the linear decrease obtained by King-Hele (1971). In fact, Schuchardt & Blum (1976) showed a co-rotation ($\Lambda = 1{\cdot}0$) of the exosphere with the Earth in the altitude range (500–600 km).

In general, the measurements of Λ, perturbed by luni-solar oscillations and geopotential harmonics (J_3) are all subject to some degree of uncertainty. While it is now generally recognized that Λ increases with altitude in the 200–400 km region, its behaviour at higher altitudes is not very well known.

The effects of the atmospheric rotation on the orbital inclination i are a source of possible trajectory dispersion. It is thus important to include the proper relative velocity of the satellite v_R with respect to the ambient air in the drag equation. We have:

$$\underset{\rightarrow R}{v} = \underset{\rightarrow}{v} - \underset{\rightarrow A}{v},$$

where $\underset{\rightarrow R}{v}$ is the velocity of the satellite relative to the ambient air. Here, $\underset{\rightarrow}{v}$ and $\underset{\rightarrow A}{v}$ are inertial velocities of the satellite and the atmosphere respectively. The most expeditious and elegant method of considering this problem has been worked out by King-Hele (1964). Assuming the path of the satellite to be horizontal where the atmospheric rotation has its main influence (*i.e.*, around perigee), he shows that the inclusion of a factor F in the drag equation approximates the effect of this rotation:

$$\left|\underset{\rightarrow D}{a}\right| \approx \frac{C_D A}{2m} \rho F v^2,$$

where
$$F = \left(1 - \frac{r_{p0}}{v_{p0}} \omega_A \cos i_0\right)^2, \tag{53}$$

and obviously,

$$v_R^2 \approx F v^2.$$

F is usually in the interval (0·9, 1·1).

A more exact approach would include the actual velocity of the spacecraft relative to the ambient air. This case becomes complicated since $\underset{\rightarrow R}{v}$ depends on the orientation of $\underset{\rightarrow}{v}$ with respect to $\underset{\rightarrow A}{v}$. This orientation is, in general, continually changing.

In addition to the west-to-east rotation of the atmosphere, sporadic winds in other directions exist (*e.g.*, Bedinger 1970) but these are usually neglected due to their small magnitude compared to the main easterly movement.

5. Models of satellite properties

As noted previously, the two most important satellite parameters are the projected area A and the drag coefficient C_D. These will be considered in succession.

5.1 *Projected area*

Because of its attitude motion, a non-spherical satellite presents a variable effective area to the flow. In order to consider this dispersion parameter (associated with the specific perturbations), some considerations and assumptions, based on (King-Hele 1964), are briefly summarized here with some additions.

We have pointed out that uncontrolled satellites have a tendency to spin about their axis of maximum moment of inertia which is approximately fixed in space, apart from a small precession due to perturbing torques. Since these assumptions usually hold true in our problem, it is possible to determine the effective cross-section area of a satellite or, at least, its average value during one orbital revolution. The method used, which will be referred to as the averaging technique for the projected area, is simple. Two extreme cases are usually considered: (a) spin axis parallel to $\underset{\rightarrow R}{v}$, (b) spin axis perpendicualr to $\underset{\rightarrow R}{v}$. In each of these two cases, the average cross-section area is computed over one rotation of the satellite about its spin axis. Cases (a) and (b) thus yield two average values of A, $A_{\parallel}$ and $A_{\perp}$ respectively. Since the actual orientation of the spin axis is usually unknown, all orientations are assumed to be equally likely. Therefore, a second average, of $A_{\parallel}$ and $A_{\perp}$, represents the mean value taken over an orbital revolution of the satellite:

$$A_{\mathrm{av}} = (A_{\parallel} + A_{\perp})/2. \tag{54}$$

Equation (54) is very well suited to satellites with their spin axis in the orbital plane. In this case, since the direction of $\underset{\rightarrow R}{v}$ changes at every perigee passage while the spin direction is approximately fixed, the two extreme values, $A_{\parallel}$ and $A_{\perp}$, have an equal probability. This averaging method is now applied to cylindrical near-spherical and spherical satellites, which represent the most common type of satellites launched so far.

5.1*a* *Rod-shaped satellites* ($l/d > 2$). This class contains almost 50% of the satellites. Their spin axis is the transverse axis. In case (a), the satellite is tumbling in the flow like an aircraft propeller while in case (b), it is tumbling end over end. The two corresponding averaged areas are (King-Hele 1964):

$$A_{\parallel} = ld; \quad A_{\perp} = \frac{2}{\pi}\left(ld + \pi\frac{d^2}{4}\right)$$

and consequently,

$$A_{\mathrm{av}} = ld\left(\frac{1}{2} + \frac{1}{\pi}\right) + \frac{d^2}{4}. \tag{55}$$

The maximum error in A_{av} is between 6% ($l/d = 2$) and 15% ($l/d = 8$). This average method is thus very adequate for cylindrical satellites with $l/d > 2$. In the particular case where the spin axis lies in the plane of the orbit, the argument given

earlier implies that A_{av} cannot be equal to $A_{\parallel}$ nor $A_{\perp}$ and the above errors are thus smaller in this case. Note that case (b) here is equivalent to the "preferred tumbling" method (Benson *et al* 1968).

5.1*b* *Disk-shaped satellites* ($l/d < 1/2$). Only a few satellites are disk-shaped. In this case, the spin is about the axis of symmetry. Tumbling (a) ($A_{\parallel}$) and tumbling (b) ($A_{\perp}$) yield the following averaged values:

$$A_{\parallel} = \pi d^2/4; \quad A_{\perp} = ld,$$

so that
$$A_{av} = \frac{\pi d^2}{8} + \frac{ld}{2}. \tag{56}$$

The theory here is less accurate. Indeed, when $l \to 0$, $A_{\perp} \to 0$ and $A_{\perp}$ becomes very different from $A_{\parallel}$. Thus the averaged value A_{av} can differ from the actual value by up to 22% (King-Hele 1964) for $l/d = 0{\cdot}5$.

5.1*c* *Near-spherical* ($1/2 < l/d < 2$) *and spherical satellites.* For these types of satellites, no axis has a moment of inertia substantially greater than any other axis. It can thus be assumed that the spin will take place about an arbitrary axis of the body. Consequently, since nearly all orientations of the satellite are equally probable (unknown), a relatively accurate value for A_{av} is (Ladner & Ragsdale 1964):

$$A_{av} = \tfrac{1}{4} \text{ (total surface area)}. \tag{57}$$

This is the "equivalent sphere" method (Benson *et al* 1968). The error may be of the same order as for cylindrical satellites and tends to zero as the shape becomes more and more spherical.

The above analysis provides the assumptions and arguments which allow the analyst to avoid a detailed formulation of this specific perturbation, namely, the variation of A with the attitude motion and the particular configuration of the satellite. However it must be kept in mind that trajectory predictions and lifetime estimations obtained with $A = A_{av}$ given above are subject to some degree of uncertainty, the magnitude of which depends upon the specific satellite. The value of A would certainly be more accurate were the spin axis orientation given, but this information is usually not available.

5.2 *Drag coefficient*

The drag coefficient is probably the most poorly known dispersion parameter. Due to the fact that C_D depends on so many variables, experimental evaluations must be resorted to. The actual surface-interaction phenomena, the molecular speed ratio and the fraction of diffuse and specular reflections are among many unknown factors to be defined more accurately. The evaluation of a proper value for C_D cannot be left aside in orbital theory since variations in this parameter lead to substantial

dispersions in the trajectory of a satellite (Lubowe 1970; Hunziker 1970). Comparisons of solutions with variable and constant C_D are given in Hunziker (1970).

In general, C_D is affected by the following:

(i) the physical and dynamic properties of the surrounding medium, namely its flow conditions, density, mean free path λ, composition and temperature. The first three are very closely related,
(ii) the nature of the interaction between the surface of the body and the medium (specular or diffuse reflections, absorption, etc.),
(iii) the satellite's attitude motion (angle of attack α_a), its form and dimensions, the properties of its surface (temperature, material, etc.).

We will first consider the flow conditions existing around the body. At satellite altitudes, one might expect free molecular flow conditions but some quantitative assessment is required. Free molecular flow is present when the mean free path of the ambient air molecules λ_0 is much larger than a typical linear dimension l of the satellite. The Knudsen number, K is defined as:

$$K = \lambda_0/l, \tag{58}$$

and the condition is thus $K \gg 1$. For satellite velocities, this condition needs some refinement. In fact, near the surface of the body, the density ρ increases due to the re-emitted particles. Since the mean free path is inversely proportional to ρ, the effective mean free path is shorter (Cook 1965). The actual mean free path λ is thus:

$$\lambda = \lambda_0 \frac{v_r}{v_i}, \tag{59}$$

where v_r = velocity of re-emitted molecules, and v_i = velocity of incident molecules.

The condition becomes:

$$\lambda_0/l \gg v_i/v_r,$$

or

$$\frac{v_r}{v_i} K \gg 1. \tag{60}$$

The smallest possible value for v_r/v_i has been shown to be approximately 1/13 (Cook 1965), assuming the molecules to be re-emitted at the body temperature.

The value of λ_0 is inversely proportional to the density as given by:

$$\lambda_0 = \frac{0{\cdot}707\, m^*}{\pi\rho\sigma^2} = \frac{0{\cdot}707}{\pi N\sigma^2} \tag{61}$$

where m^* = mass of one molecule, ρ = density, σ = diameter of a molecule, N = number of molecules per unit volume, $= \rho/m^*$. The values of λ_0 given in table 5 are taken from USSA (1976):

Table 5. Mean free path as a function of altitude

Altitude (km)	λ_0 (m)	Altitude (km)	λ_0 (km)
100	0·142	300	2·6
120	3·31	400	16
140	18	500	77
160	53	600	280
180	120	700	730
200	240	800	1400

From this table, it can be deduced that above 200 km altitude, all existing spacecraft are in free-molecular flow. Even at 180 km altitude, only the large rocket stages are likely to have entered the transition regime (Cook 1965). In the transition region, the collisions between incident and re-emitted particles lead to a reduction in C_D. For a satellite with its greatest dimension approximately one meter, free molecule flow conditions are always existent until re-entry occurs.

The analysis of C_D in free-molecular flow requires knowledge of the actual surface interaction between the body and air molecules. The mechanism of this interaction is the gray area in the analysis. The two classical extremes that are usually specified are:

(a) Specular (mirror-like) reflection, where only the normal linear momentum is altered (it is reversed) and an impulse normal to the surface of the body is transmitted. The tangential linear momentum component is conserved.

(b) Diffusive reflection, in which the incident particles are absorbed by the surface, imparting all their momentum to the body. After a period of time known as the relaxation time (very poorly known), the particles are re-emitted in a direction completely independent of the incident direction it had. This direction of re-emission is given by the Knudsen cosine law: the number of molecules re-emitted between θ and $\theta + d\theta$ is proportional to $\cos\theta\, d\theta$ (θ being the reflection angle). The speed of the re-emitted particle is proportional to its temperature (or energy) which is the same as that of the surface (T_w) it is leaving. We then speak of the "accommodation" of the incident particle to the surface of the body. All velocity distributions are usually assumed to be Maxwellian.

Although the exact nature of the interaction is difficult to model, one usually recognizes that a mixture of both situations, (a) and (b), would occur. An empirical approach to this is to consider a fraction f of the molecules to be diffusely reflected and a fraction $(1-f)$ to be re-emitted specularly. Then,

$$C_D = fC_{D_1} + (1-f)C_{D_2},$$

C_{D_1} : drag coefficient for diffuse reflection,

C_{D_2} : drag coefficient for specular reflection.

Early experiments suggested that f was close to unity but these experiments were not for high speed flows.

For a better quantitative consideration of this " partial accommodation " the following parameters prove to be useful (Ladner & Ragsdale 1964):

— the thermal or energy accommodation coefficient α
— the normal momentum accommodation coefficient σ'
— the tangential momentum accommodation coefficient σ

$$\alpha = \frac{E_i - E_r}{E_i - E_w}\ ;\ \sigma' = \frac{P_i - P_r}{P_i - P_w}\ ;\ \sigma = \frac{\tau_i - \tau_r}{\tau_i}, \tag{62}$$

(P = normal momentum, τ = tangential momentum),
where subscripts i, r and w denote incident particles, re-emitted particles (actual value) and re-emitted particles if complete accommodation occurs, respectively. Subscript w is thus the value when the temperature equilibrium between the molecule and the surface is reached. Strictly speaking an energy accommodation coefficient should be defined for each type of molecule energy (rotational, vibrational and translational) but the last one, being the most important, is usually the only one kept. From the kinetic theory of gases, the most probable velocity of a particle is:

$$v_m = (2RT/M)^{1/2}, \tag{63}$$

where M is the gas molecular weight. Equation (63) assumes a Maxwellian velocity distribution.

The most probable energy for a particle is:

$$E = \tfrac{1}{2} m^* v_m^2 = m^* \frac{RT}{M}.$$

The energy is thus directly proportional to the temperature, yielding the alternate relation (King-Hele 1964):

$$\alpha = \frac{T_i - T_r}{T_i - T_w}.$$

Above approximately 250 km, the temperature of the satellite, T_w, is mainly controlled by solar radiations, aerodynamic heating being negligible at these altitudes (Cook 1960). As the satellite crosses the Earth's shadow, T_w is expected to undergo large variations which depend on the absorptivity and the emissivity of its material. Variations of $\pm$ 50°C are likely to occur (Cook 1960). Once T_w (or its averaged value) is known, a knowledge of α and T_i is required before T_r can be calculated. Given T_r, the most probable velocity of re-emission, v_{mr}, can be obtained:

$$v_{mr} = (2RT_r/M)^{1/2}. \tag{64}$$

This variable is directly related to the molecular speed ratio of re-emission, denoted s_r, which is a fundamental variable in the evaluation of C_D, as we shall see.

In the homogeneous air, experimental work suggests that the thermal accommodation coefficient is very close to unity but for other gas mixtures, it can be as low as 0·1 (Ladner & Ragsdale 1964). Since it is now recognized that, above about 85 km, the mixture of the different species constituting the atmosphere starts to vary (due to diffusive separation), it is wise to expect some variations in α. The following range of uncertainty for the three parameters is investigated by Ladner & Ragsdale (1964):

$$0 < \alpha < 1{\cdot}0, \quad 0{\cdot}85 < \sigma < 1{\cdot}0, \quad \sigma' \simeq 1{\cdot}0. \tag{65}$$

Figure 3, taken from Ladner & Ragsdale (1964), shows the value of C_D for 85% accommodation of the tangential momentum component, 100% accommodation of the normal momentum component and α is taken in the range $0 < \alpha < 1{\cdot}0$. The dependence of C_D on the altitude is also shown. This dependence is very weak for the altitude range considered (150–500 km). The graph is summarized in table 6.

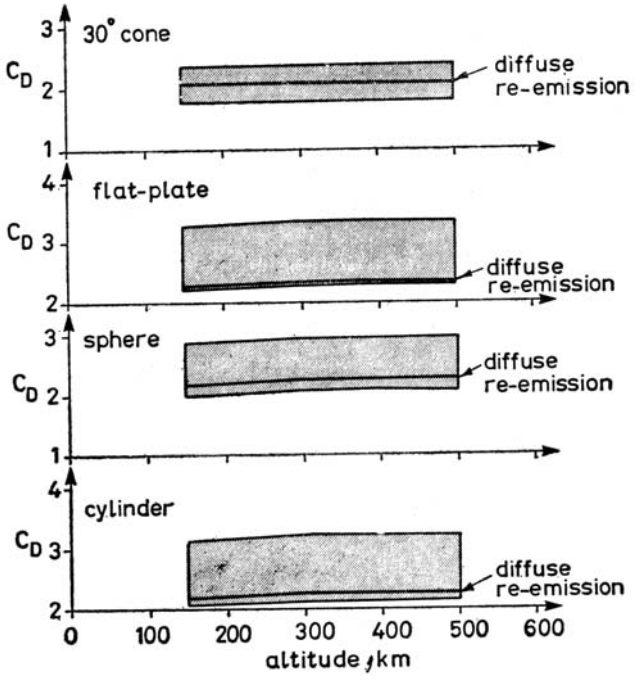

Figure 3. Drag coefficient as a function of altitude.

Table 6. Values of C_D for $0 < \alpha < 1{\cdot}0$

30° Cone	$1{\cdot}7 \lesssim C_D \lesssim 2{\cdot}4$
Flat plate	$2{\cdot}2 \lesssim C_D \lesssim 3{\cdot}4$
Sphere	$2{\cdot}0 \lesssim C_D \lesssim 3{\cdot}0$
Cylinder	$2{\cdot}0 \lesssim C_D \lesssim 3{\cdot}2$

An elementary theoretical expression for α is obtained when the incident and surface atoms are considered smooth and hard spheres (Cook 1965):

$$\alpha = \frac{4\gamma}{(1+\gamma)^2}, \quad \text{for } \gamma < 1, \tag{66}$$

where $$\gamma = \frac{\text{mass of incident gas atom}}{\text{mass of surface atom}}.$$

This classical expression is only valid when a single, head-on collision takes place. For an oblique collision, we have

$$\alpha = \frac{4\gamma}{(1+\gamma)^2}\cos^2\phi, \tag{67}$$

where ϕ is the angle between the velocity vector of the incident particle and the line joining the centres of the atoms. In many papers, α is taken as unity; even if 20% of the molecules were specularly reflected, the error in the drag coefficient is smaller than 5% for a cylinder and negligible for a sphere (Cook 1960). Furthermore, the uncertainty in α is probably negligible compared to that introduced by assuming an incorrect mode of rotation for a spinning satellite (Cook 1960), which introduces dependence of C_D upon the attitude of the body; specifically, the angle of attack, α_a, for a flat plate, or the angle between the axis and the free stream direction for a cylinder, γ_a.

Since the variable orientation of a satellite induces variations in both the reference area A and the drag coefficient C_D it is customary to adopt a constant reference value for A (A_{ref}), and allow C_D to vary for both. Williams (1972, 1975) used the following basic steps to obtain an averaged value for C_D in such a situation:

(a) Find (experimentally) the curve of C_D versus α_a (or γ_a), with $A = A_{\text{ref}} =$ constant.
(b) Fit this curve with a series of the form:

$$C_D(\alpha_a) = \sum_{j=0}^{N} b_j \cos^j \alpha_a, \quad b_j = \text{constant}.$$

(c) Knowing the attitude history of the satellite, find α_a as a function of the orbital elements.
(d) Find $\langle C_D \rangle$, the averaged value of $C_D(\alpha_a)$ over one orbital revolution.

This method makes the following assumptions (which are not always valid in the science of lifetime prediction):

(i) geometrical symmetry and simplicity of shape,
(ii) knowledge of the attitude of the spacecraft.

If one assumes an uncontrolled satellite tumbling about the axis of largest moment of inertia, the problem may be simplified. Indeed, we are back to the averaging technique due to King-Hele except that now, average values for the product $C_D A$ are obtained.

The dependence of C_D on the angle of attack is also derived by Nocilla (1972) and Cook (1960) for satellite applications. It is worth noting that Nocilla rejected the hypothesis of the accommodation coefficients since he claimed re-emission lobes obtained under these assumptions are not similar to those obtained experimentally. The following analytic expressions for C_D are taken from Cook (1960). (Subscript 1 is used to denote the value of C_D under diffuse reflection conditions while subscript 2 denotes the specular reflection case.)

(i) Flat plate, at angle of attack α_a, diffuse reflection:

$$C_{D_1} = \frac{2}{\sqrt{\pi}\, s} \exp(-s^2 \sin^2 \alpha_a) + 2 \sin \alpha_a \left(1 + \frac{1}{2s^2}\right) \operatorname{erf}(s \sin \alpha_a) + \frac{\sqrt{\pi} \sin^2 \alpha_a}{s_r}. \tag{68}$$

(ii) Flat plate, $\alpha_a = 90°$, specular reflection:

$$C_{D_2} = 4\left(1 + \frac{1}{2s^2}\right) \operatorname{erf}(s) + \frac{4}{\sqrt{\pi}\, s} \exp(-s^2). \tag{69}$$

(iii) Sphere, diffuse:

$$C_{D_1} = 2\left(1 + \frac{1}{s^2} - \frac{1}{4s^4}\right) \operatorname{erf}(s) + \frac{2s^2 + 1}{\sqrt{\pi}\, s^3} \exp(-s^2) + \frac{2\sqrt{\pi}}{3s_r}. \tag{70}$$

(iv) Sphere, specular:

$$C_{D_2} = 2\left(1 + \frac{1}{s^2} - \frac{1}{4s^4}\right) \operatorname{erf}(s) + \frac{2s^2 + 1}{\sqrt{\pi}\, s^3} \exp(-s^2). \tag{71}$$

(v) Circular cylinder, inclined at angle γ_a, diffuse reflection:

$$C_{D_1} = \frac{2rl}{A}\left\{\frac{\sqrt{\pi}}{s} \exp\left(-\frac{s^2 \sin^2 \gamma_a}{2}\right)\left[I_0\left(\frac{s^2 \sin^2 \gamma_a}{2}\right) + \left(\frac{1 + 2s^2}{2}\right) \sin^2 \gamma_a \left(I_0\left[\frac{s^2 \sin^2 \gamma_a}{2}\right] + I_1\left[\frac{s^2 \sin^2 \gamma_a}{2}\right]\right)\right] + \frac{\sqrt{\pi^3}}{4s_r} \sin^2 \gamma_a\right\} + \pi r^2 \frac{2}{\sqrt{\pi}\, s}\left[\exp(-s^2 \cos^2 \gamma_a) + \sqrt{\pi}\, s \cos \gamma_a \left(1 + \frac{1}{2s^2}\right) \operatorname{erf}(s \cos \gamma_a) + \frac{\pi s}{2s_r} \cos^2 \gamma_a\right] \tag{72}$$

where A = area perpendicular to the direction of motion, l = cylinder length, r = cylinder radius.

(vi) Circular cylinder, perpendicular to the stream, specular reflection:

$$C_{D_2} = \frac{4\sqrt{\pi}}{3s} \exp\left(-\frac{s^2}{2}\right) \left\{ I_0\left(\frac{s^2}{2}\right) + \left(\frac{1+2s^2}{2}\right)\left[I_0\left(\frac{s^2}{2}\right) + I_1\left(\frac{s^2}{2}\right)\right]\right\}$$

$$C_{D_2} \approx \frac{8}{3}\left[1 + \frac{1}{s^2}\right] \text{ for large } s \tag{73}$$

where s = molecular speed ratio, s_r = molecular speed ratio of re-emission, I_n = Bessel function of imaginary argument of order n,

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\, dt.$$

For a rod-shaped satellite (*e.g.*, Explorer I), Cook (1960) gives the fluctuations in C_D brought about by a variable angle of attack, γ_a. These fluctuations were shown to be as high as a factor of 5. Nocilla (1972) gives some graphs of C_D as a function of the angle of attack as well.

The molecular speed ratio s used in the above equations is given by:

$$s = \frac{\text{satellite speed}}{\text{most probable molecular speed}} = \frac{v}{v_m}, \tag{74}$$

where $$v_m^2 = \tfrac{2}{3} C^2 = 2RT/M, \tag{63a}$$

where C is the root mean square speed, equal to $(3RT/M)^{\frac{1}{2}}$, under the usual assumption of Maxwellian velocity distribution. From (20), v_m is also obtained by the relation:

$$v_m^2 = 2gH.$$

The molecular speed ratio for re-emitted particles is simply given by:

$$s_r = \frac{v}{v_{mr}} = s\,(T/T_r)^{\frac{1}{2}}, \tag{75}$$

where v_{mr} was given in (64). Taking v at perigee height, Cook (1960) showed that s varies between 6 and 9 where perigee altitudes between 185 and 370 km are considered, with eccentricity in the range 0 to 0·2. This was later confirmed in Cook (1965) where it is shown that $s > 5$ for perigee heights below 700 km. For a perigee height near 1500 km, $s < 4$ and the random thermal motion of the air molecules

must then be considered. Assuming s to vary between 5 and 10, values as shown in table 7 for C_D were obtained in Cook (1960). In all the cases, C_D decreases with increasing values of s. The above ranges for $C_D$are generated by $(T/T_r)^{\frac{1}{2}}$ between 1 and 2.

The influence of the satellite velocity on C_D is also treated in Bailey & Hiatt (1972); and Henderson (1976). A mathematical form used by Hunziker (1970) is the following.

$$C_D = C_{D_0} + kv_m/v,$$

where C_{D_0} = constant, k = constant parameter, v_m = air thermal speed, and v = satellite speed. Other relations between C_D and s are given in Nocilla (1972).

As a final note, it must be remembered that uncertainty is also introduced by the assumptions of a clean and smooth surface, which is not to be expected in reality. The effect of surface roughness was treated in Cook (1960) and the importance of the surface lattice structure was stressed by Ladner & Ragsdale (1964) and Cook (1965).

Despite all these uncertainties, a widely used value for satellite drag coefficient is that given by King-Hele (1964):

$$C_D = 2{\cdot}2 \pm 5\%.$$

Cook (1965) noted that between 140 and 400 km altitude during low solar activity, or 140 and 600 km altitude during high solar activity, the drag coefficient is almost independent of height and the above value seems still acceptable with an uncertainty of 10 to 15%. Above 400 km (low activity) or 600 km (high activity) the drag coefficient is likely to increase with height (probably as high as 3 (King-Hele 1978)) since the molecular weight of the atmosphere decreases and thus, both the degree of energy transfer and s decreases. At low altitudes, when transition flow occurs, C_D is likely to decrease for reasons given earlier.

Other average values for C_D used in the literature are:

$C_D = 2{\cdot}5 \pm 20\%$ (Shamberg 1959),

$C_D = 2{\cdot}12$ (Cook 1960) (average for Sputnik 3).

Table 7. Values of $C_D(s)$ from Cook (1960)

(i) Flat plane ($\alpha_a = 90°$)	
	$2{\cdot}1 < C_D < 2{\cdot}4$ (diffuse)
	$4{\cdot}0 < C_D < 4{\cdot}1$ (specular)
(ii) Sphere	
	$2{\cdot}1 < C_D > 2{\cdot}3$ (diffuse)
	$2{\cdot}0 < C_D < 2{\cdot}1$ (specular)
(iii) Infinite circular cylinder ($\gamma_a = 90°$)	
	$2{\cdot}1 < C_D < 2{\cdot}4$ (diffuse)
	$2{\cdot}7 < C_D < 2{\cdot}8$ (specular).

Modelling of the drag coefficient in free-molecular flow is still a subject of active research. In any case, an extremely precise value of C_D cannot solve a problem in which the density may vary by a factor of 10 to 15% in an unpredictable way. It needs only to be remembered that the value of C_D included in the equations may be a little off the required value and this uncertainty is carried out in the lifetime predictions (Ladner & Ragsdale 1964; Lubowe 1970; Hunziker 1970).

In this section, the three main environmental perturbations were discussed. Under realistic assumptions, variations in A were treated in terms of a constant value, A_{av}, with a few percent error. Similar average values for C_D were also given. Lift was ruled out under other meaningful assumptions. Dealing with small dense satellites minimizes the perturbing torques due to gravity gradients and solar pressure. Finally, mass loss, break-up and ablation are most likely to occur during re-entry, not during orbital decay. It is thus clear that the complexity brought about by the Specific Perturbations (table 3) can be realistically avoided.

6. Analytical approaches and decay stage

In this section, we consider the various ways of predicting the orbit contraction and lifetime of Earth satellites. But first, a short discussion on the main theoretical methods is given, as well as the definition of " decay stage ". The three main analytical approaches will then be presented.

6.1 *Theoretical methods*

The equations of motion for a satellite subject to air drag and gravitational perturbations have, in general, no closed-form exact solution. One requires some mathematical simplifying assumptions in order to get an approximate closed-form expression. The most common simplifications encountered in this survey are: (i) neglect of small order terms, (ii) orbital averaging, (iii) asymptotic expansion, (iv) variation of parameters.

The first assumption is the most widely used. For example, terms in e^n, $n > 2$ for small eccentricity orbits are usually neglected (King-Hele 1964). For near-circular orbits, v^2 is usually taken as μ/r (Mills 1959). The Earth's flattening is usually neglected above the second order: $f^3 \simeq 0$. Similarly, small variations in the orbital elements (a, e, i, Ω, ω) and, in some cases, in the density ρ are often assumed negligible over a small interval of time (usually the orbital period) (Perkins 1958; King-Hele 1964; Egorov 1971).

This leads us to the orbital averaging techniques (Uphoff 1973) which assume mean orbital elements within each revolution, varying slowly from one revolution to the next. This method is also widespread since the changes in (a, e, i, ω, Ω) due to orbital decay are slow compared to the orbital period of the satellite. The motion can thus be divided into a " fast motion " (θ) and a " slow-motion " (a, e, i, Ω, ω), along the lines of the K-B (Krylov and Bogoliubov) method of averaging. This approach fails to give oscillation within an orbit (King-Hele 1964; Santora 1975, 1976).

The asymptotic method is an extension of the K-B method by Bogoliubov and Mitropolsky. It is a powerful tool especially shaped for celestial mechanics problems

(Bogoliubov & Mitropolsky 1961). For example, Zee used it in many papers (Zee 1963, 1969, 1971b, 1972, 1973); Barry *et al* (1971) and Brofman (1967) used a modified version of it. This method gives closed-form approximate solutions for both the mean elements and their short-period oscillations within the orbit. Linearization about a reference solution can also give acceptable solutions for a small range of applicability (Perkins 1958).

The variation of parameters method (devised by Euler in the mid-eighteenth century), when applied to the equations of motion of a satellite, results in the well-known Lagrange Planetary Equations which give the rate of change of the orbital elements due to perturbing forces. The solution is then obtained after these first-order differential equations are satisfied.

An unusual technique, due to Morduchow & Volpe (1973) could be called a "reverse process": instead of using an atmospheric model and then solve the equations of motion to get the trajectory, they supposed a trajectory and derived an expression for the required atmospheric model that would result in their trial trajectory. This technique has the advantage of giving exact solutions but with the assumption of poor atmosphere models. One model, however, was found to be quite realistic.

6.2 *Decay stage*

The majority of the papers surveyed analyze the orbital decay of a satellite but not its re-entry phase. In order to clarify this assumption, definitions of decay and re-entry stages follow.

The decay stage lasts from the initial conditions until re-entry occurs. The perigee height decreases from its initial value down to about 120 to 150 km. This represents the largest portion of the lifetime. Small variations of the classical elements (a, e, i, Ω, ω) occur within one orbit. One can distinguish the "slow motion" of these variables from the "rapid motion" of the satellite around its orbit (θ). The satellite is still considered to follow a sustained orbit. As the orbit contracts, the "slow motion" becomes more and more rapid until re-entry occurs.

The re-entry stage comprises the period of time from re-entry until end-of-life. It usually lasts a few minutes. The perigee height decreases from about 120–150 km until impact or complete burn-up occurs. The classical orbital elements vary so much that the satellite can no longer be taken to be on a sustained orbit. Usually, it follows instead a ballistic trajectory (when there is no lift). The variation in the orbital elements is no longer "slow" compared to the motion around the orbit. Break-up and ablation are likely to occur during this stage.

The critical perigee height (120–150 km) at which the transition to the re-entry stage occurs will depend on the dispersion parameters and the shape of the orbit (eccentricity, argument of perigee). Very often, the border-line between the two stages is what is called the re-entry. However, it is difficult to define when re-entry actually happens. It is a state that comes into being only gradually and one may assume a short transition period between the two stages. The only exception is when re-entry is deliberately initiated.

6.3 *Analytical solutions*

The theories gathered in the literature survey tend to approach the orbital decay

prediction in three different ways. For clarity in the upcoming discussion, the following three classes will be analyzed separately: (a) osculating orbital elements approach, (b) spatial curve approach, (c) graphical approach.

6.3*a Osculating orbital elements approach.* This approach is based on the Lagrange planetary equations which give the rate of change in the orbital elements due to the perturbing forces acting on the satellite. The smallest time interval considered is the orbital period Υ, within which all fluctuations in the "slow" orbital elements are assumed to be negligible. Changes of the orbital elements are obtained by integration of these equations, when the proper perturbation models have been inserted. The lifetime (from now on denoted by L) is usually taken as one of the following conditions: (a) time required for e to become zero in elliptical orbits; (b) time required for $r_p < r_c$ where r_c is a critical perigee radius (120–150 km or r_E), (c) time required for Υ to be $\simeq$ 87 min (equivalent to $a_L \simeq$ 6510 km (King-Hele 1964)). Condition (a) is mainly used for elliptic orbits and (c) for circular orbits. Condition (b) may be used in either case.

It can be seen from these conditions that a and e as function of the number of revolutions N are the important variables. From a and e, r_p and Υ can be obtained and conditions (a), (b) or (c) apply for lifetime prediction. The elements a and e are the two variables that are directly affected by air drag; from this fact comes their importance. Other elements (Ω, ω, i) vary more slowly and are also affected by gravitational perturbations. In many theories, they are assumed constant and only planar considerations of the orbit decay follow.

The simplest case in treating decaying orbits by this approach assumes a stationary, spherically atmosphere with exponential density distribution (equation (19)) and a simple $1/r$ potential field. Luni-solar perturbations are excluded.

For small eccentricities, Parkyn (1958) obtained the average rates of change for a, e and r_p and showed that when the eccentricity is approximately zero, 5 or 6 hours were left in the satellite lifetime. His approach gave an upper limit for the lifetime which may be in error by up to 25%.

For intermediate eccentricities (when $\rho_{APO} \neq 0$ and $\rho_{APO} \neq \rho_p$), Newton (1962) uses the simple atmospheric model given by (29). Through variation of parameters, he fits an ellipse to the decaying orbit and derived the equations for e, a, and ω describing this ellipse which has the form:

$$1/r = k(1 + a\cos\theta - b\sin\theta) = k[1 + e\cos(\theta + \omega)], \tag{76}$$

where k, a and b are functions of μ, θ, h_u (unperturbed angular momentum) and e_u where subscript u denotes values for the unperturbed ellipse. The atmospheric matching coefficients, σ and λ (equation (29)) are also included in a, b and k. Holding all quantities constant except θ, Newton derives the rate of change of the eccentricity the semi-major axis and the argument of perigee of varying ellipse that best fits the decaying orbit.

$$\begin{aligned} da/d\theta &= -2\lambda a^2 - [2\sigma a(1 + e^2)/(1 - e^2)] \\ de/d\theta &= -e[\lambda a(1 - e^2) + 2\sigma] \\ d\omega/d\theta &= -[(2 + e)(\lambda h_u^2/\mu) + 2\sigma][\lambda h_u^2/\mu + 2\sigma]/e. \end{aligned} \tag{77}$$

This derivation is useful because of its simplicity. The decaying trajectory is well-defined. However, the atmospheric model is too simple to be realistic for long-range predictions. It also shows that the deacy in the semi-major axis is proportional to the density at the average altitude, a, rather than at perigee altitude. In fact, neglecting e^2 terms in the first of (77) yields:

$$da/d\theta = -2a^2(\lambda + \sigma/a) = -2a^2\rho(a). \tag{78}$$

Still with a pure $1/r$ potential field and a spherical atmospheric density model but this time, exponentially decreasing with altitude, King-Hele did an extensive work in his book, " Theory of satellite orbits in an atmosphere " (King-Hele 1964). The basic steps in this book are:

(i) Use equations for da/dt and de/dt (from Lagrange).
(ii) Express them as da/dE and de/dE, functions of a, e, E, ρ, δ.
(iii) Define $x = ae$ and find dx/dE.
(iv) Incorporate the density model (exponential form, equation (19)).
(v) Integrate over one revolution using the integral form of the Bessel functions of imaginary argument and get Δa and Δx.
(vi) Approximate $\Delta a/\Upsilon$ and $\Delta x/\Upsilon$ by da/dt and dx/dt and integrate over the lifetime of the satellite, L.
(vii) Compute $r_p = r_p(e)$, $\Upsilon = \Upsilon(e)$, $e = e(t)$, $L = L(e)$ and then $r_p = r_p(t)$, $\Upsilon = \Upsilon(t)$, $\dot{\Upsilon} = \dot{\Upsilon}(t)$, $L = L(\dot{\Upsilon})$.

Here, $$\delta = \frac{C_D A_{av} F}{m} \tag{79}$$

where F is given by (53).

Different mathematical tools were used for different situations:

(i) Small e and $3 < \beta x < 30$ (*i.e.*, $e > 0{\cdot}02$): Phase 1.
(ii) Small e and $0 < \beta x < 3$ (*i.e.*, $e < 0{\cdot}02$): Phase 2.
(iii) Circular orbits.
(iv) High eccentricity orbits: $0{\cdot}2 < e < 0{\cdot}9$.

Small e refers to eccentricity less than 0·2.

For small e, terms involving e^n, $n > 2$ or 3 are usually discarded. Phase 1 allows one to use the asymptotic expansion in the Bessel functions of imaginary argument but they must be kept unchanged in Phase 2. For high eccentricity orbits, a new variable is defined and other low-order terms are neglected.

Initially, all lifetimes are defined as the time taken for e to drop to zero (condition (a) for lifetime definition) except for circular orbits where condition (c) is used. Then, corrections are brought: in Phase 2, condition (a) implies an infinite drop in perigee height and consequently, condition (b) makes more sense, with r_i in 120–150 km or, equivalently, when $e = 0{\cdot}1\, e_0$. In Phase 1, L is the time required for Phase 2 to be reached plus the lifetime in Phase. 2. For high-eccentricity orbits, L is the time required for Phase 2 to be reached plus the lifetimes in Phase 2 and Phase 1.

Atmospheric oblateness is found to be an important perturbing factor for small eccentricity orbits in some particular orientations. High eccentricity orbits are virtually unaffected.

Variable scale-height complicates the theory but some rules have been derived by King-Hele to simplify the expressions. The expected lifetime is altered significantly when $e > 0{\cdot}2$. Subtraction of a constant coefficient can give an approximate lifetime in this case. When $e < 0{\cdot}02$, King-Hele used a " constant-H " formula where the value of H is evaluated at a height kH above or below perigee and this constant H value is substituted into the lifetime expressions derived in the constant-H theory. The parameter k takes different value for different situations.

Finally, an analysis of the variation in the orbital elements Ω, i and ω due to air drag and atmospheric rotation is included.

The various results are listed in tables 8, 9 and 10 where $Z = \beta x = a e/H$.

An advantage of King-Hele's approach is the independence of his expressions from day-to-day and yearly density variations. Indeed, since L and $\dot{\Upsilon}$ are both affected by the value $\delta\rho_{po} = C_D A_{av} F\rho_{po}/m$, expressing L as function of $\dot{\Upsilon}$ allows cancellation of this erroneous factor. The same cancellation occurs when a, r_p and Υ are expressed as functions of eccentricity. However, eccentricity versus time is dependent on $\delta\rho_{po}$. For long-lived satellites, the accuracy of this theory is not better than 10%, as King-Hele himself claims. For short-term prediction, these equations

Table 8. Spherical atmosphere (King-Hele)

$0{\cdot}02 < e_0 < 0{\cdot}2$	(Phase 1)	$L = -\dfrac{e_0 \Upsilon_0}{\dot{\Upsilon}_0}\,[F(e_0) + 0\,(H/2a_0)]$
$0 < \beta a_0 e_0 \leqslant 3$	(Phase 2)	$L = -\left(\dfrac{3e_0 \Upsilon_0}{\dot{\Upsilon}_0}\right)\dfrac{I_0(Z_0)}{I_1(Z_0)}\left[1 + 2e_0\dfrac{I_1(Z_0)}{I_0(Z_0)} - \dfrac{9e_0 Z_0}{40} + 0\,(0{\cdot}008)\right]$
$e_0 > 0{\cdot}2$		$L = -\dfrac{e_0 \Upsilon_0 F'(e_0)}{\dot{\Upsilon}_0}\left[1 + 0\left(\dfrac{4H}{5r_{po}}\right)\right]$
$e_0 = 0$		$L = \dfrac{\eta H \Upsilon_0}{2\pi\,\delta\,\rho_0\,a_0^2}$

where

$$F(e_0) = \tfrac{3}{4}\left[1 + \frac{7e_0}{6} + \frac{5e_0^2}{16} + \frac{H}{2a_0 e_0}\left(1 + \frac{11e_0}{12} + \frac{3H}{4a_0 e_0} + \frac{3H^2}{4a_0^2 e_0^2}\right) + 0\left(e_0^3, \frac{H^4}{2a_0^4 e_0^4}\right)\right.$$

$$F'(e_0) = \left[\frac{3 + e_0}{(1 + e_0)(1 - e_0)^{1/2}} - 3 - \frac{1}{\sqrt{2}}\ln\left(\frac{\sqrt{2} + (1 - e_0)^{1/2}}{(\sqrt{2} + 1)(1 + e_0)^{1/2}}\right)\right]$$

$$\times\left(\frac{3(1 - e_0)^{1/2}(1 + e_0)^2}{8e_0^2}\right)\left[1 - \frac{H(8e_0 - 3e_0^2 - 1)}{8r_{p0}\,e_0\,(1 + e_0)}\right]$$

$\eta = 1 - \exp[-\beta(a_0 - a_L)]$ $a_L \approx 6510$ km,

I_n = Bessel function of imaginary argument,

δ is given in (79).

Table 9. Oblate atmosphere (King-Hele)

Phase 1	$L \approx (L)_{\text{SPH. ATM.}} \left[1 + \frac{2f}{e_0} \sin^2 i \cos 2\bar{\omega}\right] \exp\left[c \cos^2 \omega_0 - \cos^2 \bar{\omega}\right]$
Phase 2	$L \approx (L)_{\text{SPH. ATM.}} \left[1 + c \cos 2\bar{\omega} \frac{I_2(Z_0)}{I_0(Z_0)} + 0\left(\frac{c^2}{2}\right)\right]$
$e = 0$	$L = (L)_{\text{SPH. ATM.}} \left(\frac{1}{1 + c^2/4}\right)$
Large e	$L \approx (L)_{\text{SPH. ATM.}}$

where $(\ )_{\text{SPH. ATM.}}$ = spherical atmosphere value

$\bar{\omega}$ = average value of arguments of perigee

f = flattening, (equation (32))

c = $1/2 f \beta r_{po} \sin^2 i$

Table 10. Variable scale-height.

Phase 1	$L = -\frac{3e_0 \Upsilon_0}{\dot{\Upsilon}_0}\left[1 + \frac{7e_0}{6} + \frac{H_{po}}{2a_0 e_0} - \gamma\left(\frac{7}{16} - \frac{3H_{po}}{4e_0 a_0}\right) + 0\left(\frac{5e_0^2}{16}, \frac{3H^2}{8a_0^2 e_0^2}, \gamma^2\right)\right]$
Phase 2	$L = -\left(\frac{3e_0 \Upsilon_0}{4\dot{\Upsilon}_0}\right)\frac{I_0(Z_0)}{I_1(Z_0)}\left[1 - \gamma\left\{2 + Z_0 - \frac{Z_0^2}{20} - (Z_0^2 + Z_0/2)\left(y_0 - \frac{1}{y_0}\right)\right\} + 0(\gamma^2, e)\right]$
$e = 0$	$L = \frac{H_0^* T_0 \eta}{2\pi \delta a_0^2 \rho_0}\left[1 + \gamma\left(\frac{a_0 - a_L}{H_0^*}\right)^2\left(\frac{1}{\eta} - 1\right)\right]$

where y_0 = $I_0(Z_0)/I_1(Z_0)$

γ = $dH/dr \simeq 0{\cdot}1$

*denotes values evaluated at one scale-height below initial height,

i.e., $a_0^* = a_0 - H_0^*$ and $H_0^* = H(r = a_0^*)$.

furnish a good estimate of the probable lifetime. However, assuming a $1/r$ potential field may occasionally lead to large discrepancies as pointed out by Santora (1976). Secular variations in Ω and oscillations in ω induced by the geopotential have a substantial influence on the magnitude of air drag. Furthermore, for light balloon-type satellites, King-Hele's theory needs refinement as suggested by Jupp (1976). As a last remark, King-Hele's theory is quite good when lifetime predictions are concerned but the actual trajectory is not obtained with great accuracy, mainly because it neglects oscillations within one orbit.

Later Cook and King-Hele extended this theory to include the diurnal bulge (Cook & King-Hele 1965) according to (41). The degree of severity of the day-to-night effect is shown to be proportional to F/Z, where F is given by (42) and $Z = ae/H$.

The changes in inclination of a satellite orbit due to a rotating atmosphere including day-to-night variations in air density have been analyzed by King-Hele & Walker (1970) where resonance conditions are also treated. It is demonstrated that $\Delta i/\Delta\Upsilon$ is directly proportional to Λ (equation (52)) and (sin i).

So far, gravitational perturbations have not been included in the above theories. Santora (1975) approaches this problem with a semi-analytic method. In his paper, he first finds the decay rates in Υ, e and ω for an oblate, rotating atmosphere including the diurnal density effects using equations (19), (35), (41) and (53). Auxiliary equations are derived to take into account the gravitational perturbations in perigee height and perigee argument and also for the proper selection of the scale-height H when diurnal effects are considerable. A new scale-height equation, applicable to diurnal density, follows the same variation as the density:

$$H = H_0 + S \cos \phi$$

where $$H_0 = \frac{(H_{\max} - H_{\min})}{2} = H_{\text{av}},$$

$$S = \frac{(H_{\max} - H_{\min})}{H_{\max} + H_{\min}},$$

ϕ = angular displacement from the bulge centre.

A modified Jacchia atmosphere model is used and this theory is valid for $e < 0{\cdot}2$. The "kH" (or constant-H) theory due to King-Hele is shown to be inadequate when oblateness and bulge effects are combined.

Santora's paper combines many perturbing forces together in a single mathematical development. A particular case of this unified theory considers near-circular orbits (Santora 1976). This special case is treated separately because of some particular effects (like the short-term gravitational perturbations and the different density profile involved). An exponential, oblate and diurnal atmosphere model and a (7, 0) gravity field are assumed.

Santora's work represents one of the most realistic approaches to the determination of the actual trajectory of a decaying satellite using osculating orbital elements. His equations can be directly used in lifetime predictions. Comparison with a numerically-integrated trajectory showed 2% discrepancy with his equations taking 1/50th the computer time needed for numerical integration.

Zee (1969, 1971b) presents a theory in which he derives the equations of a spatial curve in spherical coordinates and then translates his results in terms of orbital elements. Strictly speaking, this does not follow an osculating orbital elements approach and these papers will be discussed in the next section.

Finally, Lubowe (1970) contributed to the advancement of the theory by showing the importance of oblateness-drag coupling (J_2D coupling) in the orbit prediction of a satellite.

6.3*b Spatial curve approach.* Contrary to the osculating orbital elements approach, the spatial curve approach supposes approximations applied directly to the equations of motion, rather than to Lagrange's equations. Therefore, spatial curves representing the trajectory of the satellite are obtained. They usually have the form:

$$r = r(t)$$

$$r = r(\theta), \quad \theta = \theta(t),$$

or $$r = r(e), \quad e = e(\theta). \tag{80}$$

Some papers consider planar motion only and the complexity of the theory increases when three-dimensional motion is analyzed.

Assuming a spherically symmetrical atmosphere model, where density varies according to (28), Egorov (1971) has derived the time required for an orbit to decay from an initial radius r_0 down to a radius r. The velocity of the satellite is assumed to be that of a circular orbit at all times and the geopotential follows the simple $1/r$ model. He obtains

$$\hat{t} = \frac{1 - \hat{r}^{\alpha+1/2}}{(\alpha + 1/2)B}. \tag{81}$$

where $\hat{t}$ = nondimensional time $= (\mu/r_0^3)^{1/2}t$
$\hat{r}$ = nondimensional radius $= r/r_0$,
$B = C_D A \rho_0 r_0/m$,
α = defined in equation (28).

Equation (81) gives a reliable approximation under the following conditions:

— trajectory always nearly circular,
— bulge effects negligible (bulge line perpendicular to orbital plane),
— small altitude variations,
— short-term predictions.

The " layer-technique " can improve the reliability of (81). In any case, due to the great number of approximations, it is not very close to the real situation but its simplicity makes it a useful estimate when raw predictions are needed for particular orbits.

Returning to a spherically symmetrical, exponential atmosphere, Zee (1963), using the techniques of the variation of parameters, obtained the equation of an ellipse in which e and h (angular momentum) are functions of the true anomaly:

$$\frac{1}{r} = \frac{c}{h^2} + \frac{1}{K_1 h} \cos(\theta + K_2), \quad (K_1, K_2, c \text{ are constants}) \tag{82}$$

where h decreases with time (or θ). The first term is the mean motion. The second term is an oscillation about the mean path with an amplifying coefficient, $1/K_1 h$. Equation (82) shows that even an intially circular orbit does not remain strictly circular and that oscillations about the mean path increase with time. This theory is valid for low-eccentricity orbits. Extending this theory, Zee (1971b) later considered the oblateness in the mass distribution of the Earth. Using similar techniques, a spatial curve is obtained for initially low-eccentricity orbits. A similar theory is also given in Zee (1969) in which case, low radial thrust is the perturbing force.

A detailed analysis of the decay of nearly circular orbits is due to Brofman (1967). In this model the perturbing forces are simple: spherical Earth, spherical and statio-

nary atmosphere, exponential density decay and planar motion only are considered. This work is mentioned because it uses a modified Mitropolsky method, obtaining a solution of the form

$$r = \sum_{n=0}^{N} K^n r_n, \tag{83}$$

where r_n is the nth-order solution to the planar equations of motion. Both mean and oscillatory trajectories are shown in this solution.

Variational equations are used by Perkins (1958) to demonstrate that air drag makes the satellite spiral (with a sinusoidal modulation):

$$\Delta r = -K(\hat{t} - \sin \hat{t}) \quad (e_0 = 0), \tag{84}$$

where $\hat{t} = v_0 t/r_0 =$ nondimensional time; $K = g_E (C_D A/w) \rho_0 r_0^2$; w = weight. For an initially elliptic trajectory, he obtains:

$$\Delta r = A(1 - \cos \hat{t}) - Ke^{-c}\left[a(\hat{t} - \sin \hat{t}) + b/2 (\sin \hat{t} - \hat{t} \cos \hat{t})\right.$$

$$\left. + \frac{d}{2^2 + 1}\left(\sin \hat{t} - \frac{\sin 2\hat{t}}{2}\right) + \frac{f}{3^2 - 1}\left(\sin \hat{t} - \frac{\sin 3\hat{t}}{3}\right) + \ldots\right] \tag{85}$$

($A, a, b, c, d, f =$ constants).

Here, the lifetime of circular orbits is presented in the form

$$\hat{L} = \frac{\rho_0}{K} \int_{h_L}^{h_0} \frac{dh}{\rho}. \tag{86}$$

This expression is very useful since the integral can be obtained graphically if no atmosphere model is judged adequate. However, if exponential density variation is assumed, (86) reduces to:

$$\hat{L} = \frac{1}{\beta K}\{1 - \exp\left[\beta (h_L - h_0)\right]\}, \tag{87}$$

and for the power law model:

$$\hat{L} = \left(\frac{1}{1 - \eta}\right) \frac{h_0}{K}\left[1 - \left(\frac{h_L}{h_0}\right)^{1 - \eta}\right]. \tag{88}$$

Because the exponential model is not valid over the whole atmospheric range (with $\beta = 1/H = \text{constant}$), Perkins used (88) for an approximate lifetime expression ($h_L = 0$):

$$\hat{L} = \left(\frac{1}{1-\eta}\right)\frac{h_0}{K}. \tag{89}$$

Periodic oscillations in the altitude of satellites subject to drag as given by (85) were also demonstrated by Karrenberg *et al* (1962):

$$\frac{r}{r_0} \approx 1 + 2\alpha\,(S - \sin S), \tag{90}$$

$$\frac{v}{v_0} \approx 1 - \alpha\,(S - \sin S), \tag{91}$$

where α = ratio of drag to gravity acceleration, and $r_0\,S$ = arc length traversed by the satellite.

From these two equations, in-track and out-of-plane dispersion expressions due to uncertainties in the mean atmospheric density are derived.

In-track angular displacement ξ:

$$\xi = \xi_0 - \tfrac{3}{2}\,\alpha\,S^2 + 4\alpha(1 - \cos S). \tag{92}$$

In-track displacement d:

$$d = -\frac{3}{2}\,r_0\,\alpha\left(\frac{v_0\,t}{r_0}\right)^2 + 4\,\alpha\,r_0\left[1 - \cos\left(\frac{v_0\,t}{r_0}\right)\right]. \tag{93}$$

Cross-track displacement y:

$$y \approx \psi\frac{d}{r_0}, \tag{94}$$

where ψ is the angle between inertial and relative (to Earth) velocities of the satellite. In this treatment, spherical Earth, spherical and stationary atmosphere and a $1/r$ field are assumed. The density is considered constant in the range where dispersions occur. This constitutes its major weakness.

Finally, the reader is referred to Mills (1959) for an interesting discussion of the so-called 'satellite paradox' where it is shown that the effect of drag on a satellite affects its motion as if the drag force were reversed and were accelerating the satellite.

6.3*c*. *Graphical approach.* This approach usually results in a series of graphs obtained by numerical analysis as in Ladner & Ragsdale (1964) or analytical techniques, as in King-Hele (1978), for instance.

Vlasov (1967) uses a combination of 4 graphs: one for constant-L curves, one for constant-eccentricity curves, one for constant-ballistic coefficient curves and one that takes into account the yearly variations in the mean atmospheric density due to the sunspot cycle. An exponential atmosphere model is used and no gravitational perturbation is included.

Extensive work based on a graphical method is given by Ladner & Ragsdale (1964). The graphs here are obtained from numerical integration (using Runge-Kutta and Simpson's Rule) of the modified Sterne's equations. The modifications are due to variable mass and variable drag coefficient caused by attitude stabilization. The theory is mainly applicable to circular orbits and the effects of various parameters (ρ, C_D, Δm, etc.) are discussed in terms of their influence on the nominal lifetime. The bulge is modelled according to Jacchia's empirical formula (equation (40)). It is shown that maximum lifetime occurs when the bulge line (from Earth's centre to bulge centre) is perpendicular to the orbital plane and it is minimum when this line lies in the orbital plane. The maximum fraction loss in lifetime due to diurnal effects ranges from 4% (at 200 km) to 20% (at 500 km) for a circular orbit. The influence of solar activity (11 year cycle) is also analyzed and the random variations in ρ may lead to + 150% and − 60% uncertainty in lifetime. Because a fixed geometry is assumed in this paper (fixed orbital plane and fixed Sun), the predictions it yields are valid for short periods only. A detailed analysis of the drag coefficient is given and the effect of mass loss on lifetime is assessed. Uncontrolled and attitude-stabilized spacecraft are considered. The results are:

$$L = L^* \left[1 + \Delta(11 \text{ year}) - \Delta(\text{bulge}) - \Delta(\Delta m)\right] \pm U, \tag{95}$$

where L^* = nominal lifetime, $\Delta(11 \text{ year})$ = variations due to the solar cycle, $\Delta(\text{bulge})$ = variations due to the day-to-night effect, $\Delta(\Delta m)$ = variations due to mass loss, U = uncertainty due to C_D and solar flares. All these variations and L^* are given in graphs. Consequently, reading four graphs and adding a few numbers give the estimated lifetime.

In a recent paper, King-Hele (1978) describes graphical procedures for obtaining relatively accurate lifetime estimates. These graphs are based on the equations derived in King-Hele (1964) which are reported here in tables 8, 9 and 10. The effects of Earth oblateness, atmosphere flattening and varying scale-height are included in this formulation while corrections for the 11-year solar cycle, the semi-annual and the day-to-night density variations are presented in a graphical form. The equations are given in terms of the mean motion n and its rate of change $\dot{n}$. They are thus independent of the mass-to-area ratio of the satellite which is unknown for about 90% of the newly-launched satellites. Furthermore, values of n and $\dot{n}$ are readily available from observation.

The main topics treated in this paper are briefly summarized here in six subsections.

(a) Under the assumptions of a spherical Earth, a spherically symmetric and static atmosphere and an exponential density variation with altitude, the basic equation for the lifetime is:

$$L = 3\, en/4\dot{n}. \tag{96}$$

To this fundamental equation are added various factors so that a better accuracy for L is obtained. Table 8 gives such factors. Note that in this table $\Upsilon_0/\Upsilon_0 = -n/\dot{n}$.

(b) Because of the dependence of H on the altitude, correction factors are introduced in the lifetime equations. These are:

$$
\begin{aligned}
&-\gamma\left[\frac{1}{4}-\frac{1}{2Z}\right] \quad \text{for } 0{\cdot}03 < e < 0{\cdot}2 \text{ (to be added)},\\
&[1-\gamma J] \quad \text{for } e < 0{\cdot}03,\\
&[1-\gamma] \quad \text{for } e \to 0,\\
&\left[1-\frac{1}{4\gamma}\right] \quad \text{for } e > 0{\cdot}2,
\end{aligned}
\tag{97}
$$

where $\gamma = dH/dh$ and J is a function of $Z = ae/H$. Graphs of H and γ are given as functions of the altitude for low and high solar activity. For perigee height between 200 and 400 km (*i.e.*, for most lifetime predictions), γ is usually taken as 0·1. When this correction is carried out, the resulting lifetime equations (analogous to those given in table 10, apart from some revision) are accurate to about 2%, under the conditions of a spherical time-invariant atmosphere and spherical Earth.

(c) Defining a function Q such that

$$L = Q/\dot{n}, \tag{98}$$

a graph of Q versus e for $0 \leqslant e \leqslant 0{\cdot}8$ is obtained from the above equations. Separate graphs are then given for the particular cases $e \leqslant 0{\cdot}01$, $0 \leqslant e \leqslant 0{\cdot}03$, $0{\cdot}03 \leqslant e \leqslant 0{\cdot}2$ and $0{\cdot}2 \leqslant e \leqslant 0{\cdot}8$. It is interesting to note that for small-eccentricity orbits, Q is more sensitive to r_p (or n) than to e and Q is thus given versus n instead. In this range, Q is also quite affected by solar activity. For intermediate eccentricity ($0{\cdot}03 \leqslant e \leqslant 0{\cdot}2$), Q increases almost linearly with e. Finally, for $e > 0{\cdot}4$, one must be cautious in using the figures because luni-solar gravitation becomes significant.

(d) Then, corrections due to the assumption of spherical Earth are added. The perigee height is the most affected variable. From (30) we have

$$h_p = r_p - r_E,$$

where (1) r_p varies because of the odd zonal harmonics in the geopotential and luni-solar perturbations, (2) r_E varies with latitude (Earth's flattening). The amplitude of the third body perturbations being of the order of 1 km for $e < 0{\cdot}4$, these are neglected. The amplitude of the odd zonal harmonics perturbations is of the order of 10 km for most inclinations. Similar to (34), the equation for r_E is;

$$r_E = 6378{\cdot}1 - 21{\cdot}4 \sin^2 i \sin^2 \omega, \text{ in km},$$

so that $h_p - h_{p0} = 21{\cdot}4 \sin^2 i \sin^2 \omega - 10 \sin \omega$. A graph for Δh_p versus ω for different i is given. The required correction for these variations in h_p are introduced through $\dot{\bar{n}}$ where:

$$\dot{\bar{n}} = \dot{n} \exp\,[(h_p - \bar{h}_p)/H], \tag{99}$$

$\dot{n}$ = current decay rate, $\dot{\bar{n}}$ = corrected decay rate to be inserted into the lifetime equation (98), $\bar{h}_p$ = mean value of h_p, averaged over the remaining lifetime. Equation (99) follows from the direct proportionality between ^{n} and ρ_p. If ω is expected to go through many cycles, then one gets:

$$\dot{\bar{n}} = n \exp[(h_p - h_{p0} - 10{\cdot}7 \sin^2 i)/H]. \tag{100}$$

This correction to $\dot{n}$ is not always required. A graph of ω versus i gives the 'safe' region where it is not necessary.

(e) Due to the flattening of the atmosphere (f), the following correction factors are multiplied to the lifetime estimations obtained from the graphs:

$$\left[1 + \frac{f r_p I_2(Z)}{2H I_0(Z)} \sin^2 i \cos 2\bar{\omega}\right] \quad \text{for } Z < 5, \tag{101}$$

$$\left[1 + \frac{2f \sin^2 i \cos 2\omega}{e}\right] \quad \text{for } Z > 5,\ Z = ae/H, \tag{102}$$

where $\cos 2\bar{\omega}$ is the mean value of $\cos 2\omega$ in the remaining lifetime. This correction factor is usually of the order of 2%.

(f) The time-dependent variations of the density are considered under the following dispersion parameters:

— 11-year solar cycle variations,
— day-to-night variations,
— semi-annual variations,
— irregular day-to-day variations.

11-year cycle variations:

The correction is of the form:

$$L = \frac{Q}{\dot{\bar{n}}}\left[\frac{\rho_p}{\bar{\rho}_p}\right] \tag{103}$$

where ρ_p = current density at perigee, and $\bar{\rho}_p$ = mean of ρ_p over a solar cycle. The curves of $\rho = \rho(h)$ for low and high solar activity are given along with that of $\bar{\rho} = \bar{\rho}(h)$. This last one almost coincides with the night-time solar maximum curve. For lifetime greater than about 3 solar cycles, $\dot{n}$ is very small and difficult to

evaluate. An alternate method in which L is expressed as a function of m/A is given:

$$L\left(\frac{A}{m}\right) = \frac{Q}{365\, G(h_p)\, \Sigma(e)} \quad \text{in years,} \tag{104}$$

where $$G(h_p) = 10^6\, \rho_p\, \sqrt{H}\left[1 + \frac{h_p - 200}{6570}\right]^{-2\cdot5}, \tag{105}$$

and $$\Sigma(e) = 1{\cdot}362\,(1 - e)^{2\cdot5}/E(e), \tag{106}$$

and where $E(e)$ is given in figure 7.1 of King-Hele (1964). A graph of LA/m as a function of h_p is given for $0 \leqslant e \leqslant 0{\cdot}8$.

Day-to-night variations:

Similar to the above correction, (103) is used here with $\bar{\rho}_p$ now being the value of ρ_p averaged over the day-to-night cycle if the perigee is to go through these cycles (this usually takes a few months). If many cycles are to be experienced, a measurement of $\dot{n}$ at $9h$ or $20h$ local perigee time, *i.e.* when ρ_p is near the mean of the day-to-night variation, will suffice to yield adequate estimates.

Semi-annual variatious:

In this case, the lifetime estimate is multiplied by $V/\bar{V}$ where V is a factor relating the mean density to the current density at the particular time of the year when $\dot{n}$ is measured. $\bar{V}$ is its average over the remaining lifetime of the satellite. A graph of $V/\bar{V}$ versus L is given for

$$0 < L < 160 \text{ days},$$

with different curves for different days of the year. Another graph gives $V/\bar{V}$ versus the days of the year for different constant-L curves.

These graphs are subject to some uncertainty because of the yearly variations in amplitude and phase of the semi-annual effects. For lifetime of many years, one just has to measure $\dot{n}$ near 19 February, 14 May, 17 September or 30 December, where $V/\bar{V} \simeq 1$.

Irregular day-to-day variations:

These variations are largely unpredictable. In this case, it is better to compute $\dot{n}$ over 2 or 3 days. However, the possibility of unexpected magnetic storms may dramatically affect lifetime estimates.

This graphical method can be accurate to about 10% of the remaining lifetime for predictions of a year or less. For high-eccentricity orbits, *i.e.* when luni-solar perturbations are not negligible, a section of King-Hele (1978) is devoted to numerical integration programs. This section is not discussed here.

6.4 *Comparison of the various approaches*

This section will compare the three analytical approaches described above.

The osculating orbital elements approach is very well suited to lifetime predictions. The equations expressing the estimated lifetime are usually simple and easy to evaluate. Their accuracy, given the sporadic behaviour of the density, is acceptable. Using the classical elements (a, e, i, ω, Ω, M or θ) yields better physical and geometrical understanding of the effects of the various perturbations. However, this method is strictly valid for the decay stage only, where the variations of the elements a, e, i, Ω, ω are still " slow " compared to the orbital motion of the satellite. Upon re-entry, this approach technically ceases to be valid. But as long as lifetime predictions are concerned, this is not a major drawback: usually re-entry time is only of the order of a few minutes (King-Hele & Walker 1958; King-Hele *et al* 1969) and thus adds imperceptibly to the estimated lifetime.

This approach is useful for lifetime prediction because of its relative mathematical simplicity and its clear geometrical interpretation.

In the spatial curve method, the mean trajectory of the satellite along with its oscillations about this mean path are given in three-dimensional coordinates (in the general case.) Lifetime predictions as well as position estimates can be made as accurate as the density permits in both the decay and re-entry stages of the satellite life. The main drawback to this approach is its greater analytical complexity. Generally, it is not possible to obtain the position variables as explicit functions of the time, and closed-form equations (like tables 8 to 10), for the lifetime cannot be obtained. Computer assistance is generally required in this approach. Finally, because of the requirement of defining new variables, a lack of geometrical and physical understanding may be induced (*e.g.* Zee, 1971b).

Finally, graphical techniques are the easiest to use: they can be applied without any appreciable computation. Obviously, there is a tradeoff with accuracy. Reading a graph causes further uncertainties, especially when interpolation between two constant-parameter curves is required.

As a concluding remark, one can easily obtain a rough estimate of the lifetime through graphical techniques. If a better estimate is required, with little computing time, the orbital elements approach is the answer. At the expense of more computing time and the possible lack of geometrical insights, the spatial curve approach provides position and lifetime estimations at any time during the decay and re-entry phases of the spacecraft.

6.5 *Analytical methods versus other methods*

A brief comparison between analytical methods and other methods is now presented. Most theoretical investigations follow three basic steps:

Step 1. Formulation and implementation of the method (derivation).
Step 2. Utilization of the method to get results.
Step 3. Assessment of the results (accuracy).

The comparison is organized along these lines.

6.5*a Analytical methods.* Analytical methods, which are the subject of this review, are often referred to as " general perturbation methods ". Employing a combination of series expansions, approximations and analytical integration in its implementation, this method results in closed-form solutions giving position as function of time. Sometimes, the time is not an explicit variable and the solutions are given in a series of parametric equations. In this case, computer aid may be required but no numerical integration is involved. This method has the following advantages:

(i) well suited for long-term orbit decay and lifetime predictions;
(ii) little or no computing time required to get the results (efficiency);
(iii) since the solutions are obtained after step 1, they are general enough to be applied to any particular case. Many estimations are easily obtained for different satellites;
(iv) they furnish a better visualization of the effects of the various perturbations because, even before a particular satellite is considered, the expressions obtained in step 1 can already give some insight into the physical behaviour of the satellite;
(v) relatively accurate lifetime estimates are obtained at low cost.

However, some drawbacks are identified:

(i) long and tedious analytical manipulations required in step 1 to obtain the solution;
(ii) frequent necessity of using simplistic perturbation models or restricted range of application since closed-form solutions are sought.

6.5*b Numerical integration methods.* These methods fall into the category of special perturbation methods. Various numerical integration procedures (Runge-Kutta, Fehlberg, etc.) are employed and step 1 here results in a computer program. Contrary to the analytic method, the solutions are only obtained after step 2, when the various parameters in the program have been assigned their numerical value, dependent on the particular satellite and the particular orbit considered. The advantages are:

(i) simplicity in the formulation (step 1);
(ii) high accuracy for orbits of short duration;
(iii) may include more realistic and complex perturbation models with great range of applicability.

The disadvantages are:

(i) truncation and round-off errors due to the numerical nature of the method and consequently, difficulty of applications to orbits of long duration (Barry *et al* 1971);
(ii) since the solution is obtained only in step 2, this method is particular in nature: one must choose a specific satellite and a particular orbit to obtain a solution. Consequently, many computer runs and trial trajectories are re-

quired if one wants to analyze the effects of a particular dispersion parameter on the geometrical and physical aspects of the problem;

(iii) large computer time requirements.

In view of the above remarks, it can be inferred that predictions in the decay stage using numerical methods are unnecessarily expensive. However, using them when re-entry approaches would be suitable for obtaining approximate impact points.

6.5c Satellite tracking. In itself, tracking is not a method of predicting lifetime or trajectory. It must be coupled with numerical analysis. This method is useful in giving the state of a satellite and estimating some properties (C_D, mass, A, effective density, etc.). It thus provides the data that can be inserted in step 2 of the two previous methods. Tracking may be the only way out if high accuracy, especially in the re-entry phase, is required (for predicting impact points, etc.). Incorporating the possibility of tracking into semi-analytical methods may yield improved accuracy of lifetime prediction. This approach, although promising, has not received wide attention.

6.5d On the use of the various methods. From the above considerations, the following sequence in the use of these methods seems logical.

First, the analytic method, well suited for long-term predictions, would provide the approximate re-entry time of the satellite along with the probable orbital conditions (a, e, i, Ω, ω). It would be too expensive and useless to spend extensive computing time with a numerical integration method when moderately accurate estimations could be obtained at lower cost. These predictions will indicate the appropriate point where computer integration needs to be started. In fact, as the re-entry approaches, more exact evaluations can yield better accuracy on the re-entry time and on the probable impact region. The computing time would thus be restricted to the very last part of the decay stage and all of the re-entry stage.

This successive use of the two basic methods of solution just described does not eliminate their inherent drawbacks discussed in §§ 6.5a and 6.5b. Analytic methods that rely on poor perturbation models cannot be expected to yield high-accuracy re-entry conditions after the (usually) long period of time the satellite spends in its decay stage. In particular, the coupling effects of many perturbations are practically untractable if " unified " closed-form lifetime equations are sought. Not only does one usually have to ignore gravitational perturbations in analytic lifetime studies, but some important dynamic density variations must be neglected in the formulation for the equations to remain integrable. Even the rate of change of density with altitude must sometimes be assumed constant. On the other hand, numerical methods are expected to remain relatively slow, costly and hardly applicable to long-term predictions. From this situation developed new methods that successfully combine the advantages of the two basic approaches while reducing the impact of their respective drawbacks. These methods allow the inclusion of complex, " unified " perturbation models into computer-efficient numerical algorithms. Because closed-form solutions appear at various stages in the formulation, they are called semi-analytic methods. They are often based on averaging techniques and asymptotic expansions (Bogoliubov & Mitropolsky 1961). Although semi-analytic methods

are not the subject of this review, their existence is mentioned here because they are very well suited to the characteristics of long-term orbital predictions and are thus becoming popular (Santora 1975; Barry *et al* 1971; Pimm 1971; Alford & Liu 1974).

Tracking has a complementary function. It provides the initial conditions necessary to start the prediction process. It is also useful to estimate poorly known parameters of the satellite like its drag coefficient and projected area. Density profiles can be inferred from observations and fed back into the trajectory and lifetime predictor. For example, the orbital period decay rate of a satellite can be estimated by tracking and then inserted into King-Hele's equations (tables 8, 9 and 10).

7. State of the art and existing problems

From this literature survey, it can be inferred that the general behaviour of a satellite subject to the main environmental perturbations previously defined has been relatively well analyzed. The orbit size contracts at an increasing rate while the eccentricity tends to zero, atmospheric rotation changes the orientation of the orbital plane, while gravitational perturbations cause periodic oscillations in the satellite trajectory. The diurnal bulge and the atmosphere flattening have an importance which depends upon the particular geometry of the orbit. Various time-dependent density variations can have large effects on particular predictions. Finally our poor knowledge of the drag coefficient, the almost impossible task of considering all cross-section variations of an uncontrolled satellite, and the unpredictable sporadic density variations, lead to trajectory dispersions and uncertainties in lifetime predictions.

On a more general basis, any theory has its accuracy limited by:

(a) the uncertainty associated with the mathematical model of the perturbations. No matter how sophisticated and elegant these models may be, they cannot exactly account for the partly deterministic and partly random behaviour of nature;
(b) the additional approximations and assumptions made throughout the development of the theory. Without these, one would probably never obtain closed-form solutions;
(c) the fundamental ignorance of some properties of the so-called dispersion parameters, the appropriate analytic expressions giving their numerical value, and their dependence on other variables.

Before an assessment of the current state-of-the-art can be made, we define some basic criteria for assessing the various merits of the different theories. Obviously, the science of lifetime prediction would be very well developed if all the theories were accurate, simple to implement and efficient. We thus define the criteria as: (a) complexity; (b) cost of operation; (c) accuracy. Here, complexity refers to the labour and the time required in deriving and implementing the theory (in step 1). The cost of operation represents the resources (money, time, computing capacity, etc.) needed to obtain the results in step 2. Accuracy is the degree of concordance with the actual situation.

The survey found that the analytical method ranged from low to moderate accuracy, with comparatively large complexity but low cost of operation. Numerical

integration methods are accurate with low complexity but high cost of operation. In other words, what is gained in step 1 is lost in step 2. It can also be concluded that accuracy is proportional to the amount of resources spent in step 1 and/or step 2.

Santora, King-Hele, Zee and some others present relatively accurate theories with moderate complexity and low cost of operation. Still, the results are limited by the inaccuracy of simple density models. This is not always a shortcoming of the theories as such: accuracy is increased if the properties of the atmosphere are better known. Complexity may be dealt with by symbolic manipulations on a computer and the cost of operation will certainly decrease with new computer developments.

In general, lifetime estimates cannot be obtained with an accuracy better than about 10% of the remaining lifetime (King-Hele 1978). Furthermore, the approximate impact region can be known only a few hours before end-of-life, by using tracking data and numerical integration. The current problems in the science of predicting trajectory and lifetime of satellites can be summarized as follows:

(i) Difficulty in obtaining realistic models for the orbital perturbations and their associated dispersion parameters;
(ii) increasing complexity and/or cost of operation in making the models more realistic (similarly, the problem of inaccurate estimates when costs are kept at low level);
(iii) poor knowledge of the combined effects of the various dispersion parameters from a knowledge of the individual influences.

8. Conclusions

This paper has reviewed various closed-form solutions to the problem of predicting the orbital decay and lifetime of near-Earth satellites. The environment at orbital altitudes was discussed and it was concluded that, for most satellites, the gravitational and atmospheric perturbations are dominant in lifetime studies. Models for these disturbances have revealed that those due to the atmosphere are the main source of uncertainty. Density varies with a large number of deterministic and random variables and thus represents the determinant phase in perturbation modelling. Realistic models tend to be cumbersome in orbital dynamics while analytically tractable ones are usually unrealistic. These opposing tendencies have to be resolved for low-cost, computer-efficient and accurate predictions. Variations in satellite parameters (mainly its drag coefficient and projected area) cause further uncertainty. These and other unknown parameters can be estimated from observation (*e.g.* tracking). For most satellites, attitude motion and lift generation can be neglected with reasonable accuracy.

Closed-form solutions to the orbital decay problem result in numerically efficient algorithms that are usually cheap to operate. However, they must rely on simple analytic models for the drag perturbation and further approximations in solving the equations. Hence, they usually fail to accurately predict the satellite's trajectory and sometimes, the lifetime estimates are also jeopardized. On the other hand, the numerical integration of the motion equations allows state-of-the-art perturbation models but is slow, costly and restricted to predictions of short duration. Methods

of combining realistic perturbation models with powerful methods of solution (in terms of speed and computer efficiency) will thus improve the science of long-term trajectory and lifetime predictions.

Our knowledge of the gravitational and atmospheric disturbances has considerably increased from the observations of their effects on satellite orbits. Temperature and concentration measurements of atmospheric constituents and satellite-drag data have been gathered to construct accurate models for the density (Jacchia 1970, 1971 1977). Their use in orbital dynamics has recently been made attractive through computer-efficient, semi-analytic methods of solution, such as the method of averaging (Alford & Liu 1974).

In summary, it seems that the development of 'unified' semi-analytic methods along with that of automated symbolic manipulations and numerical computer algorithms will improve the state of the lifetime prediction science. The consideration of the stochastic density variations would add some degree of significance to the predictions made with existent deterministic methods of solution.

Appendix A. The hydrostatic equation.

Under the assumption of a perfect gas, we have:

$$PV = nRT, \tag{A.1}$$

where P = pressure, V = volume considered, n = number of moles in V, R = universal gas constant, and T = temperature. The number of moles n is given by:

$$n = m/M, \tag{A.2}$$

m = mass in V, M = molecular weight of the atmospheric gas. Combining (A.2) and (A.1):

$$PV = \frac{m}{M}RT \text{ or } P = \frac{RT}{M}\rho, \tag{A.3}$$

where the density is $\rho = m/V$. Defining

$$k = = RT/M \tag{A.4}$$

yields:

$$P = k\rho. \tag{A.5}$$

We suppose that the difference in pressure between height $h + dh$ and height h is due to the weight of the air in the small volume Adh acting on the cross-section area A. This weight is:

$$w = mg = \rho Vg = \rho gAdh.$$

The difference in pressure $-dP$ is thus:

$$-dP = \frac{w}{A} = \rho g dh.$$

Using (A.5), we get:

$$k d\rho = -\rho g dh - \rho dk,$$

or
$$\frac{d\rho}{\rho} = -\frac{g}{k} dh - \frac{dk}{k}. \tag{A.6}$$

This is the hydrostatic equilibrium or barometric equation. Under various assumptions, (A.6) can be integrated and solved for ρ.

This work was financially supported by the Natural Sciences and Engineering Research Council, Ottawa, Canada under a Science 1967 Scholarship (to the first author) and by the Communications Research Centre, Ottawa, Canada under Contract No. 03SU. 36100-9-9509, OSU79-00071.

The authors wish to extend their thanks to Dr R Mamen who gave much appreciated advice and also to Mrs L Leclerc who provided valuable assistance in the classification of the papers and in typing the manuscript.

List of symbols

The following notation is used throughout the text. Some symbols of temporary use are not included here.

a	semi-major axis of the orbit
A	projected area of the satellite
A_p	geomagnetic planetary amplitude
C_D	drag coefficient
d	diameter of spherical or cylindrical satellite
e	eccentricity of the orbit
E	eccentric anomaly; energy
f	density ratio; flattening (depending on subscript)
$F_{10.7}$ or $S_{10.7}$	solar flux index
h	altitude; angular momentum
H	scale-height
i	inclination of the orbital plane
K_p	geomagnetic planetary index
l	length of a cylindrical satellite
L	lifetime
m	mass (of the satellite when no subscript present)
M	molecular weight of the atmospheric gas
N	orbit number
P	atmospheric pressure

r	geocentric radius, magnitude of $\mathbf{r}$
R	universal gas constant
$\mathbf{r}$	position vector (origin at Earth's centre)
t	time
T	temperature of atmospheric gas; tropical year
v	inertial velocity of the satellite
v_m	most probable velocity of air molecules
v_{mr}	most probable velocity of re-emission of air molecules
v_R	velocity of the satellite relative to the ambient air
v_s	velocity of sound
α	accommodation coefficient (energy)
α_a	angle of attack
β	reciprocal of scale-height
γ_a	angle between cylinder axis and flow direction
δ	geocentric latitude
θ	true anomaly
λ	mean free path; longitude
Λ	ratio of atmosphere to Earth angular rates
μ	gravitational constant
Φ	Earth's potential field; fraction of the tropical year corresponding to t
ρ	atmospheric density
σ	tangential momentum accommodation coefficient; diameter of a molecule
α'	normal momentum accommodation coefficient
τ	periodic function of ϕ (equation 47)
ϕ	angular distance of the satellite from the bulge centre
ω	argument of perigee; angular rate
Ω	right ascension of the ascending node of the orbit
Λ	ratio of atmosphere to Earth angular rates
Υ	orbital period

Subscripts

A	refers to the atmosphere
APO	refers to value at apogee
av	average value
B	refers to the diurnal bulge
E	refers to Earth values
e	denotes equatorial values
p	refers to perigee values
R	relative to the atmosphere (velocity)
sa	refers to semi-annual density variations
o	initial or reference values

References

Alford R L & Liu J J 1974 The orbital decay and lifetime (LIFTIM) prediction programme, Report M-240-1278, Northrop Services Inc.

Bailey A B & Hiatt J 1972 *AIAA J.* **10** 1436

Barry B F, Pimm R S & Rowe C K 1971 Techniques of orbital decay and long-term Ephemeris prediction for satellites in earth orbit, NASA-CR-121053

Bedinger J F 1970 *J. Geophys. Res.* **75** 683

Benson R H, Fleischman E F & Hill R J 1968 *Astron. Aeronaut.* **6** 39

Billik B H 1962a *Am. Rocket Soc. J.* **32** 1926

Billik B H 1962b *Am. Rocket Soc. J.* **32** 1641

Blitzer L 1959 *Am. J. Phys.* **27** 634

Bogoliubov N N & Mitropolsky Y A 1961 *Asymptotic methods in the theory of nonlinear oscillations* (New York: Gordon & Breach) (Engl. Trans.)

Brofman W 1967 *AIAA J.* **5** 1121

Brouwer D 1959 *Astron. J.* **64** 378

Brouwer D & Hori G I 1961 *Astron. J.* **66** 193

Ching B K 1971 *J. Geophys. Res.* **76** 197

Cook G E 1960 The aerodynamic drag of near earth satellites, Min. of Aviation, Aeron. Res. Council Current Paper No. 523

Cook G E 1965 *Planet. Space Sci.* **13** 929

Cook G E 1967 *Planet Space Sci.* **15** 627

Cook G E 1970 *Planet. Space Sci.* **18** 387

Cook G E 1971 *Planet. Space Sci.* **20** 473

Cook G E & King-Hele D G 1965 *Philos. Trans. R. Soc. (London)* **A259** 33

Croopnick S R 1972 *AIAA J.* **10** 861

Drago V J & Edgecombe D S 1974 A review of NASA orbital decay reentry debris hazard, BMI-NLVP-TM-74-1 (Battelle C.L.)

Eberst R D 1978 *J. Br. Interplanet. Soc.* **31** 196

Egorov D Ya 1971 *Cos. Res.*, May, p. 923

Escobal P R 1968 *Methods of astrodynamics* (New York: John Wiley)

Groves G V 1958 *Nature (London)* **182** 1533

Guttman P T 1965 *AIAA J.* **3** 330

Halliday D & Resnick R 1962 *Physics—Part II* (New York: John Wiley) p. 587

Harris I & Priester W 1962 Theoretical models for the solar-cycle variation of the upper atmosphere, NASA TN D-1444

Henderson C B 1976 *AIAA J.* **14** 707

Hunziker R R 1970 *Astronaut. Acta.* **15** 161

Jacchia L G 1960 *J. Geophys. Res.* **65** 2275

Jacchia L G 1965 *Smithsonian Contr. to Astrophys.* **8** 215

Jacchia L G 1970 New static models of the thermosphere and exosphere with empirical temperature profiles, *Smithsonian Astro. Obs. Spl. Report* 313

Jacchia L G 1971a *J. Geophys. Res.* **76** 4602

Jacchia L G 1971b Revised static models of the thermosphere and exosphere with empirical temperature profiles, *Smithsonian Astro. Obs. Spl. Report.* 332

Jacchia L G 1977 Thermospheric temperature, density and composition: New models, Smithsonian Astro. Obs. Spl. Report 375

Jacchia L G & Slowey J 1967 *Space research VII* (Amsterdam: North Holland) p. 1077

Jastrow R & Pearse C A 1957 *J. Geophys. Res.* **62** 413

Jupp A H 1976 *Celes. Mech.* **14** 335

Kallmann-Bijl H *et al* 1961 *COSPAR International reference atmosphere (CIRA)* (Amsterdam: North Holland)

Karrenberg H K, Levin E & Lewis D H 1962 *ARS J.* **32** 576

King-Hele D G 1959 *Nature (London)* **184** 1267

King-Hele D G 1962 *Proc. R. Soc. London* **A267** 541

King-Hele D G 1964 *Theory of satellite orbits in an atmosphere* (London: Butterworths)

King-Hele D G 1976 *Planet. Space Sci.* **24** 1
King-Hele D G 1971 Measurements of upper-atmosphere rotational speed from changes in satellite orbits, RAE Tech. Rep. No. 71171
King-Hele D G 1978 *J. Br. Interplanet. Soc.* **31** 181
King-Hele D G & Cook G E 1963 *Proc. R. Soc. (London)* **A275** 357
King-Hele D G, Cook G E & Rees J M 1963 *Geophys. J.* **8** 119
King-Hele D G & Quinn E 1966 *Planet. Space Sci.* **14** 1023
King-Hele D G & Scott D W 1966 *Planet. Space Sci.* **14** 1339
King-Hele D G & Scott D W 1967 *Nature (London)* **213** 1110
King-Hele D G, Scott D W & Walker D M C 1969 Upper-atmosphere rotational speed and its variation with height, RAE, Tech. Rep. No. 69263
King-Hele D G, Scott D W & Walker D M C 1970 *Planet. Space Sci.* **18** 1433
King-Hele D G & Walker D M C 1958 *Nature (London)* **182** 426
King-Hele D G & Walker D M C 1960a *Nature (London)* **185** 727
King-Hele D G & Walker D M C 1960b *Nature (London)* **186** 928
King-Hele D G & Walker D M C 1970 The change in satellite orbital inclination caused by a rotating atmosphere with day-to-night density variation, RAE Tech. Rep. 70208 (BR222999)
King-Hele D G & Walker D M C 1971a *Planet. Space Sci.* **19** 297
King-Hele D G & Walker D M C 1971b *Planet. Space Sci.* **19** 1637
King-Hele D G, Walker D M C & Neirinck P E L 1969 The decay of Cosmos 253 rocket over England, Tech. Memo Space 119
Kosai Y 1959 *Astron. J.* **64** 367
Kosai Y 1961a *Astron. J.* **66** 8
Kosai Y 1961b Tesseral harmonics of the potential of the earth as derived from satellite motion Smithsonian Inst. Astrophys. Obs. Spl. Report No. 72
Ladner J E & Ragsdale G C 1964 Earth orbital satellite lifetime, NASA TN-D-1995
Lee V A 1962 *ARS J.* **32** 102
Liu J J F 1974 *AIAA J.* **12** 1511
Lorell J & Liu A 1971 Method of averages expansion for artificial satellite applications, NASA CR-118267
Lubowe A G 1970 *Astronaut. Acta* **15** 143
Martin H A & Priester W 1960 *Nature (London)* **185** 600
Mills B D Jr 1959 *Am. J. Phys.* **27** 115
Morduchow M & Volpe G 1973 *AIAA J.* **11** 381
Newton G P 1970 *J. Geophys. Res.* **75** 5510
Newton R R 1962 *ARS J.* **32** 770
Newton R R, Hopfield H S & Kline R C 1961 *Nature (London)* **190** 617
Nocilla S 1972 *Astronaut. Acta* **17** 245
Nonweiler T 1958 *Nature (London)* **182** 468
Parkyn D G 1958 *Am. J. Phys.* **26** 436
Perkins F M 1958 *Astronaut. Acta.* **4**, Fasc. 2, p. 113
Pimm R S 1971 Long-term orbital trajectory determination by superposition of gravity and drag perturbations, AAS Paper No. 71-376
Rowell L N, Smith M C & Sibley W L 1962 On the motion of Echo I-type satellites, NASA RM 3185
Santora F A 1975 *AIAA J.* **13** 1212
Santora F A 1976 *AIAA J.* **14** 1196
Sarychev V A 1962 *ARS J. (Suppl.)* **32** 834
Schuchardt K G H & Blum P W 1976 *Space Res. XVI* (Berlin: Akademie-Verlag)
Shamberg R 1959 A new analytic representation of surface interaction for hyperthermal free molecule flow with application to neutral-particle drag estimates of satellites, ASTIA Doc. No. AD2153 01
Smith R J 1978 *Science* **200** 28
U.S. Standard Atmosphere (USSA) 1976 NOAA, NASA, USAF (Washington: US Govt. Printing Office)
Uphoff C 1973 *AIAA J.* **11** 1512
Velez C E & Fuchs A J 1975 *AIAA J.* **13** 12

Vlasov Yu S 1967 *Cosmic Res.* **5** 173
Volland H, Wulf-Mathies C & Prister W 1972 *J. Atmos. Terr. Phys.* **34** 1053
Walker D M C 1972a Air density at heights near 200 km from analysis of the orbit of 1969-20B, RAE Tech. Rep. No. 72071
Walker D M C 1972b *Planet. Space Sci.* **20** 2165
Walker D M C 1978 *Planet. Space Sci.* **26** 291
Willey R E & Pisacane V L 1974 *J. Astron. Sci.* **21** 230
Williams R R 1972 *J. Spacecr.* **9** 565
Williams R R 1975 *J. Spacecr.* **12** 74
Zee, Chong-Hung 1963 in *AIAA Astrod. Conf.* (Connecticut: Yale Univ.) p. 155
Zee, Chong-Hung 1969 *Astronaut. Acta.* **14** 289
Zee, Chong-Hung 1971a *Astronaut. Acta.* **16** 143
Zee, Chong-Hung 1971b *Celes. Mech.* **3** 148
Zee, Chong-Hung 1972 *Astronaut. Acta* **17** 891
Zee, Chong-Hung 1973 *Astronaut. Acta.* **18** 281